

Д. Бартоломью



СТОХАСТИЧЕСКИЕ  
МОДЕЛИ  
СОЦИАЛЬНЫХ  
ПРОЦЕССОВ

# STOCHASTIC MODELS FOR SOCIAL PROCESSES

3rd Edition

D.J. Bartholomew

Professor of Statistics  
London School of Economics and  
Political Science

JOHN WILEY & SONS  
Chichester·New York·Brisbane·Toronto

Д.Бартоломью

**СТОХАСТИЧЕСКИЕ МОДЕЛИ  
СОЦИАЛЬНЫХ ПРОЦЕССОВ**

Перевод с английского В.Б.ЕЗЕРОВА,  
А.Л.КОНИНА  
Под редакцией О.В.СТАРОВЕРОВА



Москва "Финансы и статистика" 1985

ВЫШЛИ ИЗ ПЕЧАТИ

1. Ли Ц., Джадж Д., Зельнер А. Оценивание параметров марковских моделей по агрегированным временным рядам.
2. Райфа Г., Шлейфер Р. Прикладная теория статистических решений.
3. Клейнен Дж. Статистические методы в имитационном моделировании. Вып. 1 и 2.
4. Бард Й. Нелинейное оценивание параметров.
5. Болч Б. У., Хуань К. Д. Многомерные статистические методы для экономки.
6. Иберла К. Факторный анализ.
7. Зельнер А. Байесовские методы в эконометрии.
8. Хейс Д. Причинный анализ в статистических исследованиях.
9. Пуарье Д. Эконометрия структурных изменений.
10. Драймз Ф. Распределенные лаги.
11. Мостеллер Ф., Тьюки Дж. Анализ данных и регрессия. Вып. 1 и 2.
12. Бикел П., Доксам К. Математическая статистика. Вып. 1 и 2.
13. Лимер Э. Статистический анализ неэкспериментальных данных.
14. Песаран М., Слейтер Л. Динамическая регрессия: теория и алгоритмы.
15. Дидэ Э., Боши С., Бросье Ж. и др. Методы анализа данных.

ГОТОВИТСЯ К ПЕЧАТИ

Дрейпер Н., Смит Г. Прикладной регрессионный анализ. Вып. 1 и 2.

*Редколлегия:* А. Г. Аганбегян, Ю. П. Адлер, Ю. Н. Благовещенский, А. Я. Боярский, Э. Б. Ершов, Т. В. Рябушкин, Е. М. Чегыркин

Б  $\frac{0702000000-022}{010(01)-85}$  84-85

© 1982 John Wiley & Sons Ltd.

© Перевод на русский язык, предисловие, предметный указатель, «Финансы и статистика», 1985

Книга английского ученого Д. Бартоломью «Стохастические модели социальных процессов» выдержала три издания на английском языке. Она служит за рубежом одним из основных пособий для тех, кто занимается моделированием социальной мобильности. Под социальной мобильностью понимаются изменения индивидуумами своего социального места в структуре групп, а группы представляют собой определенную совокупность людей, объединенных по каким-либо признакам. В книге Д. Бартоломью нет привычных для нас четких понятий общества, классов и т. д., даже понятием социальной группы автор иногда охватывает, например, некоторые категории больных. Такой «расплывчатой» трактовкой общества или класса объясняется применение автором аппарата марковских процессов (а в книге используется в основном он) к очень широкому кругу задач, возникающих в социальной, экономической и даже в медицинской статистике.

Социальная мобильность связана, как известно, с законами развития общества и обусловлена в значительной степени личной деятельностью людей, в том числе и личным фактором производства. Определенный уровень развития производительных сил, а человек — основная производительная сила, приводит к тому, что даже «классы оставляют всегда совершенно свободным переход отдельных личностей из одного класса в другой»\*, не говоря уже о различных социальных прослойках и группах. При социализме же, после ликвидации не только массовой безработицы, но и ее последствий, социальная мобильность населения значительно возросла. В настоящее время этому способствуют также и технический прогресс, и развитие новых методов труда, и повышение уровня образования, и овладение новыми профессиями. Поэтому соответственно растет потребность в прогнозировании подвижности людей и управлении ею.

Актуальность проблем, связанных с возрастанием роли личного фактора производства, отразилась в решениях XXV и XXVI съездов КПСС. Для развития народного хозяйства СССР в 1976—1980 гг. было необходимо «повысить эффективность использования трудовых ресурсов», «... сократить удельный вес ручного труда»\*\*. За период 1969—1980 гг. доля работающих вручную (не при машинах) снизилась с 37,4 до 32,8%, но абсолютная численность их даже возросла\*\*\*. По-

\* Ленин В. И. Полн. собр. соч., т. 2, с. 477.

\*\* Материалы XXV съезда КПСС.— М.: Политиздат, 1976, с. 167—168.

\*\*\* Социалистический труд, 1981, № 9, с. 85.

этому на XXVI съезде КПСС вопрос был поставлен в несколько иной плоскости: «Обеспечить наиболее рациональное использование... трудовых ... ресурсов ... Принять меры к достижению сбалансированности имеющихся и создаваемых рабочих мест с трудовыми ресурсами»\*. Таким образом, повышение качества прогнозов трудовых ресурсов — одна из важнейших народнохозяйственных проблем.

До шестидесятых годов в СССР вопросы расширения производства решались в основном за счет привлечения дополнительной рабочей силы. На современном этапе производственные фонды могут быть использованы лишь при обеспеченности кадрами. По этой причине экономисты вопрос ставят по-другому: неправомерно говорить о нехватке, дефиците кадров, необходимо сосредоточить внимание на избытке определенного рода производственных фондов. Следовательно, движение кадров играет основополагающую роль в развитии экономики.

В предлагаемой советскому читателю книге Д. Бартоломью очень много внимания уделено иерархическим системам, которые описывают движение кадров по квалификационным ступеням. Модели, рассмотренные автором, характерны для описания отдельного предприятия за короткий период времени и лишь в ограниченной степени для описания движения населения отдельного региона или страны за длительный период. По этой причине Д. Бартоломью исследует, как правило, открытые системы, а закрытые служат ему лишь в качестве первоначальной ступени, изучать которую часто бывает гораздо проще. Математически такая последовательность изложения материала вполне понятна, так как решение любых — разностных или дифференциальных — уравнений начинается с однородных уравнений. В предлагаемой книге рассматриваются стохастические модели, т.е. модели, в которых необходимо учитывать распределения вероятностей. А динамика распределения вероятностей численностей групп в марковских моделях описывается именно разностными или дифференциальными уравнениями.

Марковские модели привлекают пристальное внимание исследователей по очевидной причине: они являются наиболее простыми среди моделей, учитывающих взаимосвязи различных социальных слоев или групп. Кроме этой причины имеется и другая: демографические методы расчета численностей возрастных групп представляют собой частный и простейший случай марковской цепи. Но этот частный случай дает такое качество результатов, на которое до сих пор ориентируются все другие исследователи населения, считая его своего рода эталоном. Однако более внимательное изучение марковских, в том числе и демографических, моделей привело к целому ряду явлений, требующих объяснения.

Один из основных парадоксов применения марковских моделей к изучению социальной мобильности заключается в том, что, несмотря на вполне удовлетворительную с практической точки зрения точность предсказания численностей групп, дисперсии или среднеквадратические отклонения этих численностей были настолько малы, что

---

\* Материалы XXVI съезда КПСС. — М.: Политиздат, 1981, с. 141.

отклонения реальных данных от расчетных невозможно было объяснить лишь случайностью процессов. Первый обратил на это внимание, по-видимому, Дж. Поллард. Он вычислил на основе марковской модели среднеквадратические отклонения численностей возрастных групп и сравнил предсказанные численности с действительными. Такое сравнение привело к целой серии работ по объяснению столь больших отклонений, но никак не к отказу от марковских моделей.

Д. Бартоломью рассматривает и другое, обнаруженное в движении кадров, расхождение статистических данных с результатами моделирования. Он предлагает несколько методов его объяснения, не выходящих за рамки марковских моделей. Расхождение состоит в том, что межгрупповые потоки, прогнозируемые на некоторое время вперед, несколько меньше, чем следовало ожидать, исходя из анализа моделей. Автор монографии приводит некоторые причины такой ситуации. Грубо говоря, они объясняются тем, что группы плохо определены и не являются однородными по движению их членов. Например, МС-модели объединяют в одну группу и тех людей, которые обязательно двигаются (М — мобильные, «кочевники»), и тех, которые обязательно остаются (С — стоящие, «оседлые»). Но такое объединение приводит к изменению доли мобильных в группе, а поэтому и к неоднородности по времени переходных матриц. Таким образом, предлагаемый подход выводит за рамки однородных марковских моделей, хотя даже в однородных марковских моделях, если по-разному учитывать идвигающихся между группами и остающихся в них, отмеченного противоречия, возможно, не будет.

В монографии Д. Бартоломью приводится еще одна причина отклонения статистических данных от результатов расчета на основе марковской модели. Она состоит в том, что реальные сообщества людей подвержены целому ряду управляющих воздействий, которые безусловно отражаются в статистических данных, но не отражены в исходной марковской модели до тех пор, пока не начинается исследование возможностей регулирования.

Любое управление, в особенности в области экономического регулирования, изменяет численности групп людей и переходы, потоки между ними. Изменение потоков входит как в оценки переходных матриц, так и в результаты расчетов без учета управления. Следовательно, до тех пор пока управляющие воздействия постоянны, модель будет отражать реальную ситуацию. Но стоит управлению измениться, статистик, следящий лишь за переходами людей, сразу же скажет, что изменились потоки, т.е. ему будет казаться, что параметры модели, отражающей действительность, зависят от времени. Можно, конечно, полагать, что статистик должен знать об управлении все, однако реальность гораздо сложнее.

Для подтверждения этой мысли рассмотрим такой пример. Пусть статистик находится на каком-либо предприятии в районе или в области. Он безусловно будет знать об изменении условий на этом предприятии, но он может не знать об изменении внешних условий. А внешние условия оказывают существенное влияние на подвижность людей вообще и рабочей силы в частности. В свою очередь подвижность,

т.е. потоки людей, служит основой для определения параметров моделей. Таким образом, создается впечатление, что существует зависимость параметров подвижности от времени, т.е. кажется, что есть необходимость в неоднородных марковских моделях движения населения и трудовых ресурсов.

Может быть высказано предположение, что отмеченные только что проблемы будут устранены, станут разрешимыми, коль скоро будет известна цель управления. Но, к сожалению, это не совсем так. Представим себе, что имеются две стратегии управления (хотя на самом деле их гораздо больше), например, какие-либо из тех, которые приведены в предлагаемой читателю монографии. Пусть, далее, стратегии управления выбирают на основе одной и той же соответствующей действительности информации или модели два человека (или две организации) независимо друг от друга. Результат действия двух выбранных стратегий управления может приводить к совершенно другой целевой структуре групп, чем та, которую следовало получить. Поэтому возникает задача о согласовании методов управления, которая пока не нашла своего решения. Важность ее решения вполне очевидна, так как несоответствие результатов управления его цели может создать ложное убеждение или в неправомерности метода моделирования, или в самой модели, или в неверной (недостаточной) информации.

В предлагаемой книге Д. Бартоломью нашло свое отражение решение многих вопросов, возникающих в практике статистиков, занимающихся проблемами движения рабочей силы. Противоречий и неясностей в представлениях о мобильности и в статистике подвижности было вполне достаточно для того, чтобы создавать все новые и новые схемы мобильности. Но хотя «схемы сами по себе ничего доказывать не могут; они могут только *иллюстрировать* процесс, если его *отдельные элементы выяснены теоретически*»\*, Д. Бартоломью приводит довольно много моделей, полученных в результате формализации схем движения людей и на основе статистики подвижности. Как следует из отдельных замечаний автора, само множество моделей, рассматриваемых в книге, вероятнее всего появилось следующим образом. Данные о мобильности людей порождали схемы движения, а схемы приводили к математическим моделям. При исследовании моделей получали некоторые следствия, которые пытались проверять на новых данных о подвижности. Получающиеся отклонения осмысливались экономистами и социологами. Ими же давалось определенное теоретическое объяснение новых, основанных, как правило, на статистике фактов, что приводило к другому взгляду на исследуемое явление (мобильность), к схеме, отличной от прежней. Далее появлялись новые модели, соответствующие другой схеме, и все повторялось с самого начала.

Из сказанного следует, что модели вообще и те, которые приведены в монографии Д. Бартоломью, не подменяют собой теоретического, экономического или социального анализа, но в большой степени способствуют его развитию, в частности формированию представления о необходимости дифференцированно подходить к людям при исследова-

---

\* Ленин В. И. Полн. собр. соч., т. 4, с. 52.



нии их подвижности, а также других более общих характеристик движения населения. Мы надеемся, что перевод монографии Д. Бартоломью даст определенный толчок, возможно, в новых направлениях, теоретическому анализу социальной мобильности населения.

Книга Д. Бартоломью предназначена, безусловно, не математикам, которые, кстати, найдут в ней новую область приложений, а специалистам в области исследования подвижности населения, овладевшим, конечно, основами математических методов в объеме программы экономических высших учебных заведений. Эти специалисты найдут в книге много интересных, а иногда и неожиданных подходов к задачам измерения текучести кадров, им откроются, очень вероятно, новые аспекты измерения подвижности.

Действительно, однозначного ответа на стоящий перед практикой вопрос — хорошо или плохо, когда мобильность населения высокая, — нет. Слишком низкая мобильность или полное устранение ее, чего хотели бы некоторые руководители предприятий, ведет к исчезновению связей между различными группами населения, к исключению обмена опытом и т.д., т.е. к уменьшению влияния личного фактора производства, в то время как вещественный его фактор все настойчивее требует именно высокой мобильности населения. Но слишком большая подвижность населения приводит подчас к неоправданно большим экономическим потерям, к снижению производительности труда.

Работа Д. Бартоломью содержит не просто обзор многих моделей подвижности кадров, применимых в основном для уровня объединения или отдельного предприятия. Все описанные в ней модели разобраны и аргументированы. Они вполне пригодны к «непосредственному» применению при соответствующей статистической информации. Более того, из изложения материала становится ясным, какая именно информация необходима и как нужно ее обрабатывать. Таким образом, знакомство с книгой будет способствовать устранению дублирования в работе, убедит в необходимости создания основы моделей. Внимательный читатель после выбора подходящей модели (а они довольно разнообразны и по сложности, и по адаптации) сможет сразу же приступить к ее доработке, что, как правило, бывает необходимо.

В монографии помимо моделей исследуется несколько стратегий управления кадрами в иерархических системах. Основная задача управления кадрами состоит в сохранении необходимых пропорций между численностями разных групп по квалификации, профессии и т.д. Однако при этом должна быть учтена подвижность самих кадров. Вопросы достижимости целевых, т.е. необходимых для нормального производства, структур и задачи их поддержания рассмотрены автором весьма детально.

Кроме отмеченных достоинств, монография Д. Бартоломью имеет еще одно — она поможет советскому читателю лучше ориентироваться в основных направлениях исследований социальной мобильности, проводимых за рубежом. Книга может служить также своеобразным справочником: она содержит ссылки почти на 260 статей и монографий более чем 260 авторов, ведущих исследования в разных направлениях. Конечно, не все направления и глубина их разработки равноценны,

но в любом случае они будут полезны для специалистов, изучающих как моделирование движения населения, так и непосредственно движение населения.

Как уже было сказано, основной математический аппарат книги — дифференциальные или разностные уравнения, в которых и отражаются свойства марковских моделей. Такие уравнения применяются для описания динамики многих явлений природы, в совершенно не связанных с движением населения областях. Поэтому при чтении монографии Д. Бартоломью еще раз убеждаешься в том, что «единство природы обнаруживается в «поразительной аналогичности» дифференциальных уравнений, относящихся к разным областям явлений»\*. И это сказывается и в терминологии. Автор употребляет термин «социальный» (как это будет понятно при чтении книги) применительно и к социальным явлениям, и к явлениям, далеким от социальных, например, таким, которые отражаются в медицинской статистике при изучении новых методов лечения. Автор приложил много сил, чтобы в монографию не проникли формулировки из других отраслей знания, в которых используются те же дифференциальные уравнения (т.е. те же вероятностные модели), но это ему не везде удалось. Так, в книгу проникли термины: социальная физика, надежность, опасность отказов (из технологических наук), теория восстановления, устойчивость, поглощающие состояния (из математики) и т.п. Читатель, конечно, поймет, что, хотя материал книги имеет непосредственное отношение к обществу, термин «устойчивость» ни в какой мере не относится к устойчивости общества, а его следует понимать как устойчивость в том смысле, в каком его понимал А. А. Ляпунов при исследовании решений дифференциальных уравнений. Мы постарались сделать все, чтобы терминология была понятной для советского читателя. Так, термин «класс» везде заменен на социальную группу или просто группу, нигде не употребляется термин «страта», а только «слой» или «градация». В иерархических системах мы старались использовать лишь термины «ступени» или «градации» вместо «группа» или «слой». Если кадровая система (фирма, предприятие) функционирует в условиях, когда потоки людей из нее или в нее оказывают существенное влияние на ее состав или структуру и какие-либо из этих потоков нельзя контролировать, то в книге обычно говорится о случайных внешних воздействиях или случайных средах, т.е. употребляются термины, взятые отнюдь не из теории социальной мобильности, а, скорее, из теории автоматического регулирования.

И все-таки «инородные» термины оставлены в переводе книги по двум причинам. Во-первых, не для всех терминов нашлись русские эквиваленты, которые бы полностью соответствовали значениям, встречающимся в других отраслях знаний, но относились бы к социальной мобильности. Во-вторых, не хотелось до конца терять уже сложившиеся связи математических методов, применяемых для анализа социальной мобильности, с такими отраслями знаний, с которыми эти связи представляются наиболее плодотворными по существу, хотя форма проникновения одного в другое оставляет желать терми-

---

\* Ленини В. И. Полн. собр. соч., т. 18, с. 306.

нологически много лучшего. Некоторые термины в переводе не унифицированы сознательно: они должны стать более привычными для специалистов в области моделирования социальной мобильности. Например, более привычным, на наш взгляд, представляется термин «предельное поведение структуры», хотя он ничуть не уступает термину «предельный режим». Поэтому в книге можно встретить и то и другое.

В переводе книги можно встретить также термины «уровень», «ступень», а иногда и «градация», когда речь идет об упорядоченных группах в иерархических системах. Они представляют собой синонимы, несмотря на то, что слово «градация» употребляется и при описании моделей систем, в которых упорядочить группы не всегда возможно. В книге не всегда структура кадров — это относительные численности групп. Иногда под структурой понимается непосредственное число членов в каждой группе. Хотя в советской научной литературе термин «структура» относится к первому его толкованию, мы сочли возможным оставить этот термин в том смысле, в каком он понимается автором, потому что из текста ясно, о чем идет речь.

Для того чтобы в известной мере компенсировать отсутствие привычного и терминологически качественного описания предлагаемых Д. Бартоломью моделей, в конце книги приведен краткий предметный указатель. Из него становится понятным, что для пояснения некоторых процессов иногда удобно иметь один и тот же термин. Поясним это соображение примером из наиболее близкой к социальной мобильности области знаний — демографии.

Если рассматривать все население, то его увеличение идет за счет рождений. Число родившихся за год получают посредством умножения общего числа жевщин в начале года на коэффициент рождаемости, который иногда называют просто рождаемостью. Если же население разбито по возрасту и полу, то общее число родившихся за год получается после суммирования произведений числа женщин определенного возраста в начале года на повозрастные коэффициенты рождаемости; последние коэффициенты имеют несколько синонимов, употребляемых, правда, неодинаково часто. Мы имеем в виду следующие термины: плодовитость или функция плодовитости, фертильность или функция фертильности.

Почти аналогично обстоит дело в отношении смертности. Для единства терминов мы стали употреблять термин «интенсивность», который имеет следующий смысл: среднее число изменений, приходящихся на одного члена группы за единицу времени. Когда все население представляет собой одну большую группу и рассматриваются изменения, происходящие в результате смерти, то можно говорить об интенсивности смерти. Если выделены группы по возрасту, то можно говорить о повозрастных интенсивностях смерти. Вообще для произвольных, а не только возрастных групп можно говорить об интенсивности смерти в группе. Примеры исследований, где используются довольно произвольные группы, есть. Это исследования профессиональной заболеваемости и смертности в медицинской статистике. В предлагаемой книге интенсивности больше всего относятся к увольнениям или выходам из

различных групп иерархических систем и систем общего вида. Так, в монографии появляются интенсивности уходов в зависимости от квалификации (при группировке по квалификации) или от стажа работы.

Читая книгу Д. Бартоломью, не следует думать, что основные вопросы моделирования социальной мобильности решены. Многие из них только начинают возникать по мере углубления в изучение как моделирования, так и самой социальной мобильности. Предлагаемые вниманию читателя модели носят описательный характер. В них автор не делает даже попыток объяснить, из-за чего люди «двигаются». Это считается очевидным и не обсуждается. В действительности движение объясняется различием условий в разных группах. Если условия меняются, то меняются и потоки людей между группами, следовательно, меняются параметры модели — интенсивности переходов, которые должны зависеть от факторов. Задача о зависимости интенсивностей переходов от такого фактора, как побудительные мотивы перемены места, до конца не решена.

Д. Бартоломью снабдил книгу большим количеством числовых примеров. Все примеры хорошо подобраны и иллюстрируют теоретические результаты, полученные автором. Однако не следует надеяться, что благодаря примерам применение моделей и результатов становится элементарным. Ведь все рассматриваемые в книге модели имеют очень малую размерность, и в реальных ситуациях адаптация их отнюдь не проста. Большая размерность приводит к своим трудностям.

Наконец, следует отметить, что управление структурой кадров посредством найма весьма эффективно для отдельного предприятия. Но в наших условиях, условиях полной занятости населения, этот метод не представляется целесообразным для любого предприятия. Найм на одно предприятие — это увольнение с другого. Изучение вопросов перераспределения рабочей силы не вошло в книгу, такой задачи у автора не было.

В русский перевод монографии Д. Бартоломью не вошли две последние главы. Вопросы, обсуждаемые в них, не имеют непосредственного отношения к социальной мобильности.

В заключение хотелось бы поблагодарить автора книги, приславшего список опечаток, замеченных в оригинале. Они были исправлены в переводе.

Книга Д. Бартоломью будет полезна для специалистов широкого круга. Она окажет помощь экономистам, изучающим вопросы, связанные с трудовыми ресурсами, статистикам целого ряда направлений и в первую очередь тем, кто занимается статистикой труда, а также специалистам в области учета рабочей силы. Предлагаемая монография даст возможность полнее и глубже изучать одну из кардинальных проблем развития экономики — проблему трудовых ресурсов.

*Доктор экономических наук  
О. В. Староверов*

Хотя цель третьего издания книги осталась прежней, форма и содержание ее существенно отличаются от предшествующих. Второе издание было значительно расширено по сравнению с первым за счет включения довольно большого количества нового материала, охватывающего те же самые вопросы. Если бы мы придерживались того же принципа и при подготовке третьего издания, то в результате получили бы не только огромную по объему книгу, но и несбалансированное изложение вопросов, связанных с современным состоянием работ в данной области. При переработке второго издания были поставлены следующие цели:

- а) дать более широкий и сбалансированный обзор современного состояния предмета, которому посвящена книга;
- б) сделать книгу более удобной для использования ее в качестве учебного пособия для аспирантов;
- в) дать по возможности исчерпывающую библиографию.

Читатель, знакомый с предыдущими изданиями работы, найдет здесь для себя много нового. Числовые примеры и их решения остались без изменений всюду, где это было возможно. Для удобства учащихся большая часть исторических сведений, пояснений к математическим источникам, комментариев к литературе помещена в конце почти каждой главы в разделе «Дополнения». С самого начала мне хотелось включить в книгу и упражнения. Они неизбежно были бы целенаправленными и подчеркивали бы дедуктивные аспекты моделирования. Многолетний опыт преподавательской работы убедил меня в том, что для аспирантов и особенно для выпускников математических факультетов наиболее трудной является область построения моделей. Тонким искусством построения и применения моделей в целях проникновения в суть процессов реального мира, вероятно, можно овладеть в полной мере только при длительной работе с ними. Тем не менее ясно, что формальное математическое образование часто мешает тем, кто начинает заниматься прикладными задачами, и может оказаться даже опасным, если разработка темы, в которой участвуют учащиеся, наталкивается на трудности. Их следует научить применять разнообразные гипотезы, смотреть на проблему под различными углами зрения, уравнивать реализм и простоту и т.п. Вдумчивый преподаватель найдет в тексте и приложениях много способов для этого. Однако есть еще много нерешенных вопросов в обосновании практической эффективности программ решений уравнений модели.

В настоящее издание по сравнению со вторым внесены следующие изменения:

Гл. 1 написана заново.

Гл. 2 — расширенный вариант той же главы из предыдущего издания. Сюда вошел новый материал об измерении социальной мобильности. Большую часть бывшего раздела 2.3, посвященную профессиональной мобильности, можно найти теперь в гл. 4.

Гл. 3 в основном осталась без изменений, кроме изложения нового подхода к исследованию предельного поведения.

Гл. 4 — новая глава о замкнутых марковских моделях с непрерывным временем, включающая некоторый материал из гл. 2 и 5 предыдущего издания книги.

Гл. 5 тоже новая, однако она содержит много прежнего материала по открытым марковским моделям с непрерывным временем.

Гл. 6 посвящена теории управления для марковских моделей. Она существенно переработана. Большая часть старого материала о достижимости была исключена и заменена новым материалом по стохастическому управлению.

Гл. 7 — сокращенный вариант гл. 6 из предыдущего издания книги. В ней рассматриваются вопросы о длительностях и добавлен материал о распределениях размеров.

Гл. 8 основана на старых гл. 7 и 8. Большую часть строгой теории мы исключили и основное внимание уделили простым моделям, которые можно исчерпывающим образом проанализировать. Добавлен новый материал о цепях вакансий.

Поскольку теперь у специалистов существует несколько трактовок вопросов, связанных с планированием кадров, меньше места мы отвели приложениям в этой области. Это позволило рассмотреть более широкий спектр приложений.

*Лондон*

*Д. Дж. Бартоломью*

## 1.1. ФИЗИКА СОЦИАЛЬНЫХ ЯВЛЕНИЙ

В «Письмах о теории вероятностей и ее применении к этическим и политическим наукам», написанным в начале девятнадцатого века, Адольф Кетле отличал статистику от того, что он называл «социальной физикой». По его мнению, статистика занимается представлением более или менее достоверной картины относительно определенной социальной группы в некоторый конкретный момент времени. Изучение же законов общества, не зависящих от времени и «капризов людей», следует рассматривать как отдельный раздел «социальной физики». В своей «Социальной физике» Кетле установил некоторые из этих законов и их эмпирическую основу. Физика и статистика сделали огромные успехи за 150 лет, прошедшие с тех пор, и немногие статистики сего дня пожелали бы принять такую ограниченную точку зрения. Тем не менее Кетле был прав, признавая аналогию между физическими и социальными системами и пытаясь найти пути применения теории вероятностей к последним. Его искренняя убежденность в том, что общество подчиняется таким же жестким законам, как те, которые «управляют» небесными светилами, сейчас звучит наивно. Кендалл (1961) дает более реалистичную оценку. Но главная догадка Кетле остается правильной. Непредсказуемость на индивидуальном уровне в сочетании с большими числами дает регулярности, описываемые законами для газов и таблицами жизни. Понимание того, что существуют закономерности и характеристики агрегированного поведения человеческого общества, было первым шагом на пути к стохастическому моделированию социальных процессов.

Аналогия с физикой, конечно, неполная и потенциально может привести к ошибкам. Даже большие человеческие популяции мало похожи на большое число молекул воздуха, например, в велосипедном насосе. Это означает, что в социальных системах случайность не бывает полностью сглажена, а структура может быть лишь неясно различима. Кроме того, поведение людей не является совершенно непредсказуемым: люди реагируют на изменения в окружающей их среде и на взаимоотношения друг с другом. Для большинства человеческих популяций характерна значительная неоднородность, которая может приводить к различному характеру поведения в разных субпопуляциях. Лишь немногие даже из хорошо изученных закономерностей в обществе можно считать независящими от времени, хотя они могут существовать достаточно долго, оставаясь незамеченными. В отличие от физика социолог редко имеет возможность поставить эксперимент для определения

роли каждой переменной. Наблюдение за системой может влиять на то, что наблюдается. Социолог обычно довольствуется пассивным наблюдением непрерывных изменений социальных систем, и то, что можно узнать, соответственно ограничено. Если можно построить модель, воспроизводящую главные качественные особенности реальной системы, то социолог вполне удовлетворен. Если предсказания, сделанные с помощью такой модели, подтверждаются опытом, то есть основания для еще большего удовлетворения. Основная цель — сделать понятными сложные характеристики, присущие реальным социальным системам, показав, насколько они объяснимы в рамках нескольких постулатов об индивидуальном поведении.

Время и случай — две основные черты социальных процессов, указывающие на теорию стохастических процессов как на подходящий аппарат для их изучения. Такие системы развиваются во времени при меняющемся и отчасти непредсказуемом поведении их элементов. В теории это формализуется определением некоторого множества состояний, которому система может принадлежать, и вероятностными законами, управляющими движением между этими состояниями. Анализ формального процесса в данном случае позволяет нам применить понятие вероятности к возможным будущим состояниям и исследовать возможное поведение. Подобные процессы, широко распространенные в природе, послужили мощным стимулом для развития этой теории. Но лишь сравнительно недавно изучение социальных процессов стало основной областью ее применения.

Чтобы осознать потенциальные возможности идей Кетле, нужно было дождаться разработки современной теории стохастических процессов. В психологии первый пример стохастической теории обучения был приведен Бушем и Мостеллером (1955). Фундаментальная работа Коулмана (1964b) по применению математических методов в социологии показала, как относительно простые стохастические модели могут использоваться для объяснения специфических проблем. В экономике Стейндр (1965) использовал теорию стохастических процессов, чтобы показать, как некоторые простые, но характерные особенности развития фирм могут быть объяснены с помощью стохастических моделей роста. Когда в 1967 г. вышло первое издание настоящей книги, уже было ясно, что существует огромный потенциал использования быстро развивающейся теории стохастических процессов при анализе социальных явлений. Однако в то время серьезные разработки имелись лишь в некоторых областях, и эти исследования были весьма неравноценными по охвату и глубине. Большой объем и более широкий диапазон данного издания в некоторой степени являются отражением значительного прогресса, достигнутого в изучении этого вопроса более чем за десятилетие.

## 1.2. МОДЕЛИ И МОДЕЛИРОВАНИЕ

В большинстве отраслей прикладной математики мы закладываем основу теоретического исследования, конструируя математические модели. Читателя-нематематика мы познакомим с тем, как инженеры



используют действующие физические модели. Например, данные о поведении самолета выводятся на основании поведения его модели при имитации условий полета. Точность этих выводов будет зависеть от того, насколько точно модель воплощает в себе особенности реального самолета. Самолет и его модель отличаются друг от друга во многих отношениях. Они могут иметь различные размеры, могут быть сделаны из различных материалов; многие конструктивные элементы, не влияющие непосредственно на поведение в полете, могут быть исключены из модели. Важное условие для модели заключается не в том, чтобы она была точно похожа на реальный объект, а в том, чтобы она вела себя в аэродинамической трубе как самолет в полете. Таким образом, основное требование состоит в том, чтобы самолет и модель были *сходны* во всех *подходящих* отношениях.

Математическая модель используется совершенно аналогичным образом. В социальных системах, которые мы будем обсуждать, составные части взаимосвязаны. Когда меняется одна характеристика, возникают изменения в других частях системы. Если эти изменения можно оценить количественно, то такие взаимосвязи могут быть описаны — по крайней мере в своей основе — математическими уравнениями. Совокупность уравнений, предназначенная для описания поведения системы, представляет собой математическую модель. Другими словами, такая модель есть совокупность предположений о соотношениях между частями системы. О ее адекватности судят по тому, как она может предсказать последствия изменений в социальной системе, которую она описывает, и по тому, может ли она объяснить изменения, происходившие в прошлом. Модель есть абстракция реального мира, в которой интересующие нас связи между реальными элементами заменены аналогичными связями между математическими величинами.

Математические модели могут быть детерминированными или стохастическими. Если эффект любого изменения в системе может быть предсказан с определенностью, то говорят, что система является детерминированной. На практике, особенно в социальных науках, это не так. Во всяком предсказании обычно существует элемент неопределенности либо из-за того, что система не полностью определена, либо из-за непредсказуемого характера поведения человека. С этой неопределенностью можно справиться, если в модель вместо математических переменных ввести распределения вероятностей. Точнее, это означает, что уравнения модели должны включать случайные переменные. Такая модель является стохастической.

В необходимости использования стохастических моделей в данной книге можно убедиться при рассмотрении некоторых явлений, которые мы будем обсуждать. Например, в случае социальной мобильности невозможно предсказать с определенностью, пойдет ли сын по стопам своего отца. Точно так же не существует надежных средств, чтобы определить, когда человек решит сменить место работы или сдаст ли студент предстоящий экзамен. Именно внутренняя неопределенность, присущая свободе выбора, которой располагает индивидум, вынуждает нас строить наши модели в стохастической форме. Как мы увидим позже, в применении детерминированных методов для аппроксимации

стохастических моделей есть некоторое преимущество, но это не влияет на основной характер процесса. Среди социологов часто возникают споры о конкурирующих достоинствах детерминированных и стохастических методов. Из предшествующего обсуждения должно быть ясно, что любая модель, описывающая поведение человека, должна быть сформулирована в стохастической форме. Когда дело доходит до выводов из моделей, полезно воспользоваться детерминированной аппроксимацией. Большая простота детерминированного варианта модели может облегчить понимание природы изучаемого явления. Тем не менее все это тактические вопросы, которые не влияют на главный принцип, положенный в основу моделирования.

Разнообразие стохастических моделей, обсуждаемых в этой книге, несколько уже, чем предполагает данное выше определение. Мы будем рассматривать здесь то, что можно назвать «объясняющими» моделями в отличие от моделей типа «черный ящик». Такое деление несколько произвольно, и было бы трудно точно определить каждый из этих типов. Тем не менее для практических целей было бы полезно сделать такое различие. Модель типа «черный ящик» рассматривает соотношение между выходом системы и ее входом. Она является стохастической, если выход не может быть точно предсказан по входу. В этом случае модель обычно имеет форму уравнения (или уравнений) регрессии, связывающего выходные переменные с входными. Форма уравнения регрессии чисто эмпирическая и выбирается из простого класса функций (например, линейных) так, чтобы он наиболее точно соответствовал выходу. С другой стороны, объясняющая модель предназначена для описания механизма, с помощью которого вход системы преобразуется в ее выход. Модель типа «черный ящик» может быть достаточной, если наша цель — предсказать выход или управлять им, манипулируя со входом. Если мы хотим понять, как система работает, то нам нужна модель, объясняющая, что происходит внутри «черного ящика».

Для иллюстрации этого положения рассмотрим прием на работу и увольнение в некоторой фирме. В течение длительного периода времени мы, вероятно, наблюдали бы соотношение между рядами цифр найма на работу и увольнения. Оценка этого соотношения была бы моделью «черного ящика», и ее можно было бы использовать для предсказания найма на работу в будущем по прошлым данным об увольнении и найме. Объясняющая модель была бы попыткой описать перемещение кадров внутри фирмы и, в частности, зависимость увольнения и найма от таких вещей, как квалификационный уровень и стаж работы. Из последующих разделов станет понятно, что нашим целям в настоящей книге лучше служат объясняющие модели. (Мы сознаем, что то, что является объяснением на одном уровне, может быть недостаточно ясным при рассмотрении на более глубоком уровне. Тем не менее цель наших моделей — помочь нам проникнуть глубже в суть явлений.) Всесторонний анализ того, что мы назвали моделированием «черного ящика», содержится в работе Бокса и Дженкинса (1971).

Стохастические модели социальных явлений разрабатывались ранее с различными целями. Здесь полезно будет выделить четыре главные функции моделей. Первая — дать возможность *заглянуть в суть*

изучаемых явлений и *лучше понять их*. Это свойственно деятельности ученого. Исследование начинается со сбора данных о процессе и формулировки модели, которая воплощает в себе наблюдаемые особенности системы. Назовем это стадией *построения модели*. Следующий шаг — использовать модель для выработки таких предсказаний о поведении системы, которые можно было бы проверить с помощью наблюдений. Эта работа потребует математических исследований для получения результатов и будет называться стадией *выводов из модели*. Заключительный шаг — сравнить сделанные выводы с реальностью и модифицировать исходную модель, если она окажется неадекватной. Эта процедура называется стадией *проверки модели*. Исследование такого рода не заканчивается до тех пор, пока развитие системы не будет точно и всесторонне описано в математической форме и не будут получены решения соответствующих уравнений.

Вторая и третья цели построения модели относятся к сфере деятельности социолога-прикладника. Широко признанным здесь является применение модели для *прогнозирования*. Специалист по социальному планированию хочет знать, что произойдет, если будет проводиться определенная политика. Руководитель хочет знать заранее последствия различного рода политики найма на работу и продвижения кадров по службе в его фирме. С помощью модели, адекватно описывающей поведение системы, можно получить ответы на подобные вопросы. Важно подчеркнуть, что никакой прогноз не является безусловным; мы всегда прогнозируем, что произойдет, *если* будут выполнены определенные условия.

С проблемой прогнозирования тесно связан вопрос планирования социальной системы и характера ее функционирования. Некоторые модели иерархических организаций, описанные в гл. 3, 5 и 8, впервые были разработаны для того, чтобы установить принципы формирования политики найма на работу и продвижения по службе. Структура управления фирмой должна обеспечивать необходимое число лиц с нужной квалификацией и опытом на каждом уровне для эффективного выполнения функции организации. Достижение этой цели зависит от кадровой политики фирмы. Следовательно, необходимо оценить конкурирующие стратегии с точки зрения их соответствия поставленным целям.

Искусство построения модели состоит в том, чтобы знать, когда и где сделать соответствующие упрощения. Цель же состоит в том, как отмечал Коулмэн (1964b), чтобы сконцентрировать всю реальность, насколько это возможно, в простой модели. Это не только помогает понять систему в целом, но и дает основу для сопоставления поведения реальных систем. Замечания, сделанные выше, не следует понимать как оправдание безграничного упрощения. Они относятся к стадии построения модели процесса, а эта стадия не может рассматриваться изолированно от аспектов получения решения. В прошлом был сильный соблазн привести модель к форме, имеющей изящное математическое решение. В связи с широким распространением быстродействующих вычислительных машин ситуация существенно изменилась, но привычка мыслить прежними категориями осталась. Для получения

нужной информации о процессе больше нет необходимости в решениях, имеющих замкнутую форму, в виде простых функций.

Дальнейший вклад, который стохастическое моделирование может внести в социальные исследования, лежит в области измерений. Это утверждение может показаться парадоксальным в связи с высказанным ранее замечанием о том, что отсутствие подходящих измерений сдерживало развитие стохастических моделей. Как покажут следующие примеры, оба утверждения справедливы, но на различных уровнях сложности. В ряде прикладных областей научные работники убедились, что грубые измерения, которыми они пользуются, непригодны. Например, успех лечения рака часто достигается благодаря вычислению доли пациентов, выживших за данный период времени, предположим, за пять лет. Однако на этот показатель влияют не относящиеся к делу факторы, такие, как смертность по другим причинам, которая в идеальном случае должна быть исключена из рассмотрения. Фикс и Нейман (1951) показали, как это можно сделать с помощью построения стохастической модели послелечебного периода и оценки одного из его параметров (см. гл. 5). Другой пример относится к изучению текучести рабочей силы. Широко применяемый метод измерения текучести заключается в том, чтобы выразить число людей, уходящих с фирмы за единицу времени, в процентах от средней численности кадров за тот же период. Большие значения этого показателя часто трактуются как свидетельство неблагоприятного нравственного климата. Но такой показатель не является адекватным, так как он не учитывает сильную зависимость склонности к уходу от стажа работы. А эта зависимость такова, что фирма с большим числом вновь нанятых на работу будет иметь обманчиво высокий показатель текучести. При построении модели теории восстановления в гл. 8 оказалось возможным продемонстрировать ограниченность обычного показателя и определить условия, при которых его целесообразно применять. Результаты этого исследования выявили преимущества показателя текучести кадров, основанного на среднем значении или медиане стажа работы.

Одно из самых важных практических преимуществ в исследовании, когда располагают рядом стохастических моделей, заключается в том, что они значительно облегчают статистический анализ социальной информации. Это особенно полезно, когда (что бывает часто) такая информация фрагментарна. Модель помогает записать функцию правдоподобия и таким образом оценить интересующие нас параметры. Статистические методы, приемлемые для стохастических моделей в общественных науках, разработаны сравнительно слабо. В течение некоторого времени в нашем распоряжении были методы марковских цепей, и первая попытка частично охватить отдельные вопросы была предпринята Бартоломью (1977а). Подробный и доступный обзор методов, применявшихся в планировании кадров, сделан Бартоломью и Форбесом (1979).

Исследования других аспектов разбросаны по литературе, но всестороннее обсуждение их параллельно с чтением данной книги было бы своевременным вкладом в развитие темы.

### 1.3. НЕКОТОРЫЕ ОСНОВНЫЕ ВОПРОСЫ

Как уже было сказано, аналогия с естественными науками не должна довлеть над ними. Теперь более подробно обсудим проблемы, возникающие при применении теории стохастических процессов к социальным явлениям. Отдельные из них носят чисто практический характер, другие более серьезны и затрагивают вопросы о допустимости или обоснованности предпринятых действий.

На практике наблюдалась нехватка ученых, работающих в области общественных наук и имеющих математическую подготовку, необходимую для того, чтобы внести существенный вклад в исследования. Крускал (1970) и Розенбаум (1971) нашли пути решения этой проблемы десятилетие тому назад. Однако до тех пор, пока математика для специалиста в области общественных наук не станет столь же важной, как и для инженера или химика, скорость прогресса будет ограничена. Опыт показывает, что для выпускника гуманитарного факультета приобрести знания по математике и статистике гораздо труднее, чем для выпускника технического учебного заведения овладеть основами экономики и социологии. В следующем разделе приводятся некоторые рекомендации, которые помогут читателю-нематематику извлечь определенную пользу при работе с данной книгой, но это не более чем первая помощь.

Иногда считают, что при таком большом количестве нерешенных проблем, связанных с измерениями, преждевременно, если не безрассудно, заниматься такими тонкими вопросами, как стохастическое моделирование. В физике и технике пользуются сравнительно малым числом основных шкал измерений — время, масса, электрический заряд и т.п., и существуют приборы, позволяющие измерять эти величины с большой точностью. В общественных же науках мало естественных шкал измерений и нет полного согласия относительно их применения. Проблема в общественных науках, конечно, состоит не в недостатке переменных для измерений, а в их многозначности и разнообразии. Как отобрать их, скомбинировать, суммировать — вот в чем истинные трудности. Какой бы теоретический вес этот аргумент не имел, здесь это не так уж важно. Беглый просмотр следующих страниц книги покажет, что уровень измерений, необходимый для многих наших моделей, весьма элементарен и часто состоит не более чем из простой классификации по некоторым признакам. Кроме того, как будет показано при дальнейшем обсуждении измерения социальной мобильности и текучести рабочей силы, модель является необходимой предпосылкой для определения подходящей шкалы измерения, а не наоборот.

Еще одно возражение проистекает из сложности социальных явлений. Его часто выражают высказываниями о том, что социальные ситуации слишком сложны, чтобы применять к ним математику, и что игнорирование этого факта неизбежно приведет к опасному упрощению. Исходную посылку этого возражения необходимо принять. Социальные явления часто крайне сложны, и наши модели неизбежно являются упрощениями. Даже если можно допустить, что модель нужна не только для воспроизведения интересующих нас особенностей реаль-

ного процесса, решение данной проблемы остается исключительно сложным. Однако мы утверждаем, что для упрощения нет альтернативы. Основной ограничивающий фактор не математический аппарат, которым мы располагаем, а способность человеческого разума осмыслить сложную ситуацию. При построении моделей нет такого положения, последствия которого были бы выше нашего понимания. Вероятно, единственная мера предосторожности против чрезмерного упрощения состоит в том, чтобы использовать комплекс моделей, а не одну. Каждая отдельная модель будет специальным случаем более сложной модели, которая потребует для достижения наилучшего приближения к реальности. Выводы, общие для нескольких специальных случаев, заслуживают большего доверия, чем выводы для случаев, относящихся к одной произвольно выбранной модели. Часто стратегия, продиктованная такими соображениями, будет очевидной. Вместо того, чтобы делать «реалистичные» предположения и мириться с присущей им сложностью, мы делаем два или более «нереалистичных», но простых предположения. Где это возможно, они выбираются как крайности, между которыми должна лежать истинная ситуация. Если выводы по конкретному вопросу сходны, можно применить их к реальной ситуации, не выходя за рамки простой модели.

Наличие вычислительных машин ведет к третьему, более тонкому возражению против модельного подхода к предмету исследования. Сходная ситуация наблюдается в теории очередей, где на решение частных моделей были затрачены большие усилия. Утверждают, что с практической точки зрения результаты могут быть получены гораздо проще и быстрее с помощью имитационного моделирования на вычислительной машине. Такой же аргумент используется и применительно к общественным наукам. Если наша цель просто состоит в получении быстрого ответа в специфической ситуации, то имитационное моделирование, как правило, оказывается наилучшим методом. В прикладных общественных науках это норма, и специалист-практик в этой области найдет, что его интересы соотносятся с теми аспектами данной книги, где рассматривается построение моделей. С другой стороны, если при стохастическом моделировании нас больше интересует понимание сущности социальных явлений, то имитация на вычислительной машине менее удобна. По сути дела, в данном случае мы ищем общие решения, тогда как имитация дает нам решения лишь для частных случаев. Лаконичность и ясность простой формулы таковы, что усилия, затраченные на решение проблемы, полностью оправдываются, даже если бы мы могли довольствоваться приближением.

Необходимо отметить два различных смысла, в которых в настоящее время употребляется термин «имитационное моделирование». Выше мы говорили об имитационном моделировании индивидуального поведения. Для стохастической модели это включает определение каждого отдельного изменения состояний согласно данной реализации случайных событий, как это предписано вероятностными характеристиками модели. С помощью такого приема мы генерируем искусственные данные на индивидуальной основе именно такого типа, который можно наблюдать на практике. Некоторые авторы термин «имитация» исполь-

зуют также применительно к вычислению ожидаемых значений случайных переменных в данной модели. Такие расчеты говорят нам о том, как вела бы себя в среднем модель, если бы она действовала, и в этом смысле можно говорить об имитации процессов. Подобные вычисления не дают никакой информации о стохастическом поведении системы.

При изучении социальных явлений методы имитационного моделирования, особенно последнего типа, широко применялись Форрестером и его коллегами в Массачусетском технологическом институте. Их работы по промышленной динамике, динамике развития городов и динамике глобального развития хорошо известны. Одна из основных целей этих работ состояла в том, чтобы лучше понять поведение очень сложных систем. Наш подход представляет собой дополнение к их подходу, хотя и менее претенциозен в том смысле, что он направлен на получение тех же результатов с помощью математического анализа.

Главное возражение против широкого применения стохастического моделирования особенно поддерживалось социологами и экономистами, которые считают, что многие применяемые модели лишены реального социального и экономического содержания. Например, в экономике они хотели бы найти модель, объясняющую межотраслевое распределение в форме «закона пропорционального действия». Этот закон несостоятелен, так как он не выполняется для тех экономических показателей, которые, как считают экономисты, влияют на размер отрасли. «Слишком многое из того, что относится к этому вопросу», — утверждают они, — «свалено в одну кучу в стохастической компоненте модели». Это равносильно отказу от поиска правильного объяснения, каким образом экономические факторы взаимодействуют в реальном мире. В этом аргументе, конечно, есть смысл, но он до некоторой степени приводит к неправильному пониманию целей стохастического моделирования. Это очень похоже на физика, объясняющего, что в элементарных книгах по теории вероятностей при анализе экспериментов с бросанием монеты не учитывается турбулентность или вязкость воздуха, момент вращения, сообщенный монете, или характер поверхности, на которую она падает. Все эти факторы имеют отношение к эксперименту в том смысле, что они влияют на результат. Однако характер их взаимодействия и влияния на результат таковы, что ни один из них не имеет предсказуемого значения. Следовательно, распределение результатов можно описать с помощью одного параметра — вероятности выпадения орла, — в некотором смысле суммирующего совместное действие всех причинных факторов. Если заранее известно, что значение переменной может быть предсказано, то эта переменная обязательно должна быть представлена в модели таким образом, чтобы ее действие могло быть определено количественно. Другими словами, переменная должна быть включена в ту часть модели, которая содержит «ошибки», если очевидно, что действие этой переменной не может быть освобождено от влияния многих других факторов, известных или неизвестных. Рост фирмы несомненно зависит от нормы процента, иностранной конкуренции, уровня безработицы, валютного курса и множества других факторов, но действие каждого из них в отдельности часто невозможно определить. Однако с точки зрения описания агрегированного поведе-

ния совсем не надо добиваться, чтобы такие переменные были представлены в модели в явной форме. Это не исключает, конечно, изучения раздельного действия таких переменных на индивидуальном уровне. Указанные два подхода взаимно дополняют друг друга.

Это положение можно рассмотреть несколько иначе. Хорошо известно, что хаотическое поведение на одном уровне дает упорядоченное поведение на агрегированном уровне. Последнее создает основу для применения теории вероятностей. В равной степени справедливо и то, что действие детерминированных процессов может привести к явно случайным эффектам. Хорошим примером является генерирование псевдослучайных чисел в вычислительной машине. Последовательность цифр полностью определена математической формулой, но чтобы установить, откуда взялась эта формула, потребовались бы довольно длинные ряды цифр. Для практических целей годятся относительно короткие ряды, но они содержат пренебрежимо малое количество информации о законе их генерирования. Аналогично изучать аспекты человеческого или организационного поведения как случайные процессы — это не то же самое, что сказать — не существует причин, определяющих это поведение. Мы просто признаем, что наши данные слишком ограничены, чтобы можно было надеяться выражать количественно индивидуальные эффекты всех причин.

#### 1.4. ПРЕДПОСЫЛКИ, ТЕРМИНОЛОГИЯ И ОБОЗНАЧЕНИЯ

##### Предпосылки

Многие специалисты в области общественных наук сочтут, что математический уровень этой книги для них несколько высок. Имея это в виду, мы стремились сделать изложение простым, а аргументы неформальными. Следовательно, нужно было тщательно продумать способ изложения идей о природе моделей и их применении, чтобы они были ясны при чтении текста между формулами. Однако, чтобы добиться действительного понимания предмета, нет альтернативы тому, чтобы поработать над математикой. По возможности мы стремились к тому, чтобы для усвоения материала было достаточно этой книги, и отдельные дополнительные разделы ее послужат введением к последующему изучению математики и теории вероятностей. Полезным руководством по математическим и статистическим методам является «Справочник по прикладной математике» (Ледерман, 1980), на который в дальнейшем мы будем ссылаться. Одно из руководств к основным томам (Бартоломью, 1981) специально предназначено для специалистов по общественным наукам и охватывает в общих чертах большую часть материала данной книги. Читатель-нематематик должен признать, что язык математика очень лаконичен, и, следовательно, надо быть готовым затратить определенное время, чтобы уяснить смысл простой формулы. Мы настоятельно советуем просчитывать числовые примеры и вычерчивать графики до тех пор, пока смысл формул не будет полностью ясен. В некоторых случаях это сделано в тексте, но в целях экономии места мы, как



правило, излагали только наиболее важные результаты анализа. Для того, кто хочет полностью разобраться в моделях, большой помощью было бы умение программировать на языках Бэйсик, Фортран, Алгол.

## Терминология

Изложение материала в книге развивается от простого к сложному. Это относится, вообще говоря, как к математической теории, так и к социальным процессам. Некоторые используемые термины могут оказаться незнакомыми или привести к путанице из-за различного толкования, которое они имеют в различных дисциплинах.

*Состояние* системы является фундаментальным понятием в теории стохастических процессов. В общественных науках этот термин может применяться как к индивидууму, так и к системе в целом. Мы будем говорить о чем-либо «находящееся в состоянии  $i$ », имея в виду, что это принадлежит к классу, категории или группе, обозначенной через  $i$ . В применении к системе этот термин относится обычно к числу индивидуумов в различных категориях. Правильный выбор множества состояний является ключевым в искусстве построения моделей.

Социальные системы мы будем подразделять на *замкнутые* и *открытые*. Оказывается, что в социологии эти термины имеют не однозначное толкование. Здесь они используются в прямом и очевидном смысле. Принадлежность к замкнутой системе не меняется со временем; ни один элемент не покидает систему и не допускается в нее. Открытая система имеет как пополнения, так и потери. Системы с потерями, но без пополнения можно рассматривать как замкнутые, если принять условие, что элементы, покидающие систему, переходят в дополнительную категорию, добавленную в данную систему. Все открытые системы можно формально преобразовать в замкнутые с помощью приема такого же рода, но, по-видимому, в этом нет особых преимуществ.

Время можно рассматривать как *дискретное* или *непрерывное*. «Реальное» время непрерывно, но для целей моделирования бывает удобно предположить, что изменения состояния могут наблюдаться только в фиксированные моменты времени. Вполне логично считать, что, например, изменения состояния системы образования происходят в начале семестра или в начале учебного года. Кроме того, дискретная модель может дать более удобные для обработки уравнения и, следовательно, стать приемлемой аппроксимацией реального процесса.

*Марковское свойство* является основным для наших моделей. Где бы оно ни появлялось, его преимущества настолько велики, что всегда имеют смысл попытки переопределить задачу так, чтобы оно работало. Грубо говоря, марковское свойство предполагает, что знание текущего состояния системы дает нам всю информацию, необходимую для предсказания ее будущего. На первый взгляд это предположение является очень сильным ограничением, но искусство построения модели и состоит в том, чтобы так определить состояния, чтобы вся известная нам предыстория процесса была использована при определении текущего состояния.

## Обозначения

Обозначения часто становятся первопричиной трудностей для тех, у кого не хватает математической подготовки. Задача усложняется, когда материал из различных источников должен быть соединен в единое, логически связанное изложение. Поэтому весьма желательно везде иметь единые обозначения, но целесообразно придерживаться и обозначений, принятых в различных областях. Эти два положения часто противоречат друг другу, и мы вынуждены идти на компромисс. Даже если используются строчные и прописные буквы греческого алфавита, часто оказывается, что символов не хватает. Мы предпочли следующий принцип. Основные величины, встречающиеся на протяжении всей книги, обозначаются одними и теми же символами; если это возможно — начальными буквами названия величины. Наиболее распространенные из них следующие:

$T$  обозначает время, измеряемое от начала процесса, оно может быть дискретным или непрерывным;

$t$  или  $\tau$  также обозначает время, но измеряемое от другого начального момента;

$n$  обозначает число людей. Единичный индекс, например  $n_i$ , используется для обозначения этого числа в данном состоянии  $i$ ; если это число меняется со временем, число в момент  $T$  обозначается  $n_i(T)$ ; двойной индекс, например  $n_{ij}$ , обозначает поток от состояния  $i$  к состоянию  $j$  на некотором интервале времени или в некоторый момент времени, причем этот интервал или момент указывается в скобках;

$N$  обозначает общее число людей в системе;

$p$  и  $q$  используются так же, как и  $n$ , с той лишь разницей, что они относятся к вероятностям (или частотам); таким образом,  $p_{ij}$  есть вероятность отдельного перехода из состояния  $i$  в состояние  $j$ ;

$k$  — общее число состояний в системе;

$g, h, i, j, m, r$  и  $s$  часто употребляются как индексы, обозначающие состояние системы, особенно в сочетании со штрихами;  $i$  и  $j$  служат только для этой цели;

$\theta$  с индексами используется для обозначения собственных значений (характеристических чисел) матрицы. Это позволило избежать путаницы с  $\lambda$ , к которой прибегали в более ранних изданиях для обозначения как собственных значений, так и опасности ухода\*. Теперь  $\lambda$  употребляется только в последнем смысле.

Некоторые другие символы имеют различный смысл в зависимости от контекста, но они выбраны так, чтобы минимизировать риск неправильного понимания. Несколько букв, такие, как  $X, x, u$  и  $v$ , являются символами общего назначения, имеющими локальный смысл.

$f(\cdot)$  обозначает функцию плотности вероятности;  $F(\cdot)$  — соответствующая функция распределения, а  $G(\cdot) = 1 - F(\cdot)$  в некоторых приложениях называется функцией дожития.

Матрицы и векторы обозначаются одним из двух способов. Чаще всего матрица обозначается жирной прописной, а вектор — жирной

\* В оригинале — hazard function, что везде переведено как интенсивность ухода. — Примеч. ред.

строчной буквой. Другим способом обозначения матрицы является запись  $(i, j)$ -го элемента в скобках, например  $\{p_{ij}\}$ ; вектор можно записать как  $\{p_i\}$ . Второй способ мы применяем только тогда, когда абсолютно ясно, каковы элементы матрицы. Иногда будут встречаться матрицы, у которых отличны от нуля только элементы главной диагонали. Такие матрицы будут записываться либо как  $\text{diag}\{p_i\}$ , либо как  $[p]_d$ . Во втором случае диагональные элементы являются элементами вектора  $p$ . Далее  $\mathbf{1}$  обозначает единичный вектор-строку, а  $e_i$  — вектор, у которого  $\mathbf{1}$  является  $i$ -м компонентом, а остальные компоненты — нули. Размерность матрицы почти никогда не указывается, если она ясна из контекста. Векторы записываются в строку, а транспонирование векторов и матриц обозначается штрихом.

Логарифмы по основанию  $e$  записываются  $\ln$ ;  $\det$  обозначает определитель. Записи  $X \sim B(\cdot \cdot)$ ,  $\text{Poiss}(\cdot)$  или  $N(\cdot \cdot)$  являются сокращенными обозначениями того, что  $X$  (или любая случайная переменная) имеет биномиальное, пуассоновское или нормальное распределение с параметрами на месте точек в скобках.

Математическое ожидание случайной переменной обозначается с помощью  $E$  перед обозначением переменной или с помощью горизонтальной черты сверху, например  $\bar{n}(T)$ ;  $\text{var}(X)$  обозначает дисперсию  $X$ , а  $\text{cov}(X, Y)$  — ковариацию  $X$  и  $Y$ .

Следует осторожно обращаться с символами  $R$  и  $S$ , каждый из которых употребляется в различных смыслах. Латинское  $R$  относится к числу нанятых. Символом  $r = (r_1, r_2, \dots, r_k)$  обозначается вектор относительных долей принятых на работу.

Прописная жирная  $R$  обозначает вектор интенсивностей переходов с элементами  $r_{ij}$ . Символ  $r$  без индексов используется чисто локально в гл. 7. Символ  $S$  употребляется главным образом в связи с вероятностями перехода для цепей вакансий в гл. 8, но, следуя общепринятой практике, мы применяем его и для обозначения обыкновенной переходной матрицы в гл. 4 и как параметр преобразования Лапласа или производящей функции.

## 2.1. ВВЕДЕНИЕ

Самым простым способом числового описания человеческих популяций является классификация их членов на основе некоторых характеристик личности. Таким образом, мы классифицируем избирателей в соответствии с тем, за кого они голосуют, семьи — по месту жительства, рабочих — по их специальности. Когда мы начинаем изучать динамику социальных процессов, то вполне естественно начать с рассмотрения перемещения людей между категориями. Поскольку такие перемещения в значительной мере непредсказуемы на индивидуальном уровне, для построения некоторой модели необходимо описать механизм перемещения в вероятностной форме. В настоящей главе мы рассмотрим эту задачу для замкнутых систем с дискретным временем. Система с дискретным временем представляет собой такую систему, в которой изменение состояния происходит только в фиксированные моменты времени. Эти моменты не обязательно находятся на одинаковом расстоянии друг от друга, хотя часто мы будем считать их таковыми. Замкнутая система — это такая система, в которой нет ни пополнений, ни потерь. Оба эти ограничения будут ослаблены в последующих главах.

Рассматриваемую систему можно охарактеризовать двумя множествами величин, которые мы будем называть «запасы» и «потоки». Запасы в какой-либо конкретный момент времени — это численности людей (или других элементов), которые перемещаются между двумя уровнями системы за некоторый интервал времени. В общественных науках изучают очень много процессов, которые можно рассматривать таким образом. Некоторые из них мы уже упоминали, а другие перечислены в разделе «Дополнения» в конце гл. 3. Разнообразие потенциальных применений этих понятий ставит перед нами задачу их описания. С математической точки зрения наиболее удобно вести рассуждения на абстрактном уровне. Такой подход к описанию стохастических процессов можно найти в большинстве учебников, но при этом математическая строгость превалирует над практической направленностью. Другая крайность — рассматривать каждую область приложения отдельно. Однако это привело бы к многочисленным повторениям и сделало бы неясными те сходства структурного характера, которые лежат в основе многих разнообразных явлений. Мы поставили целью придерживаться середины. Большая часть наших моделей (в этой и последующих главах) будет неизменно связана с конкретной областью примене-

ния. Но мы отнюдь не будем колебаться при необходимости отклониться в другие области, если это послужит демонстрации новых характерных особенностей. Таким образом, в этой главе мы займемся социальной мобильностью, которая является одним из старейших и наиболее интересных процессов с точки зрения стохастического моделирования. Однако позже, когда мы перейдем к марковским моделям взаимодействия, более естественно будет описывать их как процессы изменения точки зрения или поведения при голосовании.

Одной из самых ранних статей, в которой социальная мобильность рассматривалась как стохастический процесс, была статья Прайса (1955 а). С тех пор появилось много литературы, на которую мы будем ссылаться. Необходимо различать межпоколенную и внутриспоколенную мобильность. Первая относится к изменениям социальной группы от отца к сыну в последующих поколениях. Здесь поколение представляет естественную дискретную единицу времени; оно является, по существу, упорядоченной переменной, и мы не делаем никаких предположений о ее значении в годах. Внутриспоколенная мобильность относится к изменениям класса, которые имеют место в течение жизни отдельного индивидуума. Вполне естественно говорить о них как о социальной мобильности, если выбранные классы отражают социальный статус. Мы предпочитаем называть это профессиональной мобильностью или мобильностью рабочей силы, потому что она обычно непосредственно связана с родом занятий, а не с общественным положением, которое им определяется. Характерной особенностью профессиональной мобильности является то, что не существует естественного интервала времени, за который происходит изменение состояния. Перемещения могут происходить в любое время, и, следовательно, такие процессы лучше моделировать в непрерывном времени. Мы займемся этим в гл. 4, где будет показано, что если такой процесс наблюдается только на фиксированных интервалах времени, то его можно изучать как дискретный. Это означает, что модели с дискретным временем, рассмотренные в этой главе, имеют более широкое применение, чем может показаться на первый взгляд. Мы увидим, что некоторые открытые системы, рассмотренные в гл. 3, можно изучать и как замкнутые.

## 2.2. МАРКОВСКИЕ МОДЕЛИ СОЦИАЛЬНОЙ МОБИЛЬНОСТИ

### Основная модель

Сначала мы рассмотрим очень простую модель развития отдельной семейной линии, а затем в последующих разделах изучим последствия устранения некоторых наименее реальных ее особенностей. Основное требование к модели состоит в том, что она должна указывать направление изменений, происходящих в социальной группе. Мы предположим, что эти изменения определяются вероятностями перехода, не зависящими от времени. Пусть  $p_{ij}$  — вероятность того, что отец принадлежит группе  $i$ , а его сын — группе  $j$ ; поскольку система замкнута, то

$$\sum_{j=1}^k p_{ij} = 1,$$

где  $k$  — число групп. Обозначим через  $\mathbf{P}$  матрицу вероятностей переходов. Если рассматривать такие семейные линии, в которых отец имеет только одного сына, то история семьи (семейной линии) будет цепью Маркова. Рассматривая общество как совокупность таких семейных линий, мы сможем делать заключения об изменении структуры общества. На практике же требование, чтобы каждый отец имел только одного сына, не выполняется. В результате некоторые семейные линии обрываются, а другие ветвятся. Однако в популяции, размер которой остается постоянным в течение некоторого периода времени, каждый отец будет иметь в среднем одного сына. Можно предположить, что наши результаты для простой модели должны быть применимы в усредненном смысле в некотором реальном обществе такого рода. Позже мы дадим более четкое обоснование этому рассуждению и покажем, что сделанное нами предположение выполняется.

Допустим, вероятность того, что первоначальный прародитель семейной линии принадлежит группе  $j$ , в нулевой момент времени равна  $p_j(0)$ . Пусть вероятность того, что потомок семьи в момент времени  $T$  ( $T = 1, 2, \dots$ ) принадлежит группе  $j$ , равна  $p_j(T)$ . Тогда вероятности  $\{p_j(T)\}$  могут быть рассчитаны рекуррентно на основании того, что

$$p_j(T+1) = \sum_{i=1}^k p_i(T) p_{ij} \quad (j = 1, 2, \dots, k). \quad (2.1)$$

В матричной форме эти уравнения можно записать следующим образом:

$$\mathbf{p}(T+1) = \mathbf{p}(T)\mathbf{P}, \quad (2.2)$$

где  $\mathbf{p}(T) = (p_1(T), p_2(T), \dots, p_k(T))$ . Повторное применение (2.2) дает

$$\mathbf{p}(T) = \mathbf{p}(0)\mathbf{P}^T. \quad (2.3)$$

Элементы  $\mathbf{p}(T)$  можно также интерпретировать как ожидаемые доли генеральной совокупности по различным группам в момент времени  $T$ . Если первоначальные группы данной семьи известны, то вектор  $\mathbf{p}(0)$  представляет собой исходную структуру групп.

Матрица  $\mathbf{P}^T$  играет основополагающую роль в теории цепей Маркова. Эту матрицу можно использовать для получения из (2.3) вероятностей «состояния», но ее элементы имеют также и непосредственную вероятностную интерпретацию. Пусть  $p_{ij}^{(T)}$  обозначают  $(i, j)$ -й элемент матрицы  $\mathbf{P}^T$ , тогда (2.3) можно переписать в виде

$$p_j(T) = \sum_{i=1}^k p_i(0) p_{ij}^{(T)} \quad (j = 1, 2, \dots, k). \quad (2.4)$$

Из такого представления ясно, что  $p_{ij}^{(T)}$  есть вероятность того, что семья переходит из группы  $i$  в группу  $j$  через  $T$  поколений. Случай  $i = j$  представляет особый интерес, поскольку вероятности  $p_{ij}^{(T)}$  можно принять за основу измерений мобильности.

Во многих приложениях данная генеральная совокупность сохраняется в течение многих поколений, так что «текущее» состояние соответствует большому значению  $T$ . Следовательно, большой практический интерес представляет исследование поведения вероятностей

$\{p_i(T)\}$  и  $\{p_{ij}^T\}$  при  $T$ , стремящемся к бесконечности. В общей теории цепей Маркова показано, что такое предельное поведение зависит от структуры матрицы  $\mathbf{P}$ . Можно показать, что если матрица  $\mathbf{P}$  *регулярна* то при  $T$ , стремящемся к бесконечности, все эти вероятности стремятся к некоторым пределам. Регулярная (конечная) цепь Маркова — это такая цепь, в которой после некоторого числа  $T$  поколений возможно любое состояние (класс) независимо от начального. Точнее, для того чтобы цепь была регулярной, необходимо и достаточно, чтобы все элементы матрицы  $\mathbf{P}^T$  для некоторого  $T$  были отличны от нуля. Все матрицы вероятностей переходов, которые могут встретиться в данной книге, регулярны, но позже мы встретим примеры, матрицы в которых не обладают таким свойством.

Как только установлено существование пределов, непосредственной задачей становится их вычисление. Таким образом, если мы обозначим  $\lim_{T \rightarrow \infty} p_j(T)$  через  $p_j$ , то из (2.2) следует, что предельная структура должна удовлетворять соотношению

$$\mathbf{p} = \mathbf{p}\mathbf{P} \quad (2.5)$$

при

$$\sum_{j=1}^k p_j = 1.$$

Таким образом, предельную структуру, или распределение, можно получить, решив некоторую систему уравнений. Важным свойством такого решения является то, что оно не зависит от начального состояния системы. Поскольку в силу наших предположений каждая сохранившаяся семейная линия достигает равновесия, заданного (2.5), вектор  $\mathbf{p}$  дает ожидаемую структуру генеральной совокупности в настоящее время. Если кроме этой структуры ничего не наблюдается, то у нас нет никаких средств восстановления переходной матрицы\*. Действительно, мы не можем вывести матрицу  $\mathbf{P}$  из двух последовательно наблюдаемых структур  $\mathbf{p}(T)$  и  $\mathbf{p}(T+1)$ , хотя Уайт (1963) и Матрас (1967) изучали вопрос о том, какую неполную информацию можно получить в подобных условиях. Предельное значение  $\mathbf{P}^T$ , обозначенное  $\mathbf{P}^\infty$ , может быть получено из (2.3). Оно должно удовлетворять соотношению

$$\mathbf{p} = \mathbf{p}(0)\mathbf{P}^\infty,$$

которое может быть выполнено, если

$$\mathbf{P}^\infty = \begin{pmatrix} p_1 & p_2 & \dots & p_k \\ p_1 & p_2 & \dots & p_k \\ \vdots & \vdots & \dots & \vdots \\ p_1 & p_2 & \dots & p_k \end{pmatrix}, \quad (2.6)$$

что означает

$$\lim_{T \rightarrow \infty} p_{ij}^T = p_j. \quad (2.7)$$

\* Это справедливо при полном обследовании, но при выборочных обследованиях измерение переходной матрицы возможно. — *Примеч. ред.*

Если наша модель дает адекватное описание реальных обществ, то предшествующий анализ показывает, что их будущее развитие зависит только от исходной структуры и матрицы вероятностей переходов. Из этих двух факторов влияние исходной структуры с течением времени постепенно исчезает. Следовательно, по истечении достаточно продолжительного времени структура общества определяется только переходной матрицей. Этот вывод означает, что исследование мобильности должно быть сосредоточено на вероятностях перехода. В частности, меры мобильности должны быть функциями элементов матрицы **P**. Прежде чем продолжать обсуждение этого положения, следует проверить адекватность модели, пользуясь реальными данными о мобильности.

Опубликовано несколько эмпирических исследований мобильности, которые содержат достаточно для частичной проверки теории количество данных. Одно из таких исследований, выполненное Глассом и Холлом (Гласс, 1954), основано на случайной выборке 3500 пар отцов и сыновей в Великобритании. Другое исследование, выполненное Н. Роговой (1953), основано на данных из заявлений на получение свидетельств о браке в округе Марион (штат Индиана). Рогова получила данные о двух периодах: один — с 1905 по 1912 г. при объеме выборки 10 253; второй — с 1938 по 1941 г. при объеме выборки 9892. Более поздние данные, относящиеся к Дании, приведены в исследовании Свалагосты (1959).

Данные Гласса и Холла использовались Прайсом (1955а), и материал настоящего раздела взят из его статьи. Гласс и Холл разделили

Т а б л и ц а 2.1. Оценки вероятностей переходов для Англии и Уэльса в 1949 г. (по данным Гласса и Холла)

Группы отца	Группы сына						
	1	2	3	4	5	6	7
1. Профессиональный руководитель высшего уровня	0,388	0,146	0,202	0,062	0,140	0,047	0,015
2. Руководитель и административный работник	0,107	0,267	0,227	0,120	0,206	0,053	0,020
3. Руководитель высокого уровня в сфере умственного труда	0,035	0,101	0,188	0,191	0,357	0,067	0,061
4. Руководитель низкого уровня в сфере умственного труда	0,021	0,039	0,112	0,212	0,430	0,124	0,062
5. Квалифицированный работник физического труда и простого умственного труда	0,009	0,024	0,075	0,123	0,473	0,171	0,125
6. Полуквалифицированный работник физического труда	0,000	0,013	0,041	0,088	0,391	0,312	0,155
7. Неквалифицированный работник физического труда	0,000	0,008	0,036	0,083	0,364	0,235	0,274



членов их выборки на семь профессиональных групп, представленных в табл. 2.1, в которой содержатся также оценки вероятностей переходов.

В первых двух столбцах табл. 2.2 приведена структура группы населения в двух последовательных поколениях. Если марковская модель адекватна и если общество достигло равновесия, можно предположить, что эти распределения одинаковы, за исключением флуктуаций в выборке. Можно также предположить, что оба эти распределения соответствуют распределению равновесия, полученному из (2.5).

Прайс (1955а) выполнил расчеты, необходимые для такого сравнения, и его результаты представлены в табл. 2.2.

Различия между этими тремя распределениями невелики, хотя, по-видимому, имеется сдвиг в сторону более низких групп по мере движения вдоль таблицы. Если этот сдвиг не является просто результатом флуктуаций в выборке, то его можно принять за показатель того, что процесс еще не достиг равновесия. Другое возможное объяснение приводится ниже. Исчерпывающего ответа на вопрос об ошибке выборки нет, но первый шаг в направлении решения этой проблемы дает теория выборки, которая будет изложена позже. Хотя мы, пользуясь этими данными, не можем выполнить полностью проверку модели, представляется, что между данными и предсказаниями теории имеется широкое соответствие.

Данные Н. Роговой приводят к аналогичным выводам относительно применимости теории марковских цепей к социальной мобильности. Она получила информацию о двух группах данных, разделенных 30-летним периодом времени, так что мы можем посмотреть, произошло ли сколько-нибудь существенное изменение вероятностей переходов за этот период. Две переходные матрицы даны в табл. 2.3.

Мы следовали Кемени и Снеллу (1976) в использовании грубой группировки вместо разбивки, приведенной Роговой. С учетом ошибки выборки видим, что за данный период времени произошли изменения в переходной матрице. Однако эти изменения невелики, и мы будем предполагать, что не будет большой ошибкой считать их постоянными для достаточно коротких периодов времени.

Никакой набор данных не позволяет нам выполнить прямую проверку марковского свойства. Это свойство требует, чтобы группа сына зависела только от группы отца и не зависела от группы его деда. Чтобы проверить это условие, нам нужны записи истории семьи, по крайней мере, за три поколения. Косвенное подтверждение этого предположения обеспечивается близостью между равновесной струк-

**Таблица 2.2. Действительное и равновесное распределения социальных групп в Англии и Уэльсе (1949 г.) оцененные по данным Гласса и Холла**

Группа	Отцы	Сыновья	Предсказанное равновесие
1	0,037	0,029	0,023
2	0,043	0,046	0,042
3	0,098	0,094	0,088
4	0,148	0,131	0,127
5	0,432	0,409	0,409
6	0,131	0,170	0,182
7	0,111	0,121	0,129

Т а б л и ц а 2.3. Вероятности переходов, оцененные по данным Н. Роговой для округа Марион, штат Индиана\*

		1905—1912 гг.			1938—1941 гг.		
		1	2	3	1	2	3
Работники	умственного труда	0,594	0,396	0,009	0,622	0,375	0,003
Работники	физического труда	0,211	0,782	0,007	0,274	0,721	0,005
Фермеры		0,252	0,641	0,108	0,265	0,694	0,042

\* Из книги: G. Kemeny, J. Laurie Snell. Finite Markov Chains.  
© 1960 by D. Van Nostrand Company. Перепечатано с разрешения издательства.

турой групп, предсказанной теорией Маркова, и наблюдаемой структурой групп.

Теоретически есть основания считать, что процессы социальной мобильности не обладают марковским свойством. Это объясняется тем фактом, что границы группы очерчены произвольно. Так, например, мы могли бы еще подразделить те семь категорий, которые использовались Глассом и Холлом. В то же время некоторые группы можно было бы объединить и получить меньшее число категорий. Известно (см. Кемени и Снелл, 1976, гл. 6), что если состояния цепи Маркова объединяются, то новая цепь, вообще говоря, не обладает марковским свойством. В наших условиях это означает, что мы не можем произвольно перестроить группы и сохранить марковское свойство. Даже если имеется одна система группировки, обладающая таким свойством, она может и не оказаться именно той системой, которую нам пришлось выбрать. Однако, как мы отмечали в гл. 1, достаточно того, чтобы наша модель отражала главные свойства процесса, не будучи точной в деталях. Итак, хотя предположения, на которых основана теория, и не вполне реалистичны, эта модель достаточно близка к реальности и ее дальнейшее использование вполне оправдано.

### Обращение времени в марковской модели

Основное разностное уравнение марковской модели для межпоколенной мобильности возникает при рассмотрении ожидаемой группы в следующем поколении, обусловленной существующей группой. С помощью такого способа оказалось возможным предсказывать будущие структуры групп и получать их предельную форму. Иногда бывает интересно пойти в обратном направлении и рассмотреть, что произойдет, если мы попытаемся обратить время вспять. Соответствующие вероятности переходов тогда обозначим как  $Pr\{\text{семья перешла в группу } i \text{ из группы } j\} = \tilde{p}_{ij}$ . Эти вероятности переходов просто соотносятся с прямыми вероятностями переходов следующим образом. Пусть  $X(T)$  временно обозначает номер группы, в которой семья находится в момент  $T$ . Тогда

$$\begin{aligned} Pr\{X(T+1) = j, X(T) = i\} &= Pr\{X(T) = i\}Pr\{X(T+1) = j | X(T) = i\} \\ &= Pr\{X(T+1) = j\}Pr\{X(T) = i | X(T+1) = j\}. \end{aligned}$$

Сравнивая две альтернативные формы правой части, имеем

$$Pr\{X(T) = i | X(T + 1) = j\} = p_{ij}p_i(T)/p_j(T + 1) = \tilde{p}_{ji}. \quad (2.8)$$

Хотя обращенный процесс является марковским, он, вообще говоря, неоднороден по времени. Однако если система достигла установившегося состояния, то  $p_i(T)$  и  $p_j(T + 1)$  примут свои предельные значения. В этом случае, меняя местами  $i$  и  $j$ , получаем

$$\tilde{p}_{ij} = p_{ji}p_j/p_i \quad (i, j = 1, 2, \dots, k). \quad (2.9)$$

Может случиться, что переходные матрицы прямой и обращенной цепи окажутся одинаковыми, т.е.  $\tilde{\mathbf{P}} = \mathbf{P}$ . Это означает, что процесс протекает одинаково, в каком бы направлении по времени его не рассматривать: такой процесс мы будем называть обратимым или обращаемым. Одно из следствий обратимости (обращаемости) заключается в следующем: если мы считаем, что некто находится в группе  $i$ , то вероятность того, что его сын находится в группе  $j$ , такая же, как и вероятность того, что его отец был в группе  $j$ . Еще один способ выразить это свойство заключается в использовании понятия ожидаемого числа изменений группы между данным поколением и следующим. Предположим, что мы имеем генеральную совокупность из  $N$  семейных линий; тогда ожидаемое число переходов из группы  $i$  в группу  $j$  от одного поколения к следующему будет  $Np_{ij}$ . Ожидаемый поток в противоположном направлении есть  $Np_{ji}$ . При условии обращаемости, полученной с помощью замены  $\tilde{p}_{ij} = p_{ji}$  в (2.9), эти два потока равны. Следовательно, обращаемость означает одинаковый обмен между группами. Поскольку система находится в равновесии, нет *никакого* изменения размера групп от одного поколения к другому. Но обращаемость требует гораздо большего: семейная линия может развиваться в новом направлении, только если существует некоторое компенсационное движение в противоположном направлении. Сказанное не представляется вполне правдоподобным для случая социальной мобильности, но, как заметили Кемени и Снелл (1976), это, по-видимому, не относится к данным Гласса и Холла. Данное положение можно проверить вычислением матрицы с элементом  $(i, j)$ , равным  $p_{ij}$ ; для обратимого процесса эта матрица будет симметрической. Если мы образуем такую матрицу для данных Гласса и Холла, приведенных в табл. 2.1 и 2.2, то получим

$$\begin{pmatrix} 0,0089 & 0,0034 & 0,0046 & 0,0014 & 0,0032 & 0,0011 & 0,0003 \\ 0,0045 & 0,0112 & 0,0095 & 0,0050 & 0,0087 & 0,0022 & 0,0008 \\ 0,0031 & 0,0089 & 0,0165 & 0,0168 & 0,0314 & 0,0059 & 0,0054 \\ 0,0027 & 0,0050 & 0,0142 & 0,0269 & 0,0546 & 0,0157 & 0,0079 \\ 0,0037 & 0,0098 & 0,0307 & 0,0503 & 0,1935 & 0,0699 & 0,0511 \\ 0,0000 & 0,0024 & 0,0075 & 0,0160 & 0,0712 & 0,0568 & 0,0282 \\ 0,0000 & 0,0010 & 0,0046 & 0,0107 & 0,0470 & 0,0303 & 0,0353 \end{pmatrix}.$$

Приблизительная симметричность этой матрицы означает, что имеет место почти равный обмен между каждой парой групп. Трудно привести какие-либо социальные причины, объясняющие, почему так должно быть, и этот вопрос, очевидно, требует дальнейших исследований.

## Изменчивость размеров группы

Развиваемая до сих пор теория позволяет в любой будущий момент времени рассчитать для каждой группы ожидаемые численности или их соотношения. До сих пор у нас не было способов определения дисперсий и ковариаций наших предсказаний. Метод вычисления вторых моментов, в том числе и смешанных моментов для групповых чисел, был разработан Поллардом (1966), и теперь мы опишем его применение к нашей задаче.

Пусть  $N$  обозначает число семейных линий в генеральной совокупности; оно остается постоянным во времени. Пусть размер  $j$ -й группы в момент  $T$  равен  $n_j(T)$  и пусть число переходов между классом  $i$  и классом  $j$  с момента  $T$  до момента  $T + 1$  равно  $n_{ij}(T)$ . Из определений следует, что

$$n_j(T + 1) = \sum_{i=1}^k n_{ij}(T) \quad (j = 1, 2, \dots, k). \quad (2.10)$$

Если мы возьмем математические ожидания для каждой части этого равенства, то придем к (2.1), поскольку  $E n_j(T + 1) = N p_j(T + 1)$  и  $E n_{ij}(T) = N p_i(T) p_{ij}$ . Рассмотрим ковариацию  $n_j(T + 1)$  с  $n_l(T + 1)$ . Для упрощения изложения допустим, как обычно, что  $\text{cov}(x_i, x_j) \equiv \equiv \text{var}(x_i)$  при  $i = j$ . Тогда из (2.10) имеем

$$\begin{aligned} \text{cov}\{n_j(T + 1), n_l(T + 1)\} &= E\{n_j(T + 1) n_l(T + 1)\} - \\ &- E n_j(T + 1) E n_l(T + 1) = \sum_{i=1}^k \sum_{i'=1}^k [E\{n_{ij}(T) n_{i'l}(T) - \\ &- E n_{ij}(T) E n_{i'l}(T)\}]. \end{aligned} \quad (2.11)$$

Чтобы оценить математические ожидания в (2.11), воспользуемся хорошо известным результатом об условных математических ожиданиях:  $E(x) = E_x E(x|X)$ . В рассматриваемом случае удобно получить математические ожидания при условии, что даны  $n_i(T)$ . Уравнения для математических ожиданий следуют из того, что при заданных  $n_i(T)$  величины  $n_{ij}(T)$  ( $j = 1, 2, \dots, k$ ) являются полиномиально распределенными случайными величинами с вероятностями  $p_{ij}$  ( $j = 1, 2, \dots, k$ ). Следовательно,

$$E\{n_{ij}(T) | n_i(T)\} = n_i(T) p_{ij} \quad (2.12)$$

и

$$\begin{aligned} E\{n_{ij}(T) n_{i'l}(T) | n_i(T), n_{i'}(T)\} &= n_i(T) n_{i'}(T) p_{ij} p_{i'l} \quad (i \neq i'), \\ E\{n_{ij}(T) n_{i'l}(T) | n_i(T)\} &= n_i(T) \{n_i(T) - 1\} p_{ij} p_{il} + \delta_{jl} n_i(T) p_{ij}, \end{aligned} \quad (2.13)$$

где

$$\delta_{jl} = \begin{cases} 1 & \text{при } j = l, \\ 0 & \text{при } j \neq l. \end{cases}$$

Безусловные математические ожидания теперь получаются из (2.12) и (2.13), если взять математические ожидания правых частей по  $n_i(T)$ . Подстановка полученных выражений в (2.11) дает

$$\begin{aligned} \text{cov} \{n_j(T+1), n_l(T+1)\} &= \sum_{i=1}^k \sum_{i'=1}^k p_{ij} p_{i'l} \text{cov} \{n_i(T), n_{i'}(T)\} + \\ &+ \sum_{i=1}^k (\delta_{ji} p_{ij} - p_{ij} p_{ii}) E n_i(T). \end{aligned} \quad (2.14)$$

Таким образом, получено рекуррентное соотношение между математическими ожиданиями и ковариациями в момент  $T$  и ковариациями в момент  $T+1$ . Так как ковариации при  $T=0$  равны нулю, совокупность первых и вторых может быть вычислена по (2.14) и (2.1). Если это необходимо, то тот же метод может быть распространен и на получение моментов более высоких порядков, в том числе и смешанных.

Поскольку математические ожидания и ковариации в момент  $T+1$  — линейные функции соответствующих величин в момент  $T$ , то можно выразить соотношения между ними в матричной форме. Для этого введем вектор, состоящий из средних значений и ковариаций, и обозначим его через  $\mu(T)$ . Этот вектор сначала содержит  $k$  средних значений, а затем  $k^2$  ковариаций, расположенных в алфавитном порядке, т. е.

$$\begin{aligned} \mu &= [E(n_1), \dots, E(n_k), \text{cov}(n_1, n_1), \text{cov}(n_1, n_2), \dots, \text{cov}(n_1, n_k), \\ &\text{cov}(n_2, n_1), \text{cov}(n_2, n_2), \dots, \text{cov}(n_2, n_k), \dots, \text{cov}(n_k, n_k)], \end{aligned}$$

здесь для краткости опущены аргументы у  $n$  и  $\mu$ . В этом перечислении имеется некоторая избыточность вследствие того, что, например, ковариация между  $n_1$  и  $n_2$  появляется как  $\text{cov}(n_1, n_2)$  и как  $\text{cov}(n_2, n_1)$ , но, сохраняя обе эти записи, мы обеспечиваем симметрию выражения. Уравнения (2.1) и (2.14) теперь можно совместить и записать в форме

$$\mu(T+1) = \mu(T) \Pi, \quad (2.15)$$

где элементы матрицы  $\Pi$  размерности  $k(k+1) \times k(k+1)$  являются функциями от  $p_{ij}$ . Матрицу  $\Pi$  можно разбить на блоки следующим образом:

$$\frac{\mathbf{P} \mid \mathbf{X}}{\mathbf{O} \mid \mathbf{Y}'}$$

где  $\mathbf{P}$  размерности  $k \times k$  является переходной матрицей,  $\mathbf{O}$  размерности  $k^2 \times k$  — нулевая матрица,  $\mathbf{X}$  размерности  $k \times k^2$  — матрица с элементами вида  $\delta_{ii} p_{ij} - p_{ij} p_{ii}$ , а  $\mathbf{Y}'$  размерности  $k^2 \times k^2$  — матрица с элементами вида  $p_{ij} p_{i'l}$ <sup>1</sup>. В случае  $\mathbf{X}$  и  $\mathbf{Y}'$  ( $j, l$ ) индексируют столбец, а  $i$  или  $(i, i')$  — строку.

Уравнение (2.15) представляет собой обобщение уравнения (2.2) и может применяться таким же образом.  $\Pi$  уже не является стохастической

<sup>1</sup>  $\mathbf{Y}'$  является прямым матричным произведением  $\mathbf{P} \times \mathbf{P}$ . Прямое матричное произведение (или произведение Кронекера) можно использовать для облегчения решения более общей задачи отыскания моментов высоких порядков и смешанных моментов для  $\{n_j\}$  (см. работу Полларда (1966)).

ческой матрицей, так как она имеет отрицательные элементы, но сумма элементов строк остается равной единице. На основании этого факта можно сделать вывод, что элементы матрицы  $\mu(T)$  стремятся к пределам при стремлении  $T$  к бесконечности. Следовательно, эти пределы должны удовлетворять уравнениям

$$\left. \begin{aligned} \mu(\infty) &= \mu(\infty) \Pi \\ \sum_{i=1}^k \bar{n}_i(\infty) &= N \end{aligned} \right\}, \quad (2.16)$$

здесь и далее черточка сверху обозначает математическое ожидание случайной величины. Имеется программа для ЭВМ, вычисляющая векторы  $\mu(T)$  при всех  $T$ . В гл. 3 мы приведем результаты некоторых расчетов для более общей модели. Далее рассмотрим один частный случай и изложим идею о типе решения, получаемого для данной прикладной задачи.

Предположим, что имеется, по терминологии Прайса, совершенно мобильное общество, т. е.  $p_{ij} = p_j$  ( $j = 1, 2, \dots, k$ ). Тогда уравнение (2.14) упрощается:

$$\begin{aligned} \text{cov}\{n_j(T+1), n_l(T+1)\} &= N\delta_{jl} p_j - N p_j p_l + \\ &+ p_j p_l \sum_{i=1}^k \sum_{i'=1}^k \text{cov}\{n_i(T), n_{i'}(T)\}. \end{aligned} \quad (2.17)$$

Теперь

$$\sum_{i=1}^k \sum_{i'=1}^k \text{cov}\{n_i(T), n_{i'}(T)\} = E \left\{ \sum_{i=1}^k (n_i - \bar{n}_i) \right\}^2 = 0,$$

так что последние слагаемые в (2.17) исчезают. Оставшаяся часть, оказывается, дает дисперсии и ковариации полиномиального распределения. Действительно, легко видеть, что  $\{n_j(T)\}$  имеет полиномиальное распределение с параметрами  $N$  и  $p_j$  ( $j = 1, 2, \dots, k$ ) для всех  $T$ . Распределение группы в момент  $T+1$  не зависит от распределения группы в момент  $T$ . Следовательно, мы можем рассматривать  $N$  семейных линий как независимо распределенные по группам с вероятностями  $p_j$  ( $j = 1, 2, \dots, k$ ), что является условием того, что это распределение должно быть полиномиальным. Таким образом, в этом случае средняя квадратичная ошибка предсказанного значения величины  $p_j$  будет  $\{p_j(1-p_j)/N\}^{\frac{1}{2}}$ .

Основное предположение при выводе этих результатов состояло в том, что индивидуумы, покидающие данную группу, распределены полиномиально по всем группам. Вероятности, управляющие этими потоками, являются элементами той строки переходной матрицы, которая соответствует интересующей нас группе. Что касается вторых моментов, то сделанное предположение находит отражение во второй части (2.13) при  $i = i'$ , и это приводит к результату (2.14). Винн и Салес (1973а) рассмотрели более общую структуру ошибки, при которой ковариации между любой парой потоков, исходящих из данной группы, считаются произвольными при условии, что общий поток фиксирован. Иначе говоря, Винн и Салес (1973а) полагают, что

$$\text{cov}\{n_{ij}(T), n_{i'l}(T)\} = 0 \text{ при } i \neq i' \text{ (как и ранее)}$$

и

$$\text{cov}\{n_{ij}(T), n_{il}(T)\} = \alpha_{ijl}n_i(T) \quad (j, l = 1, 2, \dots, k) \quad (2.18)$$

при заданном  $n_i(T)$ . Величины  $\alpha$  подчинены ограничению из-за того, что

$$\sum_{j=1}^k n_{ij}(T) = n_i(T)$$

и фиксированы  $n_i(T)$ , следовательно,

$$\begin{aligned} \text{var} \sum_{j=1}^k n_{ij}(T) &= \sum_{j=1}^k \sum_{l=1}^k \text{cov}\{n_{ij}(T), n_{il}(T)\} = \\ &= \sum_{j=1}^k \sum_{l=1}^k \alpha_{ijl} = \text{var}\{n_i(T)\} = 0. \end{aligned} \quad (2.19)$$

При выполнении ограничения (2.19) величины  $\{\alpha_{ijl}\}$  могут иметь любые значения, за исключением случая, когда  $j = l$ , и ковариация становится дисперсией, а, следовательно, она должна быть положительной. Для полиномиального распределения

$$\begin{aligned} \alpha_{ijl} &= -p_{ij}p_{il} \quad \text{при } j \neq l, \\ &= p_{ij}(1 - p_{ij}) \quad \text{при } j = l. \end{aligned} \quad (2.20)$$

Обоснование, приводящее к (2.14), по-прежнему остается в силе, но последнюю сумму в правой части следует заменить на

$$\sum_{i=1}^k \alpha_{ijl} E n_i(T).$$

В матричном представлении этих уравнений подматрица  $X$  становится матрицей размерности  $k \times k^2$  с элементами  $\{\alpha_{ijl}\}$ . Винн и Салес (1973а) установили существование предельной формы (2.16) и рассмотрели оценку ошибок; они применили этот метод к данным, относящимся к профессиональной мобильности.

### 2.3. ИЗМЕРЕНИЕ МОБИЛЬНОСТИ

Часто удобно иметь скалярную меру мобильности, которая позволяла бы нам сравнивать различные сообщества людей и характеризовать изменения мобильности в рамках одного сообщества. В прикладных работах были предложены многочисленные дескриптивные меры мобильности. Некоторые примеры приведены у Матраса (1960b), а полное их рассмотрение содержится в работе Будона (1973). Полезный обзор мер приведен в работе Бибби (1975). Мы не ставим своей целью дать полное изложение этой темы, а хотим только показать, как стохастическая теория соотносится с проблемой измерения. Будем предполагать, что мобильность адекватно описывается марковской моделью, и при этом предположении поставим вопрос о том, как надо измерять мобильность.

Поскольку движение полностью определяется переходной матрицей  $P$ , вся информация о мобильности содержится в ее элементах. Таким образом, наша задача, поставленная формально, состоит в том, чтобы отобразить множество переходных матриц на некоторый подходящий интервал так, чтобы упорядочить их по степени мобильности.

Этот вопрос был в значительной степени выяснен в двух теоретических построениях, данных Шорроксом (1978), а также Соммерсом и Конлиском (1979). Последующее рассуждение основывается главным образом на этих работах, но мы построим его с учетом некоторого различия в понятиях «мировой мобильности», по-видимому, до сих пор не замеченного. Суть вопроса может быть показана при рассмотрении специального случая  $k = 2$ . Возьмем три матрицы:

$$(a) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, (б) \begin{pmatrix} p & 1-p \\ p & 1-p \end{pmatrix}, (в) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad (0 < p < 1).$$

На одном конце этой шкалы нет неясностей. Матрица (а) соответствует некоторому сообществу, в котором нет никакого движения. Ясно, что эта матрица должна быть помещена на одном конце нашей шкалы измерений. Различие, которое мы хотим установить, становится ясным, когда мы рассматриваем другой конец шкалы. Здесь имеются две возможности, представленные матрицами (б) и (в). Матрица (б) соответствует тому, что Прайс (1955а) назвал совершенно мобильным обществом. Его отличительная черта состоит в том, что социальная группа сына не зависит от группы его отца. Для общего случая всякая матрица, имеющая одинаковые строки, будет обладать таким же свойством. Если мы измерим степень, до которой группа сына определена группой отца, то станет ясно, что (а) и (б) принадлежат к противоположным концам шкалы. Случай же (в) можно объединить с (а), так как при (в) и (а) группа сына полностью определяется группой его отца. Рассматриваемый случай мобильности часто представляет наиболее существенный интерес, поскольку именно он показывает, до какой степени благоприятные возможности сына ограничены случайностью рождения. Было бы меньше путаницы, если бы этот тип мобильности характеризовал степень зависимости между поколениями или степень глобальной временной зависимости (связи поколений).

Понятие «мировая мобильность» используется также и в другом смысле, когда говорят о количестве имеющих место перемещений. Если мы хотим измерить количество перемещений, то матрица (в) диаметрально противоположна матрице (а). В одном случае двигается каждый, в другом не двигается никто. На практике это различие не очень четкое и почти не имеет значения благодаря тому, что наблюдаемые переходные матрицы всегда лежат где-то между (а) и (б). Тем не менее, устанавливая критерии, по которым следовало бы оценивать мобильность, как это делает Шоррокс (1978), важно хорошо знать, какого рода мобильность мы намереваемся измерять.

Вторая особенность таблиц мобильности, которую часто упускают из виду, состоит в том, что группы могут иногда быть упорядочены по статусу, а иногда нет. Если группы упорядочены, то движению между



смежными группами можно придавать меньший вес, чем движению между далекими группами. Такое соображение учитывается, когда пытаются измерить количество перемещений, но часто в случае исследования связи между поколениями этого делать не стоит.

В общем случае показатели мобильности различны, тогда как при  $k = 2$  они совпадают иногда. Поэтому мы начнем с общего случая, отмечая по ходу дела специфику ситуации, когда  $k = 2$ .

### Измерение зависимости между поколениями

Сначала мы рассмотрим задачу, как измерять степень, с которой группа сына зависит от группы отца. Метод такого измерения основан на спектральном разложении переходной матрицы. Если  $\mathbf{P}$  — стохастическая матрица, то ее можно представить в виде

$$\mathbf{P} = \sum_{r=1}^k \theta_r \mathbf{A}_r. \quad (2.21)$$

Матрицы  $\{\mathbf{A}_r\}$  известны как спектральное множество, они обладают следующими свойствами:

$$\mathbf{A}_r \mathbf{A}_s = 0, \text{ при } r \neq s; \quad \mathbf{A}_r \mathbf{A}_s = \mathbf{A}_r, \text{ при } r = s; \quad \sum_{r=1}^k \mathbf{A}_r = \mathbf{I}. \quad (2.22)$$

Коэффициенты  $\{\theta_r\}$  являются собственными значениями матрицы  $\mathbf{P}$ , и так как  $\mathbf{P}$  представляет собой стохастическую матрицу, одно собственное значение равно единице, а остальные меньше единицы или равны ей по модулю. Следовательно, мы можем их упорядочить так, что  $1 = \theta_1 \geq |\theta_2| \geq \dots \geq |\theta_k|$  и записать (2.21) в виде

$$\mathbf{P} = \mathbf{1}' \mathbf{p} + \sum_{r=2}^k \theta_r \mathbf{A}_r, \quad (2.23)$$

где  $\mathbf{1}' \mathbf{p} = \mathbf{A}_1$ , а  $\mathbf{1} = (1, 1, \dots, 1)$ . Последний шаг обусловлен тем фактом, что (2.23) дает

$$\lim_{T \rightarrow \infty} \mathbf{P}^T = \mathbf{A}_1,$$

а мы знаем из (2.7), что эта предельная матрица имеет тождественные строки с элементами, равными элементам установившейся структуры матрицы  $\mathbf{P}$ .

Анализ (2.23) показывает, что если  $\theta_2 = \theta_3 = \dots = \theta_k = 0$ , то  $\mathbf{P}$  имеет тождественные строки, указывающие на отсутствие зависимости между поколениями. Следовательно, степень зависимости может быть измерена тем, насколько сильно  $\mathbf{P}$  отличается от  $\mathbf{1}' \mathbf{p}$ . Один из способов измерения этого отличия состоит в построении подходящей функции от  $\theta$ . Были предложены три таких меры, их мы теперь рассмотрим по порядку.

Первая, предложенная Шорроксом (1978), а также Sommerсом и Конлиском (1979), основана на использовании второго наибольшего

по абсолютной величине значения  $\theta$ . Обозначим его  $\theta_{\max}$ , тогда мерой будет функция

$$\mu_1(\mathbf{P}) = |\theta_{\max}|. \quad (2.24)$$

Она может изменяться между 0, когда нет зависимости, и 1. Чем больше величина  $\mu_1(\mathbf{P})$ , тем сильнее зависимость группы сына от группы отца, и, следовательно, тем меньше в этом смысле степень социальной мобильности. Этой мере могут быть даны разнообразные интерпретации. Шоррокс (1978), например, показывает, что она является функцией того, что он называет периодом полураспада, который в данном контексте является числом поколений на временном интервале (подходящим образом определенном) между текущей и установившейся структурой, поделенном пополам. Недостаток функции  $\mu_1(\mathbf{P})$ , к которой мы вернемся, состоит в том, что второе по величине собственное значение матрицы  $\mathbf{P}$  может оказаться мнимым.

Вторая мера может быть основана на среднем значении  $\theta_i$ :

$$\mu_2(\mathbf{P}) = \frac{1}{k-1} \sum_{r=2}^k \theta_r. \quad (2.25)$$

Так как сумма собственных значений матрицы  $\mathbf{P}$  равна сумме ее диагональных элементов, можно записать:

$$\mu_2(\mathbf{P}) = \left( \frac{\text{trace}(\mathbf{P}) - 1}{k-1} \right), \quad (2.26)$$

где след матрицы  $\text{trace}(\mathbf{P}) = \sum_{r=1}^k p_{rr}$ . Как и ранее, эта величина равна нулю, когда нет зависимости между поколениями, и возрастает до единицы, когда сын обязательно идет по стопам отца, т.е. в неподвижном сообществе людей. Эта мера может также принимать отрицательные значения, причем минимум имеет место, когда каждая группа сына отличается от группы его отца. Знак  $\mu_2(\mathbf{P})$  может, таким образом, быть использован для различия между тем, что можно было бы назвать «положительной» и «отрицательной» зависимостью.

Вместо среднего арифметического значения всех  $\theta$  мы могли бы воспользоваться средним геометрическим их модулей. Это дает меру

$$\mu_3(\mathbf{P}) = |\theta_2 \theta_3 \dots \theta_k|^{1/(k-1)}. \quad (2.27)$$

Произведение всех собственных значений равно детерминанту матрицы  $\mathbf{P}$  (напомним, что  $\theta_1 = 1$ ), следовательно, можно написать:

$$\mu_3(\mathbf{P}) = |\mathbf{P}|^{1/(k-1)}. \quad (2.28)$$

За исключением возможной разницы в знаке, все эти меры эквивалентны при  $k = 2$ . Для  $k > 2$  мера  $\mu_3(\mathbf{P})$  страдает тем недостатком, что она достигает своего минимального значения, равного нулю, если какие-либо две строки являются тождественно равными. Следовательно, эта мера не может установить различия независимости между поколениями для одной пары групп и для всего их множества.

Соммерс и Конлиск (1979) предложили вариант  $\mu_1(\mathbf{P})$ , который позволяет обойти трудности, обусловленные мнимостью собственных значений. Они предложили заменить  $\mathbf{P}$  на

$$\mathbf{P}^* = \frac{1}{2} (\mathbf{P} + \mathbf{\Pi}^{-1} \mathbf{P}' \mathbf{\Pi}), \quad (2.29)$$

где  $\mathbf{\Pi} = \text{diag}\{p_i\}$ . Легко проверить, что  $\mathbf{P}^*$  имеет ту же предельную структуру, что и  $\mathbf{P}$ . Марковская цепь с матрицей  $\mathbf{P}^*$  обратима, так как удовлетворяет условию

$$\mathbf{\Pi} \mathbf{P}^* = (\mathbf{P}^*)' \mathbf{\Pi}. \quad (2.30)$$

Преимущество  $\mathbf{P}^*$  состоит в том, что ее собственные значения всегда действительны. Соммерс и Конлиск (1979) привели несколько аргументов в пользу этой модифицированной меры, но с практической точки зрения она вряд ли необходима. Мы уже отмечали, что условие обратимости почти удовлетворяется для данных Гласса и Холла, поэтому неудивительно, что в этом случае  $\mu_2(\mathbf{P}) = 0,505$  и  $\mu_2(\mathbf{P}^*) = 0,506$ . В 19 других случаях, для которых Соммерс и Конлиск вычислили эти две меры, разница между ними была минимальной.

### Меры количества перемещений

Простейшей мерой количества перемещений является ожидаемая доля семейных линий, для которых меняется группа от одного поколения к следующему. За период от  $T$  до  $T + 1$  эта величина равна

$$\sum_{i=1}^k \sum_{\substack{j=1 \\ j \neq i}}^k p_i(T) p_{ij}.$$

Однако в этом случае нарушено требование, чтобы мера зависела только от  $\mathbf{P}$ .

Поскольку  $p(T)$  изменяется с  $T$ , величина этой меры зависит от  $T$ , даже если  $\mathbf{P}$  остается постоянным. Чтобы избежать этого, можно вычислить эту меру для системы в равновесном состоянии, заменяя  $p_i(T)$  на  $p_i$ , что дает

$$m_1(\mathbf{P}) = \sum_{\substack{i=1 \\ i \neq j}}^k \sum_{j=1}^k p_i p_{ij} = 1 - \sum_{i=1}^k p_i p_{ii}. \quad (2.31)$$

Эта мера учитывает все возможные переходы и не учитывает упорядочивания групп. Если группы упорядочены, то представляется целесообразным придать переходам между удаленными группами больший вес, чем переходам между близкими группами. Так как переход из группы 1 в группу 3, например, можно представить себе выполненным в два этапа — из группы 1 в группу 2 и из группы 2 в группу 3, то предлагается использовать в качестве весов число пересечений границ групп. Тогда предлагается положить в основу меры

$$\sum_{i=1}^k \sum_{j=1}^k p_i p_{ij} |i - j|,$$

что является ожидаемым числом пересечений границ групп от данного момента времени до следующего, когда система находится в установившемся состоянии. Удобно, когда мера принадлежит интервалу  $(0, 1)$ , поэтому разделим это выражение на величину его наибольшего возможного значения. Наибольшее количество перемещений наблюдается, когда каждая семейная линия перемещается на первом шаге к одной из крайних групп. Соответственно этому они будут изменяться между двумя диаметрально противоположными группами, пересекая в этом процессе  $(k - 1)$  границ. Таким образом, мы определяем

$$m_2(\mathbf{P}) = \frac{1}{k-1} \sum_{i=1}^k \sum_{j=1}^k p_i p_{ij} |i - j|. \quad (2.32)$$

Сообщество людей, в котором нет детерминированной зависимости между поколениями, не дает единственного значения любой из мер, да иначе и не может быть. Если  $k = 2$ ,  $m_1(\mathbf{P})$  и  $m_2(\mathbf{P})$  эквивалентны, при отсутствии зависимости между поколениями легко показать, что

$$m_2(\mathbf{P}) = p_{12}(1 - p_{12}) \leq \frac{1}{4}.$$

Следующий пример дает представление о качестве  $m_1(\mathbf{P})$  и  $m_2(\mathbf{P})$  для  $k = 3$  и о разнице между ними. Предположим, генеральная совокупность I имеет переходную матрицу

$$\begin{pmatrix} 0,4 & 0,3 & 0,3 \\ 0,2 & 0,5 & 0,3 \\ 0,2 & 0,2 & 0,6 \end{pmatrix}.$$

Легко убедиться, что предельная структура есть  $(7/28, 9/28, 12/28)$ ; генеральная совокупность II имеет переходную матрицу

$$\begin{pmatrix} 0,5 & 0,3 & 0,2 \\ 0,2 & 0,5 & 0,3 \\ 0,2 & 0,3 & 0,5 \end{pmatrix}$$

с предельной структурой  $(16/56, 21/56, 19/56)$ . Для генеральной совокупности I

$$m_1(\mathbf{P}) = 1 - \frac{1}{28} [7(0,4) + 9(0,5) + 12(0,6)] = 0,4821;$$

$$m_2(\mathbf{P}) = \frac{1}{56} [7\{0,3 + 2(0,3)\} + 9\{0,2 + 0,3\} +$$

$$+ 12\{2(0,2) + 0,2\}] = \frac{18}{56} = 0,3214,$$

а для генеральной совокупности II

$$m_1(\mathbf{P}) = 1 - \frac{1}{56} [16(0,5) + 21(0,5) + 19(0,5)] = \frac{1}{2}(0,5);$$

$$m_2(\mathbf{P}) = \frac{1}{112} [16\{0,3 + 2(0,2)\} + 21\{0,2 + 0,3\} + 19\{2(0,2) + 0,3\}] = \frac{35}{112} = 0,3125.$$

Таким образом, количество перемещений в генеральной совокупности I существенно выше, чем в генеральной совокупности II при измерении с помощью  $m_2(\mathbf{P})$ , но положение меняется на противоположное при использовании  $m_1(\mathbf{P})$ . Это отражает большую степень движения к крайним группам в генеральной совокупности I.

Можно обобщить (2.32), введя вместо  $|i - j|$  произвольные веса, но тогда простая интерпретация в терминах чисел пересечения границ групп будет утрачена. Тем не менее такой подход установил бы связь с другим классом мер, основанных на идеях корреляции. Сомерс и Конлиск (1979) впервые предложили, чтобы группы были расположены на некоторой шкале статуса. Тогда  $x_i$  будет показателем статуса для всех членов группы  $i$  ( $i = 1, 2, \dots, k$ ). В этом случае возможной мерой количества перемещений является коэффициент корреляции между показателями статуса групп отца и сына. Большое положительное значение корреляции относилось бы к малому количеству перемещений, тогда как большое отрицательное значение корреляции указывало бы на высокую степень мобильности. Мера количества перемещений имеет интересную и полезную особенность, состоящую в том, что отсутствие зависимости между поколениями всегда соответствует нулевой точке на шкале. Это и неудивительно, поскольку даже в более очевидном смысле она измеряет зависимость группы сына от группы отца.

Не теряя общности, мы можем выбрать шкалу для показателя  $x$  так, чтобы он имел нулевое среднее значение и единичную дисперсию, и как прежде будем считать, что система находится в установившемся состоянии. В этом случае коэффициент корреляции равен:

$$\rho = \sum_{i=1}^k \sum_{j=1}^k p_i p_{ij} x_i x_j, \quad (2.33)$$

так как число семейных линий, у которых отец в группе  $i$ , а сын в группе  $j$ , пропорционально  $p_i p_{ij}$ . Пользуясь тождеством  $(x_i - x_j)^2 = x_i^2 + x_j^2 - 2x_i x_j$ , получаем

$$\rho = 1 - \frac{1}{2} \sum_{i=1}^k \sum_{j=1}^k p_i p_{ij} (x_i - x_j)^2. \quad (2.34)$$

Если теперь сравнить (2.34) с (2.32), то можно увидеть, что единственное существенное различие в определении меры состоит в замене  $|i - j|$  на  $(x_i - x_j)^2$ . Таким образом,  $\rho$  является естественным обобщением  $m_2(\mathbf{P})$ . Проблема использования  $\rho$  заключается, конечно, в том, чтобы найти подходящую шкалу для измерения социального статуса. Сомерс и Конлиск (1979) предложили так выбирать  $x$ , чтобы мак-

симизировать (2.34) с учетом требования нулевого среднего и единичной дисперсии. Если группы упорядочены так, что  $x_1 \leq x_2 \leq \dots \leq x_n$ , можно ввести ограничение на среднее и дисперсию в задачу максимизации, применяя другие методы (см. Барлоу и др., 1972). Привлекательность  $\rho$  несколько уменьшается из-за относительной трудности вычислений и весьма сложной интерпретации максимизированной корреляции.

## Другие подходы

Всякая попытка ввести некоторую числовую (скалярную) меру, основанную на объединении всей информации, содержащейся в переходной матрице, ведет к чрезмерному упрощению. Более подробную картину процесса перемещений можно получить, жертвуя, конечно, простотой, если ввести вместо одного числа набор чисел. Прайс (1955а) высказал некоторые предложения, как увеличить множество мер. Одно из них основано на продолжительности пребывания в каждой группе. Распределение длительности пребывания в группе  $i$  для марковской модели является геометрическим со средним значением  $\mu_i = 1/(1 - p_{ii})$ . В мобильном обществе пребывание в данной группе короче, чем в немобильном. Чтобы судить о том, является ли среднее значение большим или малым, нужен некоторый эталон для сравнения. Прайс предложил воспользоваться средними значениями для соответствующего общества без зависимости между поколениями. Он сравнил среднюю длительность пребывания для сообщества с переходной матрицей  $\mathbf{P}$  и длительность для сообщества с матрицей  $1'p$ . Для данных Гласса и Холла отношение первого к последнему варьировалось между 1,11 для группы 4 и 1,59 для группы 1.

## 2.4. НЕКОТОРЫЕ ОБОБЩЕНИЯ МАРКОВСКОЙ МОДЕЛИ

### Модели, зависящие от времени

В этом и следующем разделах изучается вопрос об ослаблении влияния некоторых из тех предположений, от которых зависела предыдущая модель. Первое из таких предположений состоит в том, что вероятности перехода однородны по времени. Данные Роговой наводят на мысль о том, что это предположение для некоторых сообществ не лишено оснований, и обобщение его весьма просто.

Предположим, что переходная матрица для поколения  $T$  есть  $\mathbf{P}(T)$ . Равенство (2.2) остается в силе, если  $\mathbf{P}$  заменить на  $\mathbf{P}(T)$ . Решая это уравнение рекурсивно, получаем вместо (2.3)

$$p(T) = p(0) \prod_{i=1}^T \mathbf{P}(i). \quad (2.35)$$

Таким образом можно исследовать действие различных изменений переходной матрицы. Вообще говоря, в этом случае не будет равновесной структуры групп, и усложнение модели затруднит выбор мер мобиль-

ности. Однако при таком обобщении одно свойство простой марковской цепи сохраняется. По мере возрастания  $T$  влияние исходной структуры  $\mathbf{p}(0)$  на  $\mathbf{p}(T)$  уменьшается и в пределе исчезает. В основе этого явления лежит общий результат, полученный относительно произведения положительных матриц, который состоит в том, что

$$\lim_{T \rightarrow \infty} \prod_{i=1}^T \mathbf{P}(i)$$

представляет собой матрицу с тождественно равными строками. Наиболее общие условия, при которых этот результат имеет место, описаны Хаджналом (1976), в его же работе есть ссылка на более раннее исследование. Результат имеет место, в частности, если элементы всех переходных матриц строго положительны. Следует отметить, что если какой-либо член последовательности  $\{\mathbf{P}(T)\}$  имеет одинаковые строки, то влияние  $\mathbf{p}(0)$  и всех предшествующих членов последовательности на  $\mathbf{p}(T)$  пропадает.

Особый тип временной зависимости был представлен Прайсом (1955а). Он предположил, что расхождение между структурой групп отцов и сыновей, проявляющееся в табл. 2.2, может быть обусловлено изменением в определениях этих групп. Предположим, что в прошлом процесс был однороден во времени и достигал равновесной структуры, как это наблюдалось бы для отцов. Затем предполагается, что изменения в следующем поколении происходят по двум направлениям. Первым является изменение «соответствия» группы сына группе отца, управляемое однородной по времени матрицей  $\mathbf{P}$ ; второе — изменение собственно группы, вызванное отличием в системе классификации. Если второе преобразование имеет матрицу  $\mathbf{R}(T)$ , то наблюдаемая переходная матрица будет  $\mathbf{PR}(T)$ , и если можно было бы оценить  $\mathbf{R}(T)$  по переписи населения или по другой информации, то можно было бы в свою очередь оценить  $\mathbf{P}$  и таким образом предсказать развитие процесса. Прайс (1955а) показал, что можно построить такую матрицу  $\mathbf{R}(T)$ , которая учитывала бы наблюдаемую разницу в распределениях групп отцов и сыновей по данным Гласса и Холла. Вскоре будет показано, что кажущаяся зависимость переходной матрицы от времени может быть обусловлена разнородностью населения или немарковским характером переходов. Иллюстрацию этой зависимости получил Ходж (1966), который рассмотрел приведенную в этом разделе модель, зависящую от времени, исследуя переходную матрицу за  $n$ -шагов.

### Модель с зависимым от группы коэффициентом рождаемости

Самым сильным и неестественным ограничением является то, что каждый отец имеет строго одного сына. Мы воспользовались весьма грубым аргументом, допустив, что наша модель будет приемлемой аппроксимацией, если каждый отец имеет в среднем одного сына. Здесь излагается теория, необходимая для того, чтобы поставить это положение на прочную основу.

Предположим, что распределение числа сыновей, родившихся у члена  $j$ -й группы, равно  $P_j(i)$  ( $i = 0, 1, 2, \dots$ ), и будем считать, что эти

вероятности не зависят от времени. Зависимость распределения от  $j$  позволяет нам ввести коэффициент рождаемости, зависящий от группы. Предположим также, что размеры семей не связаны между собой. Пусть среднее число потомков мужского пола в  $j$ -й группе обозначено через  $v_j$ . Для этой модели размер генеральной совокупности не будет оставаться неизменным от поколения к поколению, поэтому мы должны пользоваться понятием численности групп вместо пропорций. Используя простые соотношения между условными вероятностями и условными математическими ожиданиями, получаем

$$\bar{n}_j(T+1) = \sum_{i=1}^k \bar{n}_i(T) v_i p_{ij} \quad (j=1, 2, \dots, k; T \geq 0), \quad (2.36)$$

где  $\bar{n}_i(0) = n_i(0)$  — начальная численность в группе  $i$ . Это уравнение позволяет получить ожидаемый размер группы таким же образом, как получаются вероятности из уравнения (2.1). Для того чтобы выполнить сравнение с простейшей моделью, необходимо вернуться к пропорциям, записав

$$p_j(T) = \bar{n}_j(T) / \sum_{i=1}^k \bar{n}_i(T) \quad (j=1, 2, \dots, k).$$

Если подходить строго, то не совсем правильно пользоваться этим обозначением, потому что отношение математических ожиданий случайных величин, вообще говоря, не равно математическому ожиданию отношения. Однако для последующего эвристического рассуждения это различие не важно, и мы будем пользоваться определенной выше величиной  $p_j(T)$  как математическим ожиданием отношения.

Если обе части (2.36) разделить на

$$\sum_{i=1}^k \bar{n}_i(T)$$

и воспользоваться тем, что

$$\sum_{i=1}^k \bar{n}_i(T+1) = \sum_{i=1}^k \bar{n}_i(T) v_i,$$

то получим следующее выражение:

$$p_j(T+1) = \sum_{i=1}^k p_i(T) v_i p_{ij} / \bar{v}(T) \quad (j=1, 2, \dots, k), \quad (2.37)$$

где

$$\bar{v}(T) = \sum_{i=1}^k p_i(T) v_i / \sum_{i=1}^k p_i(T).$$

Это уравнение было найдено Матрасом (1960b). Если группы имеют одинаковый коэффициент рождаемости, то  $v_i$  не зависит от  $i$  и (2.37) сводится к (2.1). Следовательно, можно сделать вывод, что для математических ожиданий теория, изложенная в начале этого раздела, остается справедливой и тогда, когда каждая отдельная семейная линия заменяется более реальным ветвящимся процессом.



Влияние зависимого от группы коэффициента рождаемости легче всего проследить, если  $p_j(T)$  записать в следующей форме:

$$p_j(T) = \sum_{i=1}^k p_i(T-1) v_i p_{ij} \Big/ \sum_{i=1}^k p_i(T-1) v_i. \quad (2.38)$$

Рассмотрим сначала совершенно мобильное сообщество, в котором  $p_{ij} = p_j$ . В этом случае (2.38) дает

$$p_j(T) = p_j.$$

Это означает, что социальная структура совершенно мобильного сообщества не подвержена влиянию зависимого от группы коэффициента рождаемости. Этот вывод интуитивно ясен. И наоборот, коэффициент рождаемости имеет решающее значение в неподвижном обществе. В этом случае  $p_{ij} = 1$  при  $i = j$  и  $p_{ij} = 0$  при  $i \neq j$ . Из (2.36) находим, что

$$En_j(T+1) = v_j En_j(T) = v_j^{T+1} n_j(0) \quad (j = 1, 2, \dots, k; T \geq 0).$$

Соответствующее выражение для  $p_j(T)$  имеет вид

$$p_j(T) = v_j^T n_j(0) \Big/ \sum_{i=1}^k v_i^T n_i(0) \quad (j = 1, 2, \dots, k). \quad (2.39)$$

Из этого соотношения следует, что группа с наибольшим коэффициентом рождаемости будет, очевидно, преобладать в данной генеральной совокупности. В пределе при  $T$ , стремящемся к бесконечности, будем иметь  $p_j(\infty) = 1$ , если  $j$  относится к группе с наивысшим коэффициентом рождаемости; во всех остальных случаях  $p_j(\infty) = 0$ . Для промежуточного между двумя рассмотренными крайними уровнями мобильности (от совершенно мобильного до неподвижного) следует, вероятно, ожидать, что группы с высоким коэффициентом рождаемости будут иметь со временем тенденцию к увеличению своих размеров. Некоторые сведения относительно того, до какой степени тенденция к росту будет возможной, получают из (2.38). Правая часть этого равенства является взвешенным средним вероятностей  $p_{ij}$  ( $j = 1, 2, \dots, k$ ) с положительными весами. Следовательно,

$$\min_i p_{ij} \leq p_j(T) \leq \max_i p_{ij} \quad (j = 1, 2, \dots, k) \quad (2.40)$$

для всех  $T$ . Таким образом, как бы сильно не различались группы между собой по коэффициенту рождаемости, структура групп ограничена неравенствами (2.40). Например, по данным табл. 2.1 ожидаемая доля группы 5 (квалифицированные работники физического труда и просто умственного) не может упасть ниже 0,140 или превысить 0,473 до тех пор, пока модель остается справедливой.

Изложенная выше теория, очевидно, может быть развита дальше. Матричные методы применялись в демографии в течение многих лет. Матрас (1967) указал, что эти методы можно легко приспособить к изучению изменений социальных групп, а также изменений возрастной структуры населения.

## Неоднородность переходных вероятностей

Марковская модель мобильности дает вероятностное описание развития отдельной семейной линии. Если все семейные линии имеют одинаковые переходные вероятности, то эта теория будет описывать агрегированное поведение общества, как это мы уже отмечали, когда анализировали эмпирические данные о мобильности. Однако, несмотря на наши скромные успехи в уточнении данной модели, представляется маловероятным, чтобы все семейные линии имели одну и ту же переходную матрицу. Следовательно, интересно посмотреть, какое влияние на агрегированное поведение системы оказало бы различие переходных матриц. Такая же проблема возникает при изучении зависимости между поколениями, и именно этому вопросу посвящено большинство работ. Фундаментальной в этом отношении является работа Мак-Фарланда (1970), за которой последовало количественное изучение влияния неоднородности на предельную структуру, выполненное Моррисоном и его коллегами (1971).

Одна из форм, в которой проявляется влияние неоднородности, заключается в структуре переходной матрицы за  $T$  шагов. В однородной марковской генеральной совокупности это будет степень  $T$  переходной матрицы за 1 шаг. Изучение причин отмеченных отклонений будет предметом последующего анализа.

Предположим, что доля  $x_i(h)$  членов группы  $i$  в нулевой момент времени перемещается в соответствии с переходной матрицей  $\mathbf{P}(h)$  ( $h = 1, 2, 3, \dots$ ). Тогда ожидаемая доля переходящих в группу  $j$  от общей численности группы  $i$  за время от 0 до 1 равна:

$$p_{ij} = \sum_h x_i(h) p_{ij}(h) \quad (i, j = 1, 2, \dots, k). \quad (2.41a)$$

В матричной форме это можно записать как

$$\mathbf{P} = \sum_h \mathbf{X}(h) \mathbf{P}(h), \quad (2.41b)$$

где  $\mathbf{X}(h)$  — квадратная диагональная матрица с элементами  $x_i(h)$  на главной диагонали. После  $T$  шагов члены группы  $i$ , имеющие переходную матрицу  $\mathbf{P}(h)$ , будут перемещены в соответствии с матрицей  $\mathbf{P}^T(h)$  и, следовательно, общая переходная матрица для  $T$ -шагового перехода будет

$$\mathbf{P}^{(T)} = \sum_h \mathbf{X}(h) \mathbf{P}^T(h). \quad (2.42)$$

Эта матрица стремится к некоторому пределу при  $T \rightarrow \infty$ , так как каждая из матриц  $\mathbf{P}^T(h)$  также стремится к пределу. Но в отличие от отдельно взятой марковской цепи строки матрицы  $\mathbf{P}^{(\infty)}$ , вообще говоря, не будут равными из-за неравных весов, введенных матрицами  $\{\mathbf{X}(h)\}$ .

Если бы можно было наблюдать только заданную переходную матрицу за один шаг, то марковская модель приводила бы нас к  $\mathbf{P}^T$ , которая должна задаваться равенством

$$\mathbf{P}^T = \left\{ \sum_h \mathbf{X}(h) \mathbf{P}(h) \right\}^T. \quad (2.43)$$

Теперь возникает вопрос: чем (2.43) отличается от (2.42)? Ответ в частных случаях можно найти численным методом. Возьмем простой пример и предположим, что  $k = 3$ , а  $h$  принимает только два значения; тогда

$$\mathbf{P}(1) = \begin{pmatrix} 0,7 & 0,2 & 0,1 \\ 0,2 & 0,6 & 0,2 \\ 0,2 & 0,3 & 0,5 \end{pmatrix}, \quad \mathbf{P}(2) = \begin{pmatrix} 0,5 & 0,3 & 0,2 \\ 0,3 & 0,4 & 0,3 \\ 0,3 & 0,4 & 0,3 \end{pmatrix}.$$

Далее положим, что  $x_i(h) = \frac{1}{2}$  ( $h = 1, 2$ ;  $i = 1, 2, \dots, k$ ). Получим

$$\mathbf{P} = \frac{1}{2} [\mathbf{P}(1) + \mathbf{P}(2)] = \begin{pmatrix} 0,60 & 0,25 & 0,15 \\ 0,25 & 0,50 & 0,25 \\ 0,25 & 0,35 & 0,40 \end{pmatrix},$$

$$\mathbf{P}^2 = \begin{pmatrix} 0,460 & 0,328 & 0,213 \\ 0,338 & 0,400 & 0,263 \\ 0,338 & 0,378 & 0,285 \end{pmatrix}, \quad \mathbf{P}^4 = \begin{pmatrix} 0,394 & 0,362 & 0,244 \\ 0,379 & 0,370 & 0,252 \\ 0,379 & 0,369 & 0,252 \end{pmatrix}.$$

Переходные матрицы за 2 и 4 шага, рассчитанные по (2.42), для данного примера имеют следующий вид:

$$\mathbf{P}^{(2)} = \begin{pmatrix} 0,475 & 0,320 & 0,205 \\ 0,330 & 0,415 & 0,255 \\ 0,330 & 0,370 & 0,300 \end{pmatrix}, \quad \mathbf{P}^{(4)} = \begin{pmatrix} 0,407 & 0,357 & 0,236 \\ 0,375 & 0,375 & 0,250 \\ 0,375 & 0,371 & 0,254 \end{pmatrix}.$$

Для конкретной генеральной совокупности марковская модель недооценивает тех долей, которые останутся на том же уровне как после двух, так и после четырех временных интервалов. Справедливо ли это вообще или свойственно лишь приведенному примеру?

Покажем, что такая характеристика является типичной, но не универсальной, и попытаемся проанализировать некоторые особенности матриц  $\{\mathbf{P}(h)\}$ , вызывающие это явление.

Для простоты анализа положим  $x_i(h) = x(h)$  для всех  $i$ ; достаточно будет рассмотреть случай  $T = 2$ . Сравним элементы  $(j, j)$  матриц (2.42) и (2.43). Из (2.42) следует

$$p_{jj}^{(2)} = \sum_h x(h) \sum_{l=1}^k p_{jl}(h) p_{lj}(h), \quad (2.44)$$

а из (2.43)

$$p_{jj}^{2*} = \sum_{l=1}^k \left\{ \sum_h x(h) p_{jl}(h) \right\} \left\{ \sum_h x(h) p_{lj}(h) \right\}, \quad (2.45)$$

где  $p_{jj}^{2*}$  обозначает элемент  $(j, j)$  матрицы  $\mathbf{P}^2$ . Таким образом, разность  $p_{jj}^{(2)} - p_{jj}^{2*}$  можно записать:

$$\begin{aligned} \sum_{l=1}^k \sum_h x(h) \left\{ p_{jl}(h) - \sum_m x(m) p_{jl}(m) \right\} \left\{ p_{lj}(h) - \sum_m x(m) p_{lj}(m) \right\} = \\ = \sum_{l=1}^k \text{cov}(p_{jl}, p_{lj}), \end{aligned} \quad (2.46)$$

где ковариации вычисляются по отношению к распределению  $\{x(m)\}$ . Следовательно, условия, при которых диагональные элементы матрицы  $P^{(2)}$  будут больше диагональных элементов матрицы  $P^2$ , связаны с корреляцией потоков в противоположных направлениях между каждой парой групп. Таким образом, если сравнительно большой поток из группы  $j$  в группу  $l$  сопровождается большим потоком в обратном направлении для большинства пар групп, то наблюдаемое явление произойдет. Одна совокупность условий, при которых (2.46) определено положительно, состоит в том, что матрицы симметричны, т. е. при  $p_{jl}(h) = p_{lj}(h)$  ( $j, l = 1, 2, \dots, k$ ). Матрицы, обладающие этим свойством, называются дважды стохастическими, так как суммы элементов их строк, а также столбцов равны единицам\*. Предельные структуры для таких матриц имеют равные пропорции в каждой категории. Другой пример, который встречается в следующем разделе, относится к случаю, когда  $p_{jl}(h) = p_{lj}(h) p_l(h)/p_j(h)$  ( $l, j = 1, 2, \dots, k$ ), где  $p_l(h)$  и  $p_j(h)$  — предельные вероятности, связанные с  $P(h)$ ; матрица Гласса и Холла вполне отвечает этому требованию.

Пример, в котором элементы общей переходной матрицы за 1 шаг не отличаются от элементов симметрической, а недиагональные элементы составляющих ее переходных матриц отличаются, имеет следующий вид:

$$P(1) = \begin{pmatrix} 0,5 & 0,4 & 0,1 \\ 0,2 & 0,5 & 0,3 \\ 0,2 & 0,3 & 0,5 \end{pmatrix}, \quad P(2) = \begin{pmatrix} 0,5 & 0,1 & 0,4 \\ 0,3 & 0,5 & 0,2 \\ 0,3 & 0,2 & 0,5 \end{pmatrix}$$

при  $x_i(h) = \frac{1}{2}$  ( $h = 1, 2; i = 1, 2, 3$ ). В этом случае

$$P = \begin{pmatrix} 0,50 & 0,25 & 0,25 \\ 0,25 & 0,50 & 0,25 \\ 0,25 & 0,25 & 0,50 \end{pmatrix} \quad \text{и} \quad P^2 = \begin{pmatrix} 0,375 & 0,312 & 0,312 \\ 0,312 & 0,375 & 0,312 \\ 0,312 & 0,312 & 0,375 \end{pmatrix}.$$

Но

$$P^{(2)} = \frac{1}{2} \begin{pmatrix} 0,35 & 0,43 & 0,22 \\ 0,26 & 0,42 & 0,32 \\ 0,26 & 0,38 & 0,36 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 0,40 & 0,18 & 0,42 \\ 0,36 & 0,32 & 0,32 \\ 0,36 & 0,23 & 0,41 \end{pmatrix} = \\ = \begin{pmatrix} 0,375 & 0,305 & 0,320 \\ 0,310 & 0,370 & 0,320 \\ 0,310 & 0,305 & 0,385 \end{pmatrix}.$$

Даже при двух исходных матрицах данного примера, имеющих такие различные структуры, переходные матрицы за два шага отличаются не слишком сильно. Читателю будет трудно построить примеры, в которых  $P^{(2)}$  имеет меньшие диагональные элементы, чем  $P^2$ .

Особенно большое внимание привлеч частный случай модели, объединяющей совершенно различные переходные матрицы. Это так назы-

\* В советской литературе похожие матрицы называются бистохастическими, они не обязательно симметрические. — *Примеч. ред.*

ваемая МС-модель\*, которая возникает при  $h = 1, 2$  и  $P(1) = I$ , где  $I$  — единичная матрица,  $P(2) = P_M$  — произвольная переходная матрица. Согласно этой модели некоторые члены популяции — «соседные» — никогда не перемещаются, тогда как все остальные ее члены — «кочевники» — переходят так, как в обычной марковской модели. Вероятно, такой случай перемещений более правдоподобен, когда рассматривается профессиональная мобильность или миграция, а не изменения социальных групп; в гл. 5 мы встретимся с непрерывной по времени версией модели. Блюмен и др. (1955) воспользовались МС-моделью при анализе данных о текучести рабочей силы, а Винн и Салес (1973b) — для изучения данных о мобильности в Великобритании. Мак-Колл (1971) использовал ее в качестве модели динамики доходов.

Интуитивно чувствуется, что в модели можно ожидать большей группировки на главной диагонали  $T$ -шаговой переходной матрицы, поскольку определенная часть каждой группы никогда «не двигается». Так или иначе, но именно здесь получен тот случай, который можно исследовать методом, разработанным Сингером и Шпилерманом (1977b и 1979). Они предложили определение группировки, математически трактуемое легче, чем то, которое подразумевается в приведенном выше анализе. Далее нас интересовало, превосходят или нет все диагональные элементы матрицы  $P^T$  соответствующие элементы матрицы  $P^T$ . Сингер и Шпилерман (1977b и 1979) попытались установить, что распадение на группы имеет место всякий раз, когда  $\text{trace } P^T \geq \text{trace } P^T$ , т.е. неравенству должны удовлетворять не отдельные диагональные элементы, а их суммы. Это, очевидно, более слабое требование, так как неравенство для следов матриц может выполняться даже тогда, когда для отдельных элементов имеют место обратные неравенства. Так, в приведенном выше примере  $\text{trace } P^{(2)} = 1,130$ , а  $\text{trace } P^2 = 1,125$ . Это означает группировку в соответствии с определением Сингера и Шпилермана, мы же использовали этот пример только лишь для того, чтобы показать, что для второго диагонального элемента имеет место обратное неравенство. Тем не менее, если все диагональные элементы  $P^{(T)}$  превосходят соответствующие диагональные элементы матрицы  $P^T$ , то неравенство для следов матриц будет выполняться, а знание условий, при которых оно выполняется, несомненно позволит глубже понять суть проявления группировки.

Ключевой момент применения неравенства следов матриц для группировки заключается в том, что след матрицы  $P^T$  равен сумме  $T$ -х степеней ее собственных значений. Этот факт позволяет нам трансформировать критерий группировки в некоторое утверждение относительно собственных значений. Чтобы показать возможности такого метода, рассмотрим специальный случай МС-модели, когда одна и та же часть,  $s$ , каждой группы находится в покое:

$$P = sI + (1 - s)P_M. \quad (2.47)$$

---

\* МС-модель — в оригинале «mover-stayer» model — достаточно хорошо описана в статье Р. Будона (см.: Математика в социологии. Моделирование и обработка информации. — М.: Мир, 1977). Она не является однородной марковской моделью. — *Примеч. пер.*

Отсюда собственные значения матрицы  $\mathbf{P}$  равны  $s + (1 - s)\theta_i$  ( $i = 1, 2, \dots, k$ ), где  $\{\theta_i\}$  — собственные значения  $\mathbf{P}_M$ . Следовательно

$$\text{trace } \mathbf{P}^T = \sum_{i=1}^k \{s + (1 - s)\theta_i\}^T. \quad (2.48)$$

Для МС-модели переходная матрица за  $T$  шагов в действительности равна:

$$\mathbf{P}^{(T)} = s\mathbf{I} + (1 - s)\mathbf{P}_M^T, \quad (2.49)$$

для нее

$$\text{trace } \mathbf{P}^{(T)} = \sum_{i=1}^k \{s + (1 - s)\theta_i^T\}. \quad (2.50)$$

Неравенство для следов выполняется, если

$$\sum_{i=1}^k [s + (1 - s)\theta_i^T - \{s + (1 - s)\theta_i\}^T] > 0. \quad (2.51)$$

Достаточное условие выполнения этого неравенства заключается в том, чтобы каждый член суммы был положительным. Если  $\theta_i$  положительное, действительное и не равно единице, то этот результат следует из неравенства Иенсена. Переходные матрицы с действительными положительными собственными значениями часто встречаются на практике, так что можно ожидать появления группировки, соответствующей МС-модели.

Неравенство для следов матриц может, конечно, выполняться, когда некоторые собственные значения отрицательные или мнимые. Сингер и Шпилерман (1977b и 1979) довольно подробно исследовали эту ситуацию. Им также удалось показать, что неравенство для следов может быть и обратным, как это имеет место в следующем примере. Предположим, что  $k = 4$  и пусть

$$\mathbf{P}_M = \begin{pmatrix} 0 & \frac{1}{4} & \frac{3}{4} & 0 \\ \frac{1}{4} & 0 & 0 & \frac{3}{4} \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}$$

Тогда для  $s < 0,342$  будет заниженное предсказание диагональных элементов.

Из проведенного исследования можно сделать вывод, что не существует простого и надежного способа диагностики совпадения матриц  $\mathbf{P}^{(T)}$  и  $\mathbf{P}^T$ . Избыток вероятности на диагонали матрицы  $\mathbf{P}^{(T)}$  определенно свидетельствует о возможной неоднородности генеральной совокупности. Однако позже мы увидим, что это не единственно возмож-

ное объяснение, и противоречащие примеры предупреждают нас, что неоднородность может иметь место даже и в том случае, когда нет избытка на главной диагонали.

## Модели накопленной инерции

Мы уже видели, как неоднородность генеральной совокупности могла бы объяснить появление на главной диагонали переходной матрицы за  $T$  шагов элементов со значениями, большими, чем предсказывает теория Маркова. Это не единственный тип обобщения, способный вызывать такое явление. Одно из возможных объяснений дается моделью, первоначально разработанной Мак-Гиннисом (1968). Он предложил идею «накопленной инерции», описывающей ситуацию, когда вероятность изменения группы для индивидуума уменьшается с увеличением длительности его пребывания в данной группе. Эмпирическое свидетельство в поддержку такой гипотезы для случая мобильности внутри поколения дали Майерс и его коллеги (1967), Моррисон (1967) и Ланд (1969). Оно же имеет некоторое основание и для зависимости между поколениями. На первый взгляд это предположение нарушает марковское свойство, так как вероятность перехода теперь должна зависеть от предыстории и от текущего состояния. Эту трудность можно обойти, переопределив состояние способом, который широко применяется в теории стохастических процессов.

Вместо того, чтобы определять состояния марковской цепи через понятия групп (в частности, профессиональных), теперь они определяются не только через понятия группы, но и через длительность пребывания в данной группе. Длительность пребывания измеряется в единицах временного интервала между изменениями состояния, а члены данной группы подразделяются на подгруппы в соответствии с длительностью пребывания. Каждое состояние описывается парой чисел — группой и длительностью пребывания. Процесс, определенный на этом множестве состояний, является теперь марковской цепью, поскольку мы сделали длительность пребывания частью описания текущего состояния. Пока ничто не мешает неопределенно долгому пребыванию на данном уровне, что означает поэтому допущение о бесконечном числе состояний внутри каждой группы. Это выводит нас за рамки конечных марковских цепей, но, теоретически, это не представляет непреодолимых трудностей. На практике удобно работать с конечными марковскими цепями, имеющими весьма малое число состояний. Генри и его коллеги (1971) предложили одну модификацию модели Мак-Гинниса, с помощью которой достигается эта цель. Каждая группа разделена на подгруппы в зависимости от длительности пребывания, как и раньше, но теперь все подгруппы с длительностью пребывания выше некоторого выбранного уровня объединяются в одну подгруппу. Таким образом, категории длительности пребывания могут быть, например, один, два, три, четыре года и более четырех лет.

Переходная матрица для такой модифицированной модели может иметь следующий вид, если  $X$  обозначает ненулевые элементы:





структуры. В таких случаях можно воспользоваться теорией поглощающих марковских цепей. В противном случае матрица является регулярной, и существует предельная структура, которая не зависит от начальной структуры.

Для того чтобы продемонстрировать, что накопленная инерция может дать диагональные элементы в переходной матрице за  $T$  шагов, большие, чем предсказанные теорией Маркова, рассмотрим простой пример. В системе, состоящей из двух групп, члены каждой группы разделяются в зависимости от пребывания в группе больше или меньше одного года. Тогда в системе имеются следующие четыре состояния: состояние 1 — группа 1, пребывание меньше года; состояние 2 — группа 1, пребывание больше 1 года; состояние 3 — группа 2, пребывание меньше 1 года и состояние 4 — группа 2, пребывание больше 1 года. Предположим, что для перемещений в течение года переходная матрица имеет следующий вид:

$$P = \left( \begin{array}{cc|cc} 0 & 0,6 & 0,4 & 0 \\ 0 & 0,8 & 0,2 & 0 \\ \hline 0,3 & 0 & 0 & 0,7 \\ 0,1 & 0 & 0 & 0,9 \end{array} \right). \quad (2.52)$$

Пунктирные линии определяют границы групп. Эта матрица имеет такую же форму, как и в общем случае, представленном выше, и вероятность пребывания в некоторой группе также возрастает при возрастании длительности пребывания. Предположим сначала, что в каждом состоянии имеются равные числа членов. Затем некто наблюдает переходы за один период времени и не обращает внимания на длительность пребывания, при этом он записывает одношаговую матрицу для этих двух случаев в виде

$$P^* = \left( \begin{array}{cc} \frac{1}{2} (0,6 + 0,8) & \frac{1}{2} (0,4 + 0,2) \\ \frac{1}{2} (0,3 + 0,1) & \frac{1}{2} (0,7 + 0,9) \end{array} \right) = \begin{pmatrix} 0,7 & 0,3 \\ 0,2 & 0,8 \end{pmatrix}.$$

Тогда его оценка двухшаговой матрицы будет

$$(P^*)^2 = \begin{pmatrix} 0,55 & 0,45 \\ 0,30 & 0,70 \end{pmatrix}. \quad (2.53)$$

Однако действительные переходы будут устанавливаться матрицей (2.52), для которой

$$P^2 = \left( \begin{array}{cc|cc} 0,12 & 0,48 & 0,12 & 0,28 \\ 0,06 & 0,64 & 0,16 & 0,14 \\ \hline 0,07 & 0,18 & 0,12 & 0,63 \\ 0,07 & 0,06 & 0,04 & 0,81 \end{array} \right). \quad (2.54)$$

Для того, кто не может различить категории продолжительности пребывания, наблюдаемая переходная матрица между группами за два шага будет

$$\begin{aligned}
 & (\mathbf{P}^2)^* = \\
 & = \begin{pmatrix} \frac{1}{2} (0,12 + 0,48 + 0,06 + 0,64) & \frac{1}{2} (0,12 + 0,28 + 0,16 + 0,14) \\ \frac{1}{2} (0,07 + 0,18 + 0,09 + 0,06) & \frac{1}{2} (0,12 + 0,63 + 0,04 + 0,81) \end{pmatrix} = \\
 & = \begin{pmatrix} 0,65 & 0,35 \\ 0,20 & 0,80 \end{pmatrix}. \quad (2.55)
 \end{aligned}$$

Сравнивая с (2.53), видим, что введение накопленной инерции привело к более высоким значениям диагональных элементов, что и должна была объяснить разработанная теория.

## 2.5. МАРКОВСКИЕ МОДЕЛИ ВЗАИМОСВЯЗЕЙ

Базовая марковская модель, по существу, рассматривает поведение отдельной личности. Поведение всей генеральной совокупности затем выводится в предположении, что индивидуумы действуют независимо друг от друга. Именно это предположение позволило нам интерпретировать вероятности как ожидаемые доли. На практике индивидуумы взаимодействуют, наблюдая поведение друг друга и подвергаясь влиянию того, что они видят. Подозреваемое влияние групповых мнений на поведение голосующих при выборах является примером ситуации, в которой ощущаемое состояние всех может влиять на поведение каждого. Поведение при покупках, формирование моды и общепринятых точек зрения являются другими примерами аналогичного воздействия. Цель настоящего раздела — дать некоторое понимание широкого качественного влияния взаимодействия между индивидуумами.

Класс моделей, который мы собираемся описать, впервые был предложен, по-видимому, Матрасом (1967). Он предложил вероятности переходов между  $T$  и  $T + 1$  делать зависимыми от численности в группах в момент  $T$ . Возможно, это не вполне логично с точки зрения социальной мобильности, но при изучении миграции легко себе представить, что популярность некоторого региона может зависеть от величины его населения. Термин «модель взаимодействия» был предложен Конлиском (1976 и 1978), который и разработал в основном существующую теорию. В нашем изложении мы будем следовать главным образом этому автору.

Чтобы выделить класс интересующих нас моделей, надо начать с  $\mathbf{n}(T)$ -вектора, определяющего численность данной группы в момент  $T$ . Затем мы предположим, что элементы переходной матрицы для интервала  $(T, T + 1)$  зависят от наблюдаемой величины  $\mathbf{n}(T)$ . Тогда математическое ожидание  $\mathbf{n}(T + 1)$  при данном  $\mathbf{n}(T)$  можно записать в виде

$$E\{\mathbf{n}(T + 1)|\mathbf{n}(T)\} = \mathbf{n}(T)\mathbf{P}\{\mathbf{n}(T)\}. \quad (2.56)$$

Следовательно, безусловное математическое ожидание равно:

$$\bar{n}(T+1) = E n(T) P\{n(T)\}. \quad (2.57)$$

Математическое ожидание правой части (2.57) включает нелинейную функцию от  $n(T)$  и поэтому его определение — непростое дело. Некоторого успеха можно было бы достичь, решая задачу и «в лоб», но мы воспользуемся детерминированной аппроксимацией модели Конлисса (1976). Она сводится к замене  $n(T)$  в  $P\{n(T)\}$  на  $\bar{n}(T)$ . Следовательно, в качестве аппроксимации (2.57) имеем

$$\bar{n}(T+1) = \bar{n}(T) P\{\bar{n}(T)\}. \quad (2.58)$$

Другое возможное оправдание этого шага состоит в предположении, что индивидуумы подвержены влиянию не действительной численности, а ее математического ожидания. Это не реалистично, но при больших генеральных совокупностях можно ожидать, что погрешность аппроксимации не велика.

Вернемся теперь к математическим ожиданиям отношений, полагая  $p(T) = \bar{n}(T)/N$ , и выполним анализ уравнения (2.58) в форме

$$p(T+1) = p(T) P\{p(T)\}. \quad (2.59)$$

Первым обстоятельством, которое необходимо отметить для этого процесса, является то, что он представляет собой марковскую цепь, зависящую от времени. Это позволяет сделать вывод о том, например, что по истечении достаточного большого промежутка времени структура не будет зависеть от  $p(0)$ , если только все матрицы  $P\{p(T)\}$  удовлетворяют соответствующим условиям. Кроме того, учитывая первое обстоятельство, мало что можно получить, если рассматривать соотношение (2.59).

Ясно, что функции  $P\{p(T)\}$  должны удовлетворять некоторым условиям, которые обеспечивают стохастичность матрицы  $P$ . Если условия удовлетворяются, то рекуррентные вычисления ожидаемой структуры не представляют никаких трудностей. Мы хотели бы пойти дальше и определить, какого рода поведение возможно в рамках (2.59). В частности, полезно было бы знать, чем отличается модель взаимодействия от простой марковской цепи, особенно при  $T \rightarrow \infty$ .

Один простой результат, вытекающий из (2.59), состоит в том, что если  $P(\cdot)$  — непрерывная однозначная функция, то существует по крайней мере одна стационарная структура, удовлетворяющая уравнению

$$p = pP(p). \quad (2.60)$$

Это следует из теоремы Брауера о «неподвижной точке»: при отображении пространства в себя существует, по крайней мере, одна точка, остающаяся неизменной. Конечно, таких точек может быть и много, и теорема ничего не говорит об их устойчивости. Устойчивое равновесие такое, что если структура, соответствующая ему, подвергается небольшому возмущению, то система в конце концов возвращается к ней же. В противном случае структура неустойчивая. Для регулярной марковской цепи существует одно устойчивое равновесие, но в моделях взаимо-

действия при изменяющихся численностях могут существовать оба типа равновесия — устойчивое и неустойчивое.

Полный анализ систем, описываемых уравнением (2.59), весьма сложен, и здесь еще многое предстоит сделать. Однако имеются некоторые представляющие интерес специальные случаи, для которых возможен полный анализ. Мы приведем три примера, основывающихся на работе Конлисса, с целью изучения основных качественных характеристик систем взаимодействия.

### Модели притяжения и отталкивания

Рассмотрим простейший случай. Он предполагает, что привлекательность состояния линейно зависит от его текущего размера, т.е. чем больше численность в данном состоянии, тем больше вероятность перехода к этому состоянию от какого-либо другого состояния. Модель имеет вид

$$p_{ij}\{\mathbf{p}(T)\} = a_{ij} + p_j(T)b_i \quad (i, j = 1, 2, \dots, k). \quad (2.61)$$

Параметры должны быть ограничены так, чтобы вероятности были не отрицательными, и суммы элементов по строкам должны быть равны единицам при любом векторе  $\mathbf{p}(T)$ . Из условия положительности следует, что

$$a_{ij} \geq 0, \quad b_i > \max_j a_{ij}.$$

Сумма элементов  $i$ -й строки равна:

$$\sum_{j=1}^k p_{ij}\{\mathbf{p}(T)\} = \sum_{j=1}^k a_{ij} + b_i = 1 \quad (i = 1, 2, \dots, k). \quad (2.62)$$

С помощью этого уравнения можно исключить  $b_i$  из модели и представить условие положительности в виде

$$0 \leq \sum_{\substack{h=1 \\ h \neq j}}^k a_{ih} \leq 1 \quad (i, j = 1, 2, \dots, k). \quad (2.63)$$

Если  $b_i > 0$  (т.е.  $\sum_{j=1}^k a_{ij} < 1$ ), то мы имеем модель, в которой притяжение группы возрастает с ростом ее размера. Если  $b_i < 0$ , то без нарушения (2.63) размер оказывает отпугивающее действие.

В специальном случае, когда  $b_1 = b_2 = \dots = b_k = b$ , модель сводится к марковской цепи и, следовательно, с ней можно работать, пользуясь методами, изложенными в предыдущих разделах. Чтобы рассмотреть эту ситуацию, вернемся к (2.59), которое принимает вид

$$\mathbf{p}(T+1) = \mathbf{p}(T)\{\mathbf{A} + b\mathbf{1}'\mathbf{p}(T)\} = \mathbf{p}(T)\{\mathbf{A} + b\mathbf{I}\}, \quad (2.64)$$

где  $\mathbf{A} = \{a_{ij}\}$ . Система ведет себя так, как если бы она имела постоянную переходную матрицу  $\mathbf{P} = \mathbf{A} + b\mathbf{I}$ .

В общем случае имеем

$$\mathbf{p}(T+1) = \mathbf{p}(T)\{\mathbf{A} + \mathbf{B}\mathbf{1}'\mathbf{p}(T)\} = \mathbf{p}(T)\{\mathbf{A} + (\mathbf{I} - \mathbf{A})\mathbf{1}'\mathbf{p}(T)\}, \quad (2.65)$$

где  $\mathbf{B} = \text{diag}\{b_i\}$ . Для того чтобы без учета математических сложностей изучить характеристики такой системы, обсудим сначала случай  $k = 2$ . Здесь достаточно рассмотреть только  $p_1(T)$ , так как  $p_2(T) = 1 - p_1(T)$ . Из (2.65) имеем

$$p_1(T + 1) = a_{21} + p_1(T)\{1 + a_{11} - 2a_{21} - a_{22}\} + p_1^2(T) \times \\ \times \{a_{22} + a_{21} - a_{12} - a_{11}\}. \quad (2.66)$$

О поведении системы можно судить по графику зависимости  $p_1(T + 1)$  от  $p_1(T)$ , представленному на рис. 2.1. Очевидно, что когда  $p_1(T) = 0$ , то  $p_1(T + 1) = a_{21}$ , а когда  $p_1(T) = 1$ , то  $p_1(T + 1) = 1 - a_{12}$ . Благодаря ограничению на  $\mathbf{A}$  всякая кривая лежит в единичном квадрате. Любые стационарные значения должны встречаться в точках, где кривая, заданная уравнением (2.66), пересекает диагональ  $p_1(T + 1) = p_1(T)$ . На рисунке представлен случай для матрицы  $\mathbf{A}$  вида

$$\mathbf{A} = \begin{pmatrix} 0,5 & 0,4 \\ 0,2 & 0,3 \end{pmatrix}.$$

Переходное поведение системы можно вывести из графика следующим образом. Имея  $p_1(0)$ , мы можем по графику найти  $p_1(1)$ . Затем, полагая  $T = 1$  и пользуясь графиком, находим  $p_1(2)$  и т.д. Эту последовательность можно продолжить как ряд шагов, проиллюстрированных на данном рисунке; стрелки указывают направление движения.

Ясно, что эта последовательность всегда будет сходиться к точке  $(\frac{1}{2}, \frac{1}{2})$  для любой исходной структуры. Если  $a_{21} = 0$  или  $a_{12} = 1$ , то стационарной точкой будет соответственно  $(0, 0)$  или  $(1, 1)$ . Если  $a_{21} = 1$  и  $a_{12} = 0$  при, например,  $a_{11} = a_{22}$ , то  $p_1(T + 1) = 1 - p_1(T)$ , и система выходит на предельный цикл при  $p_1(T)$ , осциллирующем между  $p_1(0)$  и  $1 - p_1(0)$ .

В случае марковской цепи аналогичный анализ сделал бы кривую на рис. 2.1 прямой линией. Кроме того, никакой другой существенной разницы между моделями нет. Вообще, обе модели стремятся к единственному положению равновесия, не зависящему от начальной структуры.

Те же выводы справедливы и для больших значений  $k$ . Конлиск (1976) доказал, что существует единственный стационарный вектор  $\mathbf{p}$ , и позже он показал (Конлиск, 1978), что система сходится к  $\mathbf{p}$  при любой начальной точке при условии, что  $b_i \geq 0$  (для всех  $i$ ) и что матрица  $\mathbf{A}$  является примитивной ( $\mathbf{A}^T > 0^*$  для некоторого  $T$ ). Мы знаем, что сходимости также имеет место при отрицательных  $b$ , если  $k > 2$ , но общего доказательства для общего случая пока не существует. Положительные  $b$  возникают в версии «притяжение» данной модели. Стационарную структуру можно получить следующим образом. Полагая в (2.65)  $\mathbf{p}(T + 1) = \mathbf{p}(T) = \mathbf{p}$ , получаем

$$\mathbf{p}(\mathbf{I} - \mathbf{A}) = \mathbf{p}(\mathbf{I} - \mathbf{A})'\mathbf{p},$$

\* Неравенство  $\mathbf{A}^T > 0$  означает, что все элементы матрицы  $\mathbf{A}^T$  положительны. — Примеч. ред.

что удовлетворяется, если

$$pA = pA1'p. \quad (2.67)$$

Любой вектор  $p$ , удовлетворяющий уравнению  $pA = \theta p$  для некоторых  $\theta$ , будет удовлетворять также и (2.67). Такой вектор существует, поскольку  $\theta$  является собственным значением матрицы  $A$ . Единственный вектор, имеющий все положительные компоненты и есть тот вектор, который связан с максимальным собственным значением матрицы  $A$ .

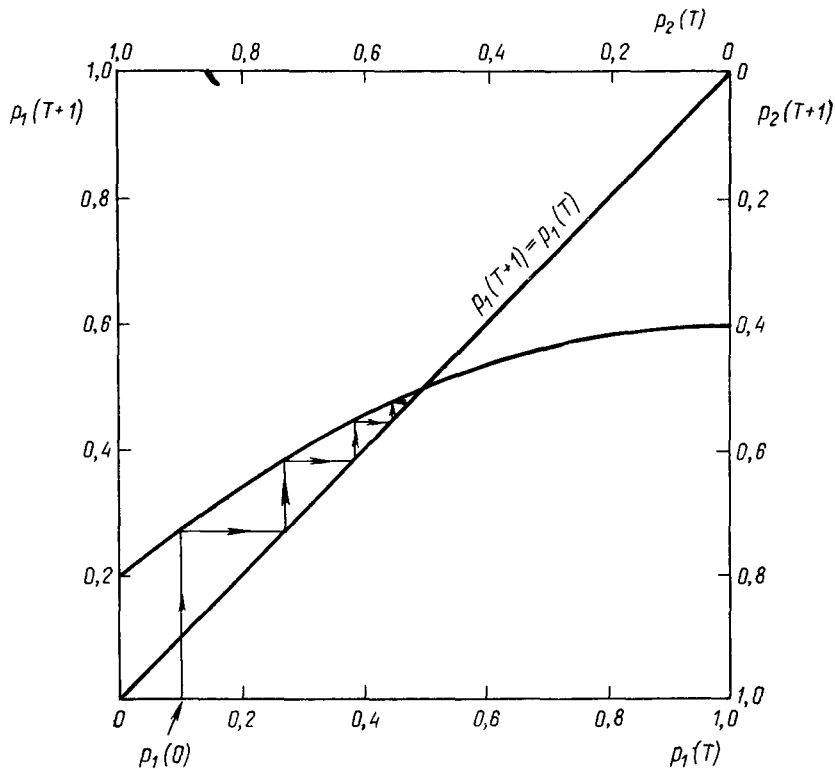


Рис. 2.1. График  $p(T+1) = p(T)P\{p(T)\}$ , иллюстрирующий приближение устойчивого состояния при  $k=2$

Следовательно, единственным решением (2.67) является левый собственный вектор матрицы  $A$ , соответствующий максимальному значению, взятому в таком масштабе, чтобы сумма его компонентов была равна единице.

### Модель неудовлетворенности

Предыдущая модель предполагает, что индивидуумы знают (или могут оценить) численности групп, включая их собственную. Легко представить себе обстоятельства, при которых каждый индивидуум реагирует только на размер своей собственной группы. Предположим,

например, что мы рассматриваем распределение заказчиков, обращающихся к поставщикам некоторого товара или услуг (гаражи, зубные врачи, розничные магазины и т.п.). Если число заказчиков, пользующихся услугами данного поставщика, возрастает, то за некоторый короткий промежуток времени качество обслуживания, вероятно, будет падать, поскольку фиксированные средства обслуживания должны обеспечивать большее число заказчиков. Более плохое обслуживание, по-видимому, должно увеличить вероятность того, что заказчик будет искать другого поставщика. В других приложениях увеличение размеров системы может иметь обратный эффект. Более крупный клуб предложит более широкий диапазон возможностей своим членам и тем самым заставит их меньше обращаться куда-нибудь еще. Оба случая охватываются модификацией линейной модели (2.61), в которой только размер собственной группы индивидуума влияет на вероятность перемещения. Мы сделаем это, выбирая

$$\begin{aligned} p_{ij} &\propto a_{ij} \quad (i \neq j), \\ p_{ii} &\propto a_{ii} + b_i p_i(T). \end{aligned} \quad (2.68)$$

Условие положительности есть  $b_i \leq -a_{ii}$  ( $i = 1, 2, \dots, k$ ), а требование равенства суммы элементов строк единице дает

$$p_{ij} = \frac{a_{ij} + \delta_{ij} b_i p_i(T)}{\sum_{j=1}^k a_{ij} + b_i p_i(T)}, \quad (2.69)$$

где  $\delta_{ij}$  — символ Кронекера. Отрицательные  $b_i$  означают, что чем больше группа, тем более вероятно, что индивидуум ее покинет; значения  $b_i$ , таким образом, являются индикатором непривлекательности размера.

При  $k = 2$

$$\begin{aligned} p_1(T+1) &= p_1(T) \left( 1 - \frac{a_{12}}{a_{11} + a_{12} + b_1 p_1(T)} \right) + \\ &+ \{1 - p_1(T)\} \left( \frac{a_{21}}{a_{21} + a_{22} + b_2 \{1 - p_1(T)\}} \right) \end{aligned} \quad (2.70)$$

и любое стационарное значение будет удовлетворять соотношению

$$p a_{12} \{a_{21} + a_{22} + b_2(1 - p)\} = (1 - p) a_{21} \{a_{11} + a_{12} + b_1 p\}. \quad (2.71)$$

Легко показать, что это уравнение имеет только один корень на интервале  $(0, 1)$ , а график, подобный рис. 2.1, покажет, что последовательность  $\{p_1(T)\}$  сходится к предельному значению из любой начальной точки. Чтобы проиллюстрировать эту модель, положим

$$A = \begin{pmatrix} 0,6 & 0,3 \\ 0,5 & 0,5 \end{pmatrix}, \quad b = (-0,4, -0,4),$$

тогда

$$\begin{aligned} p_1(T+1) &= (0,45 - 0,29p_1 + 0,2p_1^2 - 0,16p_1^3) / (0,54 + 0,12p_1 - \\ &- 0,16p_1^3), \end{aligned} \quad (2.72)$$

где  $p_1 = p_1(T)$ . Равновесное значение  $p = 0,5739$ . Анализ графика зависимости  $p_1(T + 1)$  от  $p_1(T)$  показывает, что если  $p_1(0) > 0,5739$ , то  $p_1(1) < p_1(0)$ , что и следовало ожидать на основании отпугивающего действия отрицательного  $b_1$ . После первого шага  $p_1(T)$  всегда будет лежать внутри интервала  $(0,40, 0,83)$ .

Для произвольного  $k$  Конлиску (1976) удалось показать, что существует единственное положение равновесия, хотя доказательство того, что оно достигается из любой начальной точки, у него отсутствует. Он также дает метод отыскания этого положения равновесия. Чтобы найти его, мы сначала образуем матрицу  $A_0$ , которая получается делением элементов матрицы  $A$  на сумму элементов соответствующих строк;  $A_0$  будет стохастической матрицей. Если она регулярна, то она имеет стационарную структуру, которую мы обозначим  $a_0$ . Тогда искомым равновесный вектор  $p$  равен:

$$p_i = A_i a_{0i} / (\mu A_i - b_i n_{0i}), \quad (i = 1, 2, \dots, k), \quad (2.73)$$

где  $A_i = \sum_{j=1}^k a_{ij}$ , а  $\mu$  есть единственная положительная константа, необходимая для того, чтобы  $\sum_{i=1}^k p_i = 1$ . Можно проверить, что при  $k = 2$  эта формула дает тот же ответ, что и прямой метод в приведенном выше примере. В общем случае  $\mu$  должно быть получено численными методами.

### Модель голосования

При исследовании двух приведенных моделей взаимодействия нам не удалось обнаружить их существенного отличия от простой марковской цепи. В обоих случаях система, как правило, стремилась к установившемуся состоянию, не зависящему от начальной структуры. Вычисление траектории, ведущей к положению равновесия, было несколько более сложным, но особых различий не было. Теперь перейдем к рассмотрению модели голосования, также предложенной Конлиском (1976), которая обладает совершенно новыми характеристиками.

Идею можно пояснить на примере системы, в которой принято последовательное голосование и где в каждом отдельном случае от избирателя требуется голосовать за одного из  $k$  кандидатов или предложений. После каждого тура голосования с учетом того, какую поддержку имеет каждый кандидат после него, избиратели решают, как им голосовать в следующем туре. В соответствии с моделью это происходит в два этапа. На первом этапе избиратель решает, пересматривать ли ему свою позицию. Второй этап появляется, если принято решение об изменении позиции. Тогда избиратель должен решить, который из кандидатов получит его голос в следующем туре. Главный вопрос теперь — сходится ли процедура голосования, и если сходится, то каково будет окончательное положение кандидатов. В модели Конлиска соответствующие вероятности на этих двух этапах следующие:  $P_g \{ \text{индивидуум в туре } i \text{ решает не пересматривать своего решения} \} = \alpha_i + \beta_i p_i$  ( $0 < \alpha_i < \alpha_i +$



$+ \beta_i < 1$ ),  $\text{Pr} \{ \text{индивидуум в туре } i \text{ голосует за } j | \text{ решение пересмотрено} \} \propto p_j$ .

Таким образом, данный индивидуум подвержен влиянию существующей поддержки его кандидата при решении пересматривать свою прежнюю позицию; решив изменить позицию, он отдает предпочтение тому кандидату, который уже получил достаточную поддержку. Для такой системы вероятности перехода следующие:

$$\begin{aligned} p_{ii}(\mathbf{p}(T)) &= \alpha_i + \beta_i p_i(T) + \{1 - \alpha_i - \beta_i p_i(T)\} p_i(T), \\ p_{ij}(\mathbf{p}(T)) &= \{1 - \alpha_i - \beta_i p_i(T)\} p_j(T) \quad (i \neq j). \end{aligned} \quad (2.74)$$

Для такой системы очень трудно выполнить полный анализ, но многое можно уяснить при  $k = 2$ . В этом случае имеем

$$\begin{aligned} p_1(T+1) &= p_1(T) (1 - \beta_2 - \alpha_2 + \alpha_1) + p_1^2(T) (\beta_1 + 2\beta_2 - \\ &\quad - \alpha_1 + \alpha_2) - (\beta_1 + \beta_2) p_1^3(T). \end{aligned} \quad (2.75)$$

Следовательно, точка равновесия удовлетворяет уравнению

$$(\beta_1 + \beta_2) p^3 - (\beta_1 + 2\beta_2 - \alpha_1 + \alpha_2) p^2 + (\beta_2 - \alpha_1 + \alpha_2) p = 0. \quad (2.76)$$

Легко проверить, что  $p = 0$  и  $p = 1$  являются корнями этого уравнения и, следовательно, третий корень равен:

$$p = (\beta_2 + \alpha_2 - \alpha_1) / (\beta_1 + \beta_2). \quad (2.77)$$

Рассмотрим два случая:

а) если  $\beta_2 + \alpha_2 > \alpha_1$ , а  $\beta_1 + \alpha_1 > \alpha_2$ , то  $0 < p < 1$ ;

б) в противном случае внутренней стационарной точки не существует.

Для случая (а) график  $p_1(T+1)$  представлен в левой части рис. 2.2. Из этого графика ясно, что внутреннее равновесие неустойчиво. Если  $p_1(0) \neq p$ , то структура будет сходиться к одному или другому экстремуму. В отличие от рассмотренных ранее моделей предель-

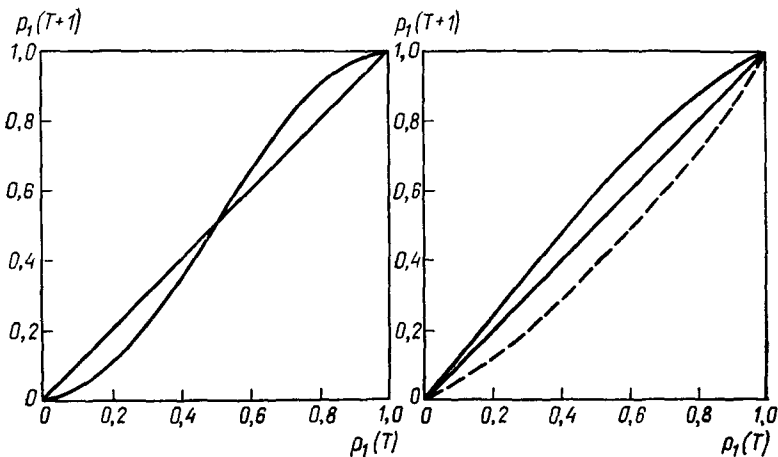


Рис. 2.2. Поведение модели голосования при  $k=2$

ное состояние здесь определяется начальным состоянием, и один из кандидатов получает все голоса. Окончательный результат зависит от того, превысит ли число голосов, поданных за кандидата 1, критическое значение, определенное (2.77). Если, например,  $\alpha_1 = \alpha_2$ ,  $\beta_1 = \frac{1}{3}$ ,  $\beta_2 = \frac{2}{3}$ , то критическое значение  $p$  равно  $\frac{2}{3}$ . Таким образом, кандидату 1 необходимо в первом туре набрать, по крайней мере, две трети голосов, чтобы стать потенциальным победителем. Система как бы имеет предубеждение против него в том смысле, что первоначальные сторонники кандидата 2 являются более преданными (т. е. они имеют большее  $\beta$ ). В случае (б) кривая примет одну из двух форм, показанных в правой части рис. 2.2. Сплошная кривая возникает, если  $\beta_2 + \alpha_2 \leq \alpha_1$ , и в этом случае кандидат 1 всегда выиграет, если он получает, по крайней мере, один голос в первом туре. Если  $\beta_1 + \alpha_1 \leq \alpha_2$ , имеет место пунктирная кривая, и кандидат 2 всегда выигрывает. Случай (б) невозможен, если  $\alpha_1 = \alpha_2$ .

Конлиском (1976) был также пролит некоторый свет на случай, когда  $k > 2$ , хотя ситуация в целом остается неясной. Предположим, мы обозначили состояния так, что  $\alpha_1 \geq \alpha_2 \geq \dots \geq \alpha_k$ . Тогда Конлиск показал, что все структуры  $\mathbf{p} = \mathbf{e}_i$  ( $i = 1, 2, \dots, k$ ) являются стационарными ( $\mathbf{e}_i$  — вектор с  $i$ -й компонентой, равной 1, остальные равны 0). Далее,  $\mathbf{p} = \mathbf{e}_i$  является локально устойчивой, если  $\alpha_i + \beta_i > \alpha_1$ ; это означает, в частности, что  $\mathbf{p} = \mathbf{e}_1$  всегда устойчиво. Если это было бы единственным устойчивым равновесием (как в случае (б)), то кандидат 1 в конечном счете получил бы все голоса, если только  $p_1(0) > 0$ . Здесь дается не полный обзор, однако он показывает возможность сходимости к экстремальной точке. Открытым остался вопрос о существовании устойчивых внутренних точек равновесия, хотя для случая  $k = 2$ , как мы знаем, внутренних точек равновесия не может быть.

### 3.1. ВВЕДЕНИЕ

В гл. 2 рассматривались замкнутые социальные системы. Это означало, либо что никакие элементы системы не покидают и не пополняют ее, либо что любые потери немедленно восстанавливаются. Нас интересовало изменение внутренней структуры системы. Предположение о замкнутости системы с практической точки зрения часто оправдано при изучении социальной и профессиональной мобильности, но нередко возникают и другие ситуации, когда потери и пополнения в системе представляют собой важную особенность процесса. В этой главе мы рассмотрим два обобщения замкнутой марковской модели, при которых потери и пополнения проявляются явно.

Характерной особенностью, общей для обеих моделей, является стохастический механизм, управляющий потерями в данной системе. Его можно изучать с помощью понятия однородной во времени вероятности потерь для каждого уровня или слоя (страта). Различие между этими двумя моделями заключено в предположениях относительно входов в систему. Чтобы пояснить это, рассмотрим указанные две модели на двух конкретных примерах, что отнюдь не ограничивает применимость этих моделей для других приложений. Примеры их применения в других областях приведены в разделе «Дополнения» в конце этой главы.

В первой модели приток в систему в момент  $T$  — либо известная величина, либо некоторая реализация известного стохастического процесса. При этих обстоятельствах конкретные численности будут случайными переменными с распределениями, зависящими от стохастических характеристик всех потоков. В качестве примера ситуации, в которой «работает» эта модель, рассмотрим систему образования, где слои (страты) представляют собой последовательные годы обучения. Модель, которую мы опишем, была применена Гани (1963) для университетской системы Австралии. Ему нужно было предсказать общее число поступивших в университеты и число ученых степеней, присуждаемых в будущем. В систему входили все обучающиеся в университетах страны студенты и аспиранты. Группами (слоями) были четыре года обучения студентов и три года аспирантуры. К концу каждого года каждый студент имел три альтернативных возможности. Он мог перейти на следующий курс, сдав необходимый экзамен, мог

остаться на второй год, если он провалился на экзамене, или мог оставить университет. Прошлые данные показали обоснованность предположения, что каждый тип перехода имеет фиксированную вероятность. Вход в систему составляют подготовленные студенты, достигшие необходимого для приема в университет возраста. Это число можно предсказать вплоть до 18 лет вперед по известным численностям каждой когорты.

На практике могут существовать известные ограничения на общий размер системы или различных ее частей, но мы пока будем ими пренебрегать. Даже если такие ограничения ввести, то модель все-таки будет полезной для прогноза потребности в числе мест в университетах. Другие примеры применения модели к системам образования даны в разделе «Дополнения».

Вторая модель находит широкое применение в области планирования кадров. Состояния системы определяют звания, возрастные группы или другие классификации работников фирмы. В этом случае мы предполагаем, что фиксирован общий размер системы, а не общее число принимаемых на работу. Тогда потребности комплектования кадров определяются потерями, а также изменением размера системы. Чтобы эта версия модели была реальной, всегда должен наблюдаться приток вновь поступающих на работу, достаточный для того, чтобы заполнить все вакансии. Другими словами, предложение должно превышать спрос. Версия модели с фиксированным общим размером системы принадлежит Янгу и Алмонду (1960), которые разработали теорию и применили эту модель при анализе структуры штата одного из университетов Великобритании. Они предполагали, что их модель будет применяться для расширяющихся систем, но при некоторых модификациях она может быть использована при любом характере изменений.

Модель Гани с фиксированным входом также может быть полезной при планировании кадров, когда поток принимаемых на работу определяется экзогенно. На практике это не часто встречающийся случай, так как немногие системы функционируют без ограничений на общий размер. Тем не менее существуют два обстоятельства, при которых имеет смысл рассмотреть модель с заданным числом принимаемых на работу. Одна ситуация — когда потребность в рабочей силе превышает ее поступление; здесь именно поток заявлений ограничивает поступления, и он определяется внешними обстоятельствами. Другая ситуация — когда мы хотим (в качестве упражнения) изучить последствия различных форм приема на работу. Пример такого рода можно найти в работах Салеса (1971) и Форбеса (1971а).

В обеих версиях модели нас будут интересовать «запасы» людей на различных «уровнях». В частности, мы будем стараться найти среднее значение и дисперсию численности людей на уровнях и понять, как структура системы зависит от устройства потоков. Позже, в гл. 6, будет рассмотрена задача управления, обратная только что сформулированной. Мы зададимся желательными значениями «запасов» и должны будем найти те «потоки» вероятностей, с помощью которых они достигаются.

## 3.2. МОДЕЛЬ СИСТЕМЫ С ЗАДАНЫМ ВХОДОМ

### Базовая модель

В качестве примера возьмем генеральную совокупность, члены которой подразделены на  $k$  слоев. Во многих приложениях слои ранжируются по старшинству, но эта особенность пока не будет рассматриваться в явной форме. Обозначения гл. 2 по возможности будут сохраняться. Так что  $n_j(T)$  обозначает число людей в группе-слое  $j$  в момент  $T$  ( $T = 0, 1, 2, \dots$ ). Начальные размеры  $n_j(0)$  ( $j = 1, 2, \dots, k$ ) групп предполагаются заданными и по определению

$$N(T) = \sum_{j=1}^k n_j(T).$$

При  $T > 0$  размеры групп представляют собой случайные переменные, и нас будут интересовать главным образом их математические ожидания. Они будут обозначаться, как и ранее, чертой над символом, представляющим данную случайную переменную; таким образом,  $\bar{n}_j(T)$  есть математическое ожидание численности группы  $j$  в момент  $T$ . Число элементов, поступающих в систему в момент  $T$ , будем обозначать  $R(T)$ . Как отмечалось, оно может быть или постоянной величиной, или случайной переменной. В последнем случае  $R(T)$  надо понимать как математическое ожидание числа элементов, поступающих в систему. Предположения относительно переходов между группами и обозначения этих переходов в данной модели социальной мобильности сохраняются. Элемент группы  $i$  переходит в группу  $j$  с вероятностью  $p_{ij}$ , но соотношение

$$\sum_{j=1}^k p_{ij} = 1$$

может больше не выполняться. В общем случае

$$\sum_{j=1}^k p_{ij} < 1,$$

поскольку в открытой системе возможны переходы, выводящие элементы за ее пределы. Вероятность потери из  $i$ -й группы в момент  $T$  обозначается  $p_{i, k+1}$ . Отметим, что

$$p_{i, k+1} = 1 - \sum_{j=1}^k p_{ij}.$$

Завершая описание модели, опишем распределение вновь поступивших по различным группам. Во многих приложениях все вновь прибывающие размещаются в самой низкой группе, но мы сделаем более общее предположение, что в  $j$ -ю группу поступает доля  $p_{0j}$ . Очевидно, что

$$\sum_{j=1}^k p_{0j} = 1,$$

и распределение  $\{p_{0j}\}$  будет называться «распределением принятых». Альтернативное предположение относительно входа в систему состоит в том, что каждый вновь поступающий распределяется в группу  $j$  с вероятностью  $p_{0j}$ . Тогда реальные числа поступающих в различные группы будут полиномиально распределенными, а не фиксированными. Наша теория охватывает оба случая, поскольку относится только к математическим ожиданиям, поэтому для нас не имеет значения,  $p_{0j}$  рассматривается как действительная доля или как ее математическое ожидание. Разница становится важной при рассмотрении дисперсий и ковариаций.

Может оказаться удобным расположить вероятности, характеризующие процесс, по стандартной форме следующим образом:

$p_{01}$	$p_{02}$	$\dots$	$p_{0k}$
$p_{11}$	$p_{12}$	$\dots$	$p_{1k} \quad p_1, \quad k+1$
$p_{21}$	$p_{22}$	$\dots$	$p_{2k} \quad p_2, \quad k+1$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$p_{k1}$	$p_{k2}$	$\dots$	$p_{kk} \quad p_k, \quad k+1$

Как и ранее,  $\mathbf{P}$  есть матрица с элементами  $\{p_{ij}\}$  ( $i, j = 1, 2, \dots, k$ );  $\mathbf{p}_0$  обозначает распределение принятых, а  $\bar{\mathbf{n}}(T)$  — вектор ожидаемого размера группы в момент  $T$ .

Теперь, когда определилось развитие системы, легко выписать уравнения, связывающие размеры групп в последовательные моменты времени. Непосредственно из определения условных математических ожиданий следует, что

$$\bar{n}_j(T+1) = \sum_{i=1}^k p_{ij} \bar{n}_i(T) + R(T+1) p_{0j} \quad (T = 1, 2, 3, \dots) \quad (3.1a)$$

$$(j = 1, 2, \dots, k)$$

или

$$\bar{\mathbf{n}}(T+1) = \bar{\mathbf{n}}(T) \mathbf{P} + R(T+1) \mathbf{p}_0. \quad (3.1b)$$

Заметим, что вновь поступающие оказывают воздействие на систему так, как если бы они все поступали в нее в момент  $T+1$ . Так как  $R(T)$  известно для всех  $T$ , это уравнение можно применять для рекуррентного вычисления ожидаемых размеров групп. Повторное применение (3.1b) дает

$$\bar{\mathbf{n}}(T) = \bar{\mathbf{n}}(0) \mathbf{P}^T + \mathbf{p}_0 \left\{ \sum_{\tau=0}^{T-1} R(T-\tau) \mathbf{P}^\tau \right\}, \quad (3.2)$$

где  $\mathbf{P}^0$  снова определяется как единичная матрица  $\mathbf{I}$ . Вероятности  $\{p_{i, k+1}\}$  не входят явным образом в эту формулу, но так как они являются дополнениями сумм элементов строк матриц  $\mathbf{P}$  и  $\mathbf{P}^\tau$ , то  $\bar{\mathbf{n}}(T)$  зависит от их значения. Когда  $R(T)$  имеет удачное математическое выражение, то оказывается возможным просуммировать матричные ряды, по-

являющиеся в (3.2), и таким образом получить решение в виде простой формулы. Это возможно, когда  $R(T)$  постоянно для всех  $T$  или, в более общем случае, если

$$R(T) = Rx^T \quad (R > 0; \quad x > 0; \quad T \geq 1). \quad (3.3)$$

Тогда имеем

$$\bar{n}(T) = \bar{n}(0) \mathbf{P}^T + Rx \mathbf{p}_0 (\mathbf{xI} - \mathbf{P})^{-1} (\mathbf{x}^T \mathbf{I} - \mathbf{P}^T) \quad (3.4)$$

при условии, что для матрицы  $\mathbf{xI} - \mathbf{P}$  существует обратная. Это имеет место, если  $x$  не равен какому-либо из собственных значений матрицы  $\mathbf{P}$ . Для матрицы с неотрицательными элементами и суммами элементов строк, строго меньшими единицы, собственные значения лежат между 0 и 1. В частности,  $x = 1$  не является собственным значением и, следовательно, (3.4) применяется для случая постоянного входа. Однако, как мы увидим, это последнее выражение играет основную роль при выяснении предельного поведения системы.

Цель следующих разделов — получить общие выводы о формах решений и их применении к исследуемым социальным процессам.

Иногда бывает удобно включить в вектор запасов математическое ожидание численности лиц, покинувших систему. Если обозначить эту величину через  $\bar{n}_{k+1}(T)$ , то можно следующим образом определить множество расширенных векторов и матриц:

$$\bar{\mathbf{n}}_a(T) = (\bar{\mathbf{n}}(T), \bar{n}_{k+1}(T)), \quad \mathbf{p}_{0a} = (\mathbf{p}_0, 0),$$

$$\mathbf{P}_a = \left( \begin{array}{c|c} \mathbf{P} & \mathbf{p}'_{k+1} \\ \hline 0 & 1 \end{array} \right).$$

Матрица  $\mathbf{P}_a$  является стохастической матрицей, соответствующей марковской цепи с одним поглощающим состоянием.

Повторение приема, ведущего к (3.1б), дает

$$\bar{\mathbf{n}}_a(T+1) = \bar{\mathbf{n}}_a(T) \mathbf{P}_a + R(T+1) \mathbf{p}_{0a}. \quad (3.5)$$

Эта форма уравнения не очень удобна в приложениях, где мы интересуемся главным образом последовательностью запасов, поскольку последний элемент  $\bar{n}_{k+1}(T)$  ведет себя иначе, чем все остальные. В частности, он растет без ограничения, даже если остальные элементы стремятся к пределам. Тем не менее эта форма, как мы покажем, полезна при изучении поведения когорт\*.

### Спектральное разложение $\bar{\mathbf{n}}(T)$

Другое возможное представление  $\bar{\mathbf{n}}(T)$ , имеющее некоторое преимущество, может быть получено при использовании стандартного результата из теории матриц, известного как теорема Сильвестра, соглас-

\* В данном разделе когорта не обязательно возрастная, она может быть числом принятых в определенном году студентов или аспирантов. — *Примеч. ред.*

но которой степень  $T$  матрицы  $\mathbf{P}$  можно представить в виде

$$\mathbf{P}^T = \sum_{r=1}^k \theta_r^T \mathbf{A}_r, \quad (3.6)$$

где  $\theta_1, \theta_2, \dots, \theta_k$  — собственные значения матрицы  $\mathbf{P}$ , которые предполагаются различными. Мы уже встречались с этим результатом для стохастической матрицы в разделе 2.3. Если среди собственных значений имеются кратные, матрицу  $\mathbf{P}^T$  все-таки можно разложить по степеням собственных значений, но выражения будут более сложными. Множество сопутствующих матриц можно найти из выражения

$$\mathbf{A}_r = \mathbf{b}'_r \mathbf{a}_r,$$

где  $\mathbf{a}_r$  — левый собственный вектор матрицы  $\mathbf{P}$ , соответствующий  $\theta_r$ , а  $\mathbf{b}'_r$  — правый собственный вектор. (Это означает, что  $\mathbf{a}_r$  и  $\mathbf{b}'_r$  удовлетворяют соответственно уравнениям  $\mathbf{a}_r \mathbf{P} = \theta_r \mathbf{a}_r$  и  $\mathbf{P} \mathbf{b}'_r = \theta_r \mathbf{b}'_r$ ; отсюда следует, что  $\mathbf{a}_r \mathbf{b}'_s = 0$  ( $r \neq s$ ) и масштаб векторов выбран так, что  $\mathbf{a}_r \mathbf{b}'_r = 1$ .) Подставляя в (3.6), получаем

$$\mathbf{P}^T = \sum_{r=1}^k \theta_r^T \mathbf{b}'_r \mathbf{a}_r = (\mathbf{b}'_1, \mathbf{b}'_2, \dots, \mathbf{b}'_k) \mathbf{D}^T \begin{pmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_k \end{pmatrix},$$

где  $\mathbf{D}^T$  — диагональная матрица с элементами  $\{\theta_r^T\}$ . Так как

$$\begin{pmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_k \end{pmatrix} (\mathbf{b}'_1, \mathbf{b}'_2, \dots, \mathbf{b}'_k) = \mathbf{I},$$

то можно записать:

$$\mathbf{P}^T = \mathbf{H} \mathbf{D}^T \mathbf{H}^{-1}, \quad (3.7)$$

где  $\mathbf{H} = (\mathbf{b}'_1, \mathbf{b}'_2, \dots, \mathbf{b}'_k)$ . Мы используем это представление в гл. 4.

Подставляя выражение для  $\mathbf{P}^T$ , данное в (3.6), в (3.2), получим

$$\bar{\mathbf{n}}(T) = \sum_{r=1}^k \left[ \theta_r^T \bar{\mathbf{n}}(0) \mathbf{A}_r + \left\{ \sum_{\tau=0}^{T-1} R(T-\tau) \theta_r^\tau \right\} \mathbf{p}_0 \mathbf{A}_r \right]. \quad (3.8)$$

Это представление особенно полезно, если сумму

$$\sum_{\tau=0}^{T-1} R(T-\tau) \theta_r^\tau$$

можно выразить простой формулой. Когда такое представление возможно, то  $\bar{\mathbf{n}}(T)$  будет суммой из  $k$  членов для всех  $T$ , и ее можно найти без необходимости рассчитывать структуру групп для всех промежуточных значений  $T$ . Уравнение (3.8) является также хорошим отправным пунктом для исследования предельного поведения  $\bar{\mathbf{n}}(T)$ . Вообще говоря, для нахождения собственных значений и множества сопут-



вующих матриц для больших  $k$  необходимо выполнить большой объем вычислений, но это обычные операции, для которых существуют стандартные программы для ЭВМ. Кроме того, в отдельных случаях, представляющих, однако, практический интерес, анализ значительно упрощается. Одним из таких примеров может служить ситуация, когда организация построена по иерархическому принципу, причем движение вниз не допускается. В следующем разделе мы рассмотрим этот частный случай и проиллюстрируем изложенную теорию.

### Иерархическая структура без понижения в должности

Предположим, что  $k$  групп упорядочены по старшинству и что переходы внутри организации допускаются только в более высокую группу. Именно такая ситуация наблюдалась при анализе австралийской университетской системы, и она является достаточно общей для того, чтобы обратить на нее особое внимание. Матрица  $\mathbf{P}$  в этом случае является верхней треугольной, т. е. с нулевыми элементами ниже главной диагонали:

$$\mathbf{P} = \begin{pmatrix} p_{11} & p_{12} \dots & p_{1k} \\ 0 & p_{22} & p_{2k} \\ \vdots & \vdots & \vdots \\ 0 & 0 & p_{kk} \end{pmatrix}.$$

Собственные значения такой матрицы равны:  $\theta_r = p_{rr}$ . Следовательно, собственные значения будут различными, если диагональные элементы матрицы  $\mathbf{P}$  различны. Здесь также можно дать явные выражения для элементов матриц  $\{A_r\}$ . Если обозначить элемент  $(i, j)$  матрицы  $A_r$  через  $A_{r,ij}$ , то

$$A_{r,ij} \begin{cases} = \sum_{m=1}^{j-i} \left( \prod_{s=i}^{i-m} p_{s,s+m} \right) \bigg/ \prod_{\substack{s=i+m-1 \\ s \neq r}}^j (p_{rr} - p_{ss}) \begin{cases} j > r > i \\ j \geq r > i \\ j > r \geq i \end{cases} \\ = 1 \text{ при } i = r = j, \\ = 0 \text{ в противном случае.} \end{cases} \quad (3.9)$$

Подстановка собственных значений и сопутствующих матриц в (3.8) дает выражение, по которому можно вычислить  $\underline{\mathbf{p}}(T)$ .

Дальнейшее упрощение можно получить, если наложить дополнительные ограничения на характер продвижений по службе, которые могут быть на практике. Предположим, например, что повышение по службе возможно только в следующую более высокую группу, т. е.  $p_{s,s+m} = 0$  для  $m > 1$ . Тогда первую сумму в (3.9) можно заменить на

$$A_{r,ij} = \prod_{s=i}^{j-1} p_{s,s+1} \bigg/ \prod_{\substack{s=i \\ s \neq r}}^j (p_{rr} - p_{ss}). \quad (3.10)$$

Если  $R(T)$  растет по закону геометрической прогрессии  $Rx^T$ , то

$$\bar{n}(T) = \sum_{r=1}^k \left\{ p_{rr}^T \bar{n}(0) \mathbf{A}_r + Rx \left( \frac{x^T - p_{rr}^T}{x - p_{rr}} \right) p_0 \mathbf{A}_r \right\} (x \neq p_{rr}). \quad (3.11)$$

Таким образом, ясно, что диагональные элементы матрицы  $\mathbf{P}$  и величины  $x$  играют решающую роль при определении развития процесса во времени.

Чтобы проиллюстрировать необходимые вычисления и получить некоторое представление о поведении иерархических систем, мы разберем численный пример с  $k = 5$ . Таблица переходных вероятностей приведена в стандартной форме:

0,75	0,25	0	0	0	
0,65	0,20	0	0	0	0,15
0	0,70	0,15	0	0	0,15
0	0	0,75	0,15	0	0,10
0	0	0	0,85	0,10	0,05
0	0	0	0	0,95	0,05

В этой таблице числа были выбраны так, чтобы отразить тот характер условий, который наблюдается при типичной иерархической системе. В случае системы образования диагональные элементы были гораздо меньше. Три четверти вновь принимаемых на работу приходят на самый низкий уровень иерархии и одна четверть — на второй снизу уровень. Вероятности увеличения,  $\{p_{i, k+1}\}$ , уменьшаются по мере продвижения к более высоким ступеням иерархии, поскольку мобильность между фирмами чаще имеет место на более низких уровнях. Увольнения с уровня 5, кроме переходов, включают уход на пенсию, так что их вероятность вполне может быть и выше 0,05, использованной в этом примере. Среднее время, проведенное на уровне  $i$ , равно  $(1 - p_{ii})^{-1}$ ; в данном примере — это три-четыре года на первых трех низших ступенях. Ожидаемое время достижения наивысшего уровня — почти 17 лет.

Матрица  $\mathbf{P}$  треугольная, и, следовательно, собственные значения равны диагональным элементам. Остается вычислить сопутствующие матрицы  $\mathbf{A}_r$ . Для этого надо подставить заданные числовые величины в (3.10). Например,

$$A_{1,11} = 1, \quad A_{1,12} = p_{12}/(p_{11} - p_{22}) = (0,20)/(-0,05) = -4,$$

$$A_{1,13} = p_{12}p_{23}/(p_{11} - p_{22})(p_{11} - p_{33}) = (0,20)(0,15)/(-0,05)(-0,10) = 6 \text{ и т. д. Окончательно матрицы имеют вид:}$$

$$\mathbf{A}_1 = \begin{pmatrix} 1 & -4 & 6 & -4,5 & 1,5 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

$$A_2 = \begin{pmatrix} 0 & 4 & -12 & 12 & -4,8 \\ 0 & 1 & -3 & 3 & -1,2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

$$A_3 = \begin{pmatrix} 0 & 0 & 6 & -9 & 4,5 \\ 0 & 0 & 3 & -4,5 & 2,25 \\ 0 & 0 & 1 & -1,5 & 0,75 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

$$A_4 = \begin{pmatrix} 0 & 0 & 0 & 1,5 & -1,5 \\ 0 & 0 & 0 & 1,5 & -1,5 \\ 0 & 0 & 0 & 1,5 & -1,5 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

$$A_5 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0,3 \\ 0 & 0 & 0 & 0 & 0,45 \\ 0 & 0 & 0 & 0 & 0,75 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Для применения (3.11) остается определить начальную структуру  $\bar{n}(0)$  по уровням и входную последовательность  $\{R(T)\}$ . Предположим, что  $n(0) = N(0) (0,40; 0,30; 0,15; 0,10; 0,05)$ , а  $R(T) = R$ . Если мы вычислим общий размер системы и размеры на отдельных уровнях в момент  $T$  ( $T > 0$ ) как кратные  $N(0)$ , нам останется только выразить  $R$  через  $N(0)$ . Для иллюстрации этих величин мы выбрали  $N(0) = 9,8333 R$ . Причина такого выбора состоит в том, что при этом  $N(\infty) = N(0)$ . Это обеспечивает отсутствие тренда в составляющей развития общего размера системы и позволяет сосредоточить внимание лишь на происходящем изменении структуры. При сделанных предположениях получаем

$$\bar{n}(T) = \sum_{r=1}^5 \left\{ p_{rr}^T n(0) A_r + R \left( \frac{1-p_{rr}^T}{1-p_{rr}} \right) p_0 A_r \right\} \quad (3.12)$$

и в пределе

$$\bar{n}(\infty) = R \sum_{r=1}^5 (1-p_{rr})^{-1} p_0 A_r. \quad (3.13)$$

В данном примере векторы  $\{p_0 A_r\}$  равны:

$$\begin{aligned} p_0 A_1 &= (0,75 \quad -3,00 \quad 4,50 \quad -3,375 \quad 1,125), \\ p_0 A_2 &= (0 \quad 3,25 \quad -9,75 \quad 9,75 \quad -3,90), \\ p_0 A_3 &= (0 \quad 0 \quad 5,25 \quad -7,875 \quad 3,9375), \\ p_0 A_4 &= (0 \quad 0 \quad 0 \quad 1,50 \quad -1,50), \\ p_0 A_5 &= (0 \quad 0 \quad 0 \quad 0 \quad 0,3375). \end{aligned}$$

Подстановка в (3.13) дает следующую предельную структуру:

$$\bar{n}(\infty) = N(0) (0,218 \quad 0,230 \quad 0,138 \quad 0,138 \quad 0,275). \quad (3.14)$$

Различие между этой и исходной структурами совершенно очевидно. Размер на высшем уровне увеличился более чем в пять раз, а размеры на более низких уровнях почти выравнялись. Эта особенность обусловлена большой средней длительностью пребывания на уровне 5, которая составляет 20 лет. Чтобы противодействовать такому чрезмерному росту на верхних ступенях, мы могли бы увеличить  $p_{rr}$  для  $r < 5$ , снизив вероятности продвижения по службе. Приведенный здесь пример иллюстрирует тот факт, что политика повышения в должности, которая кажется целесообразной, может привести к нежелательным для структуры организации последствиям. В частности, это предполагает, что избыточный рост на верхних ступенях, который наблюдается во многих организациях, может быть прямым следствием негибкой политики продвижения по должностям.

Для расчетов промежуточных структур мы должны сначала вычислить векторы  $\{n(0) A_r\}$ . Опустим детали и перейдем непосредственно к табл. 3.1, в которой представлены значения  $\bar{n}(T)$  для выбранных моментов времени.

Т а б л и ц а 3.1. Значения  $\bar{n}_j(T)/N(0)$  для примера в тексте

Уровень	T					
	0	1	2	5	10	$\infty$
1	0,400	0,336	0,295	0,239	0,220	0,218
2	0,300	0,315	0,313	0,280	0,243	0,230
3	0,150	0,158	0,165	0,174	0,159	0,138
4	0,100	0,107	0,115	0,136	0,154	0,138
5	0,050	0,058	0,065	0,091	0,137	0,275
$\sum_j \bar{n}_j(T)/N(0)$	1,000	0,974	0,953	0,920	0,913	1,000

Приближение к положению равновесия дает такую же картину, какую мы уже наблюдали в других примерах. На низких уровнях предельные математические ожидания достигаются сравнительно быстро, но на уровне 5 стремление к пределу очень незначительное. Спустя 10 лет ожидаемый размер на уровне 5 составляет лишь половину равновесного значения, тогда как на уровнях 1 и 2 почти достигнут предел.

Ожидаемый общий размер системы устойчиво уменьшается за период до 10 лет, но в конечном счете он достигает своего первоначального значения. Незначительное стремление к равновесию на более высоких уровнях и общего размера системы объясняется множителем, включающим  $(0,95)^T$  в выражении (3.12). При коэффициенте — 0,6864 его значением нельзя пренебрегать, пока  $T$  не достигнет величины порядка 100. В случаях, подобных этому, предельная структура системы не представляет большого интереса, хотя именно она показывает общее направление изменений. Иногда больший интерес могут представлять ожидаемые относительные величины численности, а не сами ожидаемые численности. Их легко получить из табл. 3.1, но они существенно не меняют общую картину.

### Предельное поведение $\bar{n}(T)$

Мы рассмотрели предельное поведение модели в особом случае с постоянным входом. Теперь исследуем это положение более подробно. В открытой системе вопрос о поведении ее структуры на длительном интервале времени осложняется наличием входного воздействия  $\{R(T)\}$ . Интуиция подсказывает, что если эта последовательность устанавливается на постоянной величине, скажем,  $R$ , то последовательность ожидаемых векторов запасов будет также стремиться к установившемуся состоянию. С другой стороны, если последовательность  $\{R(T)\}$  неопределенно возрастает, то запасы также будут возрастать и не будет установившегося состояния такого типа, как мы обсуждали в гл. 2. Однако, возможно, что относительные размеры ожидаемых запасов станут устойчивыми. Если мы определим вектор относительного запаса как

$$q(T) = \bar{n}(T) / \bar{n}(T) \mathbf{1}',$$

где  $\mathbf{1} = (1, 1, \dots, 1)$ , то уместно поставить вопрос, стремится ли последовательность  $\{q(T)\}$  к пределу при  $T \rightarrow \infty$ . Аналогично, если  $R(T)$  уменьшается со временем до нуля, то вектор запаса в пределе также обращается в нуль, но может случиться, что  $q(T)$  имеет отличное от нуля предельное значение.

Чтобы исследовать эти предельные возможности, рассмотрим уже встречавшийся специальный случай, когда  $R(T) = Rx^T$ . Он включает постоянный вход, а также охватывает возрастающие и убывающие входные последовательности.

Подставляя  $R(T) = Rx^T$  в (3.2) и деля полученное выражение на  $x^T$ , находим

$$n(T)/x^T = n(0) (P^T/x^T) + p_0 R \left\{ \sum_{\tau=0}^{T-1} (P/x)^\tau \right\}. \quad (3.15)$$

Устремим  $T \rightarrow \infty$  в правой части. Вспоминая спектральное разложение  $P^T$  в (3.6), получаем

$$P^T/x^T = \sum_{r=1}^k (\theta_r/x)^T A_r.$$

Если  $x > \max_r \theta_r$ , то это выражение в пределе обращается в нуль. При таком условии ряды во втором слагаемом (3.15) сходятся к  $(\mathbf{I} - \mathbf{P}/x)^{-1}$  и, следовательно,

$$\lim_{T \rightarrow \infty} \bar{\mathbf{n}}(T)/x^T = R \mathbf{p}_0 x (\mathbf{xI} - \mathbf{P})^{-1}. \quad (3.16)$$

Другое возможное, но эквивалентное выражение следует из (3.8):

$$\lim_{T \rightarrow \infty} \bar{\mathbf{n}}(T)/x^T = R \mathbf{p}_0 x \sum_{r=1}^k (x - \theta_r)^{-1} \mathbf{A}_r. \quad (3.17)$$

Таким образом, если  $x > \theta = \max_r \theta_r$ , то ожидаемые величины запасов будут расти по закону геометрической прогрессии вместе с входом. Предельные относительные величины запасов равны

$$\mathbf{q}(\infty) \propto \mathbf{p}_0 (\mathbf{xI} - \mathbf{P})^{-1}, \quad (3.18)$$

причем коэффициент пропорциональности определяется требованием  $\mathbf{q}(\infty) \mathbf{I}' = 1$ . Заметим, что предельная структура, определяемая (3.18), не зависит от  $\mathbf{n}(0)$  или от  $R$ . Если  $x$  очень велико, то

$$\mathbf{q}(\infty) \sim \mathbf{p}_0.$$

Это означает, что вектор приема кадров преобладает над структурой.  $\bar{\mathbf{n}}(T)$  имеет предел в единственном случае — когда  $x = 1$ , что означает постоянный вход.

Предыдущий довод «не работает», если  $x \leq \theta$ . Для матрицы  $\mathbf{P}$  известно (см., например, работу Кокса и Миллера (1965, с. 120)), что

$$0 < \theta \leq \max_{j=1}^k p_{ij},$$

где знак равенства имеет место только тогда, когда все суммы элементов строк равны. Наибольшее собственное значение будет, таким образом, меньше единицы, если все суммы элементов строк матрицы  $\mathbf{P}$  не равны единице. Для открытой системы это невозможно, поэтому мы приходим к выводу, что  $\theta < 1$ . Условие  $x > \theta$  означает, что (3.18) всегда выполняется, если вход и, следовательно, общий размер увеличиваются, а также при сокращающемся входе, если сокращение не слишком быстрое. Для выяснения того, что происходит при  $x \leq \theta$ , вернемся к (3.8) и разделим обе его части на  $\theta^T$ :

$$\bar{\mathbf{n}}(T)/\theta^T = \sum_{r=1}^k \left[ \mathbf{n}(0) \mathbf{A}_r (\theta_r/\theta)^T + \left\{ \sum_{\tau=0}^{T-1} R (x/\theta)^T (\theta_r/x)^\tau \right\} \mathbf{p}_0 \mathbf{A}_r \right]. \quad (3.19)$$

Рассмотрим сначала случай, когда  $x < \theta$ . При  $T \rightarrow \infty$  каждое слагаемое первой суммы обращается в нуль, за исключением того, для которого  $\theta_r = \theta$ , так что эта часть дает  $\mathbf{n}(0) \mathbf{A}_1$ , если мы обозначим собственные значения так, что наибольшее из них имеет индекс 1. Сумма в фигурных скобках равна

$$R \left\{ \frac{(x/\theta)^T - (\theta_r/\theta)^T}{1 - (\theta/x)} \right\}.$$

В пределе эта величина равна нулю, за исключением случая, когда  $r = 1$ ; тогда она равна

$$Rx/(\theta - x).$$

Вторая часть (3.19) дает, таким образом,

$$\frac{Rx}{(\theta - x)} \mathbf{p}_0 \mathbf{A}_1,$$

окончательно

$$\lim_{T \rightarrow \infty} \bar{\mathbf{n}}^{(T)}/\theta^T = \{\mathbf{n}(0) + Rx\mathbf{p}_0/(\theta - x)\} \mathbf{A}_1. \quad (3.20)$$

Следовательно, запасы уменьшаются с несколько меньшей скоростью, чем вход.

Осталось рассмотреть случай, когда  $x = \theta$ . Первое слагаемое в правой части (3.19) не меняется, а второе принимает вид

$$\sum_{r=1}^k \sum_{\tau=0}^{T-1} R (\theta_r/\theta)^\tau \mathbf{p}_0 \mathbf{A}_r \text{ при } T \rightarrow \infty.$$

Геометрические ряды сходятся при  $T \rightarrow \infty$ , за исключением ситуации, когда  $r = 1$ . В этом случае сумма равна  $R\mathbf{p}_0 \mathbf{A}_1 T$ . Следовательно,

$$\lim_{T \rightarrow \infty} \frac{\bar{\mathbf{n}}(T)}{T\theta^T} = R\mathbf{p}_0 \mathbf{A}_1. \quad (3.21)$$

Предельная форма  $\mathbf{q}(T)$  такая же, что и при  $x = \theta$ , и при  $x < \theta$ . Мы уже видели, что элементы сопутствующих матриц можно записать в виде  $\mathbf{A}_r = \mathbf{b}'_r \mathbf{a}_r$ , где  $\mathbf{a}_r$  и  $\mathbf{b}_r$  являются соответственно левым и правым собственными векторами, связанными с  $\theta_j$ . Подставляя  $\mathbf{A}_1 = \mathbf{b}'_1 \mathbf{a}_1$  в (3.20) и (3.21), получаем для обоих случаев

$$\mathbf{q}(\infty) \propto \mathbf{a}_1. \quad (3.22)$$

Таким образом, мы имеем две различные формы установившейся структуры, соответствующие  $x > \theta$  и  $x \leq \theta$ .

Для иллюстрации этих результатов возьмем иерархическую систему без понижения в должности и с повышением в должности на один последующий уровень, для чего мы уже нашли сопутствующие матрицы (3.10). Собственные значения равны:  $\theta_r = p_{rr}$  ( $r = 1, 2, \dots, k$ ), и далее мы предполагаем, что  $\mathbf{p}_0 = (1, 0, 0, \dots, 0)$ . Тогда из (3.17) и (3.10) следует, что для фиксированного входа

$$\bar{n}_j(\infty) = R \prod_{r=1}^{j-1} p_{r, r+1} \left/ \prod_{r=1}^j (1 - p_{rr}) \right. \quad (j = 1, 2, \dots, k), \quad (3.23)$$

где первое произведение по определению равно 1 для  $j=1$ . Эта формула показывает, как величины запасов зависят от величин вероятностей  $\{p_{rr}\}$  оставаться на данных уровнях и, в частности, от их близости к единице. Если  $R(T) = Rx^T$ , то (3.17) дает

$$q_j(\infty) \propto \prod_{r=1}^{j-1} p_{r, r+1} \left/ \prod_{r=1}^j (x - p_{rr}) \right. \quad (j = 1, 2, \dots, k) \quad (3.24)$$

при условии, что  $x > \max_r p_{rr} = \theta$ . В противном случае обращение к (3.20) или к (3.21) и (3.8) приводит к

$$q_j(\infty) \propto \prod_{r=1}^{j-1} p_{r, r+1} / \prod_{r=1}^j (\theta - p_{rr}), \quad (3.25)$$

где в произведении в знаменателе нет множителей с  $r$ , для которых  $p_{rr} = \theta$ . Сравнение с (3.23) показывает, что в знаменателе 1 заменена на  $\theta$  и что один множитель опущен.

### Применение к когорте

Если положить  $R(T) = 0$  для всех  $T$  в модели с заданным входом, то получаются другие соотношения, которые описывают движение первоначальных членов в системе. Величины запасов в момент  $T$  теперь следующие:

$$\bar{n}(T) = n(0) P^T, \quad (3.26)$$

и очевидно, что они стремятся к нулю, когда  $T$  стремится к бесконечности. Значение этого частного случая заключается в той информации, которую он дает относительно характера карьеры каждого члена когорты.

Когорта — это группа лиц, пришедших в данную организацию одновременно. Предположим, что размер группы равен  $N$  и что все ее члены находятся на уровне  $i$ . (Здесь, как и далее в этой главе, употребляется слово «уровень» вместо слова «группа», потому что основное приложение этот пример имеет в планировании кадров, но результаты находят более широкое применение.) В данном случае

$$n(0) = Ne_i,$$

где  $e_i$  — вектор, у которого  $i$ -я компонента равна единице, а остальные компоненты равны нулю. Ожидаемое распределение когорты по уровням на различных этапах ее истории можно получить из (3.26). Одной из переменных, представляющих практический интерес, является математическое ожидание части первоначальной когорты, которая остается в организации в течение различных промежутков времени. Такую переменную легко получить суммированием элементов  $\bar{n}(T)$  и делением на  $N$ . Этот подход был развит Янгом (1971) и будет рассмотрен в гл. 7 наряду с другими вопросами, касающимися продолжительности пребывания в системе. Но теперь мы займемся определением, где это возможно, средних значений и дисперсий различных величин, характеризующих развитие карьеры.

Наш анализ основан на теории поглощающих цепей Маркова, обстоятельное изложение которой можно найти в работе Кемени и Снелла (1976). Пусть  $(k+1)$ -й уровень состоит из тех, кто ушел из системы, его можно рассматривать как единственное поглощающее состоя-



ние, а остальные состояния — как переходные. Стохастическая переходная матрица для такой цепи имеет расширенную форму:

$$\mathbf{P}_a = \left[ \begin{array}{c|c} \mathbf{P} & \mathbf{p}'_{k+1} \\ \hline 0 & 1 \end{array} \right],$$

где  $\mathbf{p}'_{k+1} = (p_{1,k+1}, p_{2,k+1}, \dots, p_{k,k+1})$ . Оказывается, что большое количество информации об истории той или иной карьеры можно получить из так называемой фундаментальной матрицы  $(\mathbf{I} - \mathbf{P})^{-1}$ . Эта матрица вновь встретится в гл. 6, где она будет использована для того, чтобы определить, будет ли сохраняться структура данного уровня.

Рассмотрим сначала индивидуума, попавшего в систему на уровень  $i$  и зададимся вопросом: какова средняя продолжительность времени, которое он проведет на уровне  $j$ ? Чтобы ответить на этот вопрос, введем случайные переменные  $\{X_{ij}^{(r)}\}$ , определенные следующим образом:

$X_{ij}^{(r)} = 1$ , если попавший на уровень  $i$  спустя  $r$  единиц времени находится на уровне  $j$ ;

$X_{ij}^{(r)} = 0$  в противном случае ( $i, j = 1, 2, \dots, k; r = 0, 1, 2, \dots$ ).  
Общее время, проведенное этим индивидуумом на уровне  $j$ , равно:

$$X_{ij} = \sum_{r=0}^{\infty} X_{ij}^{(r)} \quad (i, j = 1, 2, \dots, k).$$

(Заметим, что  $X_{ij}^{(0)} = 0$ , если  $j \neq i$ , и  $X_{ij}^{(0)} = 1$  в другом случае.) Сначала найдем математическое ожидание  $X_{ij}$ :

$$E(X_{ij}) = \sum_{r=0}^{\infty} E(X_{ij}^{(r)}). \quad (3.27)$$

Из общей теории марковских цепей хорошо известно, что

$$Pr\{X_{ij}^{(r)} = 1\} = p_{ij}^{(r)},$$

где  $p_{ij}^{(r)}$  —  $(i, j)$ -й элемент матрицы  $\mathbf{P}^r$ . Следовательно,

$$E(X_{ij}^{(r)}) = p_{ij}^{(r)},$$

и далее

$$E(X_{ij}) = \sum_{r=0}^{\infty} p_{ij}^{(r)}. \quad (3.28)$$

Если мы введем матрицу  $\mathbf{X} = \{X_{ij}\}$ , то (3.28) дает

$$E(\mathbf{X}) = \sum_{r=0}^{\infty} \mathbf{P}^{(r)} = \sum_{r=0}^{\infty} \mathbf{P}^r = (\mathbf{I} - \mathbf{P})^{-1}. \quad (3.29)$$

Этот результат устанавливает соотношение между фундаментальной матрицей и ожидаемой продолжительностью пребывания или службы

по уровням. Математическое ожидание продолжительности пребывания попавшего на уровень  $i$  во всей системе равно:

$$E(X_i) = \sum_{j=1}^k E(X_{ij}) = d_i, \quad (3.30)$$

где  $d_i$  — сумма элементов  $i$ -й строки матрицы  $(\mathbf{I} - \mathbf{P})^{-1}$ . Заметим, что все эти математические ожидания в равной мере относятся как к вновь вошедшему в организацию, так и к лицу, только что повышенному в должности. Этот факт является немедленным следствием марковского свойства, которое относится ко всем попавшим на один и тот же уровень, независимо от их прошлого. Мы можем сомневаться в правильности этого предположения с точки зрения практики, но в организациях, где все или почти все принятые на работу находятся вначале на нижнем уровне иерархии, такой трудности не возникает.

Стоун (1972) рассмотрел приложение этих результатов к одной кадровой системе и выполнил расчеты для пятиуровневой матрицы, приведенной выше. Для этой матрицы находим:

					Итого по строке	
	2,86	1,90	1,14	1,14	2,29	9,33
	0	3,33	2,00	2,00	4,00	11,33
	0	0	4,00	4,00	8,00	16,00
	0	0	0	6,67	13,33	20,00
	0	0	0	0	20,00	20,00

Анализ сумм элементов строк показывает, как ожидаемый срок службы возрастает по мере продвижения вверх по уровням иерархии. Этот результат отражает уменьшение увольнений с возрастанием старшинства: если новичок «выжил» в условиях высоких темпов увольнения на нижних уровнях, то перспективы его длительного пребывания улучшаются.

Отдельные элементы в строках матрицы  $(\mathbf{I} - \mathbf{P})^{-1}$  показывают, как ожидаемая величина общего срока службы разделяется по уровням. Таким образом, ожидается, что индивидуум, поступивший в организацию на уровень 1, проведет на этом уровне 2,86 года, 1,9 года — на уровне 2 и т. д. При перемещении на уровень 2 картина меняется, отражая тот факт, что индивидуум уже прошел через уровень 1. Диагональные элементы являются математическими ожиданиями времени пребывания на каждом уровне с момента попадания на этот уровень.

При интерпретации внедиагональных элементов надо проявлять осторожность. Например, 2,29 года есть ожидаемое время, которое поступивший на уровень 1 проработает на уровне 5. Большинство из всех поступивших на уровень 1 никогда не достигнет уровня 5 и ничего не внесет в среднее время пребывания на этом уровне. К тому времени, когда удачливый новичок достигнет порога уровня 5, ожидаемое время его службы возрастет до 20 лет. Таким образом, множество элементов  $i$ -й строки можно рассматривать как типовую ожидаемую карьеру для поступившего на этот уровень, но, как и все средние значения, эти величины претерпевают существенные вариации.

Мы уже отмечали возможность того, что поступивший на нижний уровень достигнет вершины иерархии. Вероятности этого события можно рассчитать по фундаментальной матрице. Пусть  $\pi_{ij}$  обозначает вероятность того, что поступивший на уровень  $i$  проведет некоторое время на уровне  $j$  прежде, чем покинет организацию. (Вообще говоря, индивидуум может пройти через уровень  $j$  более одного раза; в простой иерархии с одноступенчатым повышением в должности он может пройти через этот уровень не более одного раза.) Если  $\mu_{ij}$  есть  $(i, j)$ -й элемент матрицы  $(\mathbf{I} - \mathbf{P})^{-1}$ , то очевидно, что

$$\mu_{ij} = \pi_{ij}\mu_{jj} + (1 - \pi_{ij}) \times 0$$

или 
$$\pi_{ij} = \mu_{ij}/\mu_{jj} \quad (i, j = 1, 2, \dots, k). \quad (3.31)$$

Таким образом, чтобы получить множество вероятностей  $\{\pi_{ij}\}$ , мы должны разделить элементы каждого столбца матрицы  $(\mathbf{I} - \mathbf{P})^{-1}$  на диагональный элемент этого столбца. Например,

$$\{\pi_{ij}\} = \begin{pmatrix} 1 & 0,57 & 0,29 & 0,17 & 0,11 \\ 0 & 1 & 0,50 & 0,30 & 0,20 \\ 0 & 0 & 1 & 0,60 & 0,40 \\ 0 & 0 & 0 & 1 & 0,67 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Диагональные элементы, очевидно, будут единичными. Наддиагональные элементы дают вероятность достижения уровня, соответствующего номеру данного столбца, при условии, что мы попали на уровень, соответствующий номеру строки. Таким образом, например, поступив на уровень 2, индивидуум с вероятностью 0,5 достигнет следующего уровня, с вероятностью 0,3 повысится в должности на два уровня и с вероятностью 0,2 достигнет вершины иерархии.

В случае матрицы, подобной только что рассмотренной, в которой  $\mathbf{P}$  имеет ненулевые элементы только на диагонали и выше ее,  $(\mathbf{I} - \mathbf{P})^{-1}$  можно найти алгебраически (см. уравнение (6.8)). Легко находим

$$\pi_{ij} = \prod_{r=i}^{j-1} p_{r, r+1}/(1 - p_{r, r}) \quad (j > i). \quad (3.32)$$

Этот результат можно было бы вывести из того, что  $p_{r, r+1}/(1 - p_{r, r})$  есть вероятность повышения в должности с уровня  $r$ , если индивидуум не останется на этом уровне. Эти формулы можно, конечно, использовать и для того, чтобы получить переходную матрицу для планирования необходимых переходов по службе, характеризующихся вероятностями  $\pi_{ij}$  и средним временем переходов  $\mu_{ij}$ .

Другой набор показателей, который можно использовать для описания развития карьеры, есть вероятности ухода с конкретных уровней. Пусть  $a_{ij}$  — вероятность того, что индивидуум, находящийся на уровне  $i$ , уйдет из организации с уровня  $j$ . Метод заключается в том, чтобы ввести одно из поглощающих состояний, соответствующих уxo-

дящим с каждого уровня, так чтобы переходная матрица была следующей:

$$\begin{array}{c|c} \mathbf{P} & \mathbf{Q} \\ \hline \mathbf{0} & \mathbf{I} \end{array},$$

где  $\mathbf{Q}$  имеет диагональные элементы, равные элементам вектора  $\mathbf{p}_{k+1}$ , а остальные элементы нулевые. Это упрощает проведение анализа, подобного изложенному выше, применительно к работникам, ограничивающим свою карьеру заданным уровнем.

Рассматриваемый метод обычно применяют для нахождения ожидаемой длительности пребывания в каждой группе ( $\mu_{ij}$ ), его легко распространить на определение их дисперсий. Действительно,

$$E(X_{ij}^2) = E\left(\sum_{r=0}^{\infty} X_{ij}^{(r)}\right)^2 = \sum_{r=0}^{\infty} \sum_{s=0}^{\infty} E(X_{ij}^{(r)} X_{ij}^{(s)}). \quad (3.33)$$

Если  $r = s$ , то

$$E(X_{ij}^{(r)})^2 = E(X_{ij}^{(r)}) = p_{ij}^{(r)}.$$

В противном случае

$$\begin{aligned} E(X_{ij}^{(r)} X_{ij}^{(s)}) &= Pr\{X_{ij}^{(r)} = 1, X_{ij}^{(s)} = 1\} = Pr\{X_{ij}^{(r)} = 1\} Pr\{X_{ij}^{(s)} = \\ &= 1 | X_{ij}^{(r)} = 1\} = p_{ij}^{(r)} p_{ij}^{(s-r)} \text{ при } s \geq r. \end{aligned} \quad (3.34)$$

Если  $s < r$ , то мы просто меняем местами  $r$  и  $s$ . Подставляя в (3.33) имеем

$$\begin{aligned} E(X_{ij}^2) &= \sum_{r=0}^{\infty} p_{ij}^{(r)} + 2 \sum_{r=0}^{\infty} \sum_{s=r+1}^{\infty} p_{ij}^{(r)} p_{ij}^{(s-r)} = \\ &= \mu_{ij} + 2 \sum_{r=0}^{\infty} p_{ij}^{(r)} (p_{ij}^{(1)} + p_{ij}^{(2)} + \dots) = \\ &= \mu_{ij} + 2\mu_{ij}(\mu_{jj} - 1) = 2\mu_{ij}\mu_{jj} - \mu_{ij}. \end{aligned} \quad (3.35)$$

Следовательно,

$$\begin{aligned} \text{var}(X_{ij}) &= 2\mu_{ij}\mu_{jj} - \mu_{ij} - \mu_{ij}^2 = \mu_{ij}^2 (2\pi_{ij} - \pi_{ij}^2) - \\ &- \mu_{jj} \pi_{ij} \quad (i, j = 1, 2, \dots, k). \end{aligned} \quad (3.36)$$

Особый интерес представляет специальный случай, когда  $i = j$ ,  $\pi_{ii} = 1$  и

$$\text{var}(X_{ii}) = \mu_{ii}^2 - \mu_{ii} = \mu_{ii}(\mu_{ii} - 1) \quad (i = 1, 2, \dots, k). \quad (3.37)$$

Этот результат показывает, в частности, что длительности пребывания на каждом уровне будут иметь большой разброс. Если  $\mu_{ii}$  довольно велико, то стандартное отклонение продолжительности пребывания будет почти таким же большим, как и математическое ожидание.

## Дисперсии и ковариации размеров групп

Теорию замкнутых систем, изложенную в гл. 2, можно обобщить на открытые системы. Для открытой системы (2.10) должно быть заменено на

$$n_j(T+1) = \sum_{i=1}^k n_{ij}(T) + n_{0j}(T+1) \quad (j=1, 2, \dots, k), \quad (3.38)$$

где  $n_{0j}(T+1)$  — число вновь поступивших в организацию на уровень  $j$  в момент  $T+1$ . Операция взятия математических ожиданий от обеих частей этого уравнения возвращает нас к (3.1). Если предположить, что входное воздействие стохастически не зависит от движения внутри системы, то

$$\begin{aligned} \text{cov} \{n_j(T+1), n_l(T+1)\} &= \text{cov} \left\{ \sum_{i=1}^k n_{ij}(T), \sum_{i=1}^k n_{il}(T) \right\} + \\ &+ \text{cov} \{n_{0j}(T+1), n_{0l}(T+1)\}. \end{aligned} \quad (3.39)$$

Первое слагаемое в правой части (3.39), по существу, такое же, как в (2.11). Его оценку можно провести так же, как в гл. 2, без какой-либо модификации. Все, что мы должны сделать, чтобы результат (2.15) был применим к рассматриваемому случаю, это добавить в (3.39) в качестве последнего слагаемого ковариацию вновь пришедших. Таким образом, в матричных обозначениях имеем

$$\mu(T+1) = \mu(T) \Pi + \mu_0(T+1), \quad (3.40)$$

где  $\mu(T)$  и  $\Pi$  определены в гл. 2. Первые  $k$  элементов вектора  $\mu_0(T+1)$  являются средними чисел поступающих на каждый уровень в момент  $T+1$ , расположенными в порядке убывания номеров уровней; остальные  $k^2$  элементов являются ковариациями этих чисел в алфавитном порядке. В наших примерах вход будет однородным по времени, так что мы не будем записывать аргумент у  $\mu_0$ , а будем писать  $\mu_0$  вместо  $\mu_0(T+1)$ . Тогда (3.40) дает

$$\mu(T) = \mu(0) \Pi^T + \mu_0 (\mathbf{I} - \Pi)^{-1} (\mathbf{I} - \Pi^T). \quad (3.41)$$

Обратная матрица для  $\mathbf{I} - \Pi$  существует потому, что наибольшее собственное значение матрицы  $\Pi$  меньше единицы. Если сейчас использовать более общий вид ошибки, рассмотренной в гл. 2, то окажется, что при соответствующих изменениях в  $\Pi$  справедливо такое же уравнение.

Вид  $\mu_0$  будет зависеть от стохастической природы входа и метода распределения по уровням. Проиллюстрируем это положение с помощью двух примеров, что приведены в начале раздела. Сначала предположим, что  $\{R(T)\}$  — последовательность независимых и одинаково распределенных случайных величин и что в момент  $T$  доля  $p_{0j}$  идет на уровень  $j$ . В этом случае первые  $k$  элементов  $\mu_0$  есть  $\bar{R}(T) p_{0j}$  ( $j=1, 2, \dots, k$ ), где  $\bar{R}(T) = ER(T)$ . Остальные  $k^2$  элементов ковариации принимают вид

$$\text{var} \{R(T)\} p_{0j} p_{0l} \quad (j, l = 1, 2, \dots, k).$$

Второе предположение относительно распределения вновь принимаемых на работу заключается в том, что каждый прибывающий должен быть отнесен к  $j$ -му уровню с вероятностью  $p_{0j}$ . В этом случае, как и ранее,  $k$  средних в векторе  $\mu_0$  равны  $\bar{R}(T)p_{0j}$  ( $j = 1, 2, \dots, k$ ), но ковариации теперь будут следующими:

$$\begin{aligned} \text{cov} \{n_{0j}(T), n_{0l}(T)\} &= \text{var} \{R(T)\} p_{0j} p_{0l} + \\ &+ \bar{R}(T) (\delta_{jl} p_{0j} - p_{0j} p_{0l}) \quad (i, j = 1, 2, \dots, k). \end{aligned} \quad (3.42)$$

Этот результат получается определением условных математических ожиданий при заданном  $R(T)$  и последующим их усреднением по распределению  $R(T)$ . При любом предположении относительно распределения вновь прибывших по уровням величины вектора  $\mu(T)$  определяются непосредственно. Рассмотрим сначала предельное поведение системы, а некоторые примеры обсудим позже.

Если вход системы однороден по времени, то  $\mu_0$  не зависит от  $T$ . При таких обстоятельствах можно исследовать предельную форму вектора  $\mu(T)$  при  $T \rightarrow \infty$ . Мы уже знаем, что предел существует для первых  $k$  элементов вектора, но можно показать, что этот результат справедлив и для всех остальных. Отсюда из (3.40) следует, что предельный вектор должен удовлетворять соотношению

$$\mu(\infty) = \mu(\infty) \Pi + \mu_0$$

или

$$\mu(\infty) = \mu_0 (\mathbf{I} - \Pi)^{-1}. \quad (3.43)$$

Как только матрица  $(\mathbf{I} - \Pi)^{-1}$  вычислена, сразу можно сравнить воздействие различных типов входов. Если  $R(T)$  — фиксированная величина  $R$ , то  $\text{var} \{R(T)\} = 0$  и  $\mu_0$  существенно упрощается.

Чтобы проиллюстрировать изложенное ранее, приведем расчеты для двух различных типов входа. В первом случае будем предполагать, что  $R(T)$  фиксировано и равно  $R$  независимо от  $T$ , а во втором случае — что  $\{R(T)\}$  — последовательность независимых пуассоновских случайных величин с постоянным средним значением, равным  $R$ . Эти предположения отражают крайние степени изменчивости входа. Для иллюстративных целей будем предполагать, что каждый вновь принимаемый на работу распределяется на уровень  $j$  с вероятностью  $p_{0j}$ . Ковариация элементов  $\mu_0$  получается из (3.42), где  $\text{var} \{R(T)\}$  равна нулю для постоянного входа и равна  $R$  для пуассоновского входа. Расчеты в табл. 3.2 относятся к обсуждавшемуся ранее в этом разделе примеру, для которого ожидаемые размеры системы на различных уровнях приведены в табл. 3.1. Далее мы предполагаем, что первоначальный размер организации равен 590 и  $\bar{R} = 60$ .

Для низших уровней иерархии приближение к равновесным значениям очень быстрое, а для наивысшего уровня — очень медленное. При входах обоих типов наблюдается значительная неопределенность в прогнозе размера на уровне 5 в отдаленном будущем. Различия между двумя типами входов сильнее всего проявляется на самых низких уровнях. До  $T = 5$  вход не оказывает никакого воздействия на верхний уровень, а далее оно едва заметно. На самом низком уровне пуас-

Таблица 3.2. Ковариационные матрицы численностей различных уровней иерархии из примера, приведенного в разделе 3.2, при (а) постоянном и (б) пуассоновском входах, причем  $N(0) = 9,833 \bar{R}(T)$ ,  $\bar{R}(T) = 60$

T	Фиксированный вход					Пуассоновский вход				
1	64,94	-41,93	0	0	0	98,69	-30,68	0	0	0
	—	86,18	-18,59	0	0	—	89,93	-18,59	0	0
2	83,82	-47,68	-4,09	0	0	131,83	-26,92	-2,99	0	0
	—	115,16	-21,51	-1,95	0	—	125,25	-20,78	-1,95	0
5	80,20	-38,20	-7,69	-1,22	-0,07	137,86	-5,70	-2,82	-0,67	-0,05
	—	124,69	-17,15	-5,07	-0,57	—	149,79	-11,25	-4,24	-0,52
10	71,54	-34,71	-6,73	-1,51	-0,24	129,97	-0,19	-0,25	-0,21	-0,07
	—	111,69	-11,99	-4,68	-1,16	—	142,27	-1,53	-1,57	-0,65
$\infty$	70,13	-34,58	-6,58	-1,43	-0,24	128,57	0	0	0	0
	—	104,79	-11,10	-3,97	-1,05	—	135,71	0	0	0
	—	—	—	-4,35	-1,86	—	—	81,43	0	0
	—	—	—	76,84	-3,74	—	—	—	81,43	0
Ожидаемые размеры уровней для $T = \infty$						128,6	135,7	81,4	81,4	162,3

\* Это значение для  $T=100$ , при котором предельное значение еще не достигнуто. Расчеты для этой таблицы выполнены Дж. Х. Поллардом.

соновский вход примерно вдвое увеличивает дисперсию по сравнению с фиксированным входом. Все ковариации либо равны нулю, либо отрицательны, но при пуассоновском входе они обращаются в нуль в пределе. Этот факт в сочетании с равенством предельных средних значений и дисперсий означает, что размеры для уровней распределены асимптотически подобно пуассоновским переменным. Поллард (1967) доказал, что это справедливо и для общего случая, когда вход состоит из последовательности независимых случайных пуассоновских величин.

Иной подход, опирающийся на производящие функции, был предложен Стаффом и Ваголкар (1971), он приводит к тому же результату. В принципе этот подход можно применять для нахождения совместных распределений размеров на различных уровнях и их моментов, но на практике слишком большая сложность вычислений снижает значение этого метода. Стафф и Ваголкар получили несколько наглядных результатов для очень простых ситуаций, но, как правило, бывает достаточно определить моменты первого и второго порядков, и их более удобно получить методом Полларда. Исключение из этого замечания — случай, когда входы  $\{R(T)\}$  представляют собой последовательность независимых пуассоновских случайных величин. В этом случае, как отмечалось выше, совместное распределение является асимптотически пуассоновским. Мы проиллюстрируем метод производящих функций на примере вывода этого результата.

Пусть  $p_j(T - \tau)$  ( $j = 1, 2, \dots, k + 1$ ) есть вероятность того, что лицо, принятое на работу в момент  $\tau$ , находится в момент  $T$  на уровне  $j$ . Будем считать, что все покинувшие систему находятся на уровне  $k + 1$ . Пусть число принятых в момент  $T - \tau$  равно  $R(T - \tau) = X$ . Тогда при заданном  $X$  все принятые в момент  $T - \tau$  будут распределены полиномиально в момент  $T$  по  $k + 1$  уровням в соответствии с вероятностями  $\{p_j(T - \tau)\}$ . Производящая функция совместных вероятностей этих чисел будет иметь вид

$$g_{\tau}(z_{1\tau}, z_{2\tau}, \dots, z_{k+1, \tau} | X) = \left( \sum_{j=1}^{k+1} p_j(T - \tau) z_{j\tau} \right)^X. \quad (3.44)$$

Безусловная производящая функция является математическим ожиданием (3.44) по отношению к распределению  $X$ . Если  $X$  имеет производящую функцию  $l(z)$ , то искома производящая функция равна:

$$g_{\tau}(z_{1\tau}, z_{2\tau}, \dots, z_{k+1, \tau}) = l \left( \sum_{j=1}^{k+1} p_j(T - \tau) z_{j\tau} \right). \quad (3.45)$$

Общая численность по уровням складывается из сохранившейся части первоначального состава сотрудников и из когорт, принятых в моменты  $1, 2, \dots, T - 1$ . Так как все  $R(T)$  являются независимыми, общее число уцелевших из этих когорт, не считая уже работавших, получается в виде произведения функций (3.45) по  $\tau$ , причем  $z_{j\tau} = z_j$  для всех  $j$ . Это дает производящую функцию

$$g(z_1, z_2, \dots, z_{k+1}) = \prod_{\tau=1}^{T-1} l \left( \sum_{j=1}^{k+1} p_j(T - \tau) z_j \right) \quad (3.46)$$



для распределения  $\mathbf{n}(T) - \mathbf{n}(0) \mathbf{P}^T$ . При больших  $T$  величиной  $\mathbf{P}^T$  можно пренебречь и тогда (3.46) можно рассматривать как производящую функцию для  $\mathbf{n}(T)$ . Это равносильно тому, что  $T$  должно быть настолько велико, чтобы почти все первоначальные члены системы смогли ее покинуть.

Если вход состоит из последовательности независимых пуассоновских случайных величин со средним значением  $R$ , то

$$l(s) = e^{R(s-1)}.$$

Следовательно,

$$\begin{aligned} g(z_1, z_2, \dots, z_{k+1}) &= \exp \left[ R \left\{ \sum_{\tau=0}^{T-1} \sum_{j=1}^{k+1} p_j(T-\tau) z_j - 1 \right\} \right] = \\ &= \exp \left\{ \sum_{j=1}^{k+1} (z_j - 1) R \sum_{\tau=0}^{T-1} p_j(T-\tau) \right\}. \end{aligned} \quad (3.47)$$

Это показывает, что размеры по уровням, за исключением оставшихся от первоначального штата, распределены независимо друг от друга с математическим ожиданием  $j$ -го уровня, равным:

$$R \sum_{\tau=0}^{T-1} p_j(T-\tau) = R \sum_{\tau=1}^T p_j(\tau).$$

Вероятности  $p_j(\tau)$  получаются из соотношения

$$\mathbf{p}(\tau) = \mathbf{p}_0 \mathbf{P}^{\tau-1}, \quad (3.48)$$

которое приводит к распределению достигнутого уровня сотрудника, принятого на работу в момент  $T - \tau$ . В пределе при  $\tau \rightarrow \infty$  вектор математических ожиданий стремится к

$$\bar{\mathbf{n}}(\infty) = R \mathbf{p}_0 \sum_{\tau=1}^{\infty} \mathbf{P}^{\tau-1} = R \mathbf{p}_0 (\mathbf{I} - \mathbf{P})^{-1}, \quad (3.49)$$

что совпадает с результатом, полученным ранее для общего случая. Теперь видно, что для пуассоновского входа эти величины полностью определяют распределение.

Метод отклонения показывает, что скорость стремления к асимптотическому распределению определяется тем, как долго первоначальные сотрудники остаются в организации. Легко заметить, что результат обобщается на случай, когда среднее значение входа зависит от времени, и это позволяет нам включать тренды или колебания во входную последовательность. Строго этот вопрос рассмотрен у Мелмана (1977а).

### Пример с периодической переходной матрицей

До сих пор мы предполагали, что переходные вероятности и распределение принимаемых на работу не зависят от времени. Но и без этого предположения (3.16) справедливо, хотя дальнейший анализ становится более сложным. В частности, могут существовать несколько предельных структур, или вовсе не существовать ни одной. Интересный

пример неоднородного процесса привел Гани в исследовании приема студентов в университет штата Мичиган. В этом университете академический год состоял из трех семестров, и переходы происходили в конце каждого семестра. Необоснованно предполагать, что вероятности перехода и вероятности распределения одинаковы для всех семестров данного года. Однако можно считать эти вероятности одинаковыми, например, для первого семестра данного и предыдущего годов. Таким образом, потребовалось считать  $\mathbf{P}$  и  $\mathbf{p}_0$  функциями времени, причем удовлетворяющими соотношениям

$$\mathbf{P}(T+3) = \mathbf{P}(T)$$

и

$$\mathbf{p}_0(T+3) = \mathbf{p}_0(T) \quad (T = 0, 1, 2).$$

Пусть  $T = 0$  относится к первому семестру первого академического года и

$$\left. \begin{aligned} \mathbf{P}(3T+j) &= \mathbf{P}_{j+1} \\ \mathbf{p}_0(3T+j) &= \mathbf{p}_{0, j+1} \end{aligned} \right\} (j=0, 1, 2; T=0, 1, 2, \dots).$$

Тогда  $\mathbf{P}_j$  и  $\mathbf{p}_{0j}$  относятся к  $j$ -му семестру любого года. В этих условиях структуру уровней можно рассчитать по следующим разностным уравнениям:

$$\left. \begin{aligned} \bar{\mathbf{n}}(3T+1) &= \bar{\mathbf{n}}(3T) \mathbf{P}_1 + R(3T+1) \mathbf{p}_{02} \\ \bar{\mathbf{n}}(3T+2) &= \bar{\mathbf{n}}(3T+1) \mathbf{P}_2 + R(3T+2) \mathbf{p}_{03} \\ \bar{\mathbf{n}}(3T+3) &= \bar{\mathbf{n}}(3T+2) \mathbf{P}_3 + R(3T+3) \mathbf{p}_{01} \end{aligned} \right\} (T=0, 1, 2, \dots). \quad (3.50)$$

Эти уравнения связывают структуру данного семестра со структурой предыдущего. Для некоторых целей удобно связать структуры одинаковых семестров последовательных лет. Так, например, для первого семестра (3.50) дает

$$\bar{\mathbf{n}}(3T+3) = \bar{\mathbf{n}}(3T) \mathbf{P}_1 \mathbf{P}_2 \mathbf{P}_3 + R(3T+1) \mathbf{p}_{02} \mathbf{P}_2 \mathbf{P}_3 + R(3T+2) \mathbf{p}_{03} \mathbf{P}_3 + R(3T+3) \mathbf{p}_{01}. \quad (3.51)$$

Аналогичные выражения можно получить для вторых и третьих семестров.

Предельное поведение системы можно исследовать с помощью методов из предыдущих разделов. Наибольший интерес представляет случай, когда  $R(T)$  постоянно или стремится к некоторому пределу, скажем,  $R$ . Тогда предельная структура уровней в первом семестре дается соотношением

$$\bar{\mathbf{n}}(\infty) = (\mathbf{I} - \mathbf{P}_1 \mathbf{P}_2 \mathbf{P}_3)^{-1} \{ \mathbf{p}_{02} \mathbf{P}_2 \mathbf{P}_3 + \mathbf{p}_{03} \mathbf{P}_3 + \mathbf{p}_{01} \} R, \quad (3.52)$$

где  $R$  — предельное значение входа. Соответствующие выражения для второго и третьего семестров можно получить перестановкой индексов в (3.50) или повторив описанную выше процедуру.

## Уравнения модели

Встречающийся на практике пример, когда возникает такая модель, описан в начале этой главы. Модель для такого примера отличается от описанной в предыдущем разделе только тем, что фиксирован общий размер организации, а не вход. Здесь вместо заданной входной последовательности  $\{R(T)\}$  мы имеем последовательность общих размеров  $\{N(T)\}$ . Как и ранее, это может быть последовательность заданных чисел или реализации известного стохастического процесса. В последнем случае символ  $N(T)$  следует понимать как ожидаемый размер. Это различие не имеет значения до тех пор, пока мы не перейдем к рассмотрению распределений размеров на разных уровнях. Янг и Алмонд (1961), предложившие эту модель, интересовались развитием организаций, и в этом отношении мы будем вначале следовать их изложению. Эта теория в значительной мере применима и к более общему случаю флуктуирующих или убывающих последовательностей.

Пусть  $M(T)$  обозначает увеличение размеров, которое произошло между  $T-1$  и  $T$ , т. е.  $M(T) = N(T) - N(T-1)$ ; ( $T = 1, 2, \dots$ ). Уравнение (3.1) остается в силе, но оно не может быть использовано в прежнем виде, так как  $\{R(T)\}$  в новой постановке задачи не известно. В любой момент времени число вновь принятых на работу должно быть достаточным как для приема до нужного размера организации, так и для восполнения потерь. Ожидаемое число вновь принимаемых в момент  $T+1$  равно:

$$\bar{R}(T+1) = M(T+1) + \sum_{i=1}^k p_{i, k+1} \bar{n}_i(T) \quad (T = 1, 2, \dots). \quad (3.53)$$

Подставляя  $\bar{R}(T+1)$  из (3.53) вместо  $R(T+1)$  в (3.1a), получаем разностные уравнения для ожидаемых размеров системы на разных уровнях:

$$\bar{n}_j(T+1) = \sum_{i=1}^k (p_{ij} + p_{i, k+1} p_{0j}) \bar{n}_i(T) + M(T+1) p_{0j} \quad (j = 1, 2, \dots, k). \quad (3.54a)$$

Если обозначить  $q_{ij} = p_{ij} + p_{i, k+1} p_{0j}$ , то это уравнение можно записать в матричной форме:

$$\bar{\mathbf{n}}(T+1) = \bar{\mathbf{n}}(T) \mathbf{Q} + \mathbf{p}_0 M(T+1), \quad (3.54b)$$

где  $\mathbf{Q}$  — матрица  $\{q_{ij}\}$ . Это матричное уравнение имеет такую же форму, как (3.1б), и в результате мы получаем

$$\bar{\mathbf{n}}(T) = \mathbf{n}(0) \mathbf{Q}^T + \mathbf{p}_0 \left\{ \sum_{\tau=0}^{T-1} M(T-\tau) \mathbf{Q}^\tau \right\}. \quad (3.55)$$

Предыдущие формулы дают все необходимое для численного исследования ожидаемой структуры. Если мы хотим изучить форму  $\bar{\mathbf{n}}(T)$  аналитическими методами, то возможны два подхода. Первый основан

на использовании спектрального представления степени матрицы, как это делалось в разделе 3.2, но в данном случае преимущества этого метода не столь велики, как при фиксированном входе. Тогда удалось получить ряд наглядных результатов для частного, но весьма важного случая треугольной матрицы  $P$ . Это было обусловлено тем, что собственные значения этой матрицы были равны ее диагональным элементам. Из данного выше определения элементов  $q_{ij}$  ясно, что матрица  $Q$ , вообще говоря, не будет треугольной и, следовательно, простота метода сразу теряется. Другой подход основан на том, что  $Q$  — стохастическая матрица, а последнее следует из соотношения

$$\sum_{j=1}^k q_{ij} = \sum_{j=1}^k p_{ij} + p_{i, k+1} \sum_{j=1}^k p_{0j} = 1 \quad (i = 1, 2, \dots, k).$$

Этот факт позволяет нам воспользоваться известными результатами относительно степеней стохастических матриц, и в частности их предельного поведения. Мы используем второй подход в следующем разделе.

Особый интерес представляет случай, когда  $M(T) = 0$  для всех  $T$ . Тогда наши уравнения описывают открытую систему с постоянным размером, но по форме они идентичны уравнениям, использованным при изучении замкнутых систем в гл. 2. Таким образом, они формально оправдывают наше замечание по поводу того, что открытую систему, в которой пополнения равны потерям, можно рассматривать как замкнутую. Каждому лицу, покидающему систему, можно поставить в соответствие вновь поступившего и эти два изменения рассматривать как одно. Таким образом, переход с уровня  $i$  на уровень  $j$  может происходить либо внутри системы, либо может быть обусловлен увольнением с уровня  $i$  и поступлением на уровень  $j$  с общей вероятностью  $p_{ij} + p_{i, k+1}p_{0j}$ .

### Точное решение для роста по закону геометрической прогрессии

Чтобы рассмотреть режим роста некоторой организации, будем предполагать, что он происходит по закону геометрической прогрессии, т.е.

$$M(T) = Mx^T \quad (T \geq 1), \quad (3.56)$$

где  $M$  — положительная константа, а  $x$  не отрицательно. Меняя величины  $M$  и  $x$ , можно получить большое разнообразие картин роста. Простейшим является случай, когда имеется всего два слоя. Мы начнем с этого случая.

При  $k = 2$  матрицу  $Q$  можно записать в виде

$$Q = \begin{pmatrix} 1 - \alpha & \alpha \\ \beta & 1 - \beta \end{pmatrix},$$

где  $1 - \alpha = p_{11} + p_{01}p_{13}$ , а  $1 - \beta = p_{22} + p_{02}p_{23}$ . Хорошо известен результат, который дает

$$Q^T = \frac{1}{\alpha + \beta} \left\{ \begin{pmatrix} \beta & \alpha \\ \beta & \alpha \end{pmatrix} + (1 - \alpha - \beta)^T \begin{pmatrix} \alpha & -\alpha \\ -\beta & \beta \end{pmatrix} \right\}. \quad (3.57)$$

Подставляя  $Q^\tau$  в (3.55) и учитывая, что  $M(T - \tau) = Mx^{T-\tau}$ , находим

$$\bar{n}(T) = \frac{n(0)}{\alpha + \beta} \left\{ \begin{pmatrix} \beta & \alpha \\ \beta & \alpha \end{pmatrix} + (1 - \alpha - \beta)^T \begin{pmatrix} \alpha & -\alpha \\ -\beta & \beta \end{pmatrix} \right\} + \\ + \frac{1}{\alpha + \beta} \sum_{\tau=0}^{T-1} Mx^{T-\tau} p_0 \left\{ \begin{pmatrix} \beta & \alpha \\ \beta & \alpha \end{pmatrix} + (1 - \alpha - \beta)^\tau \begin{pmatrix} \alpha & -\alpha \\ -\beta & \beta \end{pmatrix} \right\}. \quad (3.58)$$

Суммируя геометрические прогрессии в (3.58) и учитывая, что  $n_1(0) + n_2(0) = N(0)$  и  $p_{01} + p_{02} = 1$ , получаем вектор ожидаемых размеров в виде

$$\bar{n}(T) = \frac{1}{\alpha + \beta} \left[ \left\{ N(0) + M \frac{(x - x^{T+1})}{1 - x} \right\} (\beta, \alpha) + \right. \\ \left. + (1 - \alpha - \beta)^T (\alpha n_1(0) - \beta n_2(0), -\alpha n_1(0) + \beta n_2(0)) + \right. \\ \left. + M \left\{ \frac{x^T - (1 - \alpha - \beta)^T}{1 - (1 - \alpha - \beta)x^{-1}} \right\} (\alpha p_{01} - \beta p_{02}, -\alpha p_{01} + \beta p_{02}) \right]. \quad (3.59)$$

Общий размер организации в момент  $T$  равен:

$$N(T) = N(0) + Mx(1 - x^T)/(1 - x). \quad (3.60)$$

Переменная  $T$  входит в выражение для  $\bar{n}(T)$  через  $x^T$  и  $(1 - \alpha - \beta)^T$ . Величина  $(1 - \alpha - \beta)^T$  уменьшается при возрастании  $T$ , так как  $1 - \alpha - \beta < 1$ . Если  $x > 1$ , то слагаемые с  $x^T$  становятся доминирующими, тогда как при  $x < 1$  они в пределе обращаются в 0. В промежуточном случае при  $x = 1$  размеры на уровнях неограниченно возрастают, но их относительные значения стремятся к пределу при  $T \rightarrow \infty$ . Действительно, для  $x \leq 1$  имеем

$$\bar{n}(T) \sim \frac{N(T)}{\alpha + \beta} (\beta, \alpha). \quad (3.61)$$

Чтобы полностью оценить смысл (3.59), необходимо рассмотреть специальные случаи. Один из них обсуждается ниже; читатель найдет его полезным и для других построений.

Следующий пример представляет собой двухуровневую систему, в которой все вновь поступающие приходят на уровень 1. В системе нет понижения в должности и вероятности ухода с любого уровня одинаковы. Таблица переходных вероятностей имеет вид

1	0	
1	1	1
2	4	4
0	3	1
	4	4

Для этого примера

$$1 - \alpha = 1 - \beta = \frac{3}{4}.$$

Подстановка в (3.59) дает

$$\bar{n}(T) = N(T) \left( \frac{1}{2}, \frac{1}{2} \right) + \left( \frac{1}{2} \right)^{T+2} (n_1(0) - n_2(0), n_2(0) - n_1(0)) + \\ + Mx \left\{ \frac{x^T - \left( \frac{1}{2} \right)^T}{x - \frac{1}{2}} \right\} \left( \frac{1}{2}, -\frac{1}{2} \right). \quad (3.62)$$

Если  $x$  достаточно мало, то стремление к пределу будет быстрым. Например, если  $x = \frac{1}{2}$ , то

$$n_1(T) = \frac{1}{2} \left\{ N(0) + M \left( 1 - \left( \frac{1}{2} \right)^T \right) \right\} + \left( \frac{1}{2} \right)^{T+1} \{ n_1(0) - n_2(0) \} + \\ + \left( \frac{1}{2} \right)^{T+2} TM. \quad (3.63)$$

Разница между начальной и конечной структурами, а также величина  $M$  влияют на скорость сходимости к пределу, но их влияние не велико. Если значение  $x$  близко к 1, то, как мы увидим, стремление к пределу может быть очень медленным.

Хотя при  $k > 2$  нет простого выражения для  $\mathbf{Q}^T$ , аналогичные выводы все-таки можно сделать. Если рост происходит по геометрической прогрессии, то матричные ряды в (3.55) всегда можно просуммировать. Легко убедиться в том, что

$$(x\mathbf{I} - \mathbf{Q}) \left( \sum_{\tau=0}^{T-1} \mathbf{Q}^\tau x^{T-\tau} \right) = x(x^T \mathbf{I} - \mathbf{Q}^T),$$

и, следовательно, если  $x\mathbf{I} - \mathbf{Q}$  имеет обратную матрицу, то

$$\sum_{\tau=0}^{T-1} \mathbf{Q}^\tau x^{T-\tau} = x(x\mathbf{I} - \mathbf{Q})^{-1} (x^T \mathbf{I} - \mathbf{Q}^T).$$

Таким образом,

$$\sum_{\tau=0}^{T-1} \mathbf{Q}^\tau x^{T-\tau} = x(x\mathbf{I} - \mathbf{Q})^{-1} (x^T \mathbf{I} - \mathbf{Q}^T). \quad (3.64)$$

Обратной матрицы для  $x\mathbf{I} - \mathbf{Q}$  не существует, если ее детерминант обращается в нуль. Это происходит в том случае, когда  $x$  равен собственному значению матрицы  $\mathbf{Q}$ . Для стохастической матрицы известно, что, по крайней мере, одно собственное значение равно единице, а остальные по модулю лежат между нулем и единицей. Следовательно, (3.64) выполняется для всех  $x > 1$ , а при  $x < 1$  почти для всех, за исключением самого большого  $k - 1$  значения  $x$ . Для числовых расчетов с  $x < 1$  целесообразно воспользоваться рекуррентной формулой (3.54б). Польза представления (3.64) состоит в том, что оно дает прямую информацию о переходном поведении системы. Это поведение зависит от членов с  $x^T$  и с  $\mathbf{Q}^T$ . Из общей теории марковских цепей известно, что если  $\mathbf{Q}$  невырожденная, то  $\mathbf{Q}^T$  стремится к пределу с увеличением  $T$ . Как и в

Таблица 3.3. Ожидаемые размеры уровней для организации с ростом по геометрической прогрессии и переходными вероятностями, данными в разделе 3.2, % от суммы

T	x	100 q <sub>i</sub> (T)				
		i				
		1	2	3	4	5
0	0	40	30	15	10	5
	$\frac{1}{2}$	40	30	15	10	5
	1	40	30	15	10	5
	2	40	30	15	10	5
1	0	35,6	32,2	15,8	10,8	5,8
	$\frac{1}{2}$	37,4	31,8	15,0	10,2	5,5
	1	39,1	31,5	14,3	9,8	5,2
	2	42,1	31,0	13,1	9,0	4,8
2	0	32,5	32,8	16,6	11,5	6,5
	$\frac{1}{2}$	34,8	32,8	15,7	10,7	6,1
	1	38,4	32,4	14,2	9,6	5,4
	2	46,6	31,3	10,9	7,2	4,1
5	0	27,8	31,1	18,2	13,8	9,1
	$\frac{1}{2}$	29,2	32,0	17,7	12,8	8,3
	1	36,3	33,2	14,8	9,6	6,1
	2	59,6	31,6	5,4	2,2	1,3
10	0	24,8	27,5	17,7	16,2	13,8
	$\frac{1}{2}$	25,3	28,2	17,9	15,7	12,8
	1	33,8	32,4	15,8	10,6	7,4
	2	63,6	31,9	3,9	0,6	0,1
25	0	22,3	23,8	14,6	15,5	23,8
	$\frac{1}{2}$	22,4	23,9	14,7	15,6	23,3
	1	29,3	29,4	15,9	13,0	13,0
	2	63,8	31,9	3,8	0,5	0,0
50	0	21,8	23,0	13,8	13,9	27,3
	$\frac{1}{2}$	21,8	23,1	13,9	14,0	27,3
	1	26,2	26,8	15,2	13,6	18,1
	2	63,8	31,9	3,8	0,5	0,0

случае  $k = 2$ , поведение системы при больших  $T$  существенно зависит от значений  $x$ .

Чтобы проиллюстрировать эту теорию, возьмем  $k = 5$  и воспользуемся теми же переходными вероятностями, что и в примере из раздела 3.2. Пусть темп роста будет

$$M(T) = \frac{1}{10} N(0) x^T$$

при  $x = 0$ ;  $x = \frac{1}{2}$ ;  $x = 1$  и  $x = 2$ . Поскольку общий размер организации изменяется, затабулированы (см. табл. 3.3) относительные ожидаемые размеры, заданные соотношением

$$q_j(T) = \bar{n}_j(T)/N(T) \quad (j = 1, 2, \dots, k).$$

При  $x = 0$  нет никакого развития и стремление к пределу весьма медленное, особенно на самом высоком уровне иерархии. Это обусловлено главным образом большой разницей между начальной и предельной структурами. При  $x = \frac{1}{2}$  максимальное увеличение общего размера составляет 10%, но скорость стремления к предельной структуре едва ли является существенной. Случай, когда  $x = 1$ , демонстрирует очень медленное стремление к пределу, хотя здесь снова более низкие уровни иерархии стремятся к своим предельным значениям быстрее, чем более высокие. В этом примере величина для пятого уровня через 50 лет составляет чуть больше половины своего предельного значения. При таком характере развития структура не достигнет равновесия за какой-либо интервал времени, представляющий практический интерес. И наоборот, при  $x = 2$  предел достигается очень быстро, хотя это уже другой предел. Немногие организации способны поддерживать такой высокий темп роста и в течение столь длительного времени, чтобы конечная структура сохраняла бы практическое значение.

### Предельная структура

Выражая (3.54б) через  $q$ , имеем

$$q(T+1) \{1 + M(T+1)/N(T)\} = q(T) Q + p_0 M(T+1)/N(T). \quad (3.65)$$

Следовательно, возможность существования предельной структуры  $q$  зависит от поведения  $M(T+1)/N(T)$ , в частности, эта величина должна стремиться к пределу с возрастанием  $T$ .

Пусть

$$\lim_{T \rightarrow \infty} M(T+1)/N(T) = c,$$

тогда предельная структура, если она существует, удовлетворяет уравнению

$$q(1+c) = qQ + p_0 c \quad (3.66)$$



и, следовательно,

$$\mathbf{q} = c\mathbf{p}_0\{\mathbf{I}(1+c) - \mathbf{Q}\}^{-1}, \text{ если } c > 0. \quad (3.67)$$

Если  $c = 0$ , то предельная структура точно такая же, как для модели замкнутой марковской цепи с переходной матрицей  $\mathbf{Q}$ . Если рост идет по геометрической прогрессии  $M(T) = Mx^T$ , то  $c = 0$  тогда и только тогда, когда  $x \leq 1$ , и из табл. 3.1 видно стремление к одному и тому же пределу в трех случаях:  $x = 0$ ;  $x = \frac{1}{2}$  и  $x = 1$ . При  $x > 1$   $c = x - 1$  и в этом случае предельную структуру можно получить непосредственно из (3.64):

$$\mathbf{q} \propto \mathbf{p}_0\{\mathbf{I}x - \mathbf{Q}\}^{-1}. \quad (3.68)$$

При  $x = 2$  имеем

$$100 \mathbf{q} = (63,8; 31,9; 3,8; 0,5; 0,0),$$

что совпадает с последней строкой табл. 3.3. Существование предела для  $\mathbf{q}(T)$ , когда  $\lim_{T \rightarrow \infty} M(T+1)/N(T) = c$ , может быть установлено, если записать (3.65) в виде

$$\mathbf{q}(T+1) = \mathbf{q}(T) \mathbf{P}(T) + \mathbf{p}_0 R(T+1),$$

где  $\mathbf{P}(T) = \mathbf{Q}\{1 + M(T+1)/N(T)\}^{-1}$  и  $R(T+1) = M(T+1)/N(T+1)$ , а затем воспользоваться правилами произведения положительных матриц.

Результаты этого раздела можно подытожить следующим образом, если  $M(T) \sim Mx^T$ , то установившаяся структура такая же, как и для системы с фиксированным размером при  $x \leq 1$ . Если  $x > 1$ , то она определяется с помощью (3.68). Если  $x$  становится очень большим, то  $\mathbf{q} = \mathbf{p}_0$  — это результат, который можно было бы предвидеть, заметив, что при такой высокой скорости роста большинство членов системы должно быть новичками.

### Дисперсии и ковариации размеров уровней

Метод, который использовался для получения дисперсий и ковариаций для модели с заданным входом, может быть с небольшими модификациями применен и в данном случае. Если начать с (3.38), то больше нельзя считать, что вход не зависит от внутренних перемещений, потому что они влияют на потери, которые в свою очередь определяют вход. Следовательно, к двум слагаемым в правой части (3.39) нужно добавить слагаемые, включающие ковариации между величинами  $n_{0j}(T+1)$  и внутренними потоками. Однако эту довольно тяжеловесную алгебру обойти можно следующим образом. Каждое увольнение непосредственно обуславливает прием; если увольнение происходит с уровня  $i$ , то вероятность того, что последующее пополнение произойдет на уровне  $j$ , равна  $p_{i,k+1} p_{0j}$ . Замену ушедшего лица можно рассматривать как внутренний переход. Общая вероятность перемещения

(любого типа) с уровня  $i$ , вызывающего прибавление на уровень  $j$ , таким образом, равна:

$$q_{ij} = p_{ij} + p_{i,k+1} p_{0j}. \quad (3.69)$$

То же самое, но в другой связи, мы проделали ранее (см. с. 91—92).

Если  $M(T) = 0$  для всех  $T$ , то можно воспользоваться теорией, развитой в гл. 2 для замкнутой системы. Вектор математических ожиданий, дисперсий и ковариаций дается соотношением

$$\mu(T+1) = \mu(T) \Pi', \quad (3.70)$$

где элементы  $\Pi'$  — такие же функции величин  $q_{ij}$ , какими были элементы  $\Pi$ , представляющие собой функции величин  $p_{ij}$  в (2.15).

Когда система развивается, появляются новые лица, которые заполняют вакансии, образовавшиеся в процессе развития. Если они распределяются по уровням независимо от замен и с теми же вероятностями  $\{p_{0j}\}$ , то эти потоки будут независимы от всех других потоков и, следовательно, их вклад в дисперсии и ковариации можно просто добавить в (3.70). Пусть вектор математических ожиданий, дисперсий и ковариаций, связанный с приемом, есть  $\mu'_0(T+1)$ . Тогда

$$\begin{aligned} \mu'_0(T+1) = M(T+1) \{ & p_{01}, p_{02}, \dots, p_{0k}, p_{01}(1-p_{01}), \\ & -p_{01}p_{02} \dots p_{0k}(1-p_{0k}) \}, \end{aligned} \quad (3.71)$$

ковариации потоков входящих на уровни  $i$  и  $j$  есть

$$\delta_{ij} p_{0j} - p_{0j} p_{0i} \quad (i, j = 1, 2, \dots, k).$$

Таким образом, общее выражение для разностного уравнения имеет вид

$$\mu(T+1) = \mu(T) \Pi' + \mu'_0(T+1). \quad (3.72)$$

Дисперсии и ковариации можно вычислить, пользуясь теми же компьютерными программами, что и для фиксированного входа, заменив только  $\Pi$  на  $\Pi'$  и  $R(T+1)$  на  $M(T+1)$ .

### Контрактные системы

Соображение, приводящее к (3.54а) и (3.54б), может оказаться несостоятельным, если  $M(T+1)$  отрицательно для каких-либо  $T$ . Если это так, то математическое ожидание численности потерь

$$\sum_{i=1}^k p_{i,k+1} \bar{n}_i(T)$$

недостаточно, чтобы получить уменьшение, определяемое  $M(T+1)$ . При таких обстоятельствах  $\bar{R}(T+1)$ , заданное (3.53), будет отрицательным. Поскольку модель в том виде, в каком она до сих пор описывалась, не позволяет сокращать штаты, избыточность может дать бессмысленный результат.

Если мы хотим, чтобы модель учитывала и сокращение, то ее следует обобщить. Для этого существуют две возможности. Первая — ввести вектор  $s = (s_1, s_2, \dots, s_k)$  сокращений, элементы которого — излишки,

выраженные в долях  $|R(T)|$  на каждом уровне. Вместо того чтобы прибавлять вектор  $R(T) \mathbf{p}_0$  к  $\mathbf{n}(T) \mathbf{P}$ , мы должны вычесть  $|R(T)| \mathbf{s}$ . Основное разностное уравнение для ожидаемых размеров по уровням теперь принимает одну из следующих форм в зависимости от знака  $R(T)$ :

$$\bar{n}_j(T+1) = \sum_{i=1}^k (p_{ij} + p_{i,k+1} p_{0j}) \bar{n}_i(T) + M(T+1) p_{0j} \quad (j=1, 2, \dots, k), \quad (3.73a)$$

если  $\bar{R}(T+1) \geq 0$ . Или

$$\bar{n}_j(T+1) = \sum_{i=1}^k (p_{ij} + p_{i,k+1} s_j) \bar{n}_i(T) + M(T+1) s_j \quad (j=1, 2, \dots, k), \quad (3.73б)$$

если  $\bar{R}(T+1) < 0$ . Заметим, что  $M(T+1)$  обязательно будет отрицательным в (3.73б) и может быть отрицательным в (3.73а). Следовательно, мы должны выбирать попеременно (3.73а) и (3.73б) в зависимости от того, что требует ситуация — приема на работу или сокращения штатов. Если вектор  $\mathbf{s}$  выбран таким же, как  $\mathbf{p}_0$ , то очевидно, что оба уравнения совпадают, и мы можем воспользоваться исходным уравнением (3.54б), так как если бы оно было справедливым для любого  $M(T+1)$ . Однако, поступая таким образом, мы должны на каждом шаге проверять, не стал ли какой-нибудь из элементов  $\bar{\mathbf{n}}(T+1)$  отрицательным. Если появляются отрицательные запасы, то идти дальше можно, только выбрав новый вектор  $\mathbf{s}$ .

Другой метод, позволяющий избежать появления отрицательных размеров запасов, заключается в том, чтобы рассчитать численность увольняемых по сокращению штатов как долю запасов после увольнений и повышений в должности. При отсутствии приема на работу и сокращения штатов ожидаемый размер организации на  $j$ -м уровне в момент  $T+1$  равен:

$$\bar{n}'_j(T+1) = \sum_{i=1}^k p_{ij} \bar{n}_i(T) \quad (j=1, 2, \dots, k).$$

Сокращение штатов будет необходимо, если

$$\sum_{j=1}^k \bar{n}'_j(T+1) > N(T+1),$$

и предлагаемая процедура состоит в уменьшении каждого  $\bar{n}'_j(T+1)$  в одинаковой пропорции. Если это отношение обозначить через  $x$ , то  $x$  должен удовлетворять соотношению

$$x \sum_{j=1}^k \bar{n}'_j(T+1) = N(T+1)$$

или

$$x = N(T+1) \left/ \sum_{i=1}^k (1 - p_{i,k+1}) \bar{n}_i(T) \right. \quad (3.74)$$

Если в этом отношении уменьшить каждый уровень, то результирующие численности будут

$$\bar{n}_j(T+1) = N(T+1) \sum_{i=1}^k p_{ij} \bar{n}_i(T) / \left\{ N(T) - \sum_{i=1}^k p_{i,k+1} \bar{n}_i(T) \right\}$$

$$(j=1, 2, \dots, k), \quad (3.75a)$$

или

$$n(T+1) = n(T) P\{I + M(T)/N(T)\} / \{1 - W(T)\}, \quad (3.75b)$$

где  $W(T) = \bar{n}(T) p'_{k+1}$ . При такой версии модели последним уравнением следует пользоваться всякий раз, когда  $\bar{R}(T+1)$ , заданное (3.53), оказывается отрицательным.

#### 3.4. ДОПОЛНЕНИЯ К ГЛ. 2 И 3

Наши модели были построены на основе теории марковских цепей с дискретным временем. Обзоры разной сложности по этой проблеме можно найти почти во всех современных учебниках по теории вероятностей и стохастическим процессам. Четыре примера — Феллер (1968), Моран (1968), Кокс и Миллер (1965) и Карлин и Тейлор (1975). Имеется также несколько книг, полностью посвященных цепям Маркова. Основная работа повышенной сложности — это книга Чжуна (1967), однако Кемени и Снелл (1976) (переиздание работы, вышедшей в 1960 г.) также дают прекрасное изложение теории цепей с конечным числом состояний. Позже Исааксон и Мадсен (1976) сделали хороший обзор, включающий цепи, зависящие от времени.

Изложенная теория в значительной степени основывается на таких разделах математики, как теория матриц, их собственных значений и собственных векторов. В этой области имеется также много фундаментальных трудов; основные из них — работы Фрэзера и соавторов (1946), двухтомный труд Гантмахера (1967), Ланкастера (1969). Поскольку переходные матрицы принадлежат к более широкому классу положительных матриц, очень полезна книга Сенеты (1973).

Открытые и замкнутые марковские модели находят широкое применение почти во всех областях общественных наук, им посвящена обширная литература. Ниже перечисляются работы по отраслям, они могут служить введением для тех читателей, которые хотят познакомиться с данным вопросом в рамках своей области интересов. Приведенный перечень литературы ни в коей мере не является исчерпывающим, а классификация неизбежно получилась в определенном смысле произвольной. В частности, в известном смысле очень трудно установить границы между географией и экономикой и между географией и демографией. Читатели, чьи интересы лежат в этих сферах, должны познакомиться и со смежными областями. Ссылки на литературу даются по фамилиям авторов и по годам. Более полный перечень работ приведен в конце книги, а некоторые статьи, упомянутые в тексте гл. 2 и 3, здесь не повторяются.

## Планирование системы образования и планирование кадров

Приложения к планированию кадров можно изучить по книге Бартоломью и Форбеса (1979), которая содержит достаточно полную библиографию. В некотором смысле более обширная, но менее детализированная библиография, охватывающая вопросы планирования системы образования и планирования национальных трудовых ресурсов, содержится в работе Бартоломью (1976b). Профессиональная мобильность смыкается, с одной стороны, с социальной мобильностью, а с другой — с планированием трудовых ресурсов. Этим вопросам посвящены нигде не упоминавшиеся статьи Ходжа (1966) и Стьюмана (1975a, 1975b, 1976). Специалисты по планированию системы образования широко пользовались марковскими моделями, о чем можно прочитать в работах Гани (1963), Клафа и Мак-Рейнольдса (1966), Камата (1968a и c), Тонстада (1969), Армитаж и др. (1970), Менгеса и Элстермана (1971), Джонстона и Филпа (1973), Мура (1975) (ответ Джонстону и Филпу), Бритни (1975).

### Демография

Демографы часто пользовались моделями, по существу, теми же, что и марковские цепи. В последнее время это обстоятельство стало совершенно очевидным благодаря работам Матраса (1960a, b, 1967), Роджерса (1968), Сайкса (1969), Кришнана (1977b), Фини (1973), Файхтингера (1972, 1973). Риса и Уилсона (1973), Йозефа (1974), Салкина и др. (1975).

### География

В современных географических исследованиях значительное место стал занимать пространственный анализ. Такой анализ направлен на изучение миграции населения, он охватывает также вопросы землепользования и размещения промышленных предприятий. Следующие работы затрагивают широкий спектр тем: Кларк (1965), Тарвер и Гарли (1965), Моррисон (1967), Дрюетт (1969), Браун (1970), Лонг (1970), Белл (1974), Берри (1971), Коллинз (1973), Стаффорд (1977).

### Экономика

Приложения в этой области охватывают очень широкий диапазон вопросов, в частности, такие, как распределение доходов, размеры фирм, изменения курсов акций, рынок акций и т. д. Некоторые из этих тем будут обсуждаться в гл. 7, где мы будем заниматься распределениями размеров. Неполный список работ в этой области: Харт и Прайс (1956), Адельман (1958), Престон и Белл (1961), Коллинз и Престон (1961), Дент (1967), Драйден (1969), Филиц и Баргава (1973), Шоррокс (1976).

Марковские модели нашли широкое применение в этой области, где состояния — это обычно различные виды медицинского обслуживания, примеры можно найти в работах Маршалла и Голдхамера (1955), Наварро (1969), Меридита (1973), Андерсона (1974), Дэвиса и соавторов (1975).

В дополнение к этим основным группам работ можно отметить приложения к поведению при голосовании (Андерсон, 1954; Хоукс, 1969; Миллер, 1972), взятию интервью (Ховс и Фоли, 1973; Хедж и др., 1978), бухгалтерскому учету (Сайерт и др., 1962), стоматологии (Лю, 1968), системам ЭВМ (Фоли, 1967), лингвистике (Миллер, 1952) и международным отношениям (Кац и Проктер, 1959; Майхью и Грей, 1971; Майхью, 1972).

Социальная мобильность — тема более широкая, чем наше весьма ограниченное изложение в гл. 2. Хорошим введением в ее количественный аспект является работа Будона (1973). Работа Бибби (1975) касается главным образом вопросов измерения, но включает также более общие проблемы и ссылки на литературу. В настоящее время опубликовано большое количество таблиц социальной мобильности. Некоторые из них приводит Бибби (1975), а перечень 20 таблиц содержит табл. 1 в работе Соммерса и Конлиса (1979) с указанием первоисточников.

В исследовании Шоррокса по измерению социальной мобильности дается перечисление множества требований, предъявляемых к мерам, которые представляются необходимыми. Автор анализирует, насколько эти требования соотносятся друг с другом. Один из главных его выводов состоит в том, что не существует меры, удовлетворяющей всем приведенным им требованиям. По-видимому, это отражает трудности установления отличий зависимости между поколениями от количества движения. Шоррокс вводит, например, понятие порядка согласования, для которого мера должна упорядочивать популяции одним и тем же образом независимо от периода наблюдения. Таким образом, если два сообщества имеют переходные матрицы  $P_1$  и  $P_2$  и если для некоторой меры  $\mu(P_1) > \mu(P_2)$ , то нам нужно, чтобы  $\mu(P_1^T) > \mu(P_2^T)$  для всех целочисленных  $T$ . Мера  $m_2(P)$  не обладает этим свойством. Однако  $m_2(P)$  и была введена как такая мера движения, для которой порядок согласования не подходит. Он больше согласуется с тем, что мы назвали зависимостью между поколениями.

В дополнение к мерам, рассмотренным в гл. 2, Соммерс и Конлиск обсуждают то, что они называют регрессией на средний коэффициент. Они обобщают измерение коэффициента корреляции  $\rho$ , рассматривая  $r$ -е собственное значение  $\rho$ , вычисленное для  $P^r$ ; ими установлен также ряд неравенств для различных мер.

В нашем рассуждении о мере мобильности мы подчеркивали, что все зависит от реальности исходной предпосылки относительно марковских свойств процесса. Согласно модели потоки вызывают изменения в структуре, а не наоборот. С другой стороны, если изменения в структуре порождают потоки, то следует учитывать другие соображения.

В этой связи можно отметить различие между *структурной* и *чистой* мобильностью. Структурной мобильностью называется минимальное количество движения, необходимое для получения заданного структурного изменения. Чистая мобильность — это превышение действительного количества движения над этим минимумом. Будон (1973) и другие ученые рассмотрели вопрос о том, как выделить и измерить составляющую чистой мобильности. Поскольку наши меры относятся к системе, находящейся в равновесии (когда по определению не происходит изменений структуры системы), они представляют собой меры чистой мобильности.

Явление группировки по диагонали  $T$ -й степени переходной матрицы детально изучалось Сингером и Шпилерманом (1977 b). Как мы уже отмечали, ситуация бывает весьма простой, когда переходная матрица имеет действительные положительные собственные значения, и именно такой случай чаще всего встречается на практике. Нам представляется недостаточным исследованным вопрос о том, каким реальным социологическим факторам соответствуют матрицы с действительными положительными собственными значениями. Если можно было бы найти такую характеристику, то это помогло бы нам понять удивительное явление обращаемости, которое, по-видимому, должно быть типичным для матриц мобильности, так как обратная матрица имеет действительные положительные собственные значения.

После выхода в свет второго издания нашей книги Файхтингер (1976) и Мелман (1977a) выполнили достаточно полное исследование предельного поведения открытых марковских моделей, и их результаты представлены в новом изложении гл. 3. Это изложение одновременно и проще, и более общее, чем во втором издании; кроме того, оно выстроено в единую линию с последующим анализом модели с непрерывным временем.

Делались различные попытки обобщить марковские модели в направлениях, выходящих за рамки гл. 2 и 3. Например, Мойя-Ангелер (1976) рассматривает модель планирования образования, в которой существуют емкостные ограничения, определяющие пределы размеров на различных уровнях. Как только пределы достигнуты, излишки должны перераспределяться по другим уровням. Это похоже на модели узких мест, введенных ранее Армитажем и его коллегами (1969). С аналогичной целью в некоторых исследованиях допускалось, что численности повышенных в должности должны зависеть не только от запаса на том уровне, с которого происходит повышение, как в марковской модели, но и от запаса уровня назначения. Эти и другие различные обобщения приводят к моделям промежуточного типа между марковскими моделями, рассмотренными в гл. 3, и моделями восстановления, которые будут обсуждаться в гл. 8. Шиннар и Стьюман (1978) рассматривают общий класс моделей, включающий те, которые они называют моделями с «продолжительностью последствия». В качестве специального случая они включают эффект накопленной инерции, а также эффекты неоднородности и предвидения.

Изложенные здесь результаты, касающиеся дисперсий и ковариаций объемов запасов, предполагают, что параметры моделей известны.

На практике же это бывает редко. Их надо предварительно оценивать, и это в дальнейшем вносит ошибки в любые предсказания. Еще большие вариации возникают, если сами параметры модели подвержены случайным возмущениям. Эти вопросы исследовались в работе Бартоломью (1975), где показано, например, что если  $P$  оценена по данным за один интервал времени, то дисперсии и ковариации объемов запасов на следующем интервале должны примерно удвоиться.

Мы не рассматривали статистические проблемы оценивания параметров марковских моделей. Этот вопрос освещен в работах Бартоломью (1977а) и Бартоломью и Форбеса (1979), в которых указана и другая литература. Следует проявлять осторожность при использовании таблиц социальной мобильности, поскольку не всегда ясно, как были сделаны выборки. Например, необходимо знать, опрашивались ли сыновья с целью получения сведений об их отцах или наоборот.



## 4.1. ОСНОВЫ ТЕОРИИ

Перейдем к рассмотрению непрерывных по времени версий моделей, аналогичных марковским цепям, введенным в гл. 2. На практике выбор между дискретной и непрерывной версиями модели — это вопрос, частично связанный с отношением к действительности, а частично — с практическим удобством. Основываясь обычно на отношении к действительности, мы хотим, например, моделировать перемещение людей между профессиональными группами или регионами в непрерывном времени, но на практике вычислительные преимущества часто приводят к выбору модели с дискретным временем. Модели с непрерывным временем легче поддаются математическому анализу, и это можно считать одним из их преимуществ, даже когда в действительности процессы протекают с дискретным временем. Чтобы выявить близкие параллели между этими двумя ситуациями, мы будем придерживаться такого же характера изложения, что в гл. 2. В частности, мы получим основные теоретические результаты, ограничившись лишь операциями над уже известными нам формулами. Тем самым мы подчеркиваем единство рассматриваемых моделей, хотя это и не всегда тот путь, которому следовал бы математик.

В марковских цепях с дискретным временем изменения состояния происходят на единичных интервалах времени. Существуют два способа перехода к непрерывному времени, мы рассмотрим их по очереди. Один способ заключается в предположении, что временной интервал между переходами становится все меньше и меньше. Однако только такого предположения недостаточно, так как оно лишь ускоряет процесс из-за изменения масштаба времени. Следовательно, надо допустить еще, что вероятности переходов должны меняться таким образом, чтобы ожидаемое число изменений состояния в единицу времени осталось тем же самым. В частности, вероятности переходов из каждого состояния должны уменьшаться по мере уменьшения временного интервала между изменениями состояния. Этого можно достигнуть следующим образом. Пусть величина интервалов между переходами будет  $\delta T$ , тогда

$$p_{ij} = r_{ij} \delta T \quad (i \neq j), \quad p_{ii} = 1 - \delta T \sum_{\substack{j=1 \\ j \neq i}}^k r_{ij}. \quad (4.1)$$

При  $\delta T \rightarrow 0$  вероятность перехода из любого состояния стремится к нулю, но средняя величина интервала времени между изменениями состояния остается постоянной. Чтобы это показать, заметим, что среднее число шагов пребывания в состоянии  $i$  равно  $(1 - p_{ii})^{-1}$  (см. с. 46), каждый шаг имеет длительность  $\delta T$ , так что средняя длительность пребывания (время пребывания) равна:

$$\delta T (1 - p_{ii})^{-1} = \left\{ \sum_{j \neq i} r_{ij} \right\}^{-1} = \lambda_i^{-1}.$$

Стохастический процесс, к которому мы приходим в пределе при  $\delta T \rightarrow 0$ , есть непрерывная марковская цепь. Величины  $p_{ij}$ , заданные соотношением (4.1), часто называются *инфинитезимальными вероятностями переходов*. Величины  $r_{ij}$  называются *скоростями* или *интенсивностями переходов*, они характеризуют склонность к перемещению между парами состояний. Цепь, зависящую от времени, можно определить как цепь, в которой скорости переходов являются функциями времени. Скорости переходов часто имеют различные названия в страховой работе и в исследованиях надежности; не исключено, что наиболее известными и выразительными является термин «сила смертности». В нашей терминологии — это скорость перехода от состояния «жизнь» в состояние «смерть». Надо следить за тем, чтобы не путать скорости с вероятностями. Соотношение между ними дается (4.1).

Если обозначить через  $\mathbf{P}(\delta T)$  матрицу вероятностей переходов в (4.1), то получим

$$\mathbf{P}(\delta T) = \mathbf{I} + \delta T \mathbf{R}, \quad (4.2)$$

где  $\mathbf{R}$  — матрица с элементами  $r_{ij}$  вне диагонали и с элементами  $r_{ii} = -\lambda_i = -\sum_{j \neq i} r_{ij}$  на диагонали. Такое определение  $r_{ii}$  удобно для того, чтобы иметь простое матричное представление (4.2). Основные теоретические соотношения, необходимые нам для приложений, можно теперь получить, делая предельный переход в соответствующих формулах из гл. 2.

Прежде чем перейти к пределам, рассмотрим другой способ перехода от дискретного времени к непрерывному. В модели с дискретным временем изменения состояния происходят на фиксированных интервалах времени. Эту ситуацию можно было бы обобщить, допустив, что изменения происходят на случайных интервалах времени. Тогда наш процесс определялся бы двумя стохастическими процессами. Один из них был бы марковской цепью, определяющей, какое изменение состояния происходит, когда наступает момент изменения; второй был бы точечным процессом, и его реализация давала бы моменты времени, в которые происходят изменения состояния. Тогда обыкновенная марковская цепь проявляется как частный случай, при котором точечный процесс вырождается в совокупность точек на фиксированных интервалах. Этот способ обобщения марковской модели с дискретным временем, конечно, гораздо более общий, чем первый, поскольку он позволяет получить столько моделей, сколько существует точечных процессов. В этой и следующей главах будут рассмотрены несколько про-

цессов этого семейства, а сейчас мы отметим только один, совпадающий с определенной выше марковской цепью с непрерывным временем.

Предположим, что изменения состояния происходят в соответствии с пуассоновским процессом с интенсивностью  $\lambda$ , т.е.  $Pr\{\text{изменение происходит на } (T, T + \delta T)\} = \lambda \delta T$ , откуда следует, что вероятность того, что  $n$  изменений произойдет на интервале  $(0, T)$ , имеет пуассоновское распределение

$$P_n(T) = \frac{(\lambda T)^n}{n!} e^{-\lambda T}.$$

Далее, пусть  $\mathbf{P}$  есть матрица вероятностей перехода, определяющая изменения состояний. Далее мы покажем, что такой процесс представляет собой марковскую цепь с непрерывным временем, причем

$$\mathbf{R} = \lambda (\mathbf{P} - \mathbf{I}). \quad (4.3)$$

### Основные уравнения

Обычно мы хотим знать два множества значений вероятностей, которые в гл. 2 были обозначены  $\mathbf{p}(T)$  и  $\mathbf{P}(T)$  (см. с. 30). Вектор вероятностей находится в различных состояниях в момент времени  $T$  и легко получается из (2.2) следующим образом:

$$\mathbf{p}(T + \delta T) = \mathbf{p}(T) \mathbf{P}(\delta T) = \mathbf{p}(T) \{\mathbf{I} + \delta T \mathbf{R}\},$$

причем

$$\{\mathbf{p}(T + \delta T) - \mathbf{p}(T)\} / \delta T = \mathbf{p}(T) \mathbf{R},$$

что в пределе при  $\delta T \rightarrow 0$  дает

$$\frac{d}{dT} \mathbf{p}(T) = \mathbf{p}(T) \mathbf{R}. \quad (4.4)$$

Здесь и далее оператор дифференцирования применяется к каждому элементу стоящих за ним вектора или матрицы. Форма решения (4.4) подсказывается специальным случаем, когда  $R$  и  $p(T)$  — скаляры. В этом случае известно, что решением будет

$$p(T) = p(0) e^{RT} = p(0) \sum_{i=0}^{\infty} (RT)^i / i!. \quad (4.5)$$

В матричном случае легко проверить, что ряд в правой части (4.5) удовлетворяет (4.4), если мы полагаем  $\mathbf{R}^0 = \mathbf{I}$ . Можно показать, что это и есть требуемое решение, если ряды сходятся, а это так, если  $k$  конечно. Следовательно, решение можно записать в виде

$$\mathbf{p}(T) = \mathbf{p}(0) \exp \{\mathbf{R}T\}. \quad (4.6)$$

Матрица  $\mathbf{P}(T)$ , элемент  $(i, j)$  которой есть вероятность перехода из  $i$  в  $j$  за период между 0 и  $T$ , очевидно, удовлетворяет уравнению

$$\mathbf{p}(T) = \mathbf{p}(0) \mathbf{P}(T), \quad (4.7)$$

и мы получаем

$$\mathbf{P}(T) = \exp \{\mathbf{R}T\}. \quad (4.8)$$

Для подготовки читателя к материалу, изложенному в гл. 5, мы далее будем работать не с  $\mathbf{p}(T)$ , а с вектором ожидаемых чисел  $\bar{\mathbf{p}}(T) = \mathbf{n}(0)\mathbf{P}(T)$ .

Теперь можно проверить (4.3). Для определенного ранее процесса

$$\mathbf{P}(T) = \sum_{n=0}^{\infty} P_n(T) \mathbf{P}^n,$$

поскольку  $\mathbf{P}^n$  дает вероятности переходов при условии, что  $n$  переходов уже произошли.

Заменяя  $P_n(T)$  их значениями и замечая, что  $\mathbf{1}e^{-\lambda T} = e^{-\lambda T}$ , находим

$$\mathbf{P}(T) = \sum_{n=0}^{\infty} \frac{(\lambda T)^n}{n!} e^{-\lambda T} \mathbf{P}^n = \exp\{\lambda(\mathbf{P} - \mathbf{I})T\}. \quad (4.9)$$

Сравнивая с (4.8), получаем (4.3).

Хотя оба способа перехода к марковской модели с непрерывным временем эквивалентны, существуют практические соображения, по которым в каждом конкретном случае тому или другому способу может быть отдано предпочтение. Эти соображения приобретут еще большее значение, когда мы далее в этой главе будем обобщать модель. Весьма важно знать, имеет ли смысл говорить об изменениях категории внутри одного состояния. Например, при рассмотрении профессиональной мобильности состояния системы могут задаваться отраслями промышленности. Индивидуум, который меняет работу внутри одной и той же отрасли, не реализует изменения состояния, однако поведение таких индивидуумов может отличаться от поведения тех, кто остался на прежней работе. При моделировании подобного процесса марковской цепью  $\lambda$  будет интенсивность изменения *места работы*, а  $\mathbf{P}$  — матрицей вероятностей перехода между отраслями промышленности. Тогда ненулевые диагональные элементы будут представлять собой вероятности перемещения внутри одной и той же отрасли промышленности. Заметим, что до сих пор  $\lambda$  было одним и тем же для всех состояний. При подгонке моделью такого процесса может понадобиться выразить результаты через оценки  $\lambda$  и  $\mathbf{P}$ .

С другой стороны, в большинстве случаев не имеет смысла говорить о перемещениях внутри одного и того же состояния. Именно такая ситуация во втором из примеров, приведенных в разделе 4.2, а также в гл. 5. В этих примерах более естественным представляется определение процесса через интенсивности переходов. Однако даже в этом случае может оказаться полезным так перестроить систему параметров модели, чтобы сосредоточить внимание на двух аспектах: продолжительности пребывания в каждом состоянии и изменениях состояния. Это можно сделать следующим образом. Продолжительность пребывания дискретной марковской цепи в состоянии  $i$  распределена по геометрическому закону с

$Pr\{\text{время пребывания в } i = r\} = (1 - p_{ii}) p_{ii}^{r-1} \quad (r = 1, 2, \dots) \quad (4.10)$   
 Предельный переход, выполненный в начале данного раздела, дает экспоненциальную плотность времени пребывания с параметром  $\lambda_i =$

$= \sum_{j \neq i} r_{ij}$ . Вероятность того, что система переходит из  $i$  в  $j$  ( $i \neq j$ ) при условии, что выход из  $i$  имеет место, равна:

$$m_{ij} = r_{ij}/(1 - p_{ii}) \quad (i \neq j),$$

что при подстановке (4.1) дает

$$m_{ij} = r_{ij}/\lambda_i \quad (i \neq j) \quad (4.11)$$

при  $\sum_I m_{ij} = 1$ . Таким образом, процесс можно определить через интенсивности  $\lambda_i$  и матрицу переходов  $\mathbf{M}$ , которая содержит нули на главной диагонали и элементы  $m_{ij}$  при ( $i \neq j$ ). Итак, (4.3) можно обобщить и предложить три эквивалентные параметризации

$$\mathbf{R} = \lambda (\mathbf{P} - \mathbf{I}) = \Lambda (\mathbf{M} - \mathbf{I}), \quad (4.12)$$

где  $\Lambda$  — диагональная матрица с элементами  $\lambda_i$  на диагонали. Это показывает, что если время пребывания в состоянии  $i$  зависит от  $i$ , то переходная матрица должна иметь нули на диагонали. И наоборот, если переходная матрица не имеет нулей на диагонали и время пребывания для всех состояний распределено неодинаково, то процесс не может быть марковским.

### Решение основных уравнений

Прямое определение матрицы  $\mathbf{P}(T)$  из (4.8) включало бы суммирование  $k^2$  бесконечных рядов, члены которых были бы получены из степеней матрицы  $\mathbf{R}$ . Лучший метод состоит в применении теоремы Сильвестра (см. раздел 3.2), которая позволяет представить решение в форме конечной суммы из  $k$ , слагаемых. Вообще говоря, матрица  $\mathbf{R}$  допускает спектральное представление

$$\mathbf{R} = \sum_{i=1}^k \theta_i \mathbf{A}_i,$$

где  $\{\theta_i\}$  — собственные значения матрицы  $\mathbf{R}$ , а  $\{\mathbf{A}_i\}$  — связанные с ними сопутствующие матрицы. Это представление справедливо только тогда, когда все  $\theta$  различны. Тогда из теоремы Сильвестра следует, что

$$\exp(\mathbf{R}T) = \sum_{i=1}^k e^{\theta_i T} \mathbf{A}_i = \mathbf{P}(T), \quad (4.13)$$

где использована матричная форма (4.5). Аналогично имеем

$$\bar{\mathbf{n}}(T) = \sum_{i=1}^k e^{\theta_i T} \mathbf{n}(0) \mathbf{A}_i. \quad (4.14)$$

Поведение решения, особенно для больших значений  $T$ , будет сильно зависеть от собственных значений. Прежде чем перейти к обсуждению практических шагов на пути к полному решению, получим некоторые общие результаты, касающиеся величин собственных значений.

Собственные значения получаются решением уравнения

$$\begin{vmatrix} r_{11} - \theta & r_{21} & \dots & r_{k1} \\ r_{12} & r_{22} - \theta & \dots & r_{k2} \\ \vdots & \vdots & \dots & \vdots \\ r_{1k} & r_{2k} & \dots & r_{kk} - \theta \end{vmatrix} = 0. \quad (4.15)$$

Величина определителя не изменится, если мы заменим элементы первой строки на сумму элементов столбцов. Теперь все элементы в первой строке нового определителя будут  $-\theta$ . Следовательно, уравнение (4.15) всегда выполнено, когда  $\theta = 0$ . Из (4.13) ясно, что все корни должны быть равными нулю или быть отрицательными; в противном случае нашлось бы такое  $T$ , для которого переходные вероятности не лежали бы между 0 и 1\*.

Поскольку нас интересует главным образом вектор  $\bar{n}(T)$ , мы будем рассматривать решение в форме (4.14). Из предыдущих замечаний следует, что величины  $\bar{n}_i(T)$  можно записать в виде

$$\bar{n}_i(T) = c_{i1} + \sum_{j=2}^k c_{ij} e^{\theta_j T} \quad (i = 1, 2, \dots, k), \quad (4.16)$$

где  $\theta_j < 0$  ( $j = 2, 3, \dots, k$ ). Соотношение имеет место, если все  $\theta$  различны. Если  $\theta = 0$  — кратный корень, оно также сохраняется, но при небольшой модификации, представленной ниже. Коэффициенты  $\{c_{ij}\}$  в (4.16) можно найти сразу при определении сопутствующих матриц  $A_i$ . Эквивалентный, но более краткий способ состоит в следующем. Подставляя  $\bar{n}(T)$  из (4.16) в (4.4), получаем

$$\sum_{j=1}^k c_{ij} \theta_j e^{\theta_j T} = \sum_{j=1}^k r_{ji} \sum_{h=1}^k c_{jh} e^{\theta_h T} \quad (i = 1, 2, \dots, k), \quad (4.17)$$

где  $\theta_1 = 0$ . Сравнивая коэффициенты при  $e^{\theta_h T}$ , находим, что  $c_{ij}$  должны удовлетворять уравнениям

$$\sum_{j=1}^k r_{ji} c_{jh} = \theta_h c_{ih} \quad (i = 1, 2, \dots, k; \quad h = 1, 2, \dots, k). \quad (4.18)$$

Хотя здесь представлены  $k^2$  уравнений для того же числа неизвестных, уравнения не являются независимыми. Действительно, если мы просуммируем каждую часть (4.18) по  $i$ , то получим нуль в каждой сумме. Следовательно, чтобы определить коэффициенты  $c$ , нам нужно еще  $k$  уравнений. Их можно найти, так как необходимо выполнение начальных условий. Полагая в (4.16)  $T = 0$ , имеем

$$\sum_{j=1}^k c_{ij} = n_i(0) \quad (i = 1, 2, \dots, k). \quad (4.19)$$

Позже мы будем решать эти уравнения для каждого частного случая.

\* Корни в уравнении (4.15) могут быть комплексными, но при известных условиях регулярности на  $\mathbf{R}$  действительные части  $\theta_j < 0$ . — *Примеч. ред.*

Изложенная выше теория относится к случаю, когда все  $\theta$  различны. Если среди корней встречаются кратные, то найдем форму решения с помощью соответствующего предельного перехода в (4.16). Например, предположим, что  $\theta_2 = \theta_3 = \theta$ . Коэффициенты  $\{c_{ij}\}$  являются функциями от  $\theta$ , и нам надо определить их пределы при  $\theta_2 \rightarrow \theta_3$ . Члены, содержащие  $\theta_2$  и  $\theta_3$  в экспоненте, требуют особого внимания. В типовом случае имеем

$$\left. \begin{aligned} \lim_{\theta_2 \rightarrow \theta_3 = \theta} \{c_{i2} e^{\theta_2 T} + c_{i3} e^{\theta_3 T}\} &= e^{\theta T} \lim_{\theta_2 \rightarrow \theta_3} \{c_{i2} + c_{i3} e^{(\theta_3 - \theta_2) T}\} = \\ &= e^{\theta T} \lim_{\theta_2 \rightarrow \theta_3} \{c_{i2} + c_{i3} + c_{i3} [(\theta_3 - \theta_2) T + O(\theta_3 - \theta_2)^2]\} \end{aligned} \right\} \quad (4.20)$$

Поскольку для любой содержательной задачи предел в фигурных скобках конечен, то пара слагаемых, соответствующая  $j = 2$  и  $j = 3$ , должна быть заменена одним слагаемым вида

$$(d_{i2} + d_{i3} T) e^{\theta T}. \quad (4.21)$$

В общем случае, если имеется корень  $\theta$  кратности  $m$ , то слагаемое, соответствующее этому корню в (4.16), должно быть заменено выражением

$$\sum_{j=2}^{m+1} d_{ij} T^{j-2} e^{\theta T}. \quad (4.22)$$

Особенно простым, но важным является случай, когда кратный корень равен нулю. Тогда общее решение имеет вид

$$\bar{n}_i(T) = \sum_{j=1}^m d_{ij} T^{j-1} + \sum_{j=m+1}^k c_{ij} e^{\theta_j T} \quad (i=1, 2, \dots, k). \quad (4.23)$$

Можно показать, что  $d_{ij} = 0$  для  $j > 1$  и всех  $i$ . Необходимость таких соотношений следует из анализа предела  $\bar{n}_i(T)$  при  $T \rightarrow \infty$ . Вторая сумма в (4.23) всегда обращается в нуль, а первая сумма при нарушении равенства  $d_{im} = 0$  при  $m > 1$  будет стремиться к  $\pm \infty$  в зависимости от знака  $d_{im}$ . Так как  $\bar{n}_i(T)$  имеет конечный предел, мы приходим к выводу, что  $\bar{n}_i(T)$  должно иметь вид

$$\bar{n}_i(T) = d_{i1} + \sum_{j=m+1}^k c_{ij} e^{\theta_j T}. \quad (4.24)$$

Подставив выражение (4.24) в (4.4), получим для  $\{d_{i1}\}$  и  $\{c_{ij}\}$  уравнения, такие же, как (4.18), если в последних отбросить уравнения с  $h \leq m - 1$  и заменить  $c_{im}$  на  $d_{i1}$ .

### Предельное поведение

Предельную структуру марковского процесса с непрерывным временем можно получить различными способами. Если уже найдено как описано выше, полное решение уравнений, то надо устремить

$T \rightarrow \infty$ . Учитывая, что действительные части собственных значений для  $j \neq 1$  строго отрицательны, получаем из (4.16)

$$\lim_{T \rightarrow \infty} n_i(T) = c_{i1} \quad (i = 1, 2, \dots, k). \quad (4.25)$$

В том случае, когда  $\theta = 0$  — кратный корень,  $c_{i1}$  надо заменить на  $d_{i1}$ . Если нас интересуют только предельные величины (или нет необходимости в полной информации о процессе), то можно обойтись без вычисления всех  $c_{ij}$ . Из существования предельных состояний и из (4.4) ясно, что предельный вектор  $\bar{\mathbf{n}}(\infty)$  (или  $\mathbf{p}(\infty)$ ) должен удовлетворять уравнению

$$\bar{\mathbf{n}}(\infty) \mathbf{R} = 0. \quad (4.26)$$

Эти уравнения вместе с уравнением  $\bar{\mathbf{n}}(\infty) \mathbf{1}' = N$  можно решить обычным способом, но решения, если воспользоваться (4.12), могут быть представлены в двух различных формах. Таким образом, (4.26) также эквивалентно соотношению

$$\lambda \bar{\mathbf{n}}(\infty) \mathbf{P} = \lambda \bar{\mathbf{n}}(\infty), \quad (4.27)$$

которое показывает, что предельные структуры такие же, как для модели с дискретным временем с переходной матрицей  $\mathbf{P}$ . Одно полезное следствие состоит в том, что установившуюся структуру можно получить, не зная интенсивности или скорости изменения состояния  $\lambda$ . Из (4.12) следует также, что  $\bar{\mathbf{n}}(\infty)$  удовлетворяет соотношению

$$\bar{\mathbf{n}}(\infty) \Lambda \mathbf{M} = \bar{\mathbf{n}}(\infty) \Lambda. \quad (4.28)$$

Таким образом,

$$\bar{n}_j(\infty) \propto m_j / \lambda_j \quad (j = 1, 2, \dots, k), \quad (4.29)$$

где  $\mathbf{m}$  — предельный вектор состояния, определяемый матрицей  $\mathbf{M}$ . Позже мы увидим, что этот результат почти немедленно обобщается на полумарковские процессы.

## 4.2. ПРИМЕРЫ

Применения замкнутых марковских моделей с непрерывным временем довольно редко встречается, если не считать работы Коулмэна (1964а, б). В гл. 5 будут подробно рассмотрены два примера, в которых встречаются открытые системы, так что пока мы отложим обсуждение деталей. Некоторое время мы не будем также обращать внимание и на практические трудности построения модели, которые следуют из того, что редко существует возможность непрерывно следить за процессом; этот вопрос затрагивается в разделе 4.3. Здесь мы займемся вопросами построения моделей и обсудим, почему и в какой форме модель с непрерывным временем можно считать подходящей для двух типов социальных процессов.

Коулмэн (1964б) подобрал модель с непрерывным временем по данным Гласса и Холла о социальной мобильности, которые мы рассматривали в гл. 2, и по аналогичным данным Свалагосты (1959) для Да-



нии. Выбор модели для данных Гласса и Холла требует предположения о разбросе наблюдений во времени, не соответствующего тому, которое имело место при сборе данных. Анализ, проведенный Коулмэном, привел к допущению о том, что сыновья начинают жизнь в социальной группе своих отцов и, следовательно, на протяжении некоторого фиксированного интервала времени подчиняются действию однородных по времени интенсивностей переходов. Группа, зарегистрированная за сыном, — это группа, образовавшаяся после этого фиксированного интервала, одинакового для всех индивидуумов. Для такого анализа хотелось бы предположить, что все сыновья были опрошены об их группе в одинаковом возрасте. Но их возрасты на самом деле различались, поэтому одни проявили свою склонность к перемещению за более продолжительное время, чем другие, и это не дало возможности по агрегированным данным подобрать параметры модели. Таким образом, разработка Коулмэна для этих конкретных данных весьма сомнительна, но сама модель потенциально весьма полезна, особенно при таких исследованиях мобильности, когда изменения групп наблюдаются через фиксированные интервалы времени.

Поскольку данные Гласса и Холла по всем семи социальным группам ранжированы в порядке их статуса или престижа, Коулмэн утверждает, что большинство переходов должно быть направлено в смежные группы. Это выглядит правдоподобно в предположении, что указанные группы возникли из некоего исходного континуума, внутри которого индивидуумы перемещаются непрерывным образом. Легко придумать примеры внезапных изменений групп, но такие изменения представляются настолько редкими, что попытки построить более простую модель оправданы. Матрица  $R$  будет такой, что

$$r_{ij} = 0, \text{ если } |i - j| > 1 \quad (i, j = 1, 2, \dots, k).$$

Если бы мы располагали полной информацией, то, чтобы гарантировать выбор модели, можно было бы проверить, что было более одного изменения состояния, и если они были, то достаточно ли редко эти изменения происходили. Если, как обычно, группы регистрируются в двух фиксированных моментах времени, то возможны переходы между любыми парами состояний, поскольку число переходов на конечном интервале времени ничем не ограничено. Вероятно, при более полной информации можно было бы обнаружить убывающие функции  $p_{ij}$  от  $|i - j|$ , что и предсказывает модель Коулмэна. Однако анализ табл. 2.1 показывает, что это не так, и, следовательно, нет, ничего удивительного в том, что модель Коулмэна не вполне пригодна. Как только рассмотренными методами будет оценена  $R$ , мы сравнительно просто, пользуясь методами из последнего раздела, вычислим ожидаемое время пребывания, предельную структуру и т. п.

Коулмэн пользовался марковскими моделями с непрерывным временем и для изучения изменения точек зрения. Группа индивидуумов разбивалась во времени на категории по начальным точкам зрения и регистрировались последующие изменения категорий. Строго говоря, следовало бы замерять моменты времени, когда происходят измене-

ния, но на практике гораздо проще следить за обновлением категорий за фиксированные интервалы времени. Но при любом способе измерения основная цель оценить интенсивности переходов как для описания самого процесса, так и для предвидения будущего распределения точек зрения и времени пребывания в данной категории \*. Такое описание процесса гораздо лучше, чем обычный дескриптивный анализ, часто применяемый психологами и другими специалистами. Например, если у нас для двух разных моментов времени есть таблица переходов, показывающая перемещения индивидуумов по множеству категорий, то мы можем рассчитать меру корреляции, или связи, показывающую степень изменения позиций данных индивидуумов. Хотя такая мера и позволяет получить количественную оценку изменения, она ничего не говорит о структуре самого процесса и не дает надежной базы для прогнозирования.

Интересный пример использования этой модели для исследования изменения точек зрения привел Коулмэн (1961). Этот пример позже также обсуждался Коулмэном (1964а); дальнейшее развитие он получил в работе Сингера и Шпилермана (1977 а). В октябре 1957 г. группу американских школьников старших классов попросили ответить на два вопроса: а) «Являетесь ли Вы членом «лидирующей группы?» и б) «Согласны ли Вы с тем, что, если юноша хочет войти в состав лидирующей группы, он иногда должен поступиться своими принципами?» Мальчики и девочки опрашивались отдельно, и каждый опрашиваемый в зависимости от ответов был отнесен к одной из четырех категорий. Опрос был повторен в мае 1958 г., причем таким образом, чтобы получить ответы от тех же самых лиц. Затем была подобрана модель, чтобы посмотреть, могут ли изменения точек зрения быть учтены марковской моделью. Были выбраны следующие четыре категории.

- 1) положительные ответы на оба вопроса;
- 2) положительный ответ на (а) и отрицательный ответ на (б);
- 3) отрицательный ответ на (а) и положительный ответ на (б);
- 4) отрицательные ответы на оба вопроса.

Поскольку имеются две точки зрения, Коулмэн утверждает, что только одна могла бы измениться сразу и, следовательно, скорости  $r_{14}$ ,  $r_{23}$ ,  $r_{32}$ ,  $r_{41}$  равны нулю. Результирующая усеченная модель очень хорошо соответствует данным.

Сингер и Шпилерман (1977а) утверждают, что выводы о характере изменения точек зрения можно сделать по матрице  $M$ , которая дает переходные вероятности при условии, что переход был. Например, по ответам 3260 девочек они оценили матрицу  $M$ :

$$M = \begin{pmatrix} 0 & 0,536 & 0,464 & 0 \\ 0,690 & 0 & 0 & 0,310 \\ 0,237 & 0 & 0 & 0,763 \\ 0 & 0,234 & 0,766 & 0 \end{pmatrix}.$$

\* Напомним, что в данном случае время пребывания в категории означает время, когда человек придерживается одной точки зрения, т. е. до первого момента ее изменения. — *Примеч. ред.*

Из нее немедленно можно определить наиболее вероятные переходы. Например, изменения категорий 3 и 4 правдоподобны в обоих направлениях, тогда как изменения категории 3 на 1 и 2 на 4 гораздо менее часты. Такие цифры, очевидно, дают более наглядное представление о существе процесса, чем сами по себе скорости переходов или ряд данных о числах переходов.

Это второй пример, в котором непрерывный процесс наблюдается только в двух отдельных временных точках. Мы не ответили пока на вопрос, много ли можно узнать о процессе с помощью такого способа наблюдений. Далее рассмотрим ответы, которые содержат довольно много тонкостей.

### 4.3. ДИСКРЕТНЫЕ НАБЛЮДЕНИЯ НЕПРЕРЫВНОГО ПРОЦЕССА

По причинам, указанным выше, нам было достаточно наблюдать непрерывный во времени процесс в двух или более фиксированных моментах времени. В плане обсуждения, которое мы до сих пор вели, это ставит ряд вопросов:

- 1) можно ли считать основной процесс марковским?
- 2) если это так, можно ли оценить его параметры? К этим вопросам можно добавить третий:
- 3) при каких условиях допустимо оперировать с дискретно наблюдаемым непрерывным процессом так же, как если бы он действительно был дискретным?

Последний вопрос вызван тем обстоятельством, что дискретная по времени марковская модель из гл. 2 часто подгонялась к процессу, который в действительности был непрерывным. Например, индивидуумов в модели региональной миграции часто рассматривают так, как будто изменения местожительства происходят ежегодно. Конечно, желательное знать, так ли это в действительности.

Займемся сначала последним вопросом. Если система наблюдается в моменты времени 0 и  $T$ , то  $\mathbf{P}(T)$  — переходная матрица за интервалом  $(0, T)$ . Матрицу можно оценить непосредственно по наблюдаемому количеству переходящих индивидуумов. Если рассматривать процесс как дискретный, то мы могли бы предполагать, что матрицей периода за  $r$  шагов будет  $\mathbf{P}^r(T)$ . Далее, поскольку периоды  $T$  скорее всего задаются из соображений удобства наблюдений, не связанных с самим процессом, то полученный ответ должен быть согласован с другим выбором  $T$ , т. е. хотелось бы, чтобы

$$\mathbf{P}(rT) = [\mathbf{P}(T)]^r \text{ для любых неотрицательных } r. \quad (4.30)$$

Для марковского процесса, как видно из (4.8),  $\mathbf{P}(T) = \exp \{RT\}$  и, следовательно, (4.30) всегда выполняется. Поэтому справедлив вывод, что если основной процесс — марковский и если он наблюдается через равные интервалы времени, то в эти моменты можно его рассматривать как дискретную марковскую цепь.

Чтобы ответить на первый из поставленных выше вопросов, мы должны учесть одно соображение. Для марковского процесса соотношение (4.30) выполняется для всех положительных  $r$  и, в частности,

для  $0 < r < 1$ . Тогда, если мы сможем оценить матрицу  $\mathbf{P}(T)$ , то необходимое условие того, что она порождается марковским процессом, состоит в следующем:  $\{\mathbf{P}(T)\}^r$  ( $0 < r < 1$ ) также должна быть матрицей перехода, т. е. ее элементы неотрицательны и суммы элементов строк равны единице. Как определить, обладает ли данная матрица таким свойством? Этот вопрос всесторонне был изучен Сингером и Шпилерманом (1976а). Необходимое и достаточное условие для  $k = 2$  было получено Кингманом (1962). А именно

$$\text{trase } \mathbf{P}(T) = p_{11}(T) + p_{22}(T) > 1. \quad (4.31)$$

Для  $k > 2$  условие Кингмана (1962) на практике неприменимо, но остается ряд необходимых условий, при которых некоторые не порожденные марковским процессом матрицы могут быть исключены из рассмотрения. Вероятно, простейшее из них, для которого (4.31) является частным случаем, состоит в том, что

$$\prod_{i=1}^k p_{ii}(T) \geq \det \mathbf{P}(T) > 0, \quad (4.32)$$

где  $\det$  означает детерминант. Сингер и Шпилерман (1977а) привели пример матрицы

$$\mathbf{P}(T) = \begin{pmatrix} 0,15 & 0,35 & 0,50 \\ 0,37 & 0,45 & 0,18 \\ 0,20 & 0,60 & 0,20 \end{pmatrix},$$

для которой  $\prod_{i=1}^3 p_{ii}(T) = 0,0135$  и  $\det \mathbf{P}(T) = 0,05$ . Так как (4.32) не выполняется, справедлив вывод, что  $\mathbf{P}(T)$  не может быть порождена марковским процессом. На практике элементы  $\mathbf{P}(T)$  были бы оценены, и, следовательно, они содержали бы ошибки. При отсутствии проверки на значимость надо при малых выборках с особой осторожностью относиться к слабым нарушениям неравенств. Другие необходимые условия основаны на собственных значениях матрицы  $\mathbf{P}(T)$ . Более подробное обсуждение можно найти в работе Сингера и Шпилермана (1977а).

Может создаться впечатление, что проще всего ответить на второй наш вопрос: как найти матрицу  $\mathbf{R}$  по заданной матрице  $\mathbf{P}(T)$ ? Так как эти две матрицы связаны соотношением

$$\mathbf{P}(T) = \exp \{ \mathbf{R}T \}, \quad (4.33)$$

вся проблема заключается в том, имеет ли (4.33) в требуемой форме единственное для  $\mathbf{R}$  решение. Оказывается, что это уравнение может иметь одно решение, несколько решений или не иметь их вовсе. Если решения не существует, то совершенно ясно, что матрица  $\mathbf{P}(T)$  не может быть порождена марковским процессом. Если существует несколько решений, то более одного процесса могут привести к наблюдаемой матрице  $\mathbf{P}(T)$ , и нет способов отличить их друг от друга. Только при единственном решении мы можем быть уверены в его правильности. При  $k = 2$  можно записать:

$$\mathbf{R} = \frac{1}{T} \ln \mathbf{P}(T) = \frac{\ln(p_{11}(T) + p_{22}(T) - 1)}{p_{11}(T) + p_{22}(T) - 2} \{ \mathbf{P}(T) - \mathbf{I} \}. \quad (4.34)$$

Для существования  $R$  значение  $p_{11}(T) + p_{22}(T) - 1$  должно быть действительным и строго положительным. Так как два собственных значения  $P$  равны 1 и  $p_{11}(T) + p_{22}(T) - 1$  соответственно, то предыдущее условие можно выразить через них. Выражение условия через собственные значения становится достаточным для всех значений  $k$ . Таким образом, если собственные значения матрицы  $P(T)$  действительны и положительны, то (4.33) имеет единственное решение. На практике такое предположение, по-видимому, часто выполняется. Если это так, то решение можно вычислить с помощью методов, рассмотренных Сингером и Шпилерманом (1976 а). Криди (1979) применил их метод к некоторым данным о потоках на рынке труда. В этом примере население, состоящее из белых молодых людей мужского пола в возрасте от 14 до 24 лет, в 1966 г. разделено на три категории: занятые ( $E$ ), безработные ( $U$ ) и нетрудоспособные ( $N$ ). За 1966—1967 гг. наблюдались переходные вероятности:

	$E$	$U$	$N$
$E$	8,600	0,0294	0,1028
$U$	0,6062	0,1631	0,2438
$N$	0,3341	0,0787	0,5872 .

Оценка матрицы  $R$  для них равна:

$$R = \begin{pmatrix} -0,117 & 0,005 & 0,011 \\ 0,117 & -0,170 & 0,056 \\ 0,031 & 0,020 & -0,052 \end{pmatrix}.$$

Если имеется программа для определения  $R$ , то проще применить ее, чем вычислять все собственные значения для проверки их положительности.

Если система наблюдается в трех и более моментах времени, то возможен дальнейший прогресс в идентификации процесса. Тогда можно прямо сравнивать матрицы перехода за два шага с квадратом матрицы перехода за один шаг. Примером является классическое исследование Блюмена и его соавторов (1955) (далее оно обозначается ВКМ) профессиональной мобильности в Соединенных Штатах Америки. Данные включают профессиональные категории индивидуума, наблюдавшиеся ежеквартально. Никакой информации о переходах внутри этих интервалов не было. Имелось 11 категорий, причем одна состояла из безработных, а 10 других обозначались  $A, B, \dots, K$ , как указано в табл. 4.1. Так как в распоряжении имелись данные по восьми последовательным кварталам, можно сравнить, например, восьмую степень матрицы перехода за один шаг с матрицей перехода за восемь шагов. Результат такого сравнения показан в табл. 4.1 для мужчин в возрасте 20—24 года.

Самое удивительное, что диагональные элементы наблюдаемой матрицы перехода за восемь кварталов гораздо больше, чем предсказанные с помощью возведения переходной матрицы за один квартал в восьмую степень. С аналогичным явлением мы сталкивались при анализе модели с дискретным временем в гл. 2. Там было показано, что

Таблица 4.1. Сравнение наблюдавшихся и предсказанных вероятностей перехода за восемь кварталов для мужчин в возрасте 20—24 года. Верхнее число — наблюдавшаяся доля, нижнее — предсказанная марковской теорией<sup>1</sup>

Код отрасли промышленности	A	B	C	D	E	F	G	H	J	K	U
A	0,000	0,062	0,062	0,000	0,125	0,156	0,312	0,000	0,000	0,000	0,281
	0,002	0,086	0,105	0,042	0,116	0,053	0,181	0,016	0,058	0,004	0,337
B	0,003	0,449	0,039	0,020	0,048	0,035	0,079	0,014	0,023	0,006	0,284
	0,002	0,144	0,087	0,040	0,104	0,050	0,163	0,018	0,052	0,004	0,336
C	0,002	0,037	0,461	0,023	0,046	0,021	0,101	0,007	0,022	0,002	0,278
	0,002	0,077	0,176	0,039	0,103	0,046	0,163	0,106	0,050	0,004	0,324
D	0,000	0,064	0,044	0,459	0,083	0,024	0,091	0,011	0,030	0,002	0,192
	0,001	0,070	0,080	0,218	0,099	0,046	0,141	0,015	0,047	0,003	0,279
E	0,002	0,045	0,042	0,034	0,489	0,031	0,094	0,010	0,023	0,002	0,227
	0,001	0,072	0,075	0,040	0,276	0,046	0,147	0,013	0,044	0,004	0,279
F	0,003	0,056	0,033	0,022	0,054	0,440	0,090	0,020	0,026	0,010	0,245
	0,002	0,081	0,076	0,038	0,097	0,166	0,152	0,017	0,050	0,004	0,316
G	0,002	0,047	0,051	0,025	0,046	0,038	0,491	0,020	0,044	0,002	0,235
	0,002	0,080	0,084	0,039	0,098	0,049	0,261	0,017	0,053	0,004	0,314
H	0,000	0,044	0,007	0,015	0,026	0,085	0,096	0,439	0,074	0,000	0,214
	0,001	0,077	0,077	0,035	0,090	0,048	0,170	0,158	0,052	0,004	0,287
J	0,002	0,061	0,033	0,018	0,054	0,035	0,145	0,019	0,339	0,000	0,294
	0,002	0,084	0,085	0,038	0,105	0,049	0,178	0,018	0,105	0,004	0,333
K	0,000	0,113	0,097	0,032	0,121	0,048	0,137	0,032	0,024	0,048	0,347
	0,002	0,089	0,096	0,047	0,130	0,048	0,179	0,023	0,056	0,006	0,325
U	0,001	0,069	0,068	0,035	0,077	0,040	0,153	0,018	0,055	0,004	0,482
	0,002	0,090	0,095	0,042	0,112	0,052	0,179	0,019	0,058	0,004	0,346

<sup>1</sup> Данные из работы Блюмена и др. (1955).

его можно объяснить либо неоднородностью (т.е. различными **P**), либо накопленной инерцией. Каждое из отмеченных в гл. 2 отклонений от простой модели имеет свой аналог для модели с непрерывным временем. Отклонение от диагональных элементов из-за накопленной инерции получается, если допустить, что интервалы времени между изменениями состояний имеют такое распределение, что склонности к переходу убывают по мере увеличения длительности пребывания в данном состоянии. Такие модели будут обсуждаться в разделе 4.4. Существует также третий тип отклонения, встречающийся у моделей с непрерывным временем. Здесь можно столкнуться с неоднородностью процесса, управляющего изменением состояния. Однако мы начнем анализ с непрерывной по времени МС-модели.

Как и ранее, предполагаем, что генеральная совокупность разделена на две части — «кочевников» и «оседлых». Предположим, что часть  $s_i$  категории  $i$  составляют «оседлые», что означает, что их переходная матрица есть **I**; остальные — «кочевники» с переходной матрицей **P**. Мы придерживаемся предположения, что изменения состояния происходят в точках, которые получаются из реализации пуассоновского

процесса с параметром  $\lambda$ . Таким образом, если  $\tau$  обозначает временной интервал наблюдения (в данном случае — квартал), то

$$P(\tau) = SI + \{I - S\} \sum_{m=0}^{\infty} \frac{(\lambda\tau)^m}{m!} e^{-\lambda\tau} P^m, \quad (4.35)$$

где  $S$  — диагональная матрица с элементами  $s_i$  на диагонали.

Теперь возникает вопрос: согласуется ли, например, переходная матрица за восемь кварталов  $P(8\tau)$ , предсказанная по (4.35), с матрицей, полученной по результатам наблюдений? ВКМ подогнули модель к своим данным; полученные результаты для диагональных элементов по группе 20—24-летних приведены в табл. 4.2.

**Таблица 4.2.** Сравнение наблюдавшихся и предсказанных значений диагональных элементов в переходных матрицах за восемь кварталов для мужчин в возрасте 20—24 лет<sup>1</sup>

Код отрасли промышленности	A	B	C	D	E	F	G	H	J	K	U
Наблюдение	0,000	0,449	0,461	0,459	0,489	0,440	0,491	0,439	0,339	0,048	0,482
Предсказание	0,003	0,442	0,464	0,474	0,512	0,444	0,489	0,446	0,338	0,049	0,536

<sup>1</sup> Данные из работы Блюмена и др. (1955).

Очевидно, что соответствие гораздо лучше, чем в табл. 4.1, так что, несмотря на свою грубость, МС-модель не оставляет желать лучшего. Можно было бы ввести и более тонкие формы неоднородности, но (по крайней мере в этом случае) они вряд ли необходимы.

Предыдущую модель можно описывать двумя способами. Выше мы предположили, что неоднородность заключена в различных переходных матрицах:  $I$  — в одном случае и  $P$  — в другом. Второй способ состоит в предположении, что все индивидуумы имеют одинаковую матрицу перехода, но различные скорости движения, причем они будут такими:  $\lambda$  — для «кочевников» и 0 — для «оседлых». Оба способа построения модели приводят к (4.35). Рассмотрим теперь последнюю трактовку. Предположим, что каждый индивидуум меняет свое состояние в соответствии с марковским процессом, который характеризуется распределением Пуассона с параметром  $\lambda$  и переходной матрицей  $P$ . Неоднородность связана с тем, что  $\lambda$  может меняться, обеспечивая тем самым различные скорости перемещения индивидуумов. Переходная матрица  $P(\tau)$  для интервала  $(0, \tau)$  легко определяется следующим образом. Для каждого данного индивидуума она равна  $\sum_{n=0}^{\infty} P_n(\tau; \lambda) P^n$ , где  $P_n(\tau; \lambda)$  — вероятность  $n$  перемещений на интервале  $(0, \tau)$  для распределения Пуассона. Это выражение дает искомые вероятности при данном  $\lambda$ . Чтобы найти безусловное распределение, мы

должны определить математическое ожидание относительно  $\lambda$ ; оно равно:

$$\mathbf{P}(\tau) = \int_0^{\infty} \sum_{n=0}^{\infty} P_n(\tau; \lambda) \mathbf{P}^n dF(\lambda), \quad (4.36)$$

где  $F(\lambda)$  — функция распределения  $\lambda$ . Меняя местами знаки суммирования и интегрирования, получаем

$$\mathbf{P}(\tau) = \sum_{n=0}^{\infty} P_n(\tau) \mathbf{P}^n, \quad (4.37)$$

где

$$P_n(\tau) = \int_0^{\infty} P_n(\tau; \lambda) dF(\lambda). \quad (4.38)$$

Заметим, что (4.35) оказывается частным случаем этой модели, когда распределение  $\lambda$  концентрируется в двух точках — 0 и  $\lambda$ .

Возникает вопрос: является ли типичным увеличение диагональных элементов матрицы за  $r$  шагов по сравнению с предсказанными  $r$ -й степенью переходной матрицы за 1 шаг, которое мы видели в МС-модели для последнего класса моделей? Ответ на этот вопрос будет получен в двух примерах, первый из них был построен ВКМ. Они рассмотрели состоящую из трех категорий систему, все члены которой характеризуются одной и той же переходной матрицей:

$$\mathbf{P} = \begin{pmatrix} 0,70 & 0,15 & 0,15 \\ 0,20 & 0,60 & 0,20 \\ 0,25 & 0,25 & 0,50 \end{pmatrix}. \quad (4.39)$$

Изменения состояния происходят в соответствии с пуассоновским процессом, для которого половина каждой категории имеет  $\lambda = \frac{1}{10}$ , а другая половина имеет  $\lambda = \frac{7}{10}$ . Для всей системы число изменений состояния на интервале  $(0, \tau)$  будет равно:

$$P_n(\tau) = \frac{1}{2n!} \left\{ \left( \frac{\tau}{10} \right)^n e^{-\tau/10} + \left( \frac{7\tau}{10} \right)^n e^{-7\tau/10} \right\} \quad (n = 0, 1, 2, \dots). \quad (4.40)$$

Полагая  $\tau = 1$ , получаем

$$P_0(1) = 0,701, \quad P_1(1) = 0,219, \quad P_2(1) = 0,063, \\ P_3(1) = 0,014, \quad P_4(1) = 0,002.$$

Среднее значение для этого распределения равно 0,395, дисперсия равна 0,437, что указывает на более высокий разброс, чем у пуассоновского распределения, имеющего такое же среднее значение. Подстановка этих величин и матрицы (4.39) в (4.37) дает

$$\mathbf{P}(\tau = 1) = \begin{pmatrix} 0,90 & 0,05 & 0,05 \\ 0,07 & 0,87 & 0,06 \\ 0,08 & 0,08 & 0,84 \end{pmatrix}. \quad (4.41)$$



Эта матрица могла бы быть рассчитана по данным, собранным в начале двух следующих друг за другом кварталов. Если рассматривать ее как переходную матрицу марковской цепи за 1 шаг, то можно предсказать переходную матрицу за 8 кварталов, возведя (4.41) в восьмую степень:

$$\{P(\tau = 1)\}^8 = \begin{pmatrix} 0,54 & 0,25 & 0,21 \\ 0,33 & 0,44 & 0,24 \\ 0,35 & 0,30 & 0,36 \end{pmatrix}.$$

Однако переходная матрица за 8 кварталов, которую можно ожидать для данной модели, имеет вид:

$$P(\tau = 8) = \begin{pmatrix} 0,63 & 0,21 & 0,17 \\ 0,26 & 0,55 & 0,19 \\ 0,28 & 0,23 & 0,48 \end{pmatrix}.$$

Таким образом, введенная неоднородность приводит к недооценке диагональных элементов. Именно это и обнаружили Блюмен и его соавторы (1955) на практике.

Не обязательно предполагать, что в генеральной совокупности существуют индивидуальные различия. Такой же результат можно получить, если моменты решения о переходе появляются для каждого индивидуума в соответствии с (4.40). Последнее возможно в случае, если моменты решения двух типов: первый встречается с интенсивностью  $\lambda = \frac{1}{10}$ , а второй — с  $\lambda = \frac{7}{10}$ . К сожалению, пользуясь имеющимися данными, нельзя заметить различие в этих двух моделях. При прогнозировании они идентичны; если же нам требуется полностью понять суть процесса, то нужны данные об истории каждого отдельного случая.

Шпилерман (1972b) рассмотрел ситуацию, когда  $\lambda$  распределено непрерывно, предположив, что

$$dF(\lambda) = \frac{c^\nu}{\Gamma(\nu)} \lambda^{\nu-1} e^{-c\lambda} d\lambda \quad (\lambda \geq 0; c, \nu > 0).$$

Это гибкая форма записи способна охарактеризовать различные типы распределений, начиная от J-образного распределения и кончая уни-модальным симметричным распределением. Мы воспользуемся этой записью еще раз в гл. 7. Подставляя  $dF(\lambda)$  в (4.38), временно интервала  $\tau$ , получаем:

$$P_m(\tau) = \int_0^\infty \frac{(\lambda\tau)^m}{m!} e^{-\lambda\tau} \frac{c^\nu}{\Gamma(\nu)} \lambda^{\nu-1} e^{-c\lambda} d\lambda =$$

$$= \frac{c^\nu \tau^m}{m! \Gamma(\nu)} \int_0^\infty \lambda^{m+\nu-1} e^{-\lambda(c+\tau)} d\lambda = \binom{m+\nu-1}{m} \left( \frac{c}{c+\tau} \right)^\nu \left( \frac{\tau}{c+\tau} \right)^m,$$

где по определению  $x! = \Gamma(x + 1)$  для всех целых  $x$ . Замена распределения  $P_m(\tau)$  в матричной записи (4.37) дает

$$\mathbf{P}(\tau) = \left(\frac{c}{c+\tau}\right)^v \sum_{m=0}^{\infty} \binom{m+v-1}{m} \left(\frac{\tau}{c+\tau}\right)^m \mathbf{P}^m. \quad (4.43)$$

Если бы  $\mathbf{P}$  было скаляром, то последнее выражение представляло бы собой сумму членов отрицательного биномиального распределения и в предположении сходимости было бы равно

$$\left(\frac{c}{c+\tau}\right)^v \left\{ \mathbf{I} - \frac{\tau}{c+\tau} \mathbf{P} \right\}^{-v}. \quad (4.44)$$

Конечно, мы можем определить (4.44) как сумму бесконечного ряда (4.43), но это не снимает задачи вычисления (4.44), когда  $v$  — не целое число. Чтобы обойти эту трудность, воспользуемся тем, что  $\mathbf{P}$  можно представить в виде

$$\mathbf{P} = \mathbf{H}\mathbf{D}\mathbf{H}^{-1},$$

где  $\mathbf{D}$  — матрица, у которой на диагонали стоят собственные значения матрицы  $\mathbf{P}$  и все остальные элементы равны нулю, а в матрице  $\mathbf{H}$   $i$ -й столбец является собственным вектором, соответствующим  $i$ -му собственному значению матрицы  $\mathbf{P}$ . Легко убедиться, что

$$\mathbf{P}^m = \mathbf{H}\mathbf{D}^m \mathbf{H}^{-1}.$$

Следовательно,

$$\mathbf{P}(\tau) = \left(\frac{c}{c+\tau}\right)^v \mathbf{H} \left\{ \sum_{m=0}^{\infty} \binom{m+v-1}{m} \left(\frac{\tau}{c+\tau}\right)^m \mathbf{D}^m \right\} \mathbf{H}^{-1}. \quad (4.45)$$

Бесконечная матричная сумма в скобках — это диагональная матрица, элемент  $(i, i)$  которой представляет собой скалярную сумму

$$\sum_{m=0}^{\infty} \binom{m+v-1}{m} \left(\frac{\tau}{c+\tau}\right)^m \theta_i^m = \left(1 - \frac{\tau\theta_i}{c+\tau}\right)^{-v}, \quad (4.46)$$

где  $\theta_i$  —  $i$ -е собственное значение матрицы  $\mathbf{P}$ . Ряды (4.45) и (4.46) сходятся потому, что для стохастической матрицы  $\theta_i \leq 1^*$  для всех  $i$  и  $\tau/(c+\tau) < 1$  для  $\tau > 0$ , что означает  $\tau\theta_i/(c+\tau) < 1$ . Эквивалентный способ суммирования ряда (4.43) состоит в использовании спектрального разложения  $\mathbf{P}^m$ , приведенного в гл. 3 (см. соотношение (3.6)).

Шпилерман (1972b) проделал некоторые численные сопоставления матрицы  $\{\mathbf{P}(\tau)\}^T$  с переходной матрицей за  $T$  шагов для описанного выше процесса. Он нашел, что последняя имеет большие диагональные элементы, подобно тому, как было в случае, когда  $\lambda$  распределено в двух точках.

Наконец, покажем, что такого рода результат всегда будет иметь место для модели, у которой  $\mathbf{P}(\tau)$  задана (4.36), если только собствен-

\* Более точно  $|\theta_i| \leq 1$ . — Примеч. ред.

ные значения  $\mathbf{P}$  действительные. В частности, мы найдем условие выполнения неравенства

$$\text{trace } \mathbf{P}(2\tau) \geq \text{trace } [\mathbf{P}(\tau)]^2, \quad (4.47)$$

аналогичное условию из гл. 2. Из (4.37)

$$\mathbf{P}(2\tau) = \sum_{n=0}^{\infty} P_n(2\tau) \mathbf{P}^n$$

и, следовательно,

$$\begin{aligned} \text{trace } \mathbf{P}(2\tau) &= \sum_{n=0}^{\infty} P_n(2\tau) \sum_{i=1}^k \theta_i^n = \\ &= \sum_{i=1}^k \sum_{n=0}^{\infty} \theta_i^n P_n(2\tau) = \sum_{i=1}^k E(\theta_i^n | 2\tau), \end{aligned} \quad (4.48)$$

где  $\theta_1, \theta_2, \dots, \theta_k$  — собственные значения матрицы  $\mathbf{P}$ . Далее имеем

$$\{\mathbf{P}(\tau)\}^2 = \left\{ \sum_{n=0}^{\infty} P_n(\tau) \mathbf{P}^n \right\}^2 = \sum_{n=0}^{\infty} C_n \mathbf{P}^n,$$

где  $C_n$  — коэффициент при  $\mathbf{P}^n$  в разложении квадрата.

$$\begin{aligned} \text{trace } \{\mathbf{P}(\tau)\}^2 &= \sum_{n=0}^{\infty} C_n \sum_{i=1}^k \theta_i^n = \sum_{i=1}^k \sum_{n=0}^{\infty} C_n \theta_i^n = \\ &= \sum_{i=1}^k \left\{ \sum_{n=0}^{\infty} P_n(\tau) \theta_i^n \right\}^2 = \sum_{i=1}^k E^2(\theta_i^n | \tau). \end{aligned} \quad (4.49)$$

Следовательно, условие выполнения (4.47) следующее:

$$\sum_{i=1}^k E(\theta_i^n | 2\tau) \geq \sum_{i=1}^k E(\theta_i^n | \tau). \quad (4.50)$$

Как только распределение  $n$  известно, непосредственно проверяется выполнение условия (4.50). Когда  $n$  имеет отрицательное биномиальное распределение (4.42), то легко обнаружить, что (4.50) выполняется для действительных  $\theta_i$ .

В общем случае, когда  $P_n(\tau)$  задано (4.38), имеем

$$\begin{aligned} E(s^n | \tau) &= \sum_{n=0}^{\infty} s^n \int_0^{\infty} \frac{(\lambda\tau)^n}{n!} e^{-\lambda\tau} dF(\lambda) = \\ &= \int_0^{\infty} e^{\lambda\tau(s-1)} dF(\lambda) = E(e^{\lambda\tau(s-1)}). \end{aligned}$$

Отсюда

$$E(s^n | 2\tau) - E^2(s^n | \tau) = \text{var}(e^{\lambda\tau(s-1)}) \geq 0$$

для всех действительных  $s$ . Если распределение  $\lambda$  не вырождается в единственную точку, то, полагая  $s$  равным по порядку каждому собственному значению, получаем, что (4.50) выполняется со строгими не-

равенствами. Таким образом, становится ясным, что увеличение диагональных элементов можно объяснить различной интенсивностью перемещения, но нужно помнить, что оно может быть также обусловлено и неоднородностью в переходных матрицах. Позже в этой главе мы увидим, что существует и третье возможное объяснение.

### Предельное поведение

При определенных обстоятельствах переходная матрица  $\mathbf{P}(\tau)$ , полученная из наблюдений, может дать информацию о предельном поведении основного процесса. Предположим, что все индивидуумы имеют одну и ту же переходную матрицу  $\mathbf{P}$ , но моменты осуществления этих переходов определяются произвольным стохастическим процессом. До сих пор мы рассматривали только процессы, полученные смешиванием пуассоновских процессов, но следующий результат является более общим. Обозначим через  $P_n(\tau)$  вероятность того, что на интервале  $(0, \tau)$  происходят  $n$  изменений состояния. (Хотя обозначение и похоже на смесь пуассоновских процессов, но в более общем случае вероятность может зависеть от распределения длительности интервала). Тогда очевидно, что

$$\mathbf{P}(\tau) = \sum_{n=0}^{\infty} P_n(\tau) \mathbf{P}^n. \quad (4.51)$$

Пусть  $\mathbf{p}$  есть установившийся вектор, связанный с  $\mathbf{P}$ , так что  $\mathbf{p} = \mathbf{p}\mathbf{P} = \mathbf{p}\mathbf{P}^2 = \dots$ . Умножив обе части (4.51) на  $\mathbf{p}$ , находим

$$\mathbf{p}\mathbf{P}(\tau) = \sum_{n=0}^{\infty} P_n(\tau) \mathbf{p}\mathbf{P}^n = \sum_{n=0}^{\infty} P_n(\tau) \mathbf{p} = \mathbf{p}. \quad (4.52)$$

Отсюда следует, что  $\mathbf{p}$  также является установившимся вектором для матрицы  $\mathbf{P}(\tau)$ , полученной из наблюдений. Этот результат позволяет найти предельную структуру вложенной марковской цепи. Вообще говоря, как мы увидим ниже, структура  $\mathbf{p}$  не та же, что предельная структура, наблюдаемая в некоторый отдаленный момент времени в будущем, поскольку данный метод не учитывает времени пребывания. Однако, если основной точечный процесс является таким, что среднее время пребывания одинаково для всех категорий, то  $\mathbf{p}$  в (4.52) будет одновременно и предельной структурой.

Если неоднородность содержится в переходных матрицах, а не в точечном процессе, то по наблюдению одной только  $\mathbf{P}(\tau)$  можно выяснить очень мало о предельном поведении. В данном случае ситуация, по существу, такая же, как в дискретном случае, рассмотренном в гл. 2 (см. с. 50). Те индивидуумы, которые подчиняются переходной матрице  $\mathbf{P}(h)$ , будут «оседать» на установившейся структуре, скажем,  $\mathbf{p}(h)$ , и общая предельная структура тогда будет взвешенным средним этих предельных векторов.

ВКМ взяли свою переходную матрицу за один квартал и по (4.52) нашли установившийся вектор  $\mathbf{p}$ . В табл. 4.3 приведен пример, где  $\mathbf{p}$  сравнивается с наблюдаемой структурой.

**Т а б л и ц а 4.3. Действительная и предсказанная структуры занятости, полученные с использованием марковской модели и переходной матрицы за один квартал<sup>1</sup>**

Профессиональные группы	(C, D, E)	G	(F, H)	(A, B, J, K)	U	Всего
Средний процент наблюдаемых рабочих	28,2	17,0	6,8	13,7	34,3	100
Предсказанный процент с использованием марковской модели	27,0	18,0	8,0	15,0	32,0	100

<sup>1</sup> Данные из работы Блюмена и др. (1955).

Поскольку согласование очень хорошее, возникает вопрос: позволяет ли наш анализ предельного поведения объяснить этот факт? Результат из табл. 4.3, конечно, согласуется с гипотезой о том, что неоднородность лежит в переменных интенсивностях переходов и что система была замкнута до тех пор, пока не пришла к равновесию. Установившуюся структуру теперь можно было бы интерпретировать как структуру, связанную с вложенной марковской цепью. Последнее утверждение не подтверждало бы гипотезу о том, что  $P$  меняется от индивидуума к индивидууму, если бы не было известно, что  $MC$ -модель может рассматриваться как частный случай такой, в которой  $\lambda$  переменная.

#### 4.4. ПОЛУМАРКОВСКИЕ МОДЕЛИ

В последнем разделе в ходе исследования зависимости между наблюдаемыми переходными матрицами и основным процессом мы уже коснулись нескольких обобщений марковской модели. Здесь продолжим обсуждение того же вопроса и перейдем к классу моделей, в которых склонность к переходу зависит от времени пребывания в данной категории. Это позволит нам включить в рассмотрение накопленную инерцию как форму проявления уменьшения склонности к переходу.

К определению полумарковского процесса можно подойти двумя путями, что соответствует двум уже упоминавшимся способам определения марковского процесса. Один вытекает из интенсивностей переходов, другой начинается с пуассоновского процесса изменения состояний, а заканчивается вложенной марковской цепью. Эти два подхода эквивалентны, хотя в конкретных случаях либо тот, либо другой обычно оказывается более естественным. То же самое относится и к полумарковским процессам. Здесь мы займемся тем подходом, который представляется наиболее приемлемым при моделировании профессиональной мобильности.

Обобщим марковский процесс, допустив, что интенсивности переходов являются функциями времени пребывания в данном состоянии. Тогда  $r_{ij}(t) \delta t$  ( $i \neq j$ ) — вероятность перехода из состояния  $i$  в состояние  $j$  за малый промежуток времени  $(t, t + \delta t)$ . Накопленная инерция соответствует предположению, что  $r_{ij}(t)$  есть убывающая функция  $t$ . Полумарковские модели возникают в страховом деле и при анализе

годовой смертности по разным причинам, где  $r_{ij}(t)$  могут быть коэффициентами смертности по различным причинам как функции возраста. В большинстве таких приложений требуется, чтобы система была открытой, и мы приведем примеры такого рода в гл. 5.

Для многих случаев удобно принять следующие обозначения:

$$\lambda_i(t) = \sum_{\substack{j=1 \\ j \neq i}}^k r_{ij}(t) = -r_{ii}(t),$$

$$M_{ij}(t) = r_{ij}(t)/\lambda_i(t) \quad (i \neq j), \quad (4.53)$$

где  $\lambda_i(t)$  — опасность ухода (см. гл. 7), связанная с перемещением из состояния  $i$ , а  $M_{ij}(t)$  — вероятность перехода из  $i$  в  $j$  при условии, что уход имел место в момент  $t$ . Мы положим ее равной нулю, если  $i=j$ , и тогда  $\sum_{j=1}^k M_{ij}(t) = 1$ .

При такой формулировке нет явного условия на переход из данного состояния в самое себя. В некоторых приложениях такие перемещения бывают бессмысленными, но иногда переходы такого рода не только происходят, но их можно и наблюдать. Например, в исследовании ВКМ профессиональной мобильности состояния основывались на группировках по отраслям промышленности. Если бы мы могли наблюдать переходы только при пересечении границ групп, то приведенная выше формулировка соответствовала бы действительности. С другой стороны, если мы наблюдаем за переменами мест работы, то такие перемены могут происходить и внутри одной и той же отраслевой группы. Тогда целесообразно считать, что  $\lambda_i(t)$  обозначает продолжительность пребывания на данном месте работы, а  $M_{ij}(t)$  относится к переходам с одного места работы на другое. В этом случае вовсе не обязательно иметь  $M_{ii}(t) = 0$ . Процесс, определенный таким способом, эквивалентен полумарковскому процессу, но тогда соотношения (4.53) уже не определяются через  $r_{ij}(t)$ .

Анализ модели мы начнем с того, что выразим некоторые функции вероятности, которые легко оценить, через интенсивности переходов. Плотность времени пребывания в состоянии  $i$ , для которого  $\lambda_i(t)$  — опасность ухода, равна:

$$f_i(t) = \lambda_i(t) \exp \left\{ - \int_0^t \lambda_i(x) dx \right\}. \quad (4.54)$$

В частном случае марковского процесса  $\lambda_i(t)$  постоянна, так что распределение времени пребывания является экспоненциальным. Чтобы узнать, нужно ли использовать полумарковскую модель, следует проверить, распределены ли экспоненциально действительные продолжительности пребывания в данном состоянии. Као (1974) при моделировании стадий лечения сердечно-сосудистых заболеваний выполнил такую проверку и нашел, что распределение продолжительности стадий лечения не может быть экспоненциальным. Он также обнаружил, что время, проведенное в одном состоянии (стадии лечения), зависит от конечного состояния. Если обозначить через  $f_{ij}(t)$  плотность распре-

деления вероятностей времени пребывания в состоянии  $i$  для тех, кто переходит в состояние  $j$ , то можно выразить ее через интенсивности переходов. Если  $G_i(t) = \int_t^{\infty} f_i(x) dx$ , то вероятность того, что индивидuum останется в состоянии  $i$  до момента  $t$ , а затем на интервале  $(t, t + \delta t)$  перейдет в состояние  $j$ , равна

$$G_i(t) r_{ij}(t) \delta t.$$

Интеграл от 0 до  $\infty$  последней величины представляет собой вероятность того, что индивидuum, находящийся в состоянии  $i$ , на следующем шаге перейдет в состояние  $j$ . Этот важный параметр с полным правом можно обозначить через

$$m_{ij} = \int_0^{\infty} r_{ij}(t) G_i(t) dt \quad (i \neq j). \quad (4.55)$$

Поскольку

$$f_{ij}(t) \delta t = Pr \{ \text{переход из } i \text{ в } j \text{ происходит за промежутком}$$

$(t, t + \delta t) / \text{следующий переход в } j \} \quad (i \neq j), \text{ то}$

$$f_{ij}(t) = r_{ij}(t) G_i(t) / m_{ij}. \quad (4.56)$$

Если процесс можно наблюдать непрерывно в течение достаточно длительного времени, то  $f_{ij}(t)$ ,  $f_i(t)$  и  $m_{ij}$  можно оценить непосредственно. Как только эти величины известны,  $r_{ij}(t)$  определяется по (4.56).

Важным является частный случай, когда  $f_{ij}(t)$  не зависит от  $j$ , что было обнаружено Янгом и Хуршем (1973) при исследовании характера сна. Он известен как модель *пропорционального риска*, так как это означает, что  $r_{ij}(t) / m_{ij}$  есть функция от  $i$  и  $t$ , но не от  $j$ . Запишем:

$$r_{ij}(t) = m_{ij} r_i(t) \quad (i \neq j)$$

и, следовательно,

$$\lambda_i(t) = r_i(t) \sum_{\substack{j=1 \\ j \neq i}}^k m_{ij} = r_i(t).$$

Это значит, что форма зависимости интенсивности от времени одинаковая, независимо от состояния, в которое совершается переход, но *уровень* интенсивности меняется в соответствии с состоянием, к которому делается переход. В этом случае  $M_{ij}(t) = m_{ij}$ , а последнее означает, что вероятность перехода для индивидуума, который перемещается в момент  $t$ , не зависит от момента перехода.

Полумарковская модель имеет много характерных особенностей, представляющих практический интерес. Например, Као (1974) специально изучал «траектории» больных внутри больницы, в частности, путем измерения средних значений и дисперсий времени пребывания. Далее в настоящей главе мы остановимся только на двух аспектах. Первый — какова матрица переходных вероятностей для фиксированного интервала времени. Как мы видели в последнем разделе, когда

процесс может наблюдаться только в дискретные моменты времени, нам надо знать, какие выводы можно сделать об изучаемом процессе. Второй аспект касается распределений в установившемся состоянии.

### Вероятности переходов на фиксированном интервале времени

Для полумарковского процесса эти вероятности будут зависеть от состояния системы в момент начала интервала наблюдения. Возьмем индивидуума, который только что попал в состояние  $i$ . Пусть  $p_{ij}(\tau)$  — вероятность того, что через время  $\tau$  он окажется в состоянии  $j$ . Выражение для этой вероятности можно построить следующим образом. Если  $j \neq i$ , то индивидуум должен совершить, по крайней мере, один переход на интервале  $(0, \tau)$ . Предположим, что первый переход он совершает в состояние  $h$  и это происходит на интервале  $(t, t + \delta t)$ . Тогда между  $t$  и  $\tau$  он должен перейти из  $h$  в  $j$ . Суммируя вероятности этих событий по  $h$  и  $t$ , получаем

$$p_{ij}(\tau) = \sum_{\substack{h=1 \\ h \neq i}}^k \int_0^{\tau} r_{ih}(t) G_i(t) p_{hj}(\tau - t) dt \quad (i \neq j). \quad (4.57)$$

Если  $j = i$ , то индивидуум может оставаться в состоянии  $i$  в течение всего интервала наблюдения, и вероятность этого события есть  $G_i(\tau)$ . В соответствии с определением  $f_{ij}(t)$  окончательный результат можно записать как

$$p_{ij}(\tau) = \delta_{ij} G_i(\tau) + \int_0^{\tau} \sum_{h=1}^k m_{ih} f_{ih}(t) p_{hj}(\tau - t) dt \quad (i, j = 1, 2, \dots, k), \quad (4.58)$$

где  $m_{ii} = 0$  для всех  $i$ , а  $\delta_{ij} = 1$  при  $i = j$  и нулю при  $i \neq j$ . В принципе этих уравнений достаточно для вычисления требуемых вероятностей переходов. Рассматривая время как дискретную переменную, интеграл в (4.58) можно заменить суммой и вычислить вероятности рекуррентно, как в работе Валлианта и Милковича (1977). Такие вычисления сразу не позволяют судить даже о качественном поведении системы, поэтому мы продолжим анализ (4.58), введя преобразование Лапласа. Но и применение этого сильного математического аппарата позволяет получить весьма скудные результаты.

Преобразование Лапласа от функции  $\Phi(x)$ :

$$\Phi^*(s) = \int_0^{\infty} \Phi(x) e^{-sx} dx.$$

Нам понадобится теорема о том, что преобразование Лапласа от свертки функций

$$\int_0^t \Phi_1(x) \Phi_2(t-x) dx$$

представляет собой произведение их преобразований  $\Phi_1^*(s) \Phi_2^*(s)$



Теперь, применив преобразование Лапласа к обеим частям (4.58), получаем

$$p_{ij}^*(s) = \delta_{ij} G_i^*(s) + \sum_{h=1}^k m_{ih} f_{ih}^*(s) p_{hj}^*(s), \quad (4.59)$$

Сумму в правой части можно представить как элемент  $(i, j)$  произведения двух матриц, одна из которых имеет элементы  $m_{ij} f_{ij}^*(s)$ , а другая —  $p_{ij}^*(s)$ . Следовательно, (4.59) можно переписать в виде

$$\mathbf{P}^*(s) = \mathbf{G}^*(s) + \mathbf{f}_m^*(s) \mathbf{P}^*(s), \quad (4.60)$$

где  $\mathbf{G}^*(s)$  — диагональная матрица с элементами  $G_i^*(s)$ , а  $\mathbf{f}_m^*(s)$  — матрица с элементами  $\{m_{ij} f_{ij}^*(s)\}$ . Таким образом, преобразование Лапласа от  $\mathbf{P}^*(s)$  находим сразу:

$$\mathbf{P}^*(s) = (\mathbf{I} - \mathbf{f}_m^*(s))^{-1} \mathbf{G}^*(s). \quad (4.61)$$

Несмотря на кажущуюся простоту этого выражения, получить из него явную формулу или качественную информацию о процессе нелегко, хотя в некоторых простых частных случаях все-таки извлечь кое-что удастся. Во-первых, полезно проследить, как возникают известные результаты для марковского процесса.

Для марковского процесса

$$m_{ij}(t) = r_{ij} e^{-\lambda_i t} \left( \lambda_i = \sum_{j \neq i} r_{ij} \right) \text{ и } G_i(t) = e^{-\lambda_i t}.$$

Тогда легко получить преобразования Лапласа, необходимые для (4.61):

$$m_{ij} f_{ij}^*(s) = r_{ij} / (\lambda_i + s) \text{ и } G_i^*(s) = 1 / (\lambda_i + s).$$

Подстановка в (4.61) дает

$$\mathbf{P}^*(s) = \{ \mathbf{I} - (\mathbf{R} + \Lambda_0) \Lambda^*(s) \}^{-1} \Lambda^*(s), \quad (4.62)$$

где  $\Lambda_0$  и  $\Lambda^*(s)$  — диагональные матрицы с элементами на главной диагонали  $\lambda_i$  и  $(\lambda_i + s)^{-1}$  соответственно. Замечая, что  $\{ \Lambda^*(s) \}^{-1} = \Lambda_0 + s \mathbf{I}$ , имеем

$$\mathbf{P}^*(s) = (\Lambda_0 + s \mathbf{I} - \mathbf{R} - \Lambda_0)^{-1} = (s \mathbf{I} - \mathbf{R})^{-1} = \sum_{r=0}^{\infty} \frac{\mathbf{R}^r}{s^{r+1}},$$

что представляет собой преобразование Лапласа от  $\mathbf{P}(\tau) = e^{\mathbf{R}\tau}$ .

Некоторые результаты можно получить для общего случая, если ввести в модель дополнительные ограничения или рассмотреть малые значения  $k$ . В простейшем случае при  $k = 2$  (4.61) принимает вид

$$\begin{aligned} \mathbf{P}^*(s) &= \begin{bmatrix} 1 & -m_{12} f_{12}^*(s) \\ -m_{21} f_{21}^*(s) & 1 \end{bmatrix}^{-1} \begin{bmatrix} G_1^*(s) & 0 \\ 0 & G_2^*(s) \end{bmatrix} = \\ &= (1 - m_{12} m_{21} f_{12}^*(s) f_{21}^*(s))^{-1} \begin{bmatrix} G_1^*(s) & m_{12} f_{12}^*(s) G_2^*(s) \\ m_{21} f_{21}^*(s) G_1^*(s) & G_2^*(s) \end{bmatrix}. \end{aligned} \quad (4.63)$$

Существование обратных преобразований Лапласа в любом случае зависит от вида функций  $r_{ij}(t)$ . Но предельный вид матрицы  $\mathbf{P}(t)$  при  $t \rightarrow \infty$  легко найти, воспользовавшись тем, что для любого преобразования

$$\lim_{t \rightarrow \infty} \Phi(t) = \lim_{s \rightarrow 0} s \Phi^*(s).$$

Для (4.63)

$$\mathbf{P}(\infty) = \frac{1}{\mu_1 + \mu_2} \begin{pmatrix} \mu_1 & \mu_2 \\ \mu_1 & \mu_2 \end{pmatrix}, \quad (4.64)$$

где  $\mu_1$  и  $\mu_2$  — средние значения времени пребывания в двух состояниях.

При пропорциональных рисках выражение для преобразования Лапласа упрощается, поскольку

$$\mathbf{f}_m^*(s) = \mathbf{f}^*(s) \mathbf{M},$$

где  $\mathbf{f}^*(s)$  — диагональная матрица с элементами  $\{f_i^*(s)\}$ . Вероятно, простейший пример моделей с пропорциональным риском возникает при  $r_{ij}(t) = m_{ij} a t^b$  ( $b > -1$ ), но даже в этом случае нет явных выражений для нужных преобразований Лапласа. Значительно продвинуться вперед можно, лишь сделав дополнительное предположение о том, что  $f_i^*(s) = f^*(s)$  для всех  $i$ . Это означает, что интенсивность переходов одинакова для всех состояний. В этом случае

$$\mathbf{P}^*(s) = \{\mathbf{I} - f^*(s) \mathbf{M}\}^{-1} \mathbf{G}^*(s) = \{\mathbf{I} - f^*(s) \mathbf{M}\}^{-1} \{1 - f^*(s)\}/s. \quad (4.65)$$

Последний шаг основан на том, что преобразование Лапласа от  $\int_0^t \Phi(x) dx$  есть  $\Phi^*(s)/s$ . Разлагая обратную матрицу в (4.65) в ряд, имеем

$$\begin{aligned} \mathbf{P}^*(s) &= \left\{ \frac{1 - f^*(s)}{s} \right\} \sum_{n=0}^{\infty} \{f^*(s)\}^n \mathbf{M}^n = \\ &= \sum_{n=0}^{\infty} \left( \frac{\{f^*(s)\}^n - \{f^*(s)\}^{n+1}}{s} \right) \mathbf{M}^n. \end{aligned} \quad (4.66)$$

Теперь  $\{f^*(s)\}^n$  — преобразование Лапласа плотности сумм  $n$  независимых случайных величин с функцией плотности  $f(t)$  и, в силу приведенного выше результата,  $\{f^*(s)\}^n/s$  преобразование функции распределения их сумм. Применяя обратное преобразование к обеим частям (4.66), получаем

$$\mathbf{P}(\tau) = \sum_{n=0}^{\infty} \{F_n(\tau) - F_{n+1}(\tau)\} \mathbf{M}^n. \quad (4.67)$$

Разность вероятностей  $F_n(\tau)$  и  $F_{n+1}(\tau)$  представляет собой вероятность того, что на интервале  $(0, \tau)$  происходят  $n$  изменений состояния. Следовательно, окончательно имеем

$$\mathbf{P}(\tau) = \sum_{n=0}^{\infty} P_n(\tau) \mathbf{M}^n, \quad (4.68)$$

что, очевидно, приводит к (4.51) и могло быть получено более простым способом. При этом необходимо помнить, что в настоящем разделе интервал  $(0, \tau)$ , начинается с того момента, когда индивидuum попадает в исследуемое состояние.

## Примеры

Проблема применения (4.68) состоит в определении  $P_n(\tau)$  по предполагаемым интенсивностям переходов. Сингер и Шпилерман (1979) рассмотрели случай, когда

$$r_{ij}(t) \propto \lambda^2 t / (1 + \lambda t) \text{ для всех } i \text{ и } j. \quad (4.69)$$

Соотношение (4.69) представляет собой опасность перехода, которая возрастает с увеличением длительности пребывания в данном состоянии, причем скорость возрастания опасности стремится к нулю, когда  $t$  становится большим. Сингер и Шпилерман как пример применения такой модели обсуждают работу о стереотипных моделях психологов-эволюционистов. В социологии для характеристики возрастающей скорости перехода используется термин «кумулятивный стресс» в отличие от накопленной инерции, когда эта интенсивность уменьшается. Сингеру и Шпилерману (1979) удалось показать, что для модели (4.69)

$$P(\tau) = e^{-\lambda\tau} \left[ \cosh \lambda\tau M^{\frac{1}{2}} + M^{-\frac{1}{2}} \sinh \lambda\tau M^{\frac{1}{2}} \right], \quad (4.70)$$

где гиперболические синус и косинус как функции матричного аргумента определяются обычным способом с помощью их разложений в степенные ряды. Чтобы проанализировать, чем такой процесс отличается от марковского, сравним  $P(2\tau)$  и  $\{P(\tau)\}^2$ . Возводя обе части (4.70) в квадрат, находим

$$\begin{aligned} P^2(\tau) &= P(2\tau) + \frac{1}{2} e^{-2\lambda\tau} (I - M^{-1}) (I - \cosh 2\lambda\tau M^{\frac{1}{2}}) = \\ &= P(2\tau) + \frac{1}{2} e^{-2\lambda\tau} \sum_{n=1}^{\infty} \frac{(2\lambda\tau)^{2n}}{(2n)!} (M^{n-1} - M^n). \end{aligned} \quad (4.71)$$

Ясно, что  $P^2(\tau)$  и  $P(2\tau)$  отличаются друг от друга. Чтобы понять характер различия, рассмотрим, как в гл. 2, следы этих двух матриц. Соответствующее выражение имеет вид

$$\text{trace}(M^{n-1} - M^n) = \sum_{i=1}^k \theta_i^{n-1} (1 - \theta_i),$$

где величины  $\theta$  — собственные значения матрицы  $M$ . Последнее выражение будет положительным, если все  $\theta$  действительные и положительные (мы знаем, что  $|\theta_i| < 1$ ). Следовательно, диагональные элементы переходной матрицы за два шага будут меньше, в среднем, чем пред-

сказанные марковской моделью. Более сильный результат можно получить, если рассмотреть только малые значения  $\tau$ . Тогда

$$P^2(\tau) \doteq P(2\tau) + \frac{1}{4}(1 - e^{-2\lambda\tau})(\lambda\tau)(I - M)^*,$$

откуда следует, что каждый диагональный элемент матрицы  $P(2\tau)$  меньше соответствующего элемента матрицы  $P^2(\tau)$ .

Непросто найти убывающую интенсивность перехода, для которой можно провести аналогичное исследование. Поэтому рассмотрим частный случай, когда  $k = 2$ . Для него можно найти обратное преобразование Лапласа в (4.63). Предположим, что

$$f(t) = p\lambda_1 e^{-\lambda_1 t} + (1-p)\lambda_2 e^{-\lambda_2 t}.$$

Мы встретимся с этим распределением в гл. 7. Соответствующая интенсивность перехода — убывающая функция от  $t$ . В этом случае можно записать:

$$P(\tau) = \begin{pmatrix} m_{11}(\tau) & 1 - m_{11}(\tau) \\ 1 - m_{11}(\tau) & m_{11}(\tau) \end{pmatrix},$$

где

$$m_{11}^*(s) = \frac{G^*(s)}{[1 - \{f^*(s)\}]^2} = \frac{\left[ \frac{p}{\lambda_1 + s} + \frac{p}{\lambda_2 + s} \right]}{\left[ 1 - \left\{ \frac{p\lambda_1}{\mu_1 + s} + \frac{(1-p)\lambda_2}{\mu_2 + s} \right\} \right]^2}.$$

Последнее представляет собой рациональную алгебраическую дробь, следовательно, ее можно представить в виде

$$m_{11}^*(s) = \frac{A}{s+a} = \frac{B}{s+b} + \frac{C}{s+c} + \frac{D}{s+d},$$

поэтому обратное преобразование Лапласа дает

$$m_{11}(\tau) = Ae^{-a\tau} + Be^{-b\tau} + Ce^{-c\tau} + D^{-d\tau}.$$

Диагональные элементы  $P(2\tau)$  будут больше соответствующих элементов  $\{P(\tau)\}^2$  при условии

$$m_{11}(2\tau) > m_{11}^2(\tau) + \{1 - m_{11}(\tau)\}^2. \quad (4.72)$$

Если принять  $p = \frac{1}{2}$ ;  $\lambda_1 = \frac{5}{2}$ ;  $\lambda_2 = \frac{5}{8}$ , то

$$m_{11}(\tau) = 0,5 + 0,1231 e^{-0,8049\tau} + 0,3769 e^{-3,8826\tau}; \quad (4.73)$$

Расчеты показывают, что  $m_{11}\left(\frac{1}{2}\right) = 0,6364$ ;  $m_{11}(1) = 0,5628$ ;  $m_{11}(2) = 0,5248$ , поэтому легко убедиться в том, что (4.46) выполняется для  $\tau = \frac{1}{2}$  и  $\tau = 1$ .

\* Обозначение  $\doteq$  означает приближенное равенство при достаточно малых  $\tau$ . — Примеч. ред.

Мы считаем, что накопленная инерция всегда будет приводить к увеличению диагональных элементов матрицы  $P(2\tau)$  по сравнению с предсказанными по  $P^2(\tau)$ , и что кумулятивный стресс будет приводить к их уменьшению.

Подытожим выводы, полученные в этом и предыдущем разделах, относительно поведения непрерывного по времени процесса, наблюдаемого на фиксированных интервалах времени; мы выявили три случая, когда отклонения от марковской модели могут привести к заниженному прогнозу значений диагональных элементов, а именно неоднородности в:

- 1) переходных матрицах  $P$ ;
- 2) показателе интенсивности  $\lambda$ ;
- 3) накопленной инерции, приводящей к убыванию интенсивности переходов.

Чтобы определить, с каким из этих случаев имеем дело, нужно rasполагать более подробными данными. В частности, случай 3 можно установить по характеру распределения интервалов между переходами.

### Наблюдение на интервале с произвольной начальной точкой

В предыдущем исследовании рассматривались только такие временные интервалы, которые начинаются с момента попадания индивидуума в данное состояние. Если мы начинаем наблюдать процесс в некоторой произвольной точке, то вероятности переходов на интервале будут зависеть от того, как долго индивидуум находился в определенном состоянии до начала интервала. Если такая информация есть, то легко получить необходимую модификацию. Единственное слагаемое в (4.58), подлежащее изменению, — первое, содержащее вероятность постоянного пребывания в состоянии  $i$  в течение всего интервала. Для индивидуума, который находился в состоянии  $i$  в течение времени  $x$  до момента начала наблюдения, эта вероятность равна:

$$\exp \left\{ - \int_x^{x+\tau} \lambda_i(y) dy \right\} = G_i(\tau + x) / G_i(\tau).$$

Тогда вероятность для случайно выбранного индивидуума определяют умножением на функцию плотности от  $x$  и интегрированием.

Если система находится в равновесии, то функцию плотности от  $x$  можно получить следующим образом. При выполнении принятых условий безразлично, с какого момента пребывания в данном состоянии начинать период наблюдения. Так что

$$x = yt,$$

где  $t$  имеет функцию плотности  $f_i(t)$ ,  $y$  — независимая случайная величина, равномерно распределенная на интервале  $(0, 1)$ . Следовательно, условная функция плотности распределения  $x$  однородна на интервале  $(0, t)$  и должна быть усреднена относительно распределения  $t$ . Плотность последнего не равна  $f_i(t)$ , так как, если система наблюда-

ется в случайные моменты времени, мы попадаем в большой интервал скорее, чем в малый. Таким образом, вероятность того, что этот интервал лежит в  $(t, t + \delta t)$ , будет пропорциональна  $f_i(t)\delta t$  и  $t$ , т. е.

$$f_i(x) = \int_x^{\infty} \frac{1}{t} \left\{ \frac{t f_i(t)}{\int_0^{\infty} u f_i(u) du} \right\} dt = G_i(x)/\mu_i. \quad (4.74)$$

Отсюда первое слагаемое в (4.58) принимает вид

$$\delta_{ij} \int_0^{\infty} \frac{G_i(\tau+x)}{G_i(x)} \frac{G_i(x)}{\mu_i} dx = \frac{\delta_{ij}}{\mu_i} \int_{\tau}^{\infty} G_i(x) dx, \quad (4.75)$$

где  $\mu_i = \int_0^{\infty} t f_i(t) dt$ . Учитывая, что преобразование Лапласа от  $\int_{\tau}^{\infty} G_i(x) dx$  равно  $\{1 - G_i^*(s)\}/s$ , (4.61) можно записать как

$$P^*(s) = (I - f_m^*(s))^{-1} D \{I - G^*(s)\}, \quad (4.76)$$

где  $D$  — диагональная матрица с элементами  $\left\{ \frac{1}{\mu_i s} \right\}$ .

### Предельный режим

Предельный режим полумарковской цепи в отличие от переходного поведения получить легко. Его можно вывести эвристически следующим образом. Рассмотрим достаточно большой интервал времени после того, как система достигла своего установившегося состояния, и предположим, что на этом интервале было  $K$  изменений состояния. Тогда ожидаемое число попаданий индивидуумов системы в состояние  $i$  будет равно  $K m_i$ , где  $m = (m_1, m_2, \dots, m_k)$  — установившаяся структура, соответствующая переходной матрице  $M$  с элементами, заданными (4.55). При каждом попадании одного индивидуума ожидаемое время пребывания в состоянии  $i$  будет  $\mu_i$  и, следовательно, ожидаемое общее время пребывания в этом состоянии равно  $K m_i \mu_i$ . Доля времени, проведенного в состоянии  $i$ , следовательно, равна  $m_i \mu_i / \sum_{j=1}^k m_j \mu_j$ , что можно интерпретировать как вероятность обнаружения случайно выбранного индивидуума в состоянии  $i$ .

Таким образом, для системы размера  $N$  записываем:

$$E \bar{n}_i(\infty) = N m_i \mu_i \quad (i = 1, 2, \dots, k). \quad (4.77)$$

Если  $f_i(t) = f(t)$  для всех  $i$ , то среднее значение времени пребывания будет одинаковым и предельная структура будет той же, что и для вложенной марковской цепи с переходной матрицей  $M$ . Это, по существу, результат (4.52), но в иной форме. Другие частные случаи (4.77) проявляются в (4.29) и (4.64).

## Модели, основанные на более общих точечных процессах

Вполне возможно представить себе дальнейшие обобщения, вводя в рассмотрение разные точечные процессы. Хотелось бы ослабить два специфических предположения относительно полумарковского процесса: его однородность во времени и независимость длительностей последовательного времени пребывания в различных состояниях. Так как оба обобщения сохраняют особенности вложенной марковской цепи, все свойства системы, вытекающие из этих особенностей, остаются неизменными. Например, результат о стационарной структуре, приведенный в (4.52), остается достаточно общим, даже если  $P_n(\tau)$  может зависеть не только от длины интервала наблюдения, но и от других характеристик стохастического процесса. Некоторые дальнейшие обобщения следуют почти немедленно. Предположим, например, что точечный процесс является зависящим от времени пуассоновским процессом с интенсивностью  $\lambda(T)$ . Если мы изменим масштаб времени,

$$Y(T) = \int_0^T \lambda(x) dx,$$

то во временной шкале  $Y$  изменения состояния происходят согласно однородному по времени процессу и можно пользоваться теорией, изложенной в разделах 4.2 и 4.3. Результаты, полученные в так называемом операционном масштабе времени, можно при необходимости преобразовать в «реальное» время. Серьезным препятствием на пути подобных обобщений является большое количество разнообразных данных, которые требуются для реализации таких моделей.

## 5.1. МАРКОВСКАЯ МОДЕЛЬ С ЗАДАНЫМ ВХОДОМ

Теория открытых систем — непосредственное развитие теории замкнутых систем, которые обсуждались в гл. 4. Поэтому в настоящей главе мы сосредоточим внимание только на вновь возникающих особенностях и на приложениях. Так же как и в гл. 3, будем считать систему открытой, если имеются уходы из системы, и рассмотрим два основных класса моделей, различающихся допущениями о входных потоках. При исследовании когорт применяются модели с нулевым входом. В терминах теории случайных процессов — это непрерывные марковские процессы с поглощающими состояниями.

Базовые уравнения модели выводятся из аналогичных моделей с дискретным временем при учете ограничений, приведенных в начале гл. 4. Если рассматривать изменение состояния за малый интервал времени  $\delta T$  (3.16), то

$$\bar{n}(T + \delta T) = \bar{n}(T) \mathbf{P}(\delta T) + R(T + \delta T) p_0. \quad (5.1)$$

Осуществляя подстановку  $\mathbf{P}(\delta T) = \mathbf{I} + \delta T \mathbf{R}$  и устремляя в уравнении  $\delta T \rightarrow 0$ , получаем

$$\frac{d\bar{n}(T)}{dT} = \bar{n}(T) \mathbf{R} + R(T) p_0, \quad (5.2)$$

где  $R(T)$  — мгновенная интенсивность поступлений, которую мы ввели вместо того, чтобы анализировать все поступления в систему за интервал времени от  $T$  до  $T + \delta T$ . Важно пояснить свойства матрицы  $\mathbf{R}$ , которые следуют, в нашем случае, из соответствующей матрицы  $\mathbf{P}(\delta T)$ . Диагональные элементы  $\mathbf{R}$  определяются так:

$$\begin{aligned} \delta T r_{ii} &= p_{ii}(\delta T) - 1 = \left\{ 1 - \sum_{j=1}^{k+1} p_{ij}(\delta T) - 1 \right\} = \\ &= -\delta T \sum_{\substack{j=1 \\ j \neq i}}^{k+1} r_{ij} \quad (i = 1, 2, \dots, k). \end{aligned}$$

В последней сумме  $r_{i, k+1}$  — интенсивность уходов из системы, среди элементов  $\mathbf{R}$  она отсутствует, оставаясь только составной частью диагонального элемента. Сумма элементов  $i$ -й строки  $\mathbf{R}$ , таким образом, больше не обращается в нуль, а равна  $-r_{i, k+1}$ .



Если проинтегрировать линейные дифференциальные уравнения (5.2), применяя стандартные методы, то получим

$$\bar{n}(T) = n(0) P(T) + p_0 \int_0^T R(x) P(T-x) dx. \quad (5.3)$$

Это уравнение — непрерывный аналог (3.2). К тому же результату можно прийти непосредственно при использовании вероятностных терминов. Первый член в правой части (5.3) дает ожидаемое число индивидуумов в каждом состоянии, обусловленное начальным состоянием, так же как в замкнутой модели. Во втором члене  $p_0 R(x) dx$  — «начальный» вектор, представляющий тех, кто поступает за период  $(x, x + \delta x)$ . По прошествии времени  $T - x$  распределение индивидуумов по состояниям определяется выражением  $p_0 R(x) P(T-x) dx$ . Интегрируя по  $x$  от 0 до  $T$ , находим суммарное распределение индивидуумов по состояниям. Аналогичное выражение получается для ожидаемого числа покинувших систему. Используя индекс  $k + 1$  для обозначения числа покинувших систему, запишем:

$$\bar{n}_{k+1}(T) = \sum_{i=1}^k p_{0i} \int_0^T R(x) p_{ij}(T-x) dx + n_i(0) p_{i, k+1}(T). \quad (5.4)$$

Входной поток не обязательно должен быть непрерывным. Достаточно легко рассмотреть потоки в дискретные моменты времени, в этом случае интегралы в (5.3) и (5.4) заменяются суммами.

При дискретном времени мы можем объединить (5.3) и (5.4) в одно уравнение, используя расширенные векторы и более сложные матрицы. Если обозначить через  $k + 1$  состояние, в которое включены те, кто покинул систему, то это уравнение будет иметь вид\*:

$$\bar{n}_a(T) = n_a(0) P_a(T) + p_{0a} \int_0^T R_a(x) P_a(T-x) dx. \quad (5.5)$$

В расширенном варианте (5.2) матрица  $R$  заменяется матрицей

$$R_a = \begin{pmatrix} r_{11} & r_{12} & \dots & r_{1k} & r_{1, k+1} \\ r_{21} & r_{22} & \dots & r_{2k} & r_{2, k+1} \\ \vdots & \vdots & & \vdots & \vdots \\ r_{k1} & r_{k2} & \dots & r_{kk} & r_{k, k+1} \\ 0 & 0 & \dots & 0 & 0 \end{pmatrix}. \quad (5.6)$$

В этой матрице суммы элементов строк равны нулю. Нетрудно построить модель так, как будет рассмотрено в первом примере, чтобы поглощающим было не только одно состояние. Матрица  $R_a$  сформирована из  $R$  с добавлением нулевой строки для каждого поглощающего состояния. Для того чтобы определить вектор  $\bar{n}(T)$  ожидаемых численностей, необходимо использовать (5.5). Первый шаг состоит в нахождении

\* Вектор  $\bar{n}_a$  имеет размерность  $k + 1$  и отличается от вектора  $\bar{n}$  тем, что к  $\bar{n}$  добавлена компонента  $n_{k+1}$ , которая описывается (5.4). — *Примеч. ред.*

$R_a(T)$  методом из гл. 4. Рассуждения, приводящие к (4.8), правомерны для открытой системы, если применяются расширенные матрицы  $P(T)$  и  $R$ . Вектор уходов получается сразу при подстановке в (5.5).

### Предельный режим

Нас интересует предельный режим вектора  $\bar{n}(T)$ , а он зависит, как и при дискретном времени, от вида функций интенсивностей поступления  $R(T)$ . Требуемый результат может быть получен непосредственно из решения дифференциальных уравнений, но общий результат Мелмана (1977b) больше говорит о поведении системы. Этот результат получен для функций поступлений  $R(T)$  вида  $Re^{\alpha T}$ . При  $\alpha = 0$  он включает постоянные поступления, которые нас интересуют в первую очередь. Из (5.2) следует, что выполняется уравнение

$$\bar{n}(\infty)R + Rp_0 = 0$$

и, следовательно,

$$\bar{n}(\infty) = Rp_0(-R)^{-1}. \quad (5.7)$$

Далее введем вектор  $v(T) = \bar{n}(T)e^{-\alpha T}$ , для которого имеем

$$\frac{dv(T)}{dT} = e^{-\alpha T} \frac{d\bar{n}(T)}{dT} - \alpha v(T).$$

Подставляя этот вектор в (5.2), получаем

$$\frac{dv(T)}{dT} = v(T)(R - \alpha I) + R(T)e^{-\alpha T} p_0. \quad (5.8)$$

Это то же самое, что (5.2), только  $R$  заменяется на  $R - \alpha I$  и  $R(T)$  — на  $R(T)e^{-\alpha T}$ . Следовательно, по предположению

$$\lim_{T \rightarrow \infty} R(T)e^{-\alpha T} = R,$$

и предельный вектор  $v(\infty)$  может быть получен по аналогии с (5.7):

$$v(\infty) = Rp_0(\alpha I - R)^{-1}, \quad (5.9)$$

отсюда

$$\bar{n}(\infty) \sim e^{\alpha T} Rp_0(\alpha I - R)^{-1}. \quad (5.10)$$

Приведенный результат имеет место, если существует обратная матрица, что бывает не всегда. Можно показать, что условие выполнения (5.10) заключается в том, что  $\alpha$  должно быть больше, чем наибольшая из действительных частей собственных чисел  $R$ . Поскольку эти действительные части равны нулю или отрицательны, то (5.10) выполняется, когда асимптотическая скорость увеличения потока положительная. Если темпы сокращения системы велики, то (5.10) не выполняется.

Уравнение (5.10) позволяет высказать некоторые полезные суждения об относительных размерах ожидаемых численностей, даже если размеры системы неограниченно увеличиваются или уменьшаются.

Обозначив через  $q(\infty)$  вектор относительного ожидаемого числа индивидуумов, мы получаем

$$q(\infty) \propto R p_0 (\alpha I - R)^{-1}. \quad (5.11)$$

Проведенный выше анализ является непрерывным аналогом дискретной модели Файхтингера (1976), описанной в гл. 3.

### Распределение численностей

До сих пор мы интересовались только математическими ожиданиями численностей градаций (квалификационных групп). Желательно также иметь некоторую информацию о распределениях численностей, их дисперсии и ковариации. Пока не существует непрерывного варианта метода Полларда, который мы применяли для дискретного времени, но легко обобщить метод Стаффа и Ваголкара. Этот метод позволяет получить наглядный результат, если входной поток пуассоновский, кроме того, с его помощью можно найти непрерывный аналог ранее установленного результата об условиях, при которых размеры численностей состояний распределены по закону Пуассона.

Итак, требуется получить совместное распределение численностей групп в момент времени  $T$ , если дано, что поток поступлений в систему пуассоновский с интенсивностью  $R$ . Как и прежде, определим совместное распределение численностей за исключением тех, кто находится в системе с начального момента времени. Это позволит нам получить асимптотическое распределение, считая, что в определенный момент времени все те, кто был вначале в системе, покинут ее.

Предположим, что к моменту времени  $T$   $m$  индивидуумов поступят в систему в моменты времени  $T_1, T_2, \dots, T_m$ . Найдем производящую функцию распределения числа индивидуумов по градациям в зависимости от  $T_1, T_2, \dots, T_m$  и  $m$ . Безусловную производящую функцию определим усреднением условных производящих функций относительно общих распределений  $T$  и  $m$ . Мы произведем эту операцию в два этапа, опираясь на тот факт, что при заданных  $m$  и  $T$   $T_1, T_2, \dots, T_m$  могут рассматриваться как случайная выборка из равномерного распределения с плотностью

$$f(x) = \frac{1}{T} \quad (0 \leq x \leq T),$$

а число  $m$  имеет распределение Пуассона со средним  $RT$ .

Пусть  $X_{hj} = 1$ , если индивидуум, поступивший в момент времени  $T_h$ , в момент времени  $T$  находится в  $j$ -м состоянии, и нулю в противном случае. Тогда используя переходные вероятности, можно записать:

$$Pr \{X_{hj} = 1\} = \sum_{i=1}^k p_{0i} p_{ij}(T - T_h). \quad (5.12)$$

Следовательно, производящая функция распределения вероятностей  $X_{h1}, X_{h2}, \dots, X_{h, k+1}$  при заданных  $T_h$  и  $m$  имеет вид

$$g(z_{h1}, z_{h2}, \dots, z_{h, k+1} | T_h, m) = \sum_{j=1}^{k+1} z_{hj} \sum_{i=1}^k p_{0i} p_{ij}(T - T_h), \quad (5.13)$$

а производящая функция совместного распределения вероятностей  $\{X_{hj} : h = 1, 2, \dots, m; j = 1, 2, \dots, k+1\}$  имеет вид

$$g(z_{11}, z_{12}, \dots, z_{h, k+1} | \mathbf{T}, m) = \prod_{h=1}^m \sum_{j=1}^{k+1} z_{hj} \sum_{i=1}^k p_{0i} p_{ij} (T - T_h). \quad (5.14)$$

Когда требуется условное распределение сумм  $\sum_{h=1}^m X_{hj}$ , их производящие функции получают подстановкой  $z_{hj} = z_j$ , скажем, для  $h = 1, 2, \dots, m$ :

$$g(z_1, z_2, \dots, z_{k+1} | \mathbf{T}, m) = \prod_{h=1}^m \sum_{j=1}^{k+1} z_j \sum_{i=1}^k p_{0i} p_{ij} (T - T_h). \quad (5.15)$$

Усредняя по  $\mathbf{T}$ , находим

$$\begin{aligned} g(z_1, \dots, z_{k+1} | m) &= \frac{1}{T^n} \int_0^T \dots \int_0^T g(z_1, \dots, z_{k+1} | \mathbf{T}, m) dT_1 \dots dT_m = \\ &= \prod_{h=1}^m \sum_{j=1}^{k+1} z_j \sum_{i=1}^k p_{0i} \frac{1}{T} \int_0^T p_{ij} (T - T_h) dT_h = \\ &= \left[ \frac{1}{T} \sum_{j=1}^{k+1} z_j \sum_{i=1}^k p_{0i} \int_0^T p_{ij} (T - x) dx \right]^m \end{aligned} \quad (5.16)$$

и, наконец,

$$\begin{aligned} g(z_1, z_2, \dots, z_{k+1}) &= \sum_{m=0}^{\infty} \frac{(RT)^m}{m!} e^{-RT} g(z_1, \dots, z_{k+1} | m) = \\ &= \exp \left\{ -RT + R \left[ \sum_{j=1}^{k+1} z_j \sum_{i=1}^k p_{0i} \int_0^T p_{ij} (T - x) dx \right] \right\} = \\ &= \exp \left\{ R \sum_{j=1}^{k+1} \sum_{i=1}^k p_{0i} \int_0^T p_{ij} (T - x) dx (z_j - 1) \right\}. \end{aligned} \quad (5.17)$$

Это производящая функция распределения вероятностей  $k+1$  независимых случайных величин, имеющих распределение Пуассона. Следовательно, если не учитывать тех индивидуумов, которые находились в системе в начальный момент времени, численности градаций в момент времени  $T$  будут иметь распределения Пуассона со средними

$$R \sum_{i=1}^k p_{0i} \int_0^T p_{ij} (T - x) dx \quad (j = 1, 2, \dots, k+1), \quad (5.18)$$

что соответствует (5.3).

Этот результат может быть обобщен, когда поступления представляют собой зависящий от времени пуассоновский процесс. В этом слу-

чае моменты поступлений  $T_1, T_2, \dots, T_m$  будут рассчитываться как случайные величины, распределение которых определяется плотностью

$$f(x) = \frac{R(x)}{\int_0^T R(x) dx} \quad (0 \leq x \leq T),$$

а  $m$  распределена по закону Пуассона со средним  $\int_0^T R(x) dx$ .

Мы показали, что после достаточно большого отрезка времени численности градаций имеют распределения Пуассона. Аппроксимация распределением Пуассона будет тем точнее, чем меньше первоначальная численность градаций. В иерархической системе это происходит быстрее на нижних градациях. Если поступление не представляет собой пуассоновский процесс, то каждый приход происходит в момент  $T$  и все сводится к изучению независимых величин с известным распределением, а в производящей функции степени заменяются произведениями  $\prod_{h=1}^m$ . Исследование Поллардом дискретного случая показывает, что последняя замена не меняет дела и что распределение Пуассона может оставаться хорошей аппроксимацией, однако в дальнейшем это положение требует уточнения.

## 5.2. ПРАКТИЧЕСКИЙ ПРИМЕР ОПИСАНИЯ СИСТЕМЫ С ЗАДАНЫМ ВХОДОМ

### Выживание после лечения заболевания раком. Изучение когорты

Наш первый пример — это система, в которой нет поступлений и есть два поглощающих (или конечных) состояния. Он был выбран с целью проиллюстрировать, что хорошая стохастическая модель имеет ряд достоинств по сравнению с приемами, которые иногда использовались для решения подобных задач. Это довольно упрощенный пример описания полной неопределенности, которая возникает после лечения заболевания раком. Пациент после лечения может по прошествии некоторого времени находиться в одном из множества состояний. Эти состояния могут классифицироваться, например, так: «здоров», «заболел вновь» (рецидив болезни), «мертв»; точность классификации, очевидно, зависит от целей исследования и от имеющихся возможностей по получению данных. Стохастическая модель описания жизни пациентов после лечения от заболевания раком была построена Фикс и Нейманом (1951) и обсуждалась в более общем виде Залем (1955). Фикс и Нейман применили эту модель для оценки эффективности лечения. Далее мы опишем, как они это делали. Отметим, кстати, что указанная модель достаточно общего вида, и у нее могут быть также другие приложения.

В модели Фикс и Неймана введены четыре состояния. Описание состояний и возможные переходы показаны на рис. 5.1. Авторы по-

нимали трудность определения состояния «выздоровел» и отметили, что было бы желательно некоторые из состояний разделить. Например, пациенты, находящиеся в состоянии  $S_4$ , могут быть разделены на две группы: те, кто умер по естественным (ненасильственным) причинам, и те, судьбу которых проследить не удалось.

Можно также предположить, что необходимо предусмотреть возможность перехода из состояния  $S_1$  в состояние  $S_4$ . Мы не будем отклоняться в сторону, обсуждая эти детали, так как этот пример приведен прежде всего для того, чтобы проиллюстрировать применение теории марковских процессов к описанию жизни людей.

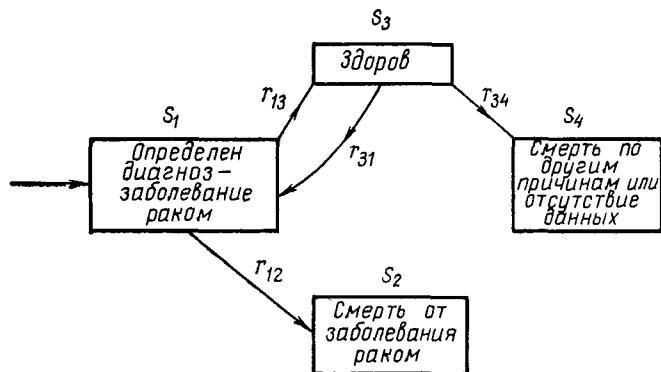


Рис. 5.1.

Первая задача в данном приложении — оценить интенсивности переходов. Для этого использовались данные о выживших, при этом сами данные были лишены недостатков, присущих в общем случае такого рода измерениям. Один из способов измерения — определение доли выживших в  $T$ -м году. Это относительное число оставшихся в живых, по крайней мере, в течение  $T$  лет от всех прошедших курс лечения. Такие измерения были бы удовлетворительными, если бы рак был единственной причиной смерти и если бы все больные наблюдались в течение полных  $T$  лет. Практически так никогда не бывает, и доля выживших в  $T$ -м году может привести к ошибочным выводам. Чтобы убедиться в неточности такого утверждения, заметим только, что измеренная интенсивность (доля) будет больше, так как следует измерить также долю тех, кто выбыл из поля зрения или умер по другим причинам, т. е. относительно большее число людей осталось бы в живых до предельного срока, если бы им суждено было умереть только от заболевания раком. Таким образом, наблюдаемые значения интенсивностей перехода зависят не только от опасности умереть от рака, но и от других причин, не имеющих отношения к заболеванию раком. Если сопоставлять по грубым интенсивностям переходов группу тех, кто прошел курс лечения, и контрольную группу, то сравнение не имело бы смысла, если бы эти две группы подвергались различным опасностям по различным причинам. Чтобы преодолеть эти естественные трудности, обычно вычисляют чистые интенсивности, которые учитывают

такие различия. Цель приведенного примера — показать, что стохастическая модель дает лучшую основу для оценки чистых интенсивностей, чем метод, используемый в страховом деле.

Интенсивности переходов между состояниями в модели Фикс и Неймана полагали постоянными величинами. Однако хорошо известно, что естественная смертность людей — непостоянная величина, и после периода младенчества она увеличивается с возрастом. В средний период жизни она не очень быстро увеличивается, и если период времени  $T$  достаточно короткий, то предположение о постоянстве будет вполне адекватно действительности. Во всяком случае, мы покажем, что можно собирать данные таким образом, чтобы можно было проверять эти предположения. Интенсивность смерти после лечения рака разных видов широко изучается. Время жизни после лечения, как было выяснено, имеет асимметричный характер, Боаг (1949), например, сделал предположение, что оно часто может быть адекватно описано с помощью асимметричного логнормального распределения. В этом случае логнормальное распределение нелегко отличить от экспоненциального, которое появляется при постоянной интенсивности смерти. Таким образом, предположение, что интенсивность смерти от рака является постоянной величиной, вероятно, достаточно реалистично. Непосредственно проанализировать факторы, влияющие на интенсивность переходов из состояния  $S_1$  в  $S_3$  (выздоровление) и из состояния  $S_3$  в  $S_1$ , не представляется возможным, но кажется правдоподобным предположение о постоянстве интенсивностей потерь по разным причинам, по крайней мере для интенсивностей выпадения пациентов из поля зрения.

В нашей модели мы предполагаем, что в нулевой момент времени в состоянии  $S_1$  находится  $N$  людей, в других состояниях людей нет. Численности людей в четырех группах в последующие моменты времени  $T$  будут случайными величинами, которые мы обозначим через  $n_j(T)$  ( $j = 1, 2, 3, 4$ );  $\bar{n}_j(T)$  — математическое ожидание случайной величины  $n_j(T)$ . Наблюдая эти случайные величины в один или несколько моментов времени, можно оценить интенсивности переходов. Затем, используя оценки, можно предсказать численности различных состояний в будущем. Наиболее важна возможность оценить эти численности, если смерть от заболевания раком будет единственной причиной.

## Применение теории

Расширенная матрица  $R_a$  в описываемом случае имеет вид

$$R_a = \begin{pmatrix} r_{11} & r_{12} & r_{13} & 0 \\ 0 & 0 & 0 & 0 \\ r_{31} & 0 & r_{33} & r_{34} \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

где  $r_{11} = -(r_{12} + r_{13})$  и  $r_{33} = -(r_{31} + r_{34})$ . Уравнение для нахождения собственных чисел матрицы есть  $|\theta I - R_a| = 0$  или

$$\theta^2 (r_{11} - \theta) (r_{33} - \theta) - \theta^2 r_{13} r_{31} = 0. \quad (5.19)$$

Очевидно, что это уравнение имеет два нулевых корня; два оставшихся корня, которые мы обозначим  $\theta_3$  и  $\theta_4$ , следующие:

$$\theta_3, \theta_4 = \frac{1}{2} \{r_{11} + r_{33} \pm \sqrt{(r_{11} - r_{33})^2 + 4r_{13}r_{31}}\}, \quad (5.20)$$

причем для расчета  $\theta_3$  возьмем положительный знак, а для  $\theta_4$  — отрицательный. Тогда, используя (4.24), получим

$$\bar{n}_i(T) = d_{i1} + c_{i3} e^{\theta_3 T} + c_{i4} e^{\theta_4 T} \quad (i = 1, 2, 3, 4). \quad (5.21)$$

Следующий шаг — записать и решить однородные уравнения для коэффициентов. Для начала положим  $i = 1$ , а  $h$  будет принимать значения 2, 3 и 4. Таким образом,

$$\left. \begin{aligned} r_{11} d_{11} + r_{31} d_{31} &= 0 \\ r_{11} c_{13} + r_{31} c_{33} &= \theta_3 c_{13} \\ r_{11} c_{14} + r_{31} c_{34} &= \theta_4 c_{14} \end{aligned} \right\} \quad (5.22)$$

Приведем три группы уравнений для  $i = 2, 3$  и 4:

$$i = 2 \quad \left. \begin{aligned} r_{12} d_{11} &= 0 \\ r_{12} c_{13} &= \theta_3 c_{23} \\ r_{12} c_{14} &= \theta_4 c_{24} \end{aligned} \right\}, \quad (5.23)$$

$$i = 3 \quad \left. \begin{aligned} r_{13} d_{11} + r_{33} d_{31} &= 0 \\ r_{13} c_{13} + r_{33} c_{33} &= \theta_3 c_{33} \\ r_{13} c_{14} + r_{33} c_{34} &= \theta_4 c_{34} \end{aligned} \right\}, \quad (5.24)$$

$$i = 4 \quad \left. \begin{aligned} r_{34} r_{31} &= 0 \\ r_{34} c_{33} &= \theta_3 c_{43} \\ r_{34} c_{34} &= \theta_4 c_{44} \end{aligned} \right\}. \quad (5.25)$$

Из уравнений немедленно следует, что  $d_{11} = d_{31} = 0$ , и, следовательно, первые уравнения в каждой группе можно опустить. Начальные условия состоят в том, что в нулевой момент времени все индивидуумы системы находятся в состоянии  $S_1$ . Предположим далее, что  $n_i(0) = 1$  и  $n_j(0) = 0$ ,  $i > 1$ . Если  $n_1(0) = N$ , то соответствующие значения  $\bar{n}_i(T)$  могут быть найдены просто умножением на  $N$  результата, полученного при предположении, что  $n_1(0) = 1$ . Тогда в добавление к записанным выше уравнениям имеем

$$\left. \begin{aligned} d_{11} + c_{13} + c_{14} &= 1 \\ d_{21} + c_{23} + c_{24} &= 0 \\ d_{31} + c_{33} + c_{34} &= 0 \\ d_{41} + c_{43} + c_{44} &= 0 \end{aligned} \right\}. \quad (5.26)$$

Для решения этих уравнений сделаем следующие преобразования. Сложим правые и левые части уравнений (5.22) и, используя начальные условия, получим

$$\theta_3 c_{13} + \theta_4 c_{14} = r_{11}. \quad (5.27)$$



Сделав аналогичные преобразования для (5.23), будем иметь

$$\theta_3 c_{23} + \theta_4 c_{24} = r_{12}, \quad (5.28)$$

но это уравнение может быть получено через  $c_{13}$  и  $c_{14}$  из уравнения (5.23), что дает

$$c_{13} + c_{14} = 1. \quad (5.29)$$

Затем можно совместно решить однородные уравнения (5.27) и (5.28), что позволяет записать:

$$c_{13} = \frac{\theta_4 - r_{11}}{(\theta_4 - \theta_3)}, \quad c_{14} = \frac{r_{11} - \theta_3}{(\theta_4 - \theta_3)},$$

и, следовательно,

$$c_{23} = \frac{r_{12}(\theta_4 - r_{11})}{\theta_3(\theta_4 - \theta_3)}, \quad c_{24} = \frac{r_{12}(r_{11} - \theta_3)}{\theta_4(\theta_4 - \theta_3)}.$$

Сделав подобные преобразования для (5.24) и (5.25), получаем

$$\left. \begin{aligned} c_{33} &= \frac{-r_{13}}{(\theta_4 - \theta_3)}, & c_{34} &= \frac{r_{13}}{(\theta_4 - \theta_3)} \\ c_{43} &= \frac{-r_{13}r_{34}}{\theta_3(\theta_4 - \theta_3)}, & c_{44} &= \frac{r_{13}r_{34}}{\theta_4(\theta_4 - \theta_3)} \end{aligned} \right\}$$

Остается определить две константы:  $d_{21}$  и  $d_{41}$ . Используя начальные условия, находим

$$d_{21} = -c_{23} - c_{24} = -r_{12}r_{33}/\theta_3\theta_4, \quad (5.30)$$

$$d_{41} = -c_{42} - c_{43} = r_{13}r_{34}/\theta_3\theta_4. \quad (5.31)$$

Сейчас рассмотрим, как использовать эти результаты, чтобы сравнить интенсивности выживания. Когда  $N = 1$ , величина  $\bar{n}_i(T)$  может быть интерпретирована как вероятность находиться в состоянии  $S_i$  в момент времени  $T$ . Таким образом,  $\bar{n}_2(T)$  и  $\bar{n}_4(T)$  представляют собой соответственно грубые интенсивности смерти вследствие заболевания раком и по естественным причинам. Однако  $\bar{n}_4(T)$  зависит также от интенсивности естественной смерти и, как мы указывали выше, это уменьшает ее величину как меру риска. На самом деле нам нужна чистая мера риска (чистая интенсивность смерти), из которой устранено влияние естественной смертности. Согласно подходу к задаче, используемому в страховом деле, чистая интенсивность смерти от рака определяется по формуле

$${}_A\bar{n}_2(T) = \bar{n}_2(T) / \left\{ 1 - \frac{1}{2} \bar{n}_4(T) \right\}. \quad (5.32)$$

Величина (5.32) должна давать среднее число смертей от заболевания раком на интервале  $(0, T)$ , если бы смертности по естественным причинам не было. Смысл уравнения (5.32) станет яснее, если его переписать:

$${}_A\bar{n}_2(T) = \bar{n}_2(T) - \frac{1}{2} \bar{n}_4(T) {}_A\bar{n}_2(T). \quad (5.33)$$

Второе слагаемое в правой части уравнения (5.33) — оценка численности людей, которые умерли бы от рака в течение рассматриваемого периода, если не умерли бы по другим естественным причинам. Оно получено в предположении, что смерть от рака, вероятность которой равна одной второй, предшествует естественной смерти по другим причинам. Предлагаемая модель предоставляет другой метод для оценки чистых интенсивностей смерти от рака. Мы можем исключить влияние естественной смертности, положив  $r_{34} = 0$ . Тогда чистая интенсивность записывается как

$$\bar{n}_2^0(T) = \frac{r_{12} r_{31}}{\theta_3^0 \theta_4^0} + \frac{r_{12} (\theta_4^0 - r_{11})}{\theta_3^0 (\theta_4^0 - \theta_3^0)} e^{\theta_3^0 T} + \frac{r_{12} (r_{11} - \theta_3^0)}{\theta_4^0 (\theta_4^0 - \theta_3^0)} e^{\theta_4^0 T}, \quad (5.34)$$

где нулевые индексы в  $\bar{n}_2(T)$ ,  $\theta_3$  и  $\theta_4$  означают, что  $r_{34}$  положена равной нулю.

Применение этих результатов может быть проиллюстрировано численными примерами. Возьмем следующие значения интенсивностей переходов:

	$r_{12}$	$r_{13}$	$r_{31}$	$r_{34}$
Пример 1	1,0	2,0	0,5	0,2
Пример 2	0,5	0,5	0,5	0,5

Подставляя эти величины в (5.20), для примера 1 находим:

$$\left. \begin{aligned} \bar{n}_2(T) &= 0,6364 - 0,3764 e^{-0,3260T} - 0,2600 e^{-3,3740T}, \\ \bar{n}_2^0(T) &= 1 - 0,7344 e^{-0,1492T} - 0,2657 e^{-3,3508T}, \end{aligned} \right\} \quad (5.35)$$

а для примера 2:

$$\left. \begin{aligned} \bar{n}_2(T) &= \frac{2}{3} - \frac{1}{2} e^{-\frac{1}{2}T} - \frac{1}{6} e^{-(3/2)T}, \\ \bar{n}_2^0(T) &= 1 - 0,7236 e^{-0,1910T} - 0,2764 e^{-1,3090T}, \end{aligned} \right\} \quad (5.36)$$

Можно выявить одну особенность, показывающую несостоятельность метода определения интенсивности смерти, принятого в страховом деле, если рассмотреть предельное поведение (5.32) при  $T \rightarrow \infty$ . Вместо того, чтобы стремиться к единице, как следовало бы ожидать от достаточно обоснованной меры, она стремится к значению, меньшему единицы в обоих примерах. Анализ (5.32) показывает, что этот результат всегда имеет место. Очевидно также, что в общем случае  $\Delta \bar{n}_2(T) < \bar{n}_2^0(T)$  при достаточно большом  $T$ . Некоторые численные значения содержатся в табл. 5.1.

Приведенный пример — хорошая иллюстрация использования стохастической модели для измерения социального явления. Он показывает также, что коррекция измерений с позиций «здорового смысла» может существенно обесценить проведенные измерения. Высказанные доводы предполагают, что модель адекватна описываемому явлению. Если в действительности интенсивности переходов не постоянны, то более простая статистическая оценка иногда предпочтительнее, потому

Т а б л и ц а 5.1. Сравнение чистых интенсивностей смерти от рака, вычисленных с помощью метода, используемого в страховом деле, и с помощью стохастической модели

	$T$	0,5	1	2	5	$\infty$
Пример 1	$\bar{n}_2(T)$	0,269	0,356	0,440	0,563	0,636
	$\bar{n}_2^0(T)$	0,269	0,358	0,455	0,652	1,000
	$A\bar{n}_2(T)$	0,272	0,370	0,477	0,656	0,778
Пример 2	$\bar{n}_2(T)$	0,199	0,326	0,474	0,626	0,667
	$\bar{n}_2^0(T)$	0,199	0,328	0,486	0,721	1,000
	$A\bar{n}_2(T)$	0,201	0,338	0,515	0,733	0,800

что она не зависит от распределения. Как будет показано, именно грубые методы эффективны при проверке адекватности модели.

При обсуждении модели предполагалось, что интенсивности переходов известны. На практике они не бывают известными, и их необходимо оценить по имеющимся данным. Общие методы оценивания упоминались в гл. 4, но для решения нашей задачи достаточно более простого метода Фикс и Неймана. В момент времени  $T$  мы можем зафиксировать численности пациентов в начальный момент в каждом из четырех состояний. Эти численности могут рассматриваться как оценки для  $\bar{n}_i(T)$ , которые в свою очередь получают при неизвестных параметрах. В обсуждаемой модели метод позволяет получить четыре уравнения для оценки неизвестных параметров. К сожалению, эти уравнения не являются линейно независимыми, так как

$$\sum_{i=1}^4 \bar{n}_i(T) = N,$$

где  $N$  — наблюдаемое число индивидуумов. Ситуация была бы еще хуже, если бы в матрице  $\mathbf{R}$  были другие ненулевые интенсивности. Такие трудности можно преодолеть, исследуя состояния системы в нескольких точках оси времени. Другой метод — рассматривать некоторые другие характеристики системы, например, по предложению Фикс и Неймана, подсчитывать число пациентов, оставшихся в состоянии  $S_1$  на интервале времени  $(0, T)$ . Если материал наблюдений достаточно обширен, то можно не только оценить все параметры, но и проверить качество модели. Предельная структура  $\bar{n}(\infty)$  может быть получена непосредственно, без проведения всех описанных вычислений, так как из (5.21) результат следует немедленно.

Из уравнений (5.30) и (5.31) получаем

$$\left. \begin{aligned} \bar{n}_2(\infty) &= d_{21} = r_{12}(r_{31} + r_{34})/\theta_3 \theta_4 \\ \bar{n}_4(\infty) &= d_{41} = r_{13} r_{34}/\theta_3 \theta_4 \end{aligned} \right\} \quad (5.37)$$

Остальные предельные значения равны нулю. Таким образом, имеется простая зависимость  $\bar{n}_2(\infty)$  и  $\bar{n}_4(\infty)$  от интенсивностей переходов. Вид этой зависимости может быть легко выявлен, если записать отношение этих величин в следующей форме:

$$\frac{\bar{n}_2(\infty)}{\bar{n}_4(\infty)} = \frac{r_{12}}{r_{13}} \left( 1 + \frac{r_{31}}{r_{34}} \right), \quad (5.38)$$

где  $r_{12}/r_{13}$  — отношение интенсивностей переходов из состояния «определен диагноз — заболел раком», и  $r_{31}/r_{34}$  — отношение интенсивностей переходов из состояния «здоров». Большая интенсивность потока выздоравливающих  $r_{13}$  способствует увеличению доли тех пациентов, кто умирает по другим естественным причинам, но этому в некоторой степени будет противодействовать возможность и большей интенсивности потока рецидивов  $r_{31}$ .

Мы уже указывали, что модель первоначально была разработана для измерения эффективности лечения. Один из способов — рассчитать  $\bar{n}_2^0(T)$  — чистую долю тех, кто умер бы от рака, при исключении влияния других причин. Фикс и Нейман приводят доводы в пользу того, что  $\bar{n}_2^0(T)$  не единственная, но, видимо, наиболее подходящая мера для оценки выживания. Обсуждение этого вопроса выходит за рамки данной книги, но мы коснулись его потому, что величины  $\bar{n}_2(T)$  и  $\bar{n}_2^0(T)$  будут полезны для построения других мер при дальнейших исследованиях. Например, Фикс и Нейман предполагают полезным рассчитывать среднюю длительность «нормальной» жизни в период  $(0, T)$  так, как если бы рак был единственной причиной смерти. Поскольку  $\bar{n}_2^0(T)$  — функция распределения длительности «нормальной» жизни при отсутствии других причин смерти, математическое ожидание может быть записано так:

$$e_2 = T \{ 1 - \bar{n}_2^0(T) \} + \int_0^T x \frac{d\bar{n}_2^0(x)}{dx} dx$$

или

$$e_2 = \int_0^T \{ 1 - \bar{n}_2^0(x) \} dx. \quad (5.39)$$

### Иерархическая кадровая система

Модели с непрерывным временем, описывающие иерархические системы, впервые были предложены Силом (1945) и Вайдой (1948). Хотя их модели немарковские, оба автора обсуждали некоторые особые случаи, которые совпадают с теми, что следуют из нашей общей теории. Рассмотрим систему, которая представлена диаграммой на рис. 5.2. Эта система имеет одно поглощающее состояние, обозначенное  $S_{h+1}$ . Продвижение возможно только на ближайшую градацию,

что изображена на схеме, а все вновь поступающие зачисляются на первую. Расширенная матрица интенсивностей переходов для описанной системы имеет вид

$$R_a = \begin{pmatrix} r_{11} & r_{12} & 0 & \dots & 0 & r_{1, k+1} \\ 0 & r_{22} & r_{23} & \dots & 0 & r_{2, k+1} \\ 0 & 0 & r_{33} & \dots & 0 & r_{3, k+1} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & . & r_{k+1, k+1} \end{pmatrix},$$

где

$$r_{ii} = -(r_{i, i+1} + r_{i, k+1}), \quad i < k,$$

$$r_{kk} = -r_{k, k+1}, \quad r_{k+1, k+1} = 0.$$

Простая треугольная структура  $R_a$  позволяет нам получить точную формулу для собственных значений и коэффициентов  $\{c_{ij}\}$ , которые есть в выражениях для определения переходных вероятностей

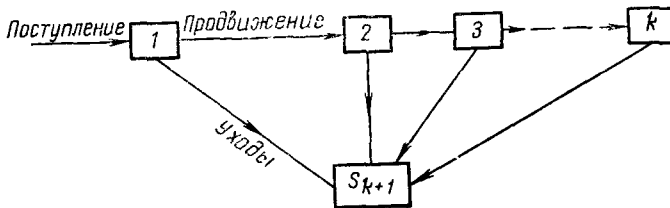


Рис. 5.2.

$\{p_{ij}(T)\}$ . Отсюда мы тотчас же находим, что  $\theta_i = r_{ii}$  ( $i=1, 2, \dots, k+1$ ). Уравнения для определения коэффициентов  $c$ , полученные из (4.19), имеют вид

$$\left. \begin{aligned} r_{i-1, i} c_{i-1, h} + r_{ii} c_{ih} &= r_{hh} c_{ih} \quad (i, h = 1, 2, \dots, k+1) \\ \sum_{h=1}^{k+1} c_{ih} &\begin{cases} = 1 & \text{при } i=1, \\ = 0 & \text{при } i > 1. \end{cases} \end{aligned} \right\} \quad (5.40)$$

Начальные условия, представленные последними двумя уравнениями, следуют из того, что все вновь прибывшие начинают свою карьеру с градации 1 — низшей ступени служебной лестницы. Решение системы уравнений (5.40) дает

$$\left. \begin{aligned} c_{ih} &= \prod_{j=1}^{i-1} r_{j, j+1} \prod_{\substack{j=1 \\ j \neq h}}^i (r_{hh} - r_{jj}) \quad (i=2, 3, \dots, k+1; h=1, 2, \dots, i) \\ c_{11} &= 1, \\ c_{ih} &= 0 \text{ во всех других случаях.} \end{aligned} \right\} \quad (5.41)$$

Представляют интерес только значения  $\bar{n}_j(T)$ , если  $j \leq k$ , в этом случае из (5.3) находим

$$\bar{n}_j(T) = \int_0^T R(x) p_{1j}(T-x) dx + \sum_{i=1}^j n_i(0) p_{ij}(T) \quad (j=1, 2, \dots, k). \quad (5.42)$$

Коэффициенты, полученные из (5.40), дают

$$p_{1j}(T) = \sum_{h=1}^k c_{jh} e^{r_h h T} \quad (5.43)$$

и выражения для них можно подставить в (5.42). Подобные выражения могут быть найдены при соответствующих начальных условиях, но они же легко могут быть выведены из выражений для  $p_{ij}(T)$ , когда имеется простая иерархическая система. Вновь поступивший, который начинает свою карьеру с  $i$ -й ступени  $k$ -уровневой системы, находится в том же состоянии, что и тот, который поступил на низшую (первую) ступень  $(k - (i - 1))$ -уровневой системы. Заменяя  $k$  на  $k - i + 1$  и переобозначая интенсивности переходов, найдем необходимые выражения. Ниже мы приведем пример. Очевидно, что верхний предел суммы в последнем члене выражения (5.42) —  $j$ .

Модель, которую мы описали, несколько более общего вида, чем марковская версия модели Вайды (1948). В последней предполагалось, что интенсивности поступлений и уходов постоянны, таким образом, результаты Вайды могут быть получены из наших, если положить  $R(x) = R$  и  $r_{i, k+1} = r_{k+1}$ , скажем, для  $i \leq k$ . Мы имеем также ожидаемые численности ступеней для любого  $T$ , а Вайда обсуждал только предельный случай.

Как мы указывали, по нескольким причинам требуется, чтобы все величины  $r_{ii}$  ( $i = 1, 2, \dots, k + 1$ ) были различны. В случае, который мы сейчас обсудим,  $r_{ii} = -(r_{i, i+1} + r_{i, k+1})$  для  $i \leq k$ , поэтому равные  $r_{ii}$  встречаются при равенстве интенсивностей уходов с различных ступеней. Случай, представляющий особый интерес, появляется тогда, когда  $r_{i, i+1} = r$  и  $r_{i, k+1} = r_{k+1}$  для  $i < k$ . Это соответствует ситуации, в которой интенсивности продвижения и интенсивности уходов одни и те же для всех ступеней, кроме последней. Соответствующее изменение общей теории может быть получено при стремлении друг к другу собственных значений  $r_{ii}$  ( $i < k$ ) в выражении (5.43). Окончательное выражение для  $p_{1j}(T)$  будет таким:

$$p_{1j}(T) = \frac{r^{j-1}}{(j-1)!} T^{j-1} e^{-(r+r_{k+1})T} \quad (j=1, 2, \dots, k-1). \quad (5.44)$$

Если  $j = k$ , то выражение имеет чуть более сложный вид, но несколько ниже мы увидим, что в нем нет необходимости. Применяя тот же метод анализа, что и ранее, находим выражение для  $p_{ij}(T)$ :

$$p_{ij}(T) = \frac{r^{j-i}}{(j-i)!} T^{j-i} e^{-(r+r_{k+1})T} \quad (i \leq j < k). \quad (5.45)$$

Подставляя в (5.42) выражения (5.44) и (5.45), можно определить  $\bar{n}_j(T)$  для  $j < k$ , а затем  $\bar{n}_k(T)$  из выражения

$$\bar{n}_k(T) = \sum_{j=1}^k \bar{n}_j(T) - \sum_{j=1}^{k-1} \bar{n}_j(T). \quad (5.46)$$

Общая численность системы в момент времени  $T$  получается при суммировании обеих частей (5.42) по  $j$  и на основании того факта, что

$$\sum_{j=1}^k p_{ij}(x) = 1 - p_{i, k+1}(x) = 1 - e^{-(r_{k+1} x)}.$$

В результате получаем

$$\sum_{j=1}^k \bar{n}_j(T) = \int_0^T R(x) e^{-r_{k+1}(T-x)} dx + N(0) e^{-r_{k+1}T}, \quad (5.47)$$

где

$$N(0) = \sum_{j=1}^k n_j(0).$$

Проиллюстрируем применение полученных результатов на примере системы, в которой имеются три ступени,  $k = 3$ :

$$r_{i, i+1} = 1 \quad (i = 1, 2), \quad r_{i, 4} = 2 \quad (i = 1, 2, 3).$$

Для этих данных

$$p_{ij}(T) = \frac{1}{(j-i)!} T^{j-i} e^{-3T} \quad (i = 1, 2; i \leq j < 3). \quad (5.48)$$

Допустим, что поступление — экспоненциальная функция времени  $R(x) = 100e^{\alpha x}$  и что  $n_1(0) = 100$ ,  $n_2(0) = 50$ ,  $n_3(0) = 20$ . Мы оставим  $\alpha$  как параметр для того, чтобы посмотреть его влияние на решение.

Из уравнения (5.42) получаем

$$\left. \begin{aligned} \bar{n}_1(T) &= \frac{100}{3+\alpha} e^{\alpha T} + 100e^{-3T} \left\{ \frac{2+\alpha}{3+\alpha} \right\} \\ \bar{n}_2(T) &= \frac{100}{(3+\alpha)^2} e^{\alpha T} + 100e^{-3T} = \left\{ T + \frac{1}{2} - \frac{T}{3+\alpha} - \frac{1}{(3+\alpha)^2} \right\} \end{aligned} \right\} \quad (5.49)$$

Вероятность  $\bar{n}_3(T)$  находят при подстановке  $\bar{n}_1(T) + \bar{n}_2(T)$  в

$$\sum_{j=1}^3 \bar{n}_j(T) = \frac{100e^{\alpha T}}{2+\alpha} + 100e^{-2T} \left\{ 1,7 - \frac{1}{2+\alpha} \right\}. \quad (5.50)$$

Из этих формул очевиден решающий характер зависимости  $\bar{n}_j(T)$  от  $\alpha$ . Если  $\alpha > 0$ , т. е. поступление увеличивается по экспоненте, то ожидаемые размеры ступеней иерархии также растут. Если  $\alpha < 0$ , то ожидаемые размеры ступеней стремятся к нулю с течением времени. Читатель может сравнить эти результаты с теми, которые получены в

гл. 3 для модели с дискретным временем и поступлениями, растущими по закону геометрической прогрессии. В случае  $\alpha = 0$  поступления постоянны и численности  $\bar{n}_j(T)$  стремятся к определенным пределам при  $T \rightarrow \infty$ . Эти пределы равны:

$$\begin{aligned}\bar{n}_1(\infty) &= 33,3, \\ \bar{n}_2(\infty) &= 11,1, \\ \bar{n}_3(\infty) &= 5,6,\end{aligned}$$

а средний размер равен 50. Последний результат может быть получен на основании того, что среднее время пребывания в системе  $r^{-1}_{i4} = \frac{1}{2}$  ( $i = 1, 2, 3$ ) и что поступление составляет 100 человек в единицу времени. Окончательный размер, конечно, не зависит от начального размера, а значения параметров, которые мы выбираем, приводят к уменьшению средних размеров с 170 до 50. Изучение (5.49) показывает, что для низших ступеней системы предельные значения достигаются быстрее, чем для средних и высших. Мы наблюдали то же явление и в модели с дискретным временем.

Интересно заметить, что зависимость относительных размеров  $\bar{n}_j(T)$  от  $\alpha$  для больших  $T$  имеет очень простую форму. Если  $\alpha > -2$ , то

$$\left. \begin{aligned}\bar{n}_1(T) &\sim \frac{100}{3+\alpha} e^{\alpha T}, \quad \bar{n}_2(T) \sim \frac{100}{(3+\alpha)^2} e^{\alpha T} \\ \bar{n}_3(T) &\sim \frac{100}{(2+\alpha)(3+\alpha)^2} e^{\alpha T}\end{aligned}\right\}. \quad (5.51)$$

Влияние увеличения интенсивности поступлений заключается в увеличении размеров двух низших ступеней относительно высших. И напротив, уменьшение интенсивности поступлений ведет к накоплению на высшей ступени иерархии. Если бы мы повторили эти вычисления для больших значений  $k$ , то нашли бы, что при больших  $T$  численности ступеней относились бы друг к другу как члены геометрической прогрессии, кроме численности  $k$ -й ступени. В особенности этот результат имеет место для случая  $\alpha = 0$ , когда поступления постоянны. Тогда имеем

$$\left. \begin{aligned}\bar{n}_j(\infty) &= \frac{R}{r} \left( \frac{r}{r+r_{k+1}} \right)^j \quad (j = 1, 2, \dots, k-1) \\ \bar{n}_k(\infty) &= \frac{R}{r_{k+1}} - \sum_{j=1}^{k-1} \bar{n}_j(\infty) = \frac{R}{r_{k+1}} \left( \frac{r}{r+r_{k+1}} \right)^{k-1},\end{aligned}\right\} \quad (5.52)$$

где  $R$  — интенсивность поступлений в систему. Таким образом, делаем вывод, что иерархическая система с постоянными интенсивностями продвижения, с постоянными интенсивностями уходов и поступлений стремится к геометрической структуре ступеней. Исключение из этого общего правила заключается в том, что численность  $k$ -й ступени будет больше, чем если бы она вычислялась как член геометрической про-



грессии при  $j = k$ . Система указанного типа стремится к структуре типа перевернутой пирамиды. Если  $R(x)$  имеет форму экспоненты, то полученный результат характерен также для ожидаемых относительных численностей ступеней. Позднее в этой главе мы увидим, что такой общий вывод верен и при невыполнении некоторых специфических допущений данной модели.

### 5.3. МОДЕЛЬ, ОПИСЫВАЮЩАЯ СИСТЕМУ УСТАНОВЛЕННОГО РАЗМЕРА

Если общие размеры системы — известная функция времени  $T$ , то ожидаемые численности состояний могут быть найдены методом, имеющим много общего с тем, который применялся при изучении модели с дискретным временем. Мы разовьем эту теорию, допустив, что общая численность системы — непрерывная функция времени  $N(T)$ , и определив скорость изменения  $M(T) = dN(T)/dT$ . Необходимые уравнения получим из (5.2) и (5.3), подставив соответствующие выражения для  $R(T)$ . Как и ранее, уравнение состоит из двух частей, одна часть для возмещения уходов, а другая — для заполнения новых мест. Таким образом, число вновь поступивших в систему за период времени  $(T, T + \delta T)$  равно:

$$R(T) \delta T = \sum_{i=1}^k \bar{n}_i(T) r'_{i,k+1} \delta T + M(T) \delta T. \quad (5.53)$$

Подставив (5.53) в (5.2), мы получаем непрерывный аналог (3.54) в виде

$$\begin{aligned} \frac{d\bar{n}(T)}{dT} &= \bar{n}(T) \mathbf{R} + \mathbf{n}(T) \mathbf{r}'_{k+1} \mathbf{p}_0 + M(T) \mathbf{p}_0 = \\ &= \bar{n}(T) \{ \mathbf{R} + \mathbf{r}'_{k+1} \mathbf{p}_0 \} + M(T) \mathbf{p}_0. \end{aligned} \quad (5.54)$$

Матрица в скобках — это непрерывный аналог матрицы  $\mathbf{Q}$  из гл. 3, она обладает тем свойством, что сумма элементов строки равна нулю. Уравнение имеет такой же вид, что и (5.2), и производя аналогичные преобразования в (5.3), получаем

$$\bar{\mathbf{n}}(T) = \bar{\mathbf{n}}(0) \mathbf{P}^+(T) = \mathbf{p}_0 \int_0^T M(x) \mathbf{P}^+(T-x) dx, \quad (5.55)$$

где  $\mathbf{P}^+(T)$  — матрица переходных вероятностей, вычисленная по матрице интенсивностей переходов методом, описанным в гл. 4. Для (5.3) и (5.5) удобно использовать преобразование Лапласа, которое для (5.3) будет следующим:

$$\mathbf{n}^*(s) = \{ \mathbf{n}(0) + \mathbf{p}_0 R^*(s) \} \mathbf{P}^*(s). \quad (5.56)$$

Для (5.55) мы только заменим  $R^*(s)$  на  $M^*(s)$  и  $\mathbf{P}^*(s)$  станет преобразованием от  $\mathbf{P}^+(T)$ , которое имеет вид

$$(s\mathbf{I} - \mathbf{R} - \mathbf{r}'_{k+1} \mathbf{p}_0)^{-1}.$$

## Предельный режим

Этот режим может быть исследован аналогично тому, как это сделано Файхтингером (1976) для дискретного времени. Из (5.53) имеем

$$R(T) = \bar{n}(T) r'_{k+1} + M(T).$$

Пусть  $\mathbf{I} = (1, 1, \dots, 1)$ , тогда

$$\mathbf{R}\mathbf{I}' = -r'_{k+1} \text{ и } \bar{n}(T) \mathbf{I}' M(T) / N(T) = M(T).$$

Подставив выражение для  $r_{k+1}$  и  $M(T)$  в (5.54), находим другой вариант уравнения:

$$\frac{d\bar{n}(T)}{dT} = \bar{n}(T) \{ \mathbf{R} + (\mathbf{I}M(T)/N(T) - \mathbf{R}) \mathbf{I}' \mathbf{p}_0 \}. \quad (5.57)$$

Оно имеет устойчивое решение, если  $N(T)$  стремится к пределу, поэтому мы рассмотрим относительные значения ожидаемых численностей, обозначив

$$\mathbf{q}(T) = \bar{n}(T) / \bar{n}(T) \mathbf{I}'.$$

В этих обозначениях имеем

$$\frac{d\mathbf{q}(T)}{dT} = \mathbf{q}(T) \left( \mathbf{R} - \frac{M(T)}{N(T)} \mathbf{I} \right) (\mathbf{I} - \mathbf{I}' \mathbf{p}_0). \quad (5.58)$$

Это уравнение можно сравнить с соответствующим уравнением марковского процесса, зависящего от времени. Если такой процесс имеет матрицу интенсивностей переходов  $\mathbf{R}(T)$ , то вероятности состояний удовлетворяют уравнению

$$\frac{d\mathbf{q}(T)}{dT} = \mathbf{q}(T) \mathbf{R}(T).$$

Если матрица в правой части (5.58) имеет тот же вид, что и матрица интенсивностей переходов, можно применить теорию процессов, зависящих от времени, для анализа поведения нашего процесса. Необходимое условие заключается в том, чтобы матрица имела отрицательные элементы по диагонали, положительные элементы везде, а суммы по строкам должны быть равны нулям. Мелман (1977b) показал, что это требование выполняется, если

$$M(T)/N(T) > \max_i (-r_{i, k+1}) \text{ для всех } T. \quad (5.59)$$

Это означает, что существует предельное значение интенсивности поступлений в систему. Для существования предела требуется, чтобы  $\lim_{T \rightarrow \infty} \mathbf{R}(T)$  существовал и это условие выполняется для (5.58), если

$$\lim_{T \rightarrow \infty} M(T)/N(T) = c.$$

В этом случае устойчивая структура удовлетворяет уравнению

$$\mathbf{q}(\infty) (\mathbf{R} - c\mathbf{I}) (\mathbf{I} - \mathbf{I}' \mathbf{p}_0) = 0 \quad (5.60)$$

или

$$\mathbf{q}(\infty) \propto \mathbf{p}_0 (c\mathbf{I} - \mathbf{R})^{-1}. \quad (5.61)$$

Константа пропорциональности выбирается такой, чтобы сумма элементов  $\mathbf{q}(\infty)$  была равна единице.

При  $c = 0$  получаем особенно важный случай, означающий, что скорость роста стремится к нулю. При этом предельная относительная структура остается неизменной, несмотря на то, что система несколько растет. В частности, такое утверждение справедливо, если роста нет совсем. Вспомнив вид преобразования Лапласа для  $\mathbf{P}(T)$ , мы можем также записать:

$$\mathbf{q}(\infty) \propto \mathbf{p}_0 \int_0^{\infty} \mathbf{P}(T) e^{-cT} dT. \quad (5.62)$$

### Числовой пример

Рассмотрим кадровую систему с  $k = 3$ ,  $p_{01} = 1$ ,  $p_{02} = p_{03} = 0$ ,  $r_{12} = r_{23} = r$  и  $r_{14} = r_{24} = r_{34} = r_4$ . Допустим также, что  $n_1(0) = N$ ,  $n_2(0) = n_3(0) = 0$ . Нахождение ожидаемой численности состояний начнем с преобразований Лапласа (5.55). Во-первых, получим преобразование  $\mathbf{P}^+(T)$ , т. е. преобразуем матрицу

$$\mathbf{R} + \mathbf{r}_{k+1} \mathbf{p}_0 = \begin{pmatrix} -r & r & 0 \\ r_4 & -r - r_4 & r \\ r_4 & 0 & -r_4 \end{pmatrix}.$$

Тогда требуемое преобразование будет

$$(\mathbf{sI} - \mathbf{R} - \mathbf{r}_{k+1} \mathbf{p}_0)^{-1} = \frac{1}{s \{s + (r + r_4)\}^2} \times \\ \times \begin{pmatrix} (s + r + r_4)(s + r_4) & r(s + r_4) & r^2 \\ r_4(s + r_4) + rr_4 & (s + r)(s + r_4) & r(s + r) \\ r_4(s + r + r_4) & rr_4 & (s + r)(s + r_4) - rr_4 \end{pmatrix}. \quad (5.63)$$

Предположим, что рассматривается случай экспоненциального роста  $M(T) = Me^{\alpha T}$  ( $M > 0$ ), когда  $M^*(s) = M/(s - \alpha)$  ( $s > \alpha$ ). Тогда векторы численностей получают из выражения

$$\bar{\mathbf{n}}^*(s) = \{(N, 0, 0) + (1, 0, 0) M/(s - \alpha)\} (\mathbf{sI} - \mathbf{R} - \mathbf{r}_{k+1} \mathbf{p}_0)^{-1}. \quad (5.64)$$

В нашем примере

$$\bar{n}_1^*(s) = (s + r_4)N(s - \alpha) + M/s(s - \alpha)(s + r + r_4). \quad (5.65)$$

Разложим (5.65) на простые дроби:

$$\bar{n}_1^*(s) = \frac{A_1}{s} + \frac{B_1}{s - \alpha} + \frac{C_1}{s + r + r_4} \quad (5.66)$$

и положим, что  $\alpha \neq 0$ . Коэффициенты разложения вычисляются следующим образом:

$$A_1 = \left( \frac{N\alpha - M}{\alpha} \right) \left( \frac{r_d}{r + r_d} \right),$$

$$B_1 = \frac{M}{\alpha} \frac{(r_d + \alpha)}{(r + r_d + \alpha)},$$

$$C_1 = \frac{r \{N(r + r_d) + N\alpha - M\}}{(r + r_d)(r + r_d + \alpha)}.$$

Можно заметить, что  $A_1 + B_1 + C_1 = N$ . Если  $\alpha = 0$ , разложение на простые дроби имеет вид

$$\bar{n}_1^*(s) = \frac{A'_1}{s} + \frac{B'_1}{s^2} + \frac{C'_1}{(s + r + r_d)}, \quad (5.67)$$

где

$$A'_1 = \frac{Mr}{(r + r_d)^2} + \frac{Nr_d}{(r + r_d)},$$

$$B'_1 = \frac{Mr_d}{(r + r_d)},$$

$$C'_1 = \frac{r \{N(r + r_d) - M\}}{(r + r_d)^2}.$$

При обратных преобразованиях исходных выражений (5.66) и (5.67) получаем

$$\left. \begin{aligned} \bar{n}_1^*(T) &= A_1 + B_1 e^{\alpha T} + C_1 e^{-(r+r_d)T} \quad (\alpha \neq 0) \\ \bar{n}_1^*(T) &= A'_1 + B'_1 T + C'_1 e^{-(r+r_d)T} \quad (\alpha = 0) \end{aligned} \right\} \quad (5.68)$$

Последние слагаемые в правых частях (5.68) стремятся к нулю с ростом  $T$ . При больших значениях  $T$  поведение  $\bar{n}_1(T)$  зависит от  $\alpha$ . Необходимо различать три случая:

$$\begin{aligned} \bar{n}_1(T) &= B_1 e^{\alpha T}, \quad \text{если } \alpha > 0, \\ &\sim B'_1 T, \quad \text{если } \alpha = 0, \\ &\sim A_1, \quad \text{если } \alpha < 0. \end{aligned}$$

$\bar{n}_2(T)$  определяется так же, как и  $\bar{n}_1(T)$ , решение для него имеет вид

$$\bar{n}_2(T) = A_2 + B_2 e^{\alpha T} + C_2 e^{-(r+r_d)T} + D_2 T e^{-(r+r_d)T} \quad (\alpha \neq 0). \quad (5.69)$$

Если  $\alpha = 0$ , второе слагаемое уравнения заменяется линейной функцией от  $T$  с различными коэффициентами. Их определение оставляем читателю. Стремление к предельному значению для  $\bar{n}_2(T)$  происходит медленнее, чем для  $\bar{n}_1(T)$  из-за множителя  $T$  в последнем слагаемом. Общий вид  $\bar{n}_2(T)$  идентичен (5.69), и для него справедливо замечание о скорости стремления к пределу.

Относительные размеры ступеней в пределе могут быть получены из выражений для  $n(T)$  или непосредственно из (5.61) или (5.62). Таким образом,

$$q(\infty) \propto (1, 0, 0) \begin{pmatrix} \alpha + r + r_4 & -r & 0 \\ 0 & \alpha + r + r_4 & -r \\ 0 & 0 & \alpha + r_4 \end{pmatrix}^{-1},$$

следовательно,

$$q(\infty) = \left\{ \frac{(r_4 + \alpha)}{(r + r_4 + \alpha)}, \frac{r(r_4 + \alpha)}{(r + r_4 + \alpha)^2}, \frac{r^2}{(r + r_4 + \alpha)^2} \right\}. \quad (5.70)$$

Изучение этого результата показывает, что экспоненциальная скорость роста приводит к увеличению относительного размера низшей ступени и уменьшению высших ступеней. Действительно, если  $\alpha$  становится очень большим, то численность низшей ступени преобладает над остальными, то же самое мы наблюдали и в модели с дискретным временем. В данном примере в пределе при  $\alpha \rightarrow \infty$ ,  $q(\infty) \rightarrow p_0$  окончательная структура будет  $(1, 0, 0)$ .

#### 5.4. СИСТЕМЫ С УСТАНОВЛЕННЫМИ ПОСТУПЛЕНИЯМИ И УХОДАМИ, ЗАВИСЯЩИМИ ОТ ПРОДОЛЖИТЕЛЬНОСТИ ПРЕБЫВАНИЯ

##### Теория

До сих пор предполагалось, что интенсивности переходов — постоянные величины, но обоснования, приводящие к (4.4), справедливы даже тогда, когда  $R$  — функция времени. Предположения, что интенсивности переходов — постоянные величины, справедливы для описания процесса выживания при заболевании раком, но почти наверняка это не так при описании потоков в кадровых системах. Потоки продвижения обычно зависят от времени пребывания в состоянии  $i$ , как мы увидим в гл. 7, склонность к уходу сильно зависит от  $T$  и стажа. В этом разделе мы обсудим последствия ослабления требования о постоянстве интенсивностей уходов. В разделе 5.5 будет введена зависимость других интенсивностей от времени пребывания в состояниях.

Существуют общие методы решения основных уравнений для марковских процессов, когда они зависят от времени. Но в частном случае, который нас будет интересовать, можно получить точное решение еще проще. Сосредоточим внимание на иерархических системах, в которых невозможен переход с верхнего на нижний уровень. Рассмотрим матрицу интенсивностей переходов, в которой  $(k + 1)$ -я категория соответствует уходам. Таким образом,

$$R_a(T) = \begin{pmatrix} r_{11}(T) & r_{12} & \dots & r_{1k} & r_{k+1}(T) \\ 0 & r_{22}(T) & \dots & r_{2k} & r_{k+2}(T) \\ \vdots & \vdots & \dots & \vdots & \vdots \\ 0 & 0 & \dots & r_{kk}(T) & r_{k+1}(T) \\ 0 & 0 & \dots & 0 & 0 \end{pmatrix}.$$

Дифференциальное уравнение для  $\mathbf{P}(T)$  может быть разделено на две части:

$$\frac{dp_{ij}(T)}{dT} = \sum_{h=1}^{j-1} r_{hj} p_{ih}(T) + r_{jj}(T) p_{ij}(T) \quad (i=1, 2, \dots, k; j=1, 2, \dots, k) \quad (5.71)$$

$$\frac{dp_{i,k+1}(T)}{dT} = \sum_{h=1}^k r_{h+1}(T) p_{ih}(T) = r_{h+1}(T) \{1 - p_{i,h+1}(T)\} \quad (i=1, 2, \dots, k). \quad (5.72)$$

Систему уравнений (5.72) можно решить при использовании начальных условий  $p_{i,h+1}(0) = 0$ ,

$$p_{i,h+1}(T) = 1 - \exp \left\{ - \int_0^T r_{h+1}(x) dx \right\} \quad (i=1, 2, \dots, k). \quad (5.73)$$

Эта вероятность представляет собой, естественно, функцию распределения времени пребывания в системе до выбытия, и она может быть вычислена непосредственно. В действительности следовало бы выбрать подходящие значения  $r_{h+1}(T)$  для каждой задачи на основе наблюдаемых распределений  $p_{i,h+1}(T)$ .

Система уравнений (5.71) не может быть решена тем же способом, который описан в гл. 4, так как не все коэффициенты постоянны. Однако она может быть преобразована в другую систему с постоянными коэффициентами. Пусть

$$\left. \begin{aligned} r'_{ij} &= r_{ij} \quad (i \neq j), \\ r'_{jj} &= r_{jj}(T) + r_{h+1}(T) \quad (j=1, 2, \dots, k) \end{aligned} \right\} \quad (5.74)$$

Заметим, что  $r'_{ij}$  постоянны для всех  $j$  при определенных значениях  $r_{jj}(T)$ . Введем новое множество вероятностей  $\{p'_{ij}(T)\}$ , обозначив их следующим образом:

$$p_{ij}(T) = p'_{ij}(T) \{1 - p_{i,h+1}(T)\} \quad (i, j=1, 2, \dots, k). \quad (5.75)$$

Вероятность  $p'_{ij}(T)$ , определенная уравнением (5.75), — условная вероятность перехода со ступени  $i$  на ступень  $j$  на интервале  $(0, T)$  при условии, что за этот интервал не было уходов из системы. Подставляя (5.75) в (5.71), получаем

$$\frac{dp'_{ij}(T)}{dT} = \sum_{h=1}^j r'_{hj} p'_{ih}(T) \quad (i, j=1, 2, \dots, k). \quad (5.76)$$

Система уравнений (5.76) может быть решена методами, описанными в разделе 4.1. Найдя  $p'_{ij}(T)$  из (5.76), можно определить  $p_{ij}(T)$  из (5.75).

Хотя вероятности  $\{p_{ij}(T)\}$  были вычислены для данной системы в предположении, что интенсивности уходов постоянны, очень просто получить вероятности для интенсивностей, зависящих от времени. Чтобы показать это, положим  $i=1$ . Тогда можно записать:

$$p'_{1j}(T) = \sum_{h=1}^k c'_{jh} e^{r'_{hh} T} \quad (j=1, 2, \dots, k). \quad (5.77)$$

Сравнивая с (5.41), видим, что коэффициенты  $\{c_{jh}\}$  те же, что и для системы с постоянными интенсивностями уходов, поскольку  $r_{jj} = r_{hh} = r'_{jj} - r'_{hh}$  и  $r_{j, j+1} = r'_{j, j+1}$ . Теперь, используя (5.75), получаем

$$p_{1j}(T) = \sum_{h=1}^k c_{jh} \exp \left\{ r'_{hh} T - \int_0^T r_{h+1}(x) dx \right\}, \quad (5.78)$$

где коэффициенты  $\{c_{jh}\}$  — те же, что были вычислены ранее при постоянных интенсивностях уходов. Единственное, что меняется при подобном обобщении, это вид экспоненты, так как  $r_{hh}(T)$  заменяется на

$$r'_{hh} T - \int_0^T r_{h+1}(x) dx.$$

### Иллюстрация теоретических результатов

Сейчас мы вполне подготовлены к оценке погрешности от предположения, что интенсивности уходов постоянны. Вернемся ко второму примеру из раздела 5.2. Положим  $k = 3$ ,  $r_{12} = r_{23} = r = 1$  и  $r_{14} = r_{24} = r_{34} = r_4 = 2$ . Ограничимся лишь постоянной скоростью поступления  $R = 100$  и рассмотрим только предельную структуру  $\bar{p}(\infty)$ . Отметим с самого начала, что ожидаемые размеры системы зависят только от среднего значения распределения полного времени пребывания в системе. Поэтому выберем такое значение  $r_4(T)$ , чтобы среднее время пребывания в системе было равно половине принятого ранее. При постоянной интенсивности уходов продолжительность пребывания в системе имеет экспоненциальное распределение. В нашем примере мы рассмотрим два крайних случая отклонений от экспоненциальности.

Для начала выберем смесь экспоненциальных распределений с плотностью

$$f(T) = \frac{1}{2} \{ \lambda_1 e^{-\lambda_1 T} + \lambda_2 e^{-\lambda_2 T} \}. \quad (5.79)$$

Чтобы среднее значение времени пребывания в системе было то же, что и в ранее рассмотренном примере, мы должны выбрать  $\lambda_1$  и  $\lambda_2$ , входящие в (5.79), такими:

$$\lambda_1^{-1} + \lambda_2^{-1} = 1. \quad (5.80)$$

Для совпадения переходных вероятностей с уже найденными при постоянных интенсивностях требуется найти такую смесь экспоненциальных распределений, чтобы выполнялось

$$\exp \left\{ - \int_0^T r_{h+1}(x) dx \right\} = \frac{1}{2} \{ e^{-\lambda_1 T} + e^{-\lambda_2 T} \}.$$

В случае, когда  $\lambda_1 = \lambda_2 = 2$ , получаем формулы для экспоненциального распределения времени пребывания в состоянии. Сделав все необходимые преобразования (5.78), находим

$$\left. \begin{aligned} \bar{n}_1(\infty) &= 50 \left\{ \frac{1}{\lambda_1 + 1} + \frac{\lambda_1 - 1}{2\lambda_1 - 1} \right\} \\ \bar{n}_2(\infty) &= 50 \left\{ \frac{1}{(\lambda_1 + 1)^2} + \frac{(\lambda_1 - 1)^2}{(2\lambda_1 - 1)^2} \right\} \\ \bar{n}_3(\infty) &= 50 - \bar{n}_1(\infty) - \bar{n}_2(\infty) \end{aligned} \right\}. \quad (5.81)$$

где  $\lambda_2$  исключено благодаря (5.80), а  $\lambda_1 \geq 1$ . Поскольку  $\lambda_1$  меняется от единицы до бесконечности, все возможные значения для  $\bar{n}(\infty)$  получают из (5.81). Можно показать, что ожидаемые размеры ступеней, определяемые по (5.81), принимают крайние значения для экспоненциального распределения при  $\lambda_1 = 2$ , а также при  $\lambda_1 = 1$  и  $\lambda_1 = \infty$ . Последние два случая эквивалентны, так как  $f(T)$  симметрична относительно  $\lambda_1$  и  $\lambda_2$ . Достижимые крайние структуры приведены в первых двух столбцах табл. 5.2. Обсудим эти результаты далее.

Предположим, что в противоположном рассмотренному в последнем разделе крайнем случае время пребывания в системе постоянно. Тогда имеем

$$\left. \begin{aligned} r_4(T) &= 0 & (T \leq r_4^{-1}) \\ r_4(T) &= \infty & (T > r_4^{-1}) \end{aligned} \right\}. \quad (5.82)$$

Когда  $r_4 = 2$ , предельные размеры ступеней определяются по формуле

$$\bar{n}_j(\infty) = 100 \int_0^{\frac{1}{2}} p_{ij}^0(x) dx \quad (j = 1, 2, 3), \quad (5.83)$$

в которой  $p_{ij}^0(x)$  получено из (5.44) при  $r_{k+1} = 0$ . Область интегрирования ограничена, так как

$$\exp \left\{ - \int_{\frac{1}{2}}^T r_4(x) dx \right\} = 0 \quad \text{при } T > \frac{1}{2}.$$

Подставляя численные значения  $r$ , находим

$$\left. \begin{aligned} \bar{n}_1(\infty) &= 100 \left( 1 - e^{-\frac{1}{2}} \right) = 39,4 \\ \bar{n}_2(\infty) &= 100 \left( 1 - \frac{3}{2} e^{-\frac{1}{2}} \right) = 9,0 \\ \bar{n}_3(\infty) &= 50 - \bar{n}_1(\infty) - \bar{n}_2(\infty) = 1,6 \end{aligned} \right\}. \quad (5.84)$$

Полученные далее результаты сведены в табл. 5.2.



**Т а б л и ц а 5.2. Предельные структуры при различных предположениях об интенсивностях уходов**

	$\lambda_1 \rightarrow 1$ или $\infty$	$\lambda_1 = 2$	Постоянное время пребывания
$\bar{n}_1(\infty)$	25,0	33,3	39,4
$\bar{n}_2(\infty)$	12,5	11,1	9,0
$\bar{n}_3(\infty)$	12,5	5,6	1,6
$\bar{N}(\infty)$	50,0	50,0	50,0

Данные табл. 5.2 в определенном смысле отражают более широкие изменения структуры, чем это может быть в действительности. Они иллюстрируют только один частный пример, однако вычисления, проведенные для других ситуаций, показывают общность характеристик. Из таблицы видно, что чем больше коэффициент вариации распределения времени пребывания в системе, тем больше численность на высших ступенях. В абсолютных значениях изменения на верхней ступени почти уравниваются изменениями на нижней, но в относительных единицах размеры высшей ступени наиболее чувствительны к допущению о постоянстве интенсивностей уходов. В действительности распределения времени пребывания в системе сильно кососимметричны, и они с успехом заменяются смешанными экспоненциальными распределениями. Поэтому можно ожидать, что при постоянных интенсивностях продвижений предельная структура должна иметь вид, более похожий на первый столбец таблицы с относительно большим числом людей на высших ступенях. Таким образом, весьма вероятно, что любой фактор, способствующий росту коэффициента вариации времени пребывания в системе при постоянном среднем, влияет на увеличение численности на высших ступенях за счет нижних.

В предыдущих рассуждениях мы полагали, что интенсивности ступеней постоянны. Мы получили бы сходные результаты, если бы рассматривали расширяющиеся организации, которые растут со скоростью  $\lim_{T \rightarrow \infty} M(T)/N(T) = 0$ , так как предельная структура в обоих случаях одинакова. Большой интерес представляет более общий случай, когда интенсивность продвижения зависит от времени пребывания в системе.

### 5.5. ПОЛУМАРКОВСКИЕ МОДЕЛИ ОТКРЫТЫХ СИСТЕМ

Как показал Гилберт (1973), обобщение теории полумарковских моделей, описанных в гл. 4, для открытых систем осуществляется очень просто. Основное соотношение (5.3) остается справедливым. Для того чтобы воспользоваться им для полумарковских моделей, требуется вычислить  $P(T)$  так, как описано в разделе 4.4. Так же как и при описании замкнутых систем, подобные вычисления могут быть довольно сложными. Далее мы продемонстрируем два подхода применительно к частному случаю простой иерархической системы, которые позво-

ляют получить вполне стоящие результаты. Первый подход заключается в непосредственном решении основной системы интегральных уравнений, а второй сводится к введению гипотетических состояний.

## Прямой подход

Хотя зависимость интенсивности продвижения от времени пребывания в системе не видна явным образом, для фирм (организаций) характерен при проведении кадровой политики учет этого времени при продвижениях. Далее в настоящем разделе попытаемся осветить влияние зависимости интенсивности продвижения от времени пребывания на полученные ранее результаты.

В гл. 2 при обсуждении модели с дискретным временем мы сталкивались с подобной проблемой, пытаясь применить принцип накопленной инерции. В данном случае мы увеличим число состояний, поэтому стаж пребывания внутри измельченной градации превратится в часть характеристики градации до ее измельчения. Это возможно, так как стаж измеряется в тех же единицах шкалы времени. Гинзберг (1971) развил эту идею, применив теорию полумарковских процессов для построения моделей с непрерывным временем. При таком подходе процесс переходов определяется с помощью матрицы вероятностей переходов, а моменты времени самих переходов — с помощью распределения вероятностей — свое для каждого из возможных переходов. Такой процесс определяется так, что интенсивности переходов являются функциями времени пребывания в данном состоянии. В нашем случае мы видоизменим подход, чтобы отразить то увеличение системы, о котором говорилось выше.

Вначале рассмотрим замкнутую систему иерархического вида, в которой продвижение осуществляется только на ближайшую находящуюся выше ступень и поступление происходит только на самый нижний уровень. Тогда нам потребуются только переходные вероятности  $p_{ij}(T)$  ( $j = 1, 2, \dots, k$ ), где  $k$  — число ступеней. Когда они будут найдены, мы можем обратиться к (5.3), чтобы получить средние численности (размеры) ступеней, которые нас интересуют. Пусть величина  $r_{j, j+1}(\tau)$  ( $j = 1, 2, \dots, k-1$ ) означает интенсивность перехода со ступени  $j$  на ступень  $j+1$  для человека, который поступил на  $j$ -ю ступень за период времени  $\tau$  до настоящего момента. Как и ранее,  $r_{j, k+1}(T)$  — интенсивность уходов индивидуумов, которые находились в системе период времени  $T$ . Уравнения для вероятностей  $p_{ij}(T)$  представляют собой частный случай тех, что приводят к (4.58), но они имеют существенно более простой вид. Это объясняется иерархической структурой системы, которую мы изучаем. Таким образом, если мы рассматриваем переход из состояния 1 в состояние  $j$ , то имеется единственно возможная последовательность переходов, а именно  $1 \rightarrow 2 \rightarrow 3 \rightarrow \dots \rightarrow j$ . Таким образом,

$$p_{1j}(T) = \int_0^T p_{11}(\tau) r_{12}(\tau) p_{2j}(T-\tau) d\tau \quad (j = 2, 3, \dots, k), \quad (5.85)$$

Вероятности  $\{p_{1j}(T)\}$  могут быть получены рекурсивно, так как  $p_{2j}(T - \tau)$  то же самое, что и  $p_{1, j-1}(T - \tau)$  для  $(k - 1)$ -уровневой системы, когда первая ступень исключается. На практике очень нелегко получить точное решение, особенно для больших  $j$ . Так как мы будем рассматривать другой, более общий метод, детально обсуждать решения интегральных уравнений нет необходимости. Главное значение прямого подхода состоит в том, что он позволяет легко получить общее решение для  $j = 1$ , которое в свою очередь ведет к определению  $\bar{n}_1(T)$ . Сейчас мы будем применять прямой подход для такой цели.

Пусть имеется максимальное значение времени, которое может провести индивидуум на низшей ступени. Если за это время не было ни продвижения, ни ухода, то автоматически осуществляется продвижение. Такое допущение необходимо, чтобы интенсивность продвижения  $r_{12}(\tau)$  была бесконечной при  $\tau = b$ , где  $b$  — максимальное время пребывания на ступени. Простейшая возрастающая функция времени пребывания, обладающая таким свойством, была предложена Вайдой (1947):

$$r_{12}(\tau) = c / (b - \tau) \quad (0 \leq \tau < b; c > 0).$$

Допустим, что функция интенсивности уходов имеет аналогичный вид

$$r_{1, k+1}(\tau) = u / (v - \tau) \quad (0 \leq \tau < v; u > 0), \quad (5.86)$$

где  $v$  — максимальное время, которое можно пробыть в организации. (Для первой ступени общий стаж работы в системе и время пребывания на первой ступени одинаковы. Следовательно, мы можем использовать для обозначения этих совпадающих понятий либо  $T$ , либо  $\tau$ .) Если принять, что справедливо (5.86), то распределение полного времени пребывания, соответствующее интенсивностям уходов (5.86), имеет плотность

$$f(\tau) = \frac{u}{v} \left(1 - \frac{\tau}{v}\right)^{u-1} \quad (0 \leq \tau < v), \quad (5.87)$$

и, таким образом, среднее время пребывания равно:

$$\mu = v / (u + 1). \quad (5.88)$$

При выполнении всех предположений мы легко находим, что

$$p_{11}(T) = \left(\frac{b-T}{b}\right) \left(\frac{v-T}{v}\right)^u \quad (0 \leq T < \min(b, v)). \quad (5.89)$$

Среднюю численность людей на первой ступени можно определить, подставив последнее равенство в (5.3). Положив интенсивность поступлений для нерархической системы, рассматриваемой в этом разделе, постоянной, имеем

$$\left. \begin{aligned} \bar{n}_1(T) &= R \int_0^T \left(1 - \frac{\lambda}{b}\right)^c \left(1 - \frac{x}{v}\right)^u dx + n_1(0) \left(1 - \frac{T}{b}\right)^c \left(1 - \frac{T}{v}\right)^u \\ &\quad (0 \leq T < \min(b, v)) \\ \bar{n}_1(T) &= R \int_0^{\min(b, v)} \left(1 - \frac{x}{b}\right)^c \left(1 - \frac{x}{v}\right)^u dx \quad (\min(b, v) \leq T) \end{aligned} \right\} (5.90)$$

Из последнего соотношения ясно, что предельное значение  $\bar{n}_1(T)$  получается при  $T = \min(b, v)$  и выражается вторым из интегралов.

Теперь можно сравнить полученные результаты с теми, которые были приведены в предположении о постоянстве интенсивностей движений и уходов. Чтобы сделать это корректно, необходимо условиться, что в марковской системе то же среднее время пребывания, что и в полумарковской. Равенства выполняются при

$$r_{1, k+1} = (u+1)/v, \quad r_{12} = (c+1)/b. \quad (5.91)$$

Если мы ограничимся сравнением предельных значений, то для марковской системы  $\bar{n}_1(\infty) = R / (r_{12} + r_{1, k+1})$ . Это выражение следует сравнить со второй формулой из (5.90), которую можно записать так:

$$\bar{n}_1(\infty) = \int_0^{\min(b, v)} \left(1 - \frac{x}{b}\right)^{r_{12}b-1} \left(1 - \frac{x}{v}\right)^{r_{1, k+1}v-1} dx. \quad (5.92)$$

Если как  $v$ , так и  $b$  стремятся к бесконечности, то

$$\bar{n}_1(\infty) \rightarrow R \int_0^{\infty} e^{-x(r_{12} + r_{1, k+1})} dx = R / (r_{12} + r_{1, k+1}),$$

что совпадает с известным результатом для постоянных интенсивностей. Наибольшее расхождение между двумя случаями, когда и  $b$  и  $v$  малы. В крайних случаях возможны равенства

$$b = r_{12}^{-1} \quad \text{и} \quad v = r_{1, k+1}^{-1}$$

(при меньших значениях  $b$  и  $v$   $u$  и  $c$  становятся отрицательными). Теперь

$$\bar{n}_1(\infty) = R \min(r_{12}^{-1}, r_{1, k+1}^{-1}). \quad (5.93)$$

Таким образом, можно сделать вывод, что для любой системы с интенсивностями переходов, определяемыми (5.85) и (5.86), справедливо

$$R / (r_{12} + r_{1, k+1}) \leq \bar{n}_1(\infty) \leq R \min(r_{12}^{-1}, r_{1, k+1}^{-1}). \quad (5.94)$$

Следовательно, изменение интенсивностей движений в сторону увеличения из-за учета времени пребывания на ступени ведет к увеличению численности первой ступени. Поскольку общие ожидаемые размеры системы не зависят от вида зависимости интенсивностей движений от времени пребывания, мы в конце концов можем сделать

вывод, что относительная ожидаемая численность на первой ступени будет увеличиваться.

Приближение к нижнему пределу неравенства (5.94), когда  $b$  и  $v$  растут, происходит довольно быстро, что видно из следующего соотношения, полученного при  $b = v$ :

$$\bar{n}_1(\infty) = R/(r_{12} + r_{1,h+1} - v^{-1}). \quad (5.95)$$

Следовательно, можно сделать вывод, что предположение о постоянстве интенсивности продвижений может быть и не столь определяющим, особенно для первой ступени. Этот же подход можно было бы применить теперь для определения численности на следующей, второй, ступени, но мы обратимся к другому методу.

### Метод введения гипотетических состояний

Метод широко используется в теории очередей, которая применяется для рассмотрения немарковских систем. Его суть в том, что рассматриваемая система заменяется более сложной марковской. Это делается следующим образом. Предположим, что уходов нет и из ряда ступеней выбраны две с постоянными интенсивностями переходов, как показано на рис. 5.3. Времена пребывания на ступенях I и II — случайные величины, распределенные по экспоненциальному закону с средними значениями, соответственно равными  $1/r'_{12}$  и  $1/r'_{23}$ . Допустим теперь, что переходы между ступенями I и II мы не можем наблюдать и что членов этих двух ступеней нельзя различить. Таким образом,

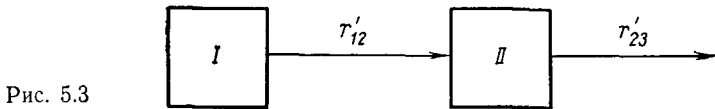


Рис. 5.3

время пребывания на такой объединенной ступени — сумма независимых случайных величин, распределенных по экспоненциальному закону. Рассмотрим с небольшой потерей общности частный случай, приняв, что  $r'_{12} = r'_{23} = r$ . Тогда время пребывания на объединенной ступени имеет плотность распределения

$$f(T) = r^2 T e^{-rT} \quad (T \geq 0). \quad (5.96)$$

Продвижение из объединенной ступени происходит при (5.96) с интенсивностями переходов

$$r_{12}(T) = f(T) \int_T^\infty f(x) dx = r \left( \frac{rT}{rT+1} \right) = 2r_{12} \left( \frac{r_{12}T}{r_{12}T + \frac{1}{2}} \right), \quad (5.97)$$

где  $r_{12}$  — постоянная интенсивность переходов, которая была бы при том же среднем времени пребывания в системе. Таким образом, мы получили марковскую систему, которая ведет себя в этом случае как немарковская с интенсивностью  $r_{12}(T)$ , определяемой уравнением (5.97).

Приведенный пример может быть обобщен. Вместо двух гипотетических ступеней можно рассмотреть  $g$ , для них:

$$f(T) = \frac{r^g}{(g-1)!} T^{g-1} e^{-rT} \quad (T \geq 0). \quad (5.98)$$

При таких условиях  $r_{12}(T)$  — всегда возрастающая функция  $T$ , скорость роста которой зависит от  $g$ , и она всегда имеет предел при  $T \rightarrow \infty$ .

Применяемый метод заключается в том, что каждая ступень действительной системы заменяется подходящим числом  $g$  гипотетических ступеней — «подступеней». Интенсивности потерь для всех подступеней должны совпадать с интенсивностью потерь из ступени, их содержащей. Средние численности подступеней определяются с помощью стандартных методов теории марковских процессов. Численности на действительных ступенях получают простым суммированием по соответствующим подступеням. Продемонстрируем этот метод на примере простой иерархической системы, которая рассматривалась в разделе 5.2, при постоянных интенсивностях продвижения. Снова ограничимся рассмотрением предельного режима. Допустим, что интенсивности уходов одинаковы со всех ступеней и обозначим их через  $r_{k+1}$ , кроме того, допустим, что поступления  $R$  постоянны. Хотя в общем случае это и не обязательно, предположим, что число  $g$  одинаково для каждой ступени. Тогда система будет заменена другой с  $gk$  ступенями с постоянными или равными интенсивностями продвижений. Чтобы облегчить сравнения с результатами, приведенными в (5.52), обозначим эти интенсивности через  $gr$ . Пусть  $\bar{z}_j(\infty)$  обозначает предельную среднюю численность  $j$ -й ступени; она может быть получена из (5.52) заменой  $r$  на  $gr$  и  $k$  на  $gk$ . Таким образом,

$$\left. \begin{aligned} \bar{z}_j(\infty) &= \frac{R}{gr} \left( \frac{gr}{gr+r_{k+1}} \right)^j \quad (j = 1, 2, \dots, gk-1) \\ \bar{z}_{gk}(\infty) &= \frac{R}{r_{k+1}} \left( \frac{gr}{gr+r_{k+1}} \right)^{gk-1} \end{aligned} \right\} \quad (5.99)$$

Тогда ожидаемая численность на  $j$ -й ступени первоначальной системы равна:

$$\left. \begin{aligned} \bar{n}_j(\infty) &= \frac{R}{gr} \sum_{i=g(j-1)+1}^{gj} \left( \frac{gr}{gr+r_{k+1}} \right)^i \\ &= \frac{R}{r_{k+1}} \left( \frac{gr}{gr+r_{k+1}} \right)^{g(j-1)} \left\{ 1 - \left( \frac{gr}{gr+r_{k+1}} \right)^g \right\} \end{aligned} \right\} \quad (5.100)$$

Если  $g = 1$ , оно превращается в (5.52). Увеличивая  $g$ , мы усиливаем зависимость продвижений от времени пребывания на ступени. В пределе при  $g \rightarrow \infty$  все продвижения осуществляются после того, как че-

ловец прослужил некоторое фиксированное время  $r^{-1}$ , в этом случае

$$\bar{n}_j(\infty) = \frac{R}{r_{k+1}} \exp \left\{ -\frac{r_{k+1}}{r} (j-1) \right\} \left[ 1 - \exp \left\{ -\frac{r_{k+1}}{r} \right\} \right] \quad (j=1, 2, \dots, k-1) \quad (5.101)$$

и

$$\bar{n}_k(\infty) = R/r_{k+1} - \sum_{j=1}^{k-1} \bar{n}_j(\infty).$$

Интересная особенность полученных результатов состоит в том, что относительные средние численности ступеней, кроме последней, образуют геометрическую прогрессию для любого  $g$ . Постоянный множитель прогрессии изменяется от значения  $(1 + r_{k+1}/r^{-1})$ , которое соответствует постоянной интенсивности продвижения, до значения  $\exp(-r_{k+1}/r)$ , которое соответствует продвижениям через постоянное время пребывания. Некоторые численные значения приведены в табл. 5.3.

**Таблица 5.3. Сравнение первых членов прогрессий, соответствующих двум противоположным правилам продвижения**

$r_{k+1}/r$	2	1	$\frac{1}{2}$	$\frac{1}{5}$	0
$\left(1 + \frac{r_{k+1}}{r}\right)^{-1}$	0,333	0,500	0,667	0,833	1,000
$\exp\left\{-\frac{r_{k+1}}{r}\right\}$	0,135	0,368	0,607	0,819	1,000

Таким образом, насколько важна гипотеза о постоянстве интенсивностей продвижения, зависит от отношения  $r_{k+1}/r$ . Если это отношение мало (продвижений гораздо больше, чем уходов), то гипотеза не влияет на результат. С другой стороны, если величина  $r_{k+1}/r$  велика, то характер гипотезы становится весьма важным. Это положение иллюстрирует табл. 5.4, в которой сравниваются структура для больших и малых значений отношения.

**Таблица 5.4. Относительная структура системы  $q(\infty)$ , когда продвижение осуществляется через постоянные или случайные промежутки времени ( $k=4$ )**

	Промежуток времени	1	2	3	4
$\frac{r_{k+1}}{r} = 2$	Случайный	0,667	0,222	0,074	0,037
	Постоянный	0,865	0,117	0,016	0,002
$\frac{r_{k+1}}{r} = \frac{1}{2}$	Случайный	0,333	0,222	0,148	0,296
	Постоянный	0,393	0,239	0,145	0,223

Интенсивности продвижений являются возрастающими функциями от времени пребывания на ступенях, поэтому относительно возрастают численности на начальных ступенях по сравнению с последующими. Был получен аналогичный результат, когда допускалась зависимость интенсивностей уходов от общего стажа. Другое сходство между этими двумя случаями заключается в том, что численность на более высоких ступенях существенно зависит от выполнения допущений. Тем не менее можно ожидать, что в марковских моделях получаются завышенные численности на более высоких ступенях и заниженные на начальных. Хотя с практической точки зрения более реалистично предполагать, что интенсивности продвижений — возрастающие функ-

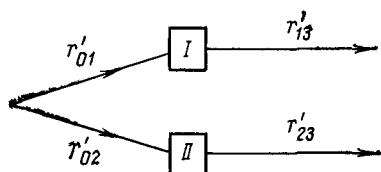


Рис. 5.4

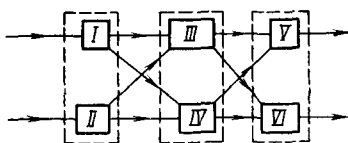


Рис. 5.5

ции времени пребывания (стажа), нельзя не отметить, что метод разбиения на гипотетические группы может быть применен и тогда, когда интенсивности — убывающие функции. Один из методов, реализующий такое разбиение, заключается в том, что марковские системы разбивают на параллельные состояния — подгруппы. Эту идею можно продемонстрировать на примере системы, показанной на рис. 5.4, с двумя параллельными гипотетическими подгруппами и постоянными интенсивностями переходов. Время пребывания в группе тех людей, которые прошли через блок I — случайная величина, распределенная по экспоненциальному закону с параметром  $r'_{13}$ . Для тех, кто прошел через блок II, параметр равен  $r'_{23}$ . Если нельзя различать эти две подгруппы, то очевидно, что плотность распределения времени пребывания будет смесью двух экспоненциальных распределений. Более точно, функция распределения будет иметь вид

$$f(T) = \left( \frac{r'_{01}}{r'_{01} + r'_{02}} \right) r'_{13} e^{-r'_{13} T} + \left( \frac{r'_{02}}{r'_{01} + r'_{02}} \right) r'_{23} e^{-r'_{23} T} \quad (T \geq 0). \quad (5.102)$$

Легко показать, что интенсивности уходов для такой комбинации подгрупп

$$f(T) \Big| \int_T^{\infty} f(x) dx$$

— это убывающая функция  $T$ .

Покажем сразу же, как этот результат может быть использован при изучении трехуровневой иерархической системы, интенсивности продвижений которой имеют вид, определяемый плотностью распределения (5.102). Рассмотрим марковскую систему, показанную на рис. 5.5, в которой возможные переходы показаны стрелками. Допустим, что



каждая пара подгрупп внутри пунктирного прямоугольника имеет одинаковые интенсивности уходов. Если прямоугольники представляют собой ступени действительной системы, эта будет марковская система с убывающими интенсивностями продвижений, в которой подгруппы объединены. Интенсивности переходов между подгруппами могут быть выбраны так, чтобы получились требуемые значения средних величин как для отдельных ступеней, так и по системе в целом. Следует заметить, что матрица интенсивностей переходов, соответствующая описанной выше системе, является треугольной, поэтому для нее при вычислениях остаются те же упрощения, которые справедливы для простой иерархической системы. Общая цель введения такого рода зависимостей заключается в том, чтобы увеличить относительную численность на последних ступенях при постоянных интенсивностях продвижений.

## 5.6. ДОПОЛНЕНИЯ К ГЛ. 4 И 5

### Теория

Теория марковских процессов с непрерывным временем часто излагается весьма кратко в элементарных учебниках и изложение нередко ограничивается специальными случаями, такими, как процессы размножения и гибели. Работа Кокса и Миллера (1965) полезна для начального более углубленного изучения, а книга Исааксона и Мадсена (1976) содержит полезную вводную главу. В книге Бартлетта (1955) кратко изложены наиболее важные положения теории, а тот, кто считает, что ему необходимо полное изложение, должен воспользоваться работой Чжуна (1967). Более глубокие исследования содержат много тонких особенностей, которые появляются при использовании непрерывного времени, например исследование вероятности бесконечно многих состояний за конечный интервал времени. Такая тщательность вряд ли уместна здесь, в прикладной работе, где в лучшем случае модель представляет собой более или менее грубое описание действительности. По этой причине важно, чтобы наши исследования были теснее связаны с исследованиями при дискретном времени, где для поведения, выходящего из обычных рамок, остается меньше возможностей.

При изложении теории полумарковских моделей снова была использована литература по теории случайных процессов, в особенности книга Кокса и Миллера (1965).

Часто упоминаются статьи Пика (1961a, b), но обзорная статья Синлара (1975) может дать наиболее полезный источник информации, она содержит ссылки на серьезную литературу, посвященную данному предмету. Для читателя, интересующегося социальными приложениями, очень полезна работа Гинзберга (1971).

По названиям работы Пика и Синлара больше относятся к марковской теории восстановлений, чем к теории полумарковских процессов. Марковские процессы восстановления и полумарковские процессы эквивалентны, и некоторые читатели воспринимают эти два термина

как синонимы. Их различие связано с тем, что они образуются по-разному, что важно для некоторых процессов. При основном на полумарковских процессах подходе главное внимание уделяется последовательности состояний и моментов времени, в которые происходят изменения. Марковские процессы восстановления связаны с общими процессами восстановлений, с ними мы встретимся в гл. 8.

## Приложения

По различным причинам модели с непрерывным временем стали применяться значительно позже, чем модели с дискретным временем. Работа Фикс и Неймана (1951) по описанию выживания после лечения заболевания раком — наиболее ранний пример марковской модели, так же как и работа Хербста (1963), которая обсуждается в гл. 7. Сравнительно недавно стало ясно, что гибкость марковских моделей предоставляет широкие возможности для исследователя. Кроме работы Криди (1979), можно назвать неопубликованную работу Плевиса (1980), посвященную построению модели для оценивания учителями успехов учащихся. Успехи представляют собой динамический процесс. Система будет замкнутой с  $k = 3$ . Плевис сумел показать, что, хотя марковская модель не очень точна при описании успехов всех 300 учащихся, она гораздо лучше подходит, когда учащихся разбивают на две группы — «основную» и «контрольную». Замкнутые модели на базе марковских процессов применялись Соренсоном и Халлинаном (1977), а также Холландом и Лейнхардтом (1977) для описания динамики групп и социальных взаимоотношений. В последнем случае определено такое понятие, как «схожесть», для каждой пары индивидуальных. Состояние системы в любой момент времени описывается в виде матрицы сфер влияния, показывающей пары, между которыми существуют отношения. Затем предполагается, что переходы из состояния в состояние описываются как марковский процесс с определенными интенсивностями переходов. Практическая ценность такого рода моделей ограничена из-за большой размерности пространства состояний, которая для них обычно требуется. В модели другого вида, предложенной Ямагучи (1980), описывается выбор дружеских отношений, математически представляющий собой процесс размножения и гибели. Эта модель обобщает более ранние результаты Коулмэна (1964b), у которого переходы в модели размножения и гибели разрешаются или нет в зависимости от того, дружеские отношения являются прямыми или косвенными. Модель открытой системы Сливки и Каннавале (1973) описывает процесс прохождения дел преступников в судебной системе. Используемые данные — случайная выборка о судьбах 1067 заключенных, арестованных полицией Филадельфии. Вводятся несколько состояний, которые образуют иерархическую систему: арест, предварительное обвинение (1), предварительный разбор дела (2), обвинительный акт (3), суд (4), тюремное заключение (5). Арест здесь — поступление в систему, поэтому он не нумеруется, состояние (6) — внешняя среда, куда могут быть уходы из состояний 2, 3 и 5. Матрица оценок интенсивностей переходов имеет вид

$$R_a = \begin{pmatrix} -1,680 & 0,0585 & 0 & 0 & 0 & 0,1095 \\ 0 & -0,0310 & 0,0308 & 0 & 0 & 0,0002 \\ 0 & 0 & -0,0110 & 0,0110 & 0 & 0 \\ 0 & 0 & 0 & -0,0190 & 0,0190 & 0 \\ 0 & 0 & 0 & 0 & -0,9600 & 0,9600 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Относительные численности в устойчивых состояниях системы при постоянных поступлениях, выражения для которых читатель может получить в качестве упражнения, имеют вид

$$n_1 = R (r_{12} + r_{16})^{-1}, \quad n_2 = R r_{12}' / (r_{12} + r_{16}) (r_{23} + r_{26}),$$

$$n_3 = R r_{12} r_{23}' / (r_{12} + r_{16}) (r_{23} + r_{26}) r_{34},$$

$$n_4 = R r_{12} r_{23}' / (r_{12} + r_{16}) (r_{23} + r_{26}) r_{45},$$

$$n_5 = R r_{12} r_{23}' / (r_{12} + r_{16}) (r_{23} + r_{26}) r_{56}.$$

Полумарковские модели выходят за рамки традиционного применения вероятностного подхода при анализе полной неопределенности и смертности по разным причинам, но они получают все большее распространение. Мы уже упоминали о работе Янга и Хурша (1973). Эти авторы использовали дискретный вариант модели, но из их теории следует, что непрерывный аналог связан с ним очень тесно.

В работе Као (1974) описано движение пациентов с сердечно-сосудистыми заболеваниями, особое внимание в ней уделено времени пребывания в различных состояниях, которое имеет распределения Вейбулла. Као рассматривает также получение оценок переходных матриц для вложенных марковских процессов. Наиболее глубокие теоретические исследования применения моделей для описания клинических экспериментов содержатся в работе Вейсса и Зелена (1965). Мак-Клин (1980) предложил полумарковскую модель для многоуровневой популяции с поступлениями по закону Пуассона. Валлиант и Милкович (1977) попытались сравнить полумарковские и марковские модели при прогнозировании численности персонала. Поскольку марковские модели — это частный случай полумарковских, можно сказать, что последние должны давать результаты, по крайней мере, не хуже, чем первые. Мнения авторов расходятся довольно сильно из-за различных методов оценивания. Во многих приложениях система наблюдалась весьма ограниченный период времени, поэтому данные о времени пребывания (стаж работы) не были достаточно полными. Такого рода проблемы были выявлены, но не решены. Томпсон (1980) предложил способ преодолеть эту проблему при изучении территориальной мобильности в Онтарио. Соответствующие материалы есть в работах Чжуна (1968), Сингера и Шпилермана (1976а), Бартоломью (1977а), а также Тумы и др. (1979).

Основное приложение полумарковских моделей для описания мобильности трудовых ресурсов в Скандинавских странах рассмотрено

в серии публикаций Гинзберга (1978а, 1978б, 1978с, 1979а, 1979б, 1979с). Серия широко связывает статистический анализ динамики систем с оценкой соответствующих параметров.

### Разновидности и обобщения

Существует много способов развития представленных теоретических результатов. Например, можно было бы использовать преобразования Лапласа на более ранней стадии. Поскольку отображение для  $\exp(\mathbf{R}T)$  — это  $(s\mathbf{I} - \mathbf{R})^{-1}$ , матрица  $\mathbf{P}(T)$  и вектор  $\bar{\mathbf{n}}(T)$  могут быть найдены через соответствующие отображения. Следует отдать должное усилиям Сингера и Шпилермана, предпринятым для решения уравнения

$$\frac{1}{T} \ln \mathbf{P}(T) = \mathbf{Q}.$$

С практической точки зрения ситуация сильно упрощается вследствие того, что большинство наблюдаемых матриц имеет вещественные положительные собственные значения, поэтому серьезных препятствий для решения не возникает.

В предыдущих изданиях этой книги приводилась другая процедура решения модели системы с фиксированными размерами (см. раздел 5.3). Там было показано, что в преобразованиях Лапласа ожидаемый вектор состояний имеет вид

$$\bar{\mathbf{n}}^*(s) = \{sn(0) \mathbf{p}_{k+1}^*(s) + M^*(s)\} \mathbf{p}_0 \mathbf{P}^*(s) (1 - s\mathbf{p}_0 \mathbf{p}_{k+1}^*(s))^{-1} + \mathbf{n}(0) \mathbf{P}^*(s).$$

Для иллюстрации приложения использован пример из раздела 5.3, поэтому приведенный результат эквивалентен полученному нами в настоящем издании работы. Здесь мы следовали результатам Мелмана (1977б), чтобы подчеркнуть параллельность соответствующих теоретических результатов для непрерывного и дискретного времени.

### 6.1. ПОСТАНОВКА ЗАДАЧИ

В предыдущих главах мы рассматривали различные варианты марковских цепей, описывающих социальные системы. В большинстве случаев исследовалось влияние постоянных переходных вероятностей на изменение структуры системы. Это давало нам возможность посмотреть, как зависит среднее поведение системы от склонности к перемещениям ее членов. Мы видели, например в гл. 3, как в кадровой системе с довольно правдоподобными интенсивностями переходов проявляется тенденция к росту численности на верхних уровнях за счет нижних. Это явление часто наблюдается на практике, где оно обычно проявляется в конце периода развития системы. Инерция, присущая кадровым системам, обеспечивает то, что интенсивности продвижений, которые возросли в период расширения системы, стремятся остаться на этом уровне и тогда, когда рост системы прекратился. В результате структура системы имеет вид пирамиды вершиной вниз. Это явление было характерно для многих университетов после их бурного роста в 60-е годы.

В ряде случаев можно оказывать некоторые управляющие воздействия на систему. Действительно, кадровое планирование и планирование образования как раз и служат тому, чтобы создавать системы и управлять ими, когда это необходимо. Обычно все начинается с предсказания того, что произойдет, если «имеющаяся в настоящее время тенденция сохранится». Указанные проблемы описаны в первых главах настоящей книги, но это только первый шаг. Структура, которая получается при прогнозировании, редко удовлетворяет кого бы то ни было, и поэтому возникает вопрос, что можно сделать, чтобы изменить положение вещей. При таком предположении возникает необходимость в теории управления. При разработке теории управления для марковских моделей мы должны двигаться в направлении, противоположном тому, в котором мы действовали в предыдущих главах. Там были даны вероятности переходов, и нам хотелось узнать, какие численности будет иметь система в будущем. Сейчас желаемые численности задаются и необходимо определить потоки людей, которые позволяют получить эту структуру.

В статической ситуации задача тривиальна. Любую желаемую структуру можно получить, если не вводить ограничений на то, когда

и где перемещать людей. На практике такие произвольные действия весьма нежелательны, а иногда и невозможны. Меры, предпринятые для управления системой, должны быть такими, чтобы серьезно учитывались желания людей, если мы хотим, чтобы система функционировала гармонично и эффективно. Далее мы должны исследовать, чего можно достичь при ограничениях, которые определяются необходимостью управлять хорошо. На практике это означает, что только некоторые из управляющих воздействий на потоки могут быть использованы, но и их применение весьма ограничено.

### Управляющие переменные

Теория управления наиболее развита для моделей с дискретным временем. За исключением короткого отступления в конце раздела 6.2, мы будем рассматривать модели открытых систем с дискретным временем, которые исследовались в гл. 3. Цель управления может быть различной. Например, может оказаться желательным, чтобы мы уложились в заданный бюджет или имеются определенные ограничения относительно численности на ступенях. Многие из этих целей можно сформулировать, используя понятие вектора ожидаемых состояний  $\mathbf{n}(T)$  для некоторых моментов  $T$ . Все возможные формулировки даются в разделе «Дополнения». Подход в настоящей главе заключается в том, что имеется желаемая структура или последовательность структур, которую мы хотели бы получить или поддерживать. Такие структуры будем обозначать звездочкой. На практике обычно эта структура не определяется точно. Вполне нормально, если имеющаяся структура будет достаточно близка к идеальной. Во всяком случае приемлемо, если отклонения случайных переменных будут малыми тогда, когда достигнута идеальная структура. Следовательно, представляется более разумным определить цель не в виде единственного значения  $\mathbf{n}^*$ , а в виде множества допустимых структур. Прежде чем приступить к исследованию, следует полнее понять ситуацию. Наша первая цель — выяснить, как функционирует система, и это можно сделать наиболее простым способом, классифицируя цели.

Поскольку управление осуществляется потоками людей, нужно рассматривать следующие понятия:

- а) увольнения (вектор  $\mathbf{w}$ , идентичный вектору уходов из гл. 3);
- б) продвижения; здесь рассматриваются как повышения в должности, так и понижения (матрица  $\mathbf{P}$ );
- в) найм на работу; он может пониматься в двух смыслах: как общее количество поступающих  $\{R(T)\}$  и как распределение их по отдельным группам  $\mathbf{r} \equiv \mathbf{r}_0$ . В этой главе вначале мы будем иметь дело с распределением (или управлением поступлениями).

Увольнения и продвижения тесно связаны в модели. Так как

$$\sum_{j=1}^k p_{ij} + w_i = 1 \quad (i = 1, 2, \dots, k),$$

любое изменение  $w_i$  или  $p_{ij}$  требует компенсирующего изменения одного или нескольких параметров, чтобы выполнялось равенство. Это не вызывает каких-либо трудностей при интерпретации приложений, но указанные случаи необходимо различать.

Поток увольнений можно менять в довольно широких пределах. Он может быть увеличен или просто увольнением в административном порядке, или созданием таких финансовых и других условий, которые стимулируют уходы, а также уменьшен путем создания лучших условий для работы и соответствующим стимулированием желания остаться в системе. Эти методы управления действуют, однако, довольно приближенно, а случаи увольнений в административном порядке вообще нежелательны.

Более определенно управление осуществляется потоками продвижений, которые являются результатами административных решений. Однако даже в этой ситуации часто есть причины, по которым изменение интенсивностей переходов осуществляется в меньших пределах. Например, увеличение интенсивности продвижений может привести к повышению по службе недостаточно квалифицированных работников и затруднит возвращение к ранее принятым значениям интенсивностей, когда будет обнаружено, что это необходимо. Точно так же уменьшение интенсивности продвижения вероятнее всего вызовет проблемы для людей, чьи надежды на ожидаемое вознаграждение исчезнут.

Потоки поступлений — наиболее привлекательная форма управления, так как решение о приеме большего или меньшего числа людей на данную ступень иерархии может быть принято без перемещения тех, кто уже работает в организации. Могут возникнуть практические трудности. При поиске новых работников соответствующей квалификации те, кто находится в системе, могут не одобрять эти действия, например, надеясь продвинуться по службе. Тем не менее из трех методов управления вектор поступлений выглядит наиболее предпочтительным. В некоторых приложениях это единственная управляющая переменная. Так, если уровни системы образуются по признаку возраста (или стажа), то ясно, что нельзя управлять процессом «продвижения» с одного уровня на другой.

Управление имеет два аспекта — *достижимость нужного и поддержание достигнутого*. Достижимость включает проверку, может ли быть цель достигнута, и если может, то каким способом. Поддержание достигнутого означает — что нужно сделать, чтобы сохранить достигнутую однажды заданную структуру. Логически достижимость должна исследоваться первой, но так как поддержание имеет значение для большего числа приложений, удобнее рассматривать их в обратном порядке. Кроме того, решая вопросы, связанные с поддержанием, мы таким образом решаем некоторые из вопросов достижимости.

В этой главе важное значение имеет различие между детерминистской и стохастической внешней средой. Сейчас необходимо подчеркнуть, что, кроме известного различия между детерминистскими и стохастическими моделями, имеет существенное значение, являются ли случайными неконтролируемые потоки. Когда марковские модели применялись для предсказания, мы могли легко переходить от детерми-

нистской к вероятностной интерпретации, и во многих случаях не было необходимости их различать. При рассмотрении задачи управления этому вопросу нужно уделять больше внимания, так как постановка задачи и использование решения зависит от того, какое допущение принято относительно характера внешней среды. Вначале в разделах 6.2 и 6.3 мы будем считать, что окружение является детерминистским. В действительности внешняя среда обычно стохастическая, но детерминистский аналог позволяет аппроксимировать более правдоподобный стохастический случай и служит основанием для дальнейшего исследования стохастического варианта.

## 6.2. ПОДДЕРЖАНИЕ СОСТОЯНИИ В ДЕТЕРМИНИСТСКОЙ ВНЕШНЕЙ СРЕДЕ

Если  $\mathbf{n}^*$  — вектор численностей групп, который необходимо поддерживать, то должны существовать значения  $\mathbf{P}$ ,  $\mathbf{w}$  и  $\mathbf{r}$ , такие, что

$$\mathbf{n}^* = \mathbf{n}^* \mathbf{P} + \mathbf{n}^* \mathbf{w}' \mathbf{r}. \quad (6.1)$$

Заметим, что в этом случае  $R(T) = \mathbf{n}^* \mathbf{w}'$ , так как если численности групп сохраняются, то необходимо, чтобы общие размеры системы оставались постоянными. Если управление осуществляется только поступлением, то  $\mathbf{P}$  и  $\mathbf{w}$  — фиксированные величины и  $\mathbf{r}$  надо определять. Это соответствует управлению поступлениями. Управление продвижением возникает тогда, когда  $\mathbf{w}$  и  $\mathbf{r}$  даны, а значения матриц  $\mathbf{P}$  выбираются.

### Управление наймом

Существует много причин, обуславливающих необходимость сохранить данную структуру и общую численность в течение определенного периода времени. Например, такая задача может возникнуть как часть кадрового планирования, когда выбраны определенные структура и численность, и требуется определить, какую политику продвижения и приема новых работников в систему необходимо проводить, чтобы обеспечить ее поддержание. И наоборот, например, существующая кадровая система может превращаться в систему, имеющую вид пирамиды вершиной вниз, и ставится вопрос, как исправить ухудшающуюся с каждым днем ситуацию. Не каждая структура может быть сохранена, и важное направление в исследованиях — отличить те структуры, которые могут быть сохранены, от тех, которые сохранить нельзя.

Так как в первую очередь нас интересует относительная структура системы, в этом разделе удобно использовать следующие переменные:

$$q_i(T) = \bar{n}_i(T)/N \quad (i = 1, 2, \dots, k),$$

где  $N$  — общая численность людей в системе. Тогда основное разностное уравнение для марковской модели имеет вид

$$\mathbf{q}(T+1) = \mathbf{q}(T) \mathbf{P} + \mathbf{q}(T) \mathbf{w}' \mathbf{r}. \quad (6.2)$$



Структура  $q(T)$  может быть сохранена, если можно найти такие значения управляющих параметров, что  $q(T+1) = q(T)$ . Другими словами, нас интересуют значения параметров, которые удовлетворяют уравнению

$$q = qP + qw'r. \quad (6.3)$$

Поскольку в рассматриваемом случае вектор  $r$  — единственный из параметров, поддающихся управлению, необходимо найти  $r$ , удовлетворяющее уравнению (6.3). Это легко сделать, разрешая (6.3) относительно  $r$ :

$$r = q(I - P)/qw'. \quad (6.4)$$

Нетрудно проверить, что вычисленные таким образом элементы вектора  $r$  в сумме дают единицу, но не все они обязательно положительные. Если компоненты положительные, то (6.4) определяет единственную политику управления посредством поступлений, если же есть отрицательные, то структура не может быть сохранена.

Иногда полезно иметь возможность выделить множество структур, которые можно поддерживать при данных  $P$  и  $w$ . Простой способ выделения области значений, которые могут быть сохранены, следует непосредственно из (6.4). Это множество  $q$ , такое, что

$$q \geq qP.$$

В целях наглядного представления множества дадим геометрическую интерпретацию задаче. Любая структура может быть представлена в виде точки в  $k$ -мерном евклидовом пространстве, где координатные оси — значения элементов  $q$ . Так как все  $q$  должны быть неотрицательны и в сумме равны единице, то все возможные структуры лежат на гиперплоскости  $\sum_{i=1}^k q_i = 1$  в положительном ортанте. Когда  $k = 2$ , они представляют собой гипотенузу прямоугольника с вершинами  $(0, 0)$ ,  $(0, 1)$ ,  $(1, 0)$ . Когда  $k = 3$ , множество сохраняемых структур образует равносторонний треугольник, показанный на рис. 6.1. При большей размерности трудно изобразить ситуацию на рисунке, но геометрическая терминология все же остается полезной. Множество возможных структур обозначим через  $\mathcal{X}$ . Поддерживаемая область  $\mathcal{M}$  будет подмножеством  $\mathcal{X}$ , и ее границы мы хотели бы определить.

Граница области  $\mathcal{M}$  находится следующим образом. Из (6.3)

$$q(I - P) = qw'r.$$

Следовательно,

$$q = qw'r(I - P)^{-1}, \quad (6.5)$$

Обращение матрицы всегда существует, как было отмечено ранее в (3.4). Вектор  $r$  может быть записан в виде  $\sum_{i=1}^k r_i e_i$ , где  $e_i$  — вектор с единицей на  $i$ -м месте и нулями на остальных. Подставляя его в (6.5), находим

$$q = qw' \sum_{i=1}^k r_i \{e_i (I - P)^{-1}\}. \quad (6.6)$$

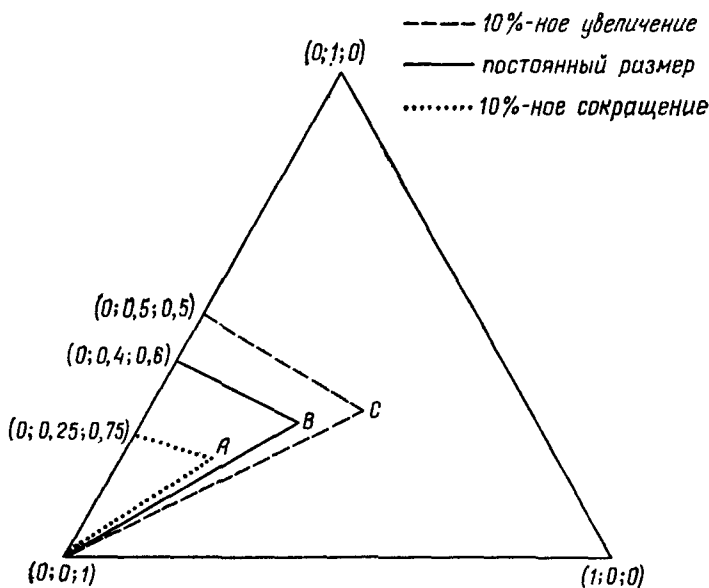
После умножения обеих частей (6.6) на единичный вектор-столбец получаем

$$1 = \mathbf{q}\mathbf{w}' \sum_{i=1}^k r_i d_i,$$

где  $d_i$  — сумма элементов  $i$ -й строки  $(\mathbf{I} - \mathbf{P})^{-1}$ . Подставляя отсюда значение  $\mathbf{q}\mathbf{w}'$  в (6.6), имеем

$$\mathbf{q} = \sum_{i=1}^k \left( \frac{r_i d_i}{\sum_{j=1}^k r_j d_j} \right) \frac{1}{d_i} \mathbf{e}_i (\mathbf{I} - \mathbf{P})^{-1}. \quad (6.7)$$

Таким образом,  $\mathbf{q}$  может быть представлена как выпуклая комбинация (т. е. средневзвешенное с неотрицательными значениями весов) точек с координатами  $d_i^{-1} \mathbf{e}_i (\mathbf{I} - \mathbf{P})^{-1}$  ( $i = 1, 2, \dots, k$ ). Для каждого



$$\hat{A} = (0,158; 0,210; 0,632), \quad B = (0,286; 0,286; 0,428), \quad C = (0,386; 0,307; 0,307)$$

Рис. 6.1. Область поддержания при управлении наймом,  $k=3$  и

$$\mathbf{P} = \begin{pmatrix} 0,5 & 0,4 & 0 \\ 0 & 0,6 & 0,3 \\ 0 & 0 & 0,8 \end{pmatrix}$$

г существует только одна комбинация значений весов и, следовательно, единственное  $\mathbf{q}$ , и если существует область изменений значений  $\mathbf{g}$ , то существует множество точек  $\mathbf{q}$ , являющихся выпуклой оболочкой, образованной точками

$$d_i^{-1} \mathbf{e}_i (\mathbf{I} - \mathbf{P})^{-1} \quad (i = 1, 2, \dots, k).$$

Таким образом, область поддержания — выпуклое множество, у которого эти точки — вершины. Координаты вершин легко вычислить, если по очереди взять строки  $(I - P)^{-1}$  и пронормировать их так, чтобы сумма каждой строки была равна единице. Если  $w$  одинаковы для всех уровней, то  $qw'$  постоянно, и результат следует немедленно из (6.6).

Итак, у нас есть метод определения вершин множества  $\mathcal{M}$  путем нормирования элементов строк  $(I - P)^{-1}$ . Когда  $k = 3$ , область может быть представлена в двумерном пространстве; для системы с матрицей

$$P = \begin{pmatrix} 0,5 & 0,4 & 0 \\ 0 & 0,6 & 0,3 \\ 0 & 0 & 0,8 \end{pmatrix}$$

Искомая область приведена на рис. 6.1. Она представляет собой выпуклое множество с вершинами  $(0; 0; 1)$ ,  $(0; 0,4; 0,6)$ ,  $(0,286; 0,286; 0,428)$ . Заметим, что все системы, которые можно поддерживать путем поступлений, имеют структуру групп вида пирамиды вершиной вниз. В приведенном примере последней точкой В диаграммы представлена структура групп, имеющая вид пирамиды вершиной вниз, в которой относительная численность верхней ступени наибольшая. В случае большей размерности не так легко на диаграмме изобразить область, но изучение координат вершин дает хорошее представление о структурах, которые можно поддерживать.

В приведенном численном примере представлен важный класс матриц переходных вероятностей, которые мы назовем супердиагональными. Они имеют вид

$$P = \begin{bmatrix} p_{11} & p_{12} & 0 & 0 & 0 \\ 0 & p_{22} & p_{23} & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdot & \cdot & \cdot & p_{k-1, k} \\ 0 & \cdot & \cdot & \cdot & p_{kk} \end{bmatrix}.$$

Если матрица имеет такой вид, то переходы возможны только на ближайшую ступень более высокого уровня; точно такие матрицы появляются также тогда, когда ступени определяются возрастными группами. Для матрицы  $(I - P)$  можно легко получить обратную; ее  $i$ -я строка равна:

$$e_i (I - P)^{-1} = \left( 0, \dots, 0, \frac{1}{1 - p_{ii}}, \frac{p_{i, i+1}}{(1 - p_{ii})(1 - p_{i+1, i+1})}, \dots \right. \\ \left. \dots, \frac{1}{1 - p_{ii}} \prod_{r=i}^{k-1} \frac{p_{r, r+1}}{(1 - p_{r+1, r+1})} \right) \quad (i = 1, 2, \dots, k-1) \quad (6.8)$$

и

$$e_k (I - P)^{-1} = \left( 0, 0, \dots, 0, \frac{1}{1 - p_{kk}} \right).$$

Выражение для расчета суммы элементов строк матрицы  $(\mathbf{I} - \mathbf{P})^{-1}$  простое, но координаты вершин множества  $\mathcal{M}$  легко вычисляются численным методом. Далее мы увидим, что структуры, соответствующие вершинам, поддерживаются в результате поступлений только на одну ступень системы. Первая строка матрицы  $(\mathbf{I} - \mathbf{P})^{-1}$  представляет особый интерес, так как она отражает структуру, которая поддерживается путем найма на нижнюю ступень иерархии. При таких поступлениях  $j$ -я ступень имеет численность, пропорциональную

$$\frac{1}{(1 - p_{jj})} \prod_{r=1}^{j-1} \frac{p_{r, r+1}}{(1 - p_{r+1, r+1})}, \quad (6.9)$$

где произведение принимается равным единице при  $j = 1$ . Выражение (6.9) нужно сравнить с (3.23), которое будет таким же, если имеются постоянные поступления на низшую ступень. Обе ситуации существенно схожи, но мы снова подчеркнем особую важность величины  $(1 - p_{jj})$  в том, какая численность людей на ступени будет поддерживаться. Другие поддерживаемые структуры — средневзвешенные от строк соответствующим образом нормированной матрицы  $(\mathbf{I} - \mathbf{P})^{-1}$ . Так как первая строка единственная, для которой поступления на низшую ступень не равны нулю, очевидно, что не может быть сохранена структура системы, численность на нижней ступени которой больше, чем  $1/d_1 (1 - p_{11})$ .

Приведенные выражения упрощаются еще больше, если равны нулю также диагональные элементы матрицы  $\mathbf{P}$ . Это получается, когда  $k$  ступеней представляют собой возрастные группы (группы с одним и тем же стажем работы), а время дискретно. В (6.8) остаются только числители, а ненулевые элементы каждой строки образуют убывающую последовательность, что чувствуется и интуитивно. Отсюда не следует, что любая поддерживаемая возрастная структура будет убывающей функцией возраста, так как нормирование и взвешивание рядов может привести к появлению «горбов». Например, очевидно, что трудности с поступлением в среднюю возрастную группу приведут к уменьшению ее численности, а это повлечет в дальнейшем уменьшение следующих возрастных групп. Если «возраст» — это время пребывания или стаж работы в системе, то поступления должны по определению осуществляться на нижнюю ступень. Таким образом, используется только первая строка, и любая сохраняющаяся в системе структура по стажу работы должна быть монотонно невозрастающей. Такой способ определения уровней выявляет важную связь между моделями, основанными на марковских цепях, и моделями теории восстановления, которые будут обсуждаться в последующих главах.

## Управление наймом при развитии и сокращении системы

Множество сохраняемых структур удивительно мало для близких к действительности матриц  $\mathbf{P}$ . В частности, если  $\mathbf{P}$  — верхняя треугольная матрица, то в  $\mathcal{M}$  обычно не допускаются типичные структуры вида пирамиды вершиной вверх для кадровой системы традиционного

типа. В этом разделе мы изучим влияние развития или сокращения системы. Исследование может проводиться с двух точек зрения. Развитие или сокращение системы может рассматриваться как еще один управляющий параметр, посредством которого может изменяться множество сохраняемых структур, или как что-то вносимое извне, что нам неподвластно. В некотором смысле мы не можем сохранить структуру системы, если она изменяется, так как численности некоторых групп неизбежно меняются. Мы можем, однако, рассмотреть, как определил Форбес (1970), квазистационарные системы\*, при которых сохраняются относительные размеры групп. Вероятно, наиболее правдоподобным является случай, когда относительные численности людей на ступенях определяются тем, *какую* работу выполняет организация, а *общая* численность зависит от того, *каков* ее объем.

Пусть  $N(T)$  — общий размер системы в момент времени  $T$  и пусть  $N(T+1) = (1 + \alpha)N(T)$ . Тогда основное уравнение для организации установленного размера (3.54а) примет вид

$$\bar{n}(T+1) = \bar{n}(T)P + \bar{n}(T)w'r + \alpha N(T)r. \quad (6.10)$$

Вводя  $q(T) = \bar{n}(T)/N(T)$ , получаем

$$(1 + \alpha)q(T+1) = q(T)P + (q(T)w' + \alpha)r. \quad (6.11)$$

Таким образом, условие, что квазистационарная структура  $q$ , сохраняется, запишется как

$$(1 + \alpha)q = qP + (q w' + \alpha)r \quad (6.12)$$

для  $r \geq 0^{**}$ , для которых  $\sum_{i=1}^k r_i = 1$ . Если такие  $r$  существуют, то они определяются из уравнения

$$r = \{q(I - P) + \alpha q\} / (q w' + \alpha). \quad (6.13)$$

Сравним это выражение с (6.4). Как и ранее, сумма элементов  $r$  всегда равна единице, но нас интересуют те из  $q$ , для которых они неотрицательны. Очевидно, что любое  $q$ , для которого  $r$  из соотношения (6.4) неотрицательны, определяет неотрицательные  $r$  в соответствии с (6.13), если  $\alpha > 0$ . Следовательно, если система расширяющаяся, то для нее область сохранения больше, чем для системы постоянного размера. Противоположным является случай, когда  $\alpha < 0$ . Таким образом, увеличение системы приводит к росту области сохраняемых структур, а сокращение — к ее уменьшению.

Вывод выражений для определения координат вершин области  $M$  для данного случая совершенно аналогичен, а соотношения для них получаются путем очевидной модификации соотношений для координат вершин  $M$  в соответствии с

$$e_i (I(1 + \alpha) - P)^{-1} \quad (i = 1, 2, \dots, k). \quad (6.14)$$

\* В советской литературе такие системы обычно называются стабильными. — Примеч. ред.

\*\* Условия  $r \geq 0$  для вектора  $r$  означают, как обычно, что все его компоненты  $r_i \geq 0$ . — Примеч. ред.

Единственное изменение заключается в том, что в выражении (6.8)  $(1 - p_{ii})$  заменяется в знаменателе на  $(1 + \alpha - p_{ii})$  для всех  $i$ . Если  $\alpha$  положительно, то очевидно, что система увеличивается. Это ведет к тому, что координаты вершин такие, что структуры сужаются к вершине. Если  $\alpha$  отрицательно, то возникает более сложная ситуация, так как видно, что  $(1 + \alpha - p_{ii})$  будет отрицательно, если  $|\alpha| > 1 - p_{ii}$ . Допустим, что в настоящее время численность на  $i$ -й ступени  $n_i(T)$ ; тогда в момент  $(T + 1)$  она будет равна  $(1 + \alpha)n_i(T)$ , если сохраняется имеющаяся структура. Однако если ожидаемое количество уходов  $w_i n_i(T)$  таково, что  $|\alpha| > w_i$ , то структура, конечно, не может быть сохранена. Область  $\mathcal{M}$  только тогда не является пустой, когда  $|\alpha| \leq w_i$  для всех  $i$  и когда максимальная интенсивность сокращения системы  $\alpha = -\min_i w_i$ , если, по крайней мере, одна структура для нее может поддерживаться.

Чтобы проиллюстрировать эти результаты, обратимся к численному примеру при  $k = 3$  (см. рис. 6.1). Координаты новых вершин, соответствующих 10%-ному увеличению и 10%-ному сокращению системы равны:

10%-ное увеличение:  $(0; 0; 1)$ ,  $(0; 0,5; 0,5)$ ,  $(0,386; 0,307; 0,307)$ ;  
 10%-ное сокращение:  $(0; 0; 1)$ ,  $(0; 0,25; 0,75)$ ,  $(0,158; 0,210; 0,632)$ .

Эти области также изображены на рисунке, из которого ясно, что изменение общих размеров системы оказывает сильное влияние на размеры области, особенно это видно при сравнении с координатами вершин, определенными ранее.

Дальнейшие вычисления могут интерпретироваться по-разному. Допустим, что нам надо исследовать влияние изменений интенсивностей увольнений на размеры области сохранения. Если прибавить  $\alpha$  к каждой интенсивности увольнений и вычесть его из каждого  $p_{ii}$ , то вектор поступлений, обеспечивающий сохранение  $q$ , будет равен:

$$r = q \{I - (P - \alpha I)\} / (qw' + \alpha). \quad (6.15)$$

Уравнение (6.15) идентично (6.13), поэтому влияние фиксированных изменений интенсивностей увольнений и влияние изменений в размерах системы эквивалентны. Но такая интерпретация может быть сделана для  $\alpha > 0$ , когда  $\alpha < \min_i p_{ii}$ .

### Управление посредством продвижений

Вектор найма в определенных обстоятельствах не поддается управлению, например, когда он ограничен рамками предложения или подчинен кадровой политике руководства. В этом случае необходимо узнать, в какой степени может быть осуществлено управление с помощью интенсивностей продвижений (понижений). Для системы постоянного размера эта задача состоит в нахождении матрицы  $P$ , удовлетворяющей уравнению

$$q = qP + qw'r. \quad (6.16)$$



и область поддержания — это множество таких  $q$ , что

$$0 \leq \sum_{j=1}^i r_j \left\{ \sum_{j=1}^k q_j \omega_j \right\} - \sum_{j=1}^i q_j \omega_j \leq q_i (1 - \omega_i) \quad (i = 1, 2, \dots, k-1). \quad (6.21)$$

Уравнение (6.20) имеет простое начальное решение, особенно когда  $r_1 = 1$ ,  $r_i = 0$  ( $i > 1$ ). В этом случае

$$p_{i, i+1} = \sum_{j=i+1}^k q_j \omega_j / q_i.$$

Словами это можно пояснить так: доля тех, кого надо продвигать с  $i$ -й ступени, должна быть равна числу людей, покинувших систему с  $(i+1)$ -й до  $k$ -й ступени, деленному на численность на  $i$ -й ступени.

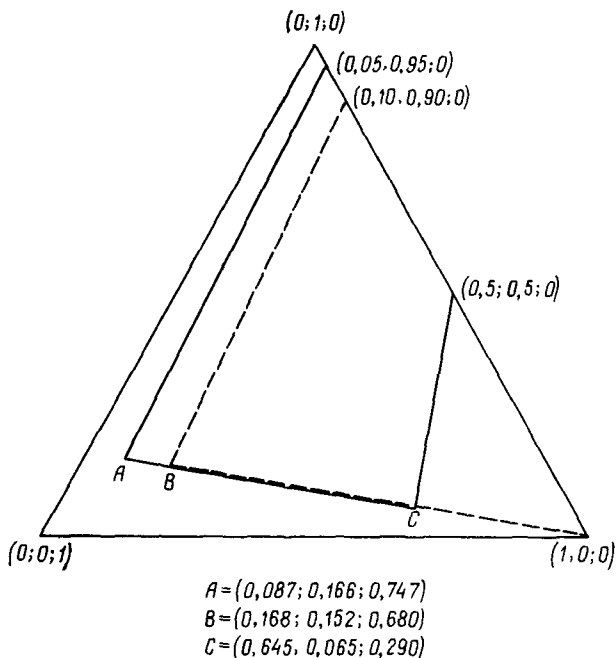


Рис. 6.2. Области поддержания при управлении продвижением,  $k=3$ , векторе уходов  $(0,1; 0,1; 0,2)$  и векторах поступлений а)  $(1; 0; 0)$  и б)  $(0,5; 0,5; 0)$ . Пунктирные границы определяет область задаваемая а), а сплошные — задаваемая б)

Последнее на самом деле, — вариант результата с дискретным временем, аналогичный тому, который мы получим при исследовании моделей восстановления позднее, в гл. 8.

Границы области поддержания, определяемые (6.21), очень интересны, особенно когда их сравнивают с соответствующими областями управления посредством найма. На рис. 6.2 показаны области для трехуровневой системы, в которой интенсивности уходов те же, что в



системе, изображенной на рис. 6.1, т. е. равны  $(0,1; 0,1; 0,2)$ , два вектора поступлений соответственно равны  $(1; 0; 0)$  и  $(0,5; 0,5; 0)$ . В противоположность ситуации, указанной на рис. 6.1, исключены из  $\mathcal{M}$  очень ярко выраженные структуры вида пирамиды вершиной вниз, а структуры вида пирамиды вершиной вверх включены. Точное решение зависит, конечно, от значений  $g$ ; значение вектора  $g = (1; 0; 0)$  для сохранения пирамидальных структур подходит больше, чем другие значения  $g$ . Если задача состоит в том, чтобы остановить тенденцию к росту численности на верхней ступени, то вероятнее всего для этого управление продвижением более эффективно, чем управление поступлением.

Если в качестве управляющего параметра используются как  $P$ , так и  $g$ , то очевидно, что может быть сохранена любая структура. Один из методов поддержания данной структуры состоит в том, чтобы принять все продвижения равными нулю и набирать на каждую ступень такое же число людей, какое покинуло систему. Существуют, конечно, и другие более разумные способы.

Нетрудно обобщить полученные результаты на случай развивающихся или сокращающихся систем. Приведем некоторые из основных результатов, а читатель без особого труда их проверит.

Область поддержания определяется неравенством

$$q \geq (qw' + \alpha)g / (1 + \alpha) \quad (6.22)$$

для любых  $P$ . Для супердиагональных  $P$  элементы  $P$  должны удовлетворять уравнениям

$$q_1 p_{11} = q_1 (1 + \alpha) - r_1 \left\{ \sum_{i=1}^k q_i \omega_i + \alpha \right\} \quad (6.23)$$

и

$$q_i p_{i, i+1} + q_{i+1} p_{i+1, i+1} = q_{i+1} (1 + \alpha) - r_{i+1} \left\{ \sum_{i=1}^k q_i \omega_i + \alpha \right\} \quad (i = 1, 2, \dots, k-1).$$

Разрешая их относительно  $p_{i, i+1}$  и используя неравенства  $0 \leq p_{i, i+1} \leq 1 - \omega_i$ , получаем следующее определение области поддержания:

$$0 \leq \sum_{j=1}^i r_j \sum_{j=1}^k q_j (\omega_j + \alpha) - \sum_{j=1}^i q_j (\omega_j + \alpha) \leq \leq q_i (1 - \omega_i) \quad (i = 1, 2, \dots, k-1). \quad (6.24)$$

### Поддержание при непрерывном времени

Многие из результатов относительно поддержания структур распространяются на модели с непрерывным временем путем небольших изменений. Это происходит потому, что области поддержания не затрагиваются, например, предельными переходами, описанными в гл. 4, которые мы применяли при переходе к моделям с непрерывным

временем на базе марковских цепей. Таким образом, условие  $q \geq qP$ , когда структура  $q$  поддерживается путем найма, запишется как

$$q > q(I + \delta TR)$$

или

$$qR > 0, \quad (6.25)$$

где  $R$  — матрица интенсивностей переходов. Координаты вершины области поддержания получаются путем нормализации строк  $(-R)^{-1}$ , которые пропорциональны  $(I - P)^{-1}$ . Легко показать, что вектор пополнений, обеспечивающий сохраняемость  $q$ , равен:

$$r = qR(qRI')^{-1}, \quad (6.26)$$

и он легко нормируется по  $qR$ . Если система развивается со скоростью  $\alpha$  так, что  $N(T) = N(0)e^{\alpha T}$ , то условие поддержания записывается следующим образом:

$$q(I\alpha - R) \geq 0. \quad (6.27)$$

Аналогично можно проанализировать и управление продвижением.

### 6.3. ДОСТИЖИМОСТЬ В ДЕТЕРМИНИСТСКОЙ ВНЕШНЕЙ СРЕДЕ

Достижимость — более сложное понятие, чем поддержание. Структура, которая может быть достигнута из одной точки, может оказаться недостижимой из другой. Число шагов, необходимых для достижения заданного состояния, и маршрут следования могут меняться, и можно как требовать, так и не требовать, чтобы общий размер системы оставался одним и тем же во всех промежуточных точках. Чтобы провести полный анализ, нужно рассмотреть все эти случаи, но для практических целей достаточно провести относительно ограниченные исследования. Начнем с рассмотрения нескольких точек-структур, которые можно пытаться достичь, но которых нет в области поддержания  $M$ . Если бы мы все же попали в такую точку-структуру, мы не могли бы оставаться в ней долго, в лучшем случае мы бы выиграли время, чтобы получить более основательные изменения в системе. Поэтому ограничимся тем, что желаемые структуры находятся в  $M$ . В этом случае получаются простые результаты, позволяющие решать наиболее важные задачи. Если принять, что «достичь» означает «подойти произвольно близко», то можно доказать, что *любая поддерживаемая структура  $q^*$  может быть достигнута из любой начальной структуры*. Этот вывод легко доказать, если обратиться к результатам относительно установленной структуры, приведенным к гл. 3. Мы приведем доказательство для случая управления наймом, так как для управления продвижением оно аналогично. Достаточно показать, что вывод справедлив, если общая численность системы остается фиксированной. Пусть мы начали из  $q(0)$ , и зададимся вопросом, что произойдет, если мы неоднократно используем вектор распределения принятых  $g^*$ , который по определению сохраняет структуру  $q^*$ . Известно, что устойчивая структура удовлетворяет в этом случае уравнению

$$q = qP + qwg^*.$$

По определению  $g^*$  также удовлетворяет уравнению

$$q^* = q^*P + q^*w^*,$$

что означает, что  $q = q^*$ . Но  $g$  существует только тогда, когда  $q^*$  принадлежит области  $M$ . Это показывает не только существование такой стратегии, но и то, что она единственна. Таким образом, достаточно делать найм в таких постоянных пропорциях, как если бы мы уже достигли желаемой структуры и хотели бы ее поддерживать. Сейчас возникает вопрос: можно ли добиться большего? На него можно было бы ответить, располагая некоторым критерием оптимальности. Существует много способов формулирования критерия оптимальности, например, в зависимости от временного интервала или стоимости. Опишем два подхода к этой задаче, применяя в обоих случаях методы математического программирования. При первом подходе, называемом управлением без временных ограничений, ставится цель вначале определить наименьшее число шагов, за которое достигается целевая структура, а затем выбрать среди всех возможных стратегий с таким числом шагов наилучшую по критерию стоимости. При втором методе ставится цель получить структуру, как можно более близкую к целевой, за заданное число шагов. Это приводит к идее применения этих стратегий, повторяя их так, чтобы максимально приблизиться к желаемой структуре.

### Управление наймом без временных ограничений

Для формулирования задачи управления в терминах математического программирования удобнее использовать понятия абсолютных численностей, а не относительных величин. Соответственно начнем с уравнения

$$n(T+1) = n(T)P + f(T+1), \quad (6.28)$$

где  $f(T+1) = R(T+1)g$  — вектор числа принятых в систему за период времени от  $T$  до  $T+1$ . В этой главе не будем обозначать поразному численности состояний и их математические ожидания. Пусть задана такая последовательность общих численностей людей в системе  $N(T)$ , которая удовлетворяет ограничениям, вводимым ниже. Задача состоит в том, чтобы найти  $T^*$  и последовательность векторов  $f(1)$ ,  $f(2)$ , ...,  $f(T^*)$ , таких, что  $T^*$  — наименьшее из всех  $T$ , для которых  $n(T) = n^*$  — целевая структура. Ограничения на неизвестные, которые получаются из модели, запишем так:

$$n(T+1) = n(T)P + f(T+1) \quad (T = 0, 1, 2, \dots, T^* - 2), \quad (6.29a)$$

$$n^* = n(T^* - 1)P + f(T^*), \quad (6.29b)$$

$$\sum_{i=1}^k n_i(T) = N(T) \quad (T = 1, 2, \dots, T^* - 1), \quad (6.29в)$$

$$n(T) \geq 0 \quad (T = 1, 2, \dots, T^* - 1), \quad (6.29г)$$

$$f(T) \geq 0 \quad (T = 1, 2, \dots, T^*).$$

Если  $T^* = 1$ , то (6.29a) и (6.29в) исключаются.

Все соотношения линейны относительно  $n_j(T)$  и  $f_j(T)$ , и, следовательно, мы можем применять стандартные методы математического программирования, чтобы найти какое-либо допустимое решение. Так как  $T^*$  неизвестно, необходимо последовательно брать  $T^* = 1, 2, \dots$  до тех пор, пока не получим первые значения допустимого решения. Заметим, что решение при ограничениях (6.29) дает не только значения  $f(T)$ , которые нам нужно найти, но и промежуточные численности  $n(T)$ , лежащие на траектории системы. Когда находят  $T$ , обычно получают много решений и, следовательно, остается возможность выбрать из множества решений одно, минимизирующее какую-то функцию, выражающую экономические или социальные цели. Одна из возможностей при рассмотрении проблем кадровой системы — минимизировать затраты на заработную плату. Если средняя заработная плата на  $i$ -й ступени  $c_i$ , то общий расход на заработную плату, соответствующий некоторой стратегии найма, равен

$$\sum_{T=1}^{T^*-1} \sum_{i=1}^k c_i n_i(T). \quad (6.30)$$

Теперь мы имеем стандартную постановку задачи математического программирования. (Последующие значения не влияют, так как затраты в будущем не принимаются в расчет.) Оптимизирующий алгоритм используется для минимизации (6.30) после того, как найдено значение  $T^*$ , для которого существует допустимое решение.

Когда  $T^* = 1$ , проверка существования допустимого решения тривиальна, но она имеет определенный практический интерес. В этом случае выполняется только (6.29б), и мы должны найти, существуют ли  $f(1) \geq 0$ , такие, что

$$n^* = n(0)P + f(1). \quad (6.31)$$

Если они существуют, то их можно получить по формуле

$$f(1) = n^* - n(0)P. \quad (6.32)$$

Условие, чтобы  $f(1)$  было неотрицательным, конечно, заключается в достижимости за один шаг. При методах с временными ограничениями, рассмотренных в следующем разделе, берут этот результат в качестве исходного.

Можно провести некоторые интересные исследования возможных стратегий наборов, которые минимизируют функцию стоимости (6.30), комбинируя численные значения неизвестных и уравнения (6.29). Всего в векторах  $n(T)$  ( $T = 1, 2, \dots, T^* - 1$ ) и  $f(T)$  ( $T = 1, 2, \dots, T^*$ )  $2kT^* - k$  значений переменных. Число уравнений может быть определено так:

$$(6.29 \text{ а}) \quad k(T^* - 1)$$

$$(6.29 \text{ б}) \quad k$$

$$(6.29 \text{ в}) \quad T^* - 1$$

$$\hline kT^* + T^* - 1$$

В соответствии с хорошо известным результатом теории линейного программирования число ненулевых переменных в допустимом базисном решении не может превышать  $kT^* + T^* - 1$ . В невырожденном случае оно равно этой величине. Если все элементы  $\mathbf{p}(0)$  положительны, то из (6.29а) очевидно, что вся последовательность векторов  $\mathbf{p}(T)$  имеет положительные компоненты и их число равно  $k(T^* - 1)$ . Остается  $(kT^* + T^* - 1) - (kT^* - k) = T^* + k - 1$  ненулевых значений (самое большое), которые, следовательно, распределены по  $kT^*$  возможным значениям векторов  $\mathbf{f}(T)$ . Следовательно, оставшиеся, по крайней мере  $(k - 1)(T^* - 1)$  компонент векторов найма, равны нулю, что означает, что найм осуществляется на несколько уровней.

Дальнейшие исследования помогут получить ответ на вопрос: как распределены нулевые компоненты? Допустим, что мы достигли момента времени  $T^* - 1$ . Тогда у нас поставлена одношаговая задача, которую мы умеем решать. Следовательно, если применить расчетную процедуру, описанную выше, можно решить задачу при  $T^* = 1$  и получить, что все элементы последнего вектора распределения нанятых положительны. Остается распределить  $T^* + k - 1 - k$  ненулевых элементов среди  $T^* - 1$  полученных векторов. Допустим, что  $N(T)$  неубывает настолько быстро из-за сокращения штатов, чтобы не было нанятых в каждый момент времени  $T$ . Итак, в каждом векторе должна быть, по крайней мере, одна ненулевая компонента. Следовательно, если остается только  $T^* - 1$  ненулевых компонент, то их придется не более чем по одной на вектор. Таким образом, типичная стратегия найма, минимизирующая (6.30), состоит в том, что поступления осуществляются только на одну ступень в каждый момент времени, кроме последнего шага, когда поступления осуществляются на все уровни.

Рассуждения следует несколько модифицировать, если встретится вырожденный случай и если в начальном векторе  $\mathbf{p}(0)$  содержатся нули (как окажется в численном примере, рассматриваемом ниже). Стратегия, которая определяется решением задачи линейного программирования, могла бы использоваться на практике, если бы найм осуществлялся всегда на один и тот же уровень. Но если стратегия требует, чтобы уровни часто менялись, то ее трудно применять. В связи с этим возникает необходимость введения дополнительных ограничений на область допустимых решений из-за условий практического применения.

Если опустить  $(T^* - 1)$  ограничений (6.29в), что означает, что мы не будем отслеживать промежуточную численность людей в организации, то может быть не более  $kT^*$  ненулевых элементов в оптимальном решении. Из них только  $kT^* - k(T^* - 1) = k$  могут быть в векторах нанятых. Следовательно, в общем случае не может быть поступлений до момента времени  $T^* - 1$ . Нелепость такой стратегии в большинстве практических ситуаций обуславливает необходимость введения ограничений вида (6.29в).

## Управление продвижением без временных ограничений

В этом случае задача состоит в том, чтобы найти последовательность матриц  $\mathbf{P}(T)$ , которые в совокупности с определенным вектором  $\mathbf{r}$  переводили бы систему из состояния  $\mathbf{n}(0)$  в  $\mathbf{n}^*$  за наиболее короткий промежуток времени. В этом случае основное уравнение может быть записано в виде

$$\mathbf{n}(T+1) = \mathbf{n}(T)\mathbf{P}(T) + \mathbf{n}(T)\mathbf{w}'\mathbf{r} + M(T+1)\mathbf{r}, \quad (6.33)$$

где  $M(T+1) = N(T+1) - N(T)$ . Здесь неизвестными являются векторы  $\mathbf{n}(T)$  и матрицы  $\mathbf{P}(T)$ , но основное уравнение не линейно относительно этих переменных. Следовательно, мы должны сформулировать задачу в новом виде, в которой ограничения линейризуются благодаря тому, что все неизвестные выражаются через потоки людей, а не через пропорции. Если временно отказаться от матричной записи, то будет легче подсчитывать уравнения и почувствовать линейную зависимость в них. Пусть  $n_{ij}(T) = n_i(T)p_{ij}$  — поток людей с  $i$ -й ступени на  $j$ -ую за интервал времени  $(T, T+1)$ . Тогда мы должны найти такое  $T^*$  и такие значения переменных  $n_{ij}(T)$  ( $i, j = 1, 2, \dots, k$ ) и  $n_i(T)$  ( $i = 1, 2, \dots, k$ ), что

$$n_j(T+1) = \sum_{i=1}^k n_{ij}(T) + r_j \left\{ \sum_{i=1}^k n_{ij}(T) \omega_i + M(T+1) \right\} \\ (j = 1, 2, \dots, k; T = 0, 1, \dots, T^* - 2), \quad (6.34a)$$

$$n_j^* = \sum_{i=1}^k n_{ij}(T^* - 1) + r_j \left\{ \sum_{i=1}^k n_{ij}(T^* - 1) \omega_i + M(T^*) \right\} \quad (j = 1, 2, \dots, k), \\ (6.34b)$$

$$\sum_{i=1}^k n_{ij}(T) = n_i(T) (1 - \omega_i) \quad (i = 1, 2, \dots, k-1; T = 0, 1, \dots, T^* - 1), \\ (6.34в)$$

$$\sum_{j=1}^k n_j(T) = N(T) \quad (T = 1, 2, \dots, T^* - 1), \quad (6.34г)$$

$$\mathbf{n}(T) \geq 0 \quad (T = 1, 2, \dots, T^* - 1), \\ \mathbf{n}_{ij}(T) \geq 0 \quad (i, j = 1, 2, \dots, k; T = 0, 1, \dots, T^* - 1). \quad (6.34д)$$

Предположим, как и ранее, что  $N(T)$  заданы и что (6.34г) опускаются при  $T^* = 1$ . Случай  $i = k$  исключен из (6.34в), так как он следует из (6.34а) вместе с (6.34в). Суммируя обе части (6.34а) по  $j$  и вычитая (6.34в), просуммированное по  $i$ , получаем опущенное уравнение.

Так же как и в случае управления наймом, полезно подсчитать количество уравнений и переменных. Число переменных зависит от того, является ли  $\mathbf{P}$  матрицей особой структуры. Подсчеты переменных для трех видов матриц приведены в следующей таблице:

Переменные	Полностью заполненные матрицы	Верхние треугольные матрицы	Супердиагональные матрицы
$n_j(T)$	$k(T^* - 1)$	$k(T^* - 1)$	$k(T^* - 1)$
$n_{ij}(T)$	$k^2 T^*$	$\frac{1}{2}k(k+1)T^*$	$(2k-1)T^*$
Всего	$k(k+1)T^* - k$	$\frac{1}{2}k(k+3)T^* - k$	$3kT^* - T^* - k$

Число уравнений равно:

Система уравнений	Число уравнений
(6.34а) (6.34б) (6.34в) (6.34г)	$k(T^* - 1)$ $k$ $T^*(k-1)$ $T^* - 1$
Всего	$2kT^* - 1$

Поэтому в общем случае в базисном допустимом решении будет  $2kT^* - 1$  ненулевых элементов. Опять-таки это меньше общего числа элементов, поэтому оптимальная стратегия (в смысле минимизации функции стоимости, такой же, как (6.29)) будет состоять из последовательности матриц с большим количеством нулей. Если найм осуществляется на каждую ступень, то вектор  $n_j(T)$  должен содержать ненулевые компоненты в каждой позиции. Остается  $(2kT^* - 1) - k(T^* - 1) = kT^* + k - 1$  элементов, которые должны быть распределены среди  $T^*$  матриц  $\mathbf{P}$ . На предпоследнем шаге, когда  $T^* = 1$ , имеется  $2k - 1$  ненулевых элемента для  $k^2$  позиций последней матрицы переходов. Остается  $k(T^* - 1)$  ненулевых элементов для  $(T^* - 1)$  матриц. Поскольку каждая матрица в каждой строке должна иметь, по крайней мере, один ненулевой элемент, в каждой строке имеется ровно один отличный от 0 элемент. Такое заключение объясняется тем, что на каждом шаге все члены данной градации, которые не покинули систему, должны или остаться там, где они были, или все вместе перейти в другую градацию. Если матрица супердиагональная, то это означает, что либо все продвигаются, либо не продвигается никто. Ясно, что подобная политика не встречается на практике. Однако есть другие допустимые решения (не базисные), при которых достигается цель за то же число шагов. Среди них можно найти хотя бы одно, которое больше подходит, но оно не приведет к минимуму функции стоимости. Если найм разрешается не в каждую градацию, то некоторые  $n_j(T)$  могут быть равны нулю, и поэтому больше ненулевых элементов

может быть в матрицах переходных вероятностей. Это не намного улучшает положение, поскольку на практике не бывает так, чтобы целые градации были время от времени пустыми.

Полученные здесь результаты отражают ограниченность использования линейного программирования для управления в том простом виде, который мы рассматривали. Однако они явно показывают, что способы изменения структуры системы путем управления продвижением будут с малой вероятностью или самыми быстрыми, или самыми дешевыми. Возможно, основная польза от применения алгоритма линейного программирования состоит в определении, может ли быть достигнута желаемая структура за относительно небольшое число шагов. Но даже этот результат должен применяться осторожно вследствие стохастической природы потоков. Алгоритм линейного программирования позволяет получить число шагов, если траектория системы развивается в среднем. В действительности число шагов — случайная величина.

### Управление наймом и продвижением с временным ограничением

На практике целевая структура может представляться не тем, что можно определить точно, а направлением, в котором необходимо двигаться. Тогда требуется получить значение, близкое к цели, за относительно короткий период времени, т. е. в действительности может быть поставлена задача — за определенный период времени сделать все, что только возможно.

Если обозначить заданный момент времени через  $T^*$ , то задача будет состоять в том, чтобы получить структуру, как можно более близкую к заданной, за  $T^*$  шагов. Чтобы сформулировать задачу в математической форме, мы должны определить близость в терминах измерения расстояния между двумя структурами. Предположим, что эта мера обозначена  $D(\mathbf{n}^*, \mathbf{n}(T^*))$  для положения системы в момент времени  $T^*$ . Задача состоит в том, чтобы получить минимум этой величины, используя соответствующие последовательности управляющих параметров и ограничений. Тогда формально мы можем сформулировать следующие две задачи.

#### 1. Управление наймом

Минимизировать

$$D(\mathbf{n}^*, \mathbf{n}(T^*))$$

при условии

$$\mathbf{n}(T+1) = \mathbf{n}(T)\mathbf{P} + \mathbf{f}(T+1) \quad (T = 0, 1, \dots, T^* - 1), \quad (6.35a)$$

$$\sum_{i=1}^k n_i(T) = N(T) \quad (T = 1, 2, \dots, T^*), \quad (6.35b)$$

$$\mathbf{n}(T) \geq 0 \quad (T = 1, 2, \dots, T^*); \quad \mathbf{f}(T+1) \geq 0 \quad (T = 0, 1, \dots, T^* - 1). \quad (6.35b)$$



2. Управление продвижением  
Минимизировать

$$D(n^*, n(T^*))$$

при условии

$$n_j(T+1) = \sum_{i=1}^k n_{ij}(T) + r_j \left\{ \sum_{i=1}^k n_{ij}(T) \omega_i + M(T+1) \right\} \\ (j=1, 2, \dots, k; T=0, 1, \dots, T^*-1), \quad (6.36a)$$

$$\sum_{j=1}^k n_{ij}(T) = n_i(T) (1 - \omega_i) \quad (i=1, 2, \dots, k-1; T=0, 1, \dots, T^*-1), \quad (6.36б)$$

$$\sum_{i=1}^k n_i(T) = N(T) \quad (T=1, 2, \dots, T^*). \quad (6.36в)$$

Насколько легко могут быть решены эти задачи, зависит от вида функции расстояния. Эта функция отражает величины потерь за то, что различные градации переполнены или не заполнены до конца. Она имеет довольно общий вид и будет основной при исследовании в этом разделе:

$$D_a = \sum_{i=1}^k W_i |n_i^* - n_i(T^*)|^a \quad (a > 0), \quad (6.37)$$

где  $W_1, W_2, \dots, W_k$  — множество неотрицательных весовых коэффициентов, отражающих важность достижения точного значения на каждом уровне. Большие значения  $a$  придают большую значимость расходам. Если  $a = 1$ , то сформулированная в виде (6.35) и (6.36) задача сводится к задаче линейного программирования, а если  $a = 2$ , то — к задаче квадратичного программирования. В оставшейся части раздела мы рассмотрим частные случаи общей проблемы, в которых практическая значимость сочетается с простотой математических выкладок, позволяющей получать точное решение.

Во-первых, допустим, что организация имеет постоянную численность. Это дает нам возможность поставить задачу в терминах пропорций градаций, что позволяет обойтись без (6.36б) и (6.36в). Во-вторых, примем, что  $T^* = 1$ . С первого взгляда такое ограничение кажется чрезмерно суровым, но нужно как следует понять, что стохастическая природа потоков требует делать пересчет стратегий на каждом шаге. Принятие ограничения означает, что на каждом шаге мы продвигаемся как можно ближе к цели. С точки зрения неопределенности внешней среды при планировании кадров было бы крайне желательным и очень удобным (вследствие применения математического аппарата) продвигаться на каждом шаге к целевой структуре как можно ближе. Третье ограничение заключается в том, что мы рассматриваем только случай управления наймом. Такое управление больше применяется, чем управление продвижением, так как оно относится еще к управлению возрастным распределением. Стратегии управления наймом, най-

денные методами программирования, с практической точки зрения также более подходят, чем стратегии управления продвижением.

Частный случай, который мы попробуем рассмотреть детально, может быть записан следующим образом:

минимизировать

$$D_a = \sum_{i=1}^k W_i |q_i^* - q_i(1)|^a \quad \text{по } \mathbf{r}(1)$$

при условии

$$\mathbf{q}(1) = \mathbf{q}(0)\mathbf{P} + \mathbf{q}(0)\mathbf{w}'\mathbf{r}(1) \quad (\mathbf{r}(1) \geq 0). \quad (6.38)$$

Рассмотрим теперь

$$\begin{aligned} q^* - q(1) &= q^* - q(0)\{\mathbf{P} + \mathbf{w}'\mathbf{r}(1)\} = q(0)\mathbf{w}' \left[ \frac{q^* - q(0)\mathbf{P}}{q(0)\mathbf{w}'} - \mathbf{r}(1) \right] \propto \\ &\propto \mathbf{y} - \mathbf{r}(1), \end{aligned}$$

где  $\mathbf{y} = \{q^* - q(0)\mathbf{P}\}/q(0)\mathbf{w}'$ . Заметим, что  $\sum_{i=1}^k y_i = 1$ . Таким образом, вместо  $D_a$  мы можем минимизировать функцию разностей  $y_i - r_i$ . Сделав соответствующие подстановки, окончательно запишем нашу задачу:

минимизировать

$$D_a = \sum_{i=1}^k W_i |y_i - r_i|^a \quad (a > 0) \quad (6.39)$$

при условии

$$\sum_{i=1}^k r_i = 1 \quad (\mathbf{r} \geq 0),$$

где для краткости мы опустили аргумент  $\mathbf{r}$ . Если цель достигнута за один шаг, то минимум  $D_a$  равен нулю и решением будет  $r_i = y_i$ . Если же случалось так, что, по крайней мере, один из  $y_i$  отрицателен, то решением будет  $\mathbf{r}$ , «ближайшее» к  $\mathbf{y}$ .

Покажем сначала, что  $r_i = 0$  при каждом  $i$ , для которого  $y_i \leq 0$ . (Это справедливо и для более широкого класса функций расстояния.) Обозначим:  $\Sigma^+$  — суммирование по тем  $i$ , для которых  $y_i > 0$ , и  $\Sigma^-$  — суммирование по оставшимся  $i$ . Тогда

$$D_a = \Sigma^+ W_i (y_i - r_i)^a + \Sigma^- W_i (|y_i| + r_i)^a.$$

$D_a$  уменьшается, если каждое  $r_i$  в первой сумме увеличивается до тех пор, пока не станет  $r_i = y_i$ , и оно увеличивается, если каждое  $r_i$  во второй сумме увеличивается. Следовательно, нам необходимо, чтобы во второй сумме значения  $r_i$  были как можно меньше. Ничто не запрещает сделать их равными нулю, потому что  $\Sigma^+ y_i > 1$ ; отсюда мы можем увеличивать значения  $r_i$  в первой сумме без выполнения условия, в соответствии с которым  $r_i > y_i$  для любого  $i$ . Следовательно, можно заменить  $D_a$  в (6.39) на  $\Sigma^+ W_i (y_i - r_i)^a$  и добавить ограничения  $r_i \leq \min(y_i, 1)$ .

Случай, когда  $a = 1$

$$D_1 = \Sigma^+ W_i y_i - \Sigma^+ W_i r_i.$$

Минимум имеет место, когда  $\Sigma^+ W_i r_i$  максимальна при условии, что  $0 \leq r_i \leq \min(y_i, 1)$ ,  $\Sigma^+ r_i = 1$ . Решение заключается в том, чтобы выбрать  $i$ , соответствующее максимальному значению коэффициента  $W_i$ , а  $r_i$  наибольшее. Если  $r_i$  оказалось меньше единицы, то выбираем  $i$ , соответствующее следующему по величине значению  $W_i$ , и берем соответствующее  $r_i$  настолько большим, насколько позволяют ограничения, и т. д. В частном случае, когда  $W_i = W$  для всех  $i$ ,  $D_1$  не зависит от значений  $r_i$  (так как  $\Sigma^+ r_i = 1$ ), и поэтому любое решение, удовлетворяющее ограничениям  $r_i \leq y_i$ , если  $y_i > 0$ , и  $r_i = 0$ , если  $y_i \leq 0$ , дает одно и то же минимальное значение целевой функции. Частный случай, когда  $r_i \propto y_i$ , если  $y_i > 0$ , из приведенного класса решений исследуется численно\*. С точки зрения практики это означает, что каждая градация, на которой вообще не осталось людей, получает некоторое число новых работников, пропорциональное ее нуждам. Такая стратегия исходит из принципа «справедливости». Он выбирается по причинам, не относящимся к математике.

Случай, когда  $a = 2$

$$D_2 = \Sigma^+ W_i (y_i - r_i)^2.$$

Это выражение минимизируется по  $r_i$  при условии, что  $\Sigma^+ r_i = 1$  и  $r_i \geq 0$  для всех  $i$ . Пусть

$$\Phi = \Sigma^+ W_i (y_i - r_i)^2 + 2\alpha \Sigma^+ r_i,$$

где  $\alpha$  — неопределенный множитель. Если  $\mathbf{r}^*$  — точка минимума среди  $\mathbf{r}$ , то

$$\left. \begin{aligned} \frac{\partial \Phi}{\partial r_i} &= 0, \text{ если } r_i^* > 0 \\ \frac{\partial \Phi}{\partial r_i} &> 0, \text{ если } r_i^* = 0 \end{aligned} \right\}. \quad (6.40)$$

Для определенной выше функции  $\Phi$  получаем

$$\left. \begin{aligned} r_i^* &= y_i - \frac{\alpha}{W_i}, \text{ если } y_i > \alpha/W_i \\ r_i^* &= 0, \text{ если } y_i \leq \alpha/W_i \end{aligned} \right\}, \quad (6.41)$$

где значение  $\alpha$  должно быть выбрано таким, чтобы  $\Sigma^+ r_i^* = 1$ . Если, например,  $W_i$  равны единицам и

$$y_1 = 0,7; y_2 = -0,5; y_3 = 0,1; y_4 = 0,4; y_5 = 0,3,$$

то можно показать, что  $\alpha = 0,133$  и, следовательно,

$$r_1^* = 0,567; r_2^* = r_3^* = 0; r_4^* = 0,267; r_5^* = 0,167.$$

Тот же результат может быть получен следующим образом.

\* Ответ получается по формуле  $r_i = \frac{y_i^+}{\Sigma y_i^+}$ , где  $y_i^+ = \max(y_i, 0)$ . — Примеч. ред.

Т а б л и ц а 6.1а. Численное сравнение стратегий управления при  $k=3$  по структурам градаций

$T_0$	Стратегия	$q_1(T_0)$	$q_2(T_0)$	$q_3(T_0)$	$T$	Стратегия	$q_1(T)$	$q_2(T)$	$q_3(T)$
1	Постоянная $S_1$ $S_2$ Оптимальная	0,200	0	0,800	5	Постоянная $S_1$ $S_2$ Оптимальная	0,300	0,253	0,447
		0,100	0,100	0,800			0,223	0,268	0,508
		0,100	0,100	0,800			0,253	0,253	0,495
		0,200	0	0,800			0,300	0,253	0,447
2	Постоянная $S_1$ $S_2$ Оптимальная	0,280	0,080	0,640	6	Постоянная $S_1$ $S_2$ Оптимальная	0,295	0,272	0,433
		0,151	0,179	0,670			0,237	0,276	0,487
		0,165	0,165	0,670			0,264	0,264	0,472
		0,280	0,080	0,640			0,295	0,272	0,433
3	Постоянная $S_1$ $S_2$ Оптимальная	0,304	0,160	0,536	7	Постоянная $S_1$ $S_2$ Оптимальная	0,291	0,281	0,428
		0,182	0,228	0,590			0,247	0,280	0,473
		0,207	0,207	0,586			0,272	0,272	0,457
		0,304	0,160	0,536			0,286	0,286	0,429
4	Постоянная $S_1$ $S_2$ Оптимальная	0,306	0,218	0,477		Целевая структура	0,286	0,286	0,428
		0,205	0,254	0,540					
		0,235	0,235	0,531					
		0,306	0,218	0,477					

Т а б л и ц а 6.1б. Численное сравнение стратегий управления вектором найма  $\gamma'(T)$  при  $k=3$

Стратегия	$T$							Относительная стоимость
	1	2	3	4	5	6	7	
$S_1$	0,500	0,559	0,641	0,719	0,782	0,831	0,868	1,06
	0,500	0,441	0,359	0,281	0,218	0,169	0,132	
	0	0	0	0	0	0	0	
$S_2$	0,500	0,639	0,747	0,827	0,883	0,922	0,949	1,05
	0,500	0,361	0,253	0,173	0,117	0,078	0,051	
	0	0	0	0	0	0	0	
Оптимальная	1	1	1	1	1	1	0,965	1,00
	0	0	0	0	0	0	0,032	
	0	0	0	0	0	0	0,003	

Вычисляем

$$\left. \begin{aligned} y'_i &= 0, \text{ если } y_i \leq 0 \\ y'_i &= y_i - (\sum^+ y_i - 1) / W_i \sum^+ W_j^{-1} \end{aligned} \right\}. \quad (6.42)$$

Если  $y'_i \geq 0$ , то  $r = y'$ ; в противном случае заменяем первоначальные  $y_i$  на полученные  $y'_i$  и повторяем процесс вычислений до тех пор, пока не найдем неотрицательный вектор. Аналогичные стратегии получают и для других значений  $a$ , и можно легко проверить, что при одинаковых  $W_i$  они эквивалентны стратегии, описанной для случая, когда  $a = 2$ . Заметим, что решение, приводимое в (6.41), также минимизирует  $D_1$ , но в нем проявляется тенденция иметь меньшее число ненулевых элементов в управляющих векторах, чем при той стратегии, в которой  $r_i \propto y_i$ .

### Численное сравнение стратегий при управлении наймом

Чтобы проиллюстрировать результаты, возьмем  $k = 3$  и значения параметров, представленные на рис. 6.1, а именно

$$P = \begin{pmatrix} 0,5 & 0,4 & 0 \\ 0 & 0,6 & 0,3 \\ 0 & 0 & 0,8 \end{pmatrix}, \quad w = (0,1; 0,1; 0,2).$$

В качестве целевой структуры возьмем  $q^* = (0,286; 0,286; 0,428)$ , которая обозначена вершиной В на рис. 6.1; это наименьшая структура вида пирамиды вершиной вниз в области поддержания. В качестве начальной структуры возьмем  $q(0) = (0; 0; 1)$ . Все точки вдоль траектории перехода из  $q(0)$  в  $q^*$  могут рассматриваться как новые начальные структуры, так что вычисления дают информацию больше, чем это кажется с первого взгляда. Постоянная стратегия, заставляющая структуру приближаться в пределе к  $q^*$ , — это  $r^* = (1; 0; 0)$ . Анализ задачи при управлении без временных ограничений показывает, что наименьшее число шагов, за которое может быть достигнута целевая структура, равно семи. В табл. 6.1а и 6.1б сравниваются характеристики следующих четырех стратегий:

- а) постоянная, использующая все время  $r^* = (1; 0; 0)$ ;
- б)  $S_1$  — адаптивная стратегия минимизации

$$D_1 \text{ с } r_i \propto y_i, \text{ если } y_i > 0;$$

- в)  $S_2$  — адаптивная стратегия минимизации  $D_2$ ;
- г) оптимальная — стратегия управления без ограничений времени, использующая стоимости градаций в отношении 1 : 2 : 3.

Основное различие, обнаруженное в таблицах, имеется между  $S_1$  и  $S_2$ , с одной стороны, и постоянной и оптимальной стратегией — с другой. Первая пара позволяет реально осуществлять найм на два нижних уровня, в то время как последняя — лишь на самый нижний. Хотя соответствующие траектории движения к цели различны, различия после семи шагов почти нет. Постоянная и оптимальная стратегии в

этом случае весьма схожи, хотя так бывает не всегда. Например, если бы задача заключалась в достижении целевой структуры из точки  $q(0) = (1; 0; 0)$ , то оптимальная и адаптивные стратегии заключались бы в принятии в систему на высшие уровни. Многочисленные расчеты для различных структур как целевых, так и начальных показывают, что одношаговая адаптивная стратегия обычно очень подходит для достижения окрестности цели. Вычисления, приведенные в книге Бартоломью (1975), свидетельствуют, что она несколько хуже стратегий без временных ограничений и стратегий с временными ограничениями. Последние обсуждаются позднее. Эти выводы будут подкреплены, когда будут рассматриваться стохастические аспекты задачи в разделах 6.4 и 6.5.

#### 6.4. ПОДДЕРЖАНИЕ СОСТОЯНИЙ ПРИ СЛУЧАЙНЫХ ВНЕШНИХ ВОЗДЕЙСТВИЯХ

Когда мы исследуем поведение системы при случайных внешних воздействиях, возникают новые весьма тонкие особенности. В детерминистском случае структура может быть или сохранена, или нет. При случайных воздействиях возможно поддержание любой структуры, если достаточно много людей покинуло систему. Тем не менее некоторые структуры поддерживать легче других. Далее мы введем понятие вероятности того, что структура может быть сохранена. Величина этой вероятности зависит от момента времени, когда принято решение о принятии в систему новых людей. Если бы контролируемые потоки существовали уже на этом этапе, то мы были бы более информированы и поэтому лучше бы поддерживали структуру. Другое различие между детерминистским и вероятностным исследованием вытекает из первого. В детерминистском случае структура, которая может быть сохранена, сохраняется неограниченно долго. При случайном внешнем воздействии лишь некоторые структуры могут сохраняться бесконечно долго, и, следовательно, мы должны изучить, как обеспечить в течение определенного периода времени структуру, наиболее близкую к целевой. Отсюда задачи поддержания и достижимости тесно связаны, и мы не можем рассматривать одну из них без связи с другой.

#### Вероятность поддержания структуры посредством найма

Структура может быть поддержана с помощью найма тогда и только тогда, когда для каждой градации сумма потоков в нее (до найма) не превышает размера, соответствующего целевой структуре. Таким образом, чтобы  $n$  была структурой, которая может быть сохранена, требуется

$$n_j \geq \sum_{i=1}^k n_{ij} \quad (j = 1, 2, \dots, k), \quad (6.43)$$

где аргумент у  $n_{ij}(T)$  для простоты опущен. Ранее, когда могли иметь место любые потоки,  $n_{ij}$  считались случайными величинами, и мы могли вычислить вероятность того, что можно сохранить определенную структуру, по формуле

$$P_M = Pr \left\{ n_j \geq \sum_{i=1}^k n_{ij}, j=1, 2, \dots, k \mid \mathbf{n} \right\}. \quad (6.44)$$

В марковских моделях потоки уходов из градаций  $i$ ,  $n_{i1}, n_{i2}, n_{ik}, n_{i, k+1}$  имеют полиномиальное распределение с параметрами  $n_i, p_{i1}, p_{i2}, \dots, \dots, p_{ik}, \omega_k$  ( $i = 1, 2, \dots, k$ ), поэтому в принципе  $P_M$  может быть определена.

Перед тем, как перейти к обсуждению, как это сделать практически, кратко рассмотрим, что произошло бы, если бы число вновь нанятых  $\{f_i\}$  задавалось до того, как стали известны значения  $n_{ij}$ . Тогда можно было бы действовать в предположении, что размер найма равен своему математическому ожиданию, и выбрать  $\mathbf{f} = \mathbf{n} - \mathbf{nP}$ . Такая политика приведет к сохранению структуры только в том случае, если сделанное предположение точно выполняется, а это маловероятно. Поэтому в данной ситуации рассматривать  $P_M$  не имеет никакого практического смысла. Таким образом, далее рассмотрим предположение, что  $n_{ij}$  известны к моменту выбора  $\mathbf{f}$ .

В некоторых специальных случаях получить оценку для  $P_M$  легко. Например, пусть  $\mathbf{n}$  представляет собой распределение по возрастам (или стажу работы) с интервалом в один год, а  $n_i$  — численность людей в возрасте (стаж)  $i$ . Тогда через год все индивидуумы или покинут систему, или будут в возрасте  $i + 1$ , и переходные вероятности равны:

$$P_{ij} = 1 - \omega_i, \text{ если } j = i + 1, \\ P_{ij} = 0 \text{ в противном случае.}$$

После всех уходов вектор  $\mathbf{n}$  заменяется вектором

$$(0, n_1 - W_1, n_2 - W_2, \dots, n_{k-1} - W_{k-1}),$$

где  $W_i$  — численность людей, покинувших систему в возрасте  $i$ . Его значение может быть восстановлено до  $\mathbf{n}$ , если выбрать следующее значение вектора нанятых:

$$\mathbf{f} = (n_1, n_2 - n_1 + W_1, n_3 - n_2 + W_2, \dots, n_k - n_{k-1} + W_{k-1}).$$

Определим вероятность того, что все элементы этого вектора неотрицательны. Поскольку  $W_i$  — независимые биномиально распределенные величины, то

$$P_M = Pr \{ \mathbf{f} \geq 0 \} = \prod_{i=1}^{k-1} Pr \{ W_i \geq \max(0, n_i - n_{i+1}) \} = \\ = \prod_{i=1}^{k-1} \sum_{W_i=m_i}^{n_j} \binom{n_i}{W_i} \omega_i^{W_i} (1 - \omega_i)^{n_i - W_i}, \quad (6.45)$$

где  $m_i = \max(0, n_i - n_{i+1})$ . Если  $n_1 \leq n_2 \leq \dots \leq n_k$ , то  $P_M = 1$ . Это означает, что, по крайней мере, некоторые структуры точно могут быть сохранены даже при случайных переменных.

В матрицах  $\mathbf{P}$  общего вида вероятность представляется не так просто, как в (6.45), но она может быть найдена, если  $N$  не слишком велико; некоторые ее значения приведены в табл. 6.2. С другой стороны, может быть найдена хорошая аппроксимация в соответствии с центральной предельной теоремой. Таким образом, пусть

$$X_j = n_j - \sum_{i=1}^k n_{ij},$$

тогда

$$E(X_j) = n_j - \sum_{i=1}^k n_i p_{ij}, \quad \text{var}(X_j) = \sum_{i=1}^k n_i p_{ij}(1-p_{ij});$$

$$\text{cov}(X_j, X_h) = - \sum_{i=1}^k n_i p_{ih} p_{ij} \quad (j, h = 1, 2, \dots, k; j \neq h).$$

$X_j$  имеет полиномиальное распределение со средними и дисперсиями, приведенными выше. С учетом поправки на непрерывность получим

$$P_M = Pr\{X_j \geq 0; j = 1, 2, \dots, k\} = \\ = Pr\left\{u_j \geq \frac{-E(x_j) - \frac{1}{2}}{\left\{\sum_{i=1}^k n_i p_{ij}(1-p_{ij})\right\}^{\frac{1}{2}}}; j = 1, 2, \dots, k\right\}, \quad (6.46)$$

где  $\mathbf{u}$  имеет стандартное нормальное распределение с такой же корреляционной матрицей, что и  $X_j$ . Эта аппроксимация особенно полезна, когда  $N$  велико, а  $k$  мало. Некоторое представление о ее точности можно получить из табл. 6.2.

Т а б л и ц а 6.2. Точное и аппроксимированное значение  $P_M$  для системы при  $k=3$

$$\mathbf{P} = \begin{pmatrix} 0,7 & 0,2 & 0 \\ 0 & 0,8 & 0,1 \\ 0 & 0 & 0,9 \end{pmatrix}$$

n			Точное	Аппрок- смирован- ное	n			Точное	Аппрок- смирован- ное
8	8	8	0,369	0,358	2	5	11	0,707	0,646
18	18	18	0,304	0,304	6	15	33	0,775	0,708
36	36	36	0,270	0,271	12	30	66	0,870	0,858
72	72	72	0,247	0,248	24	60	132	0,955	0,950

Для практических целей точность аппроксимации достаточна, даже когда размеры градаций весьма малы. Структура в левой части таблицы представляет последнюю сохраняемую структуру вида пирамиды вершиной вниз, а в правой части — середину области  $\mathcal{M}$ . Более полную картину зависимости  $P_M$  от структуры при том же значении  $\mathbf{P}$  и



$N = 90$  дает рис. 6.3. Детерминистская область поддержания обозначена на рисунке, и мы видим, что  $P_M$  мало вне области и велико внутри нее. Однако часто весьма велика вероятность того, что структуры внутри  $M$  не могут быть сохранены при случайных воздействиях внешней

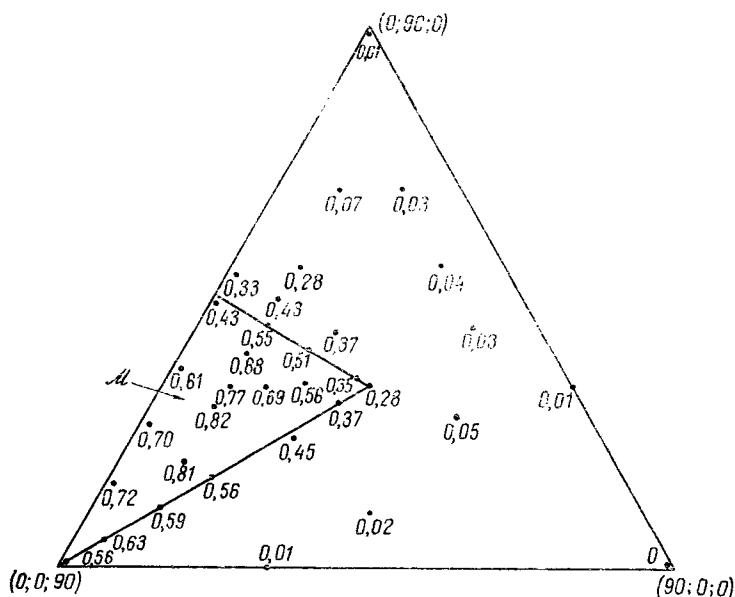


Рис. 6.3. Вероятность поддержания различных структур посредством управления наймом при

$$P = \begin{pmatrix} 0,7 & 0,2 & 0 \\ 0 & 0,8 & 0,1 \\ 0 & 0 & 0,9 \end{pmatrix}$$

среды. Как это и следовало ожидать, можно показать, что  $P_M \rightarrow 1$  внутри области  $M$  и  $P_M \rightarrow 0$  вне ее. На границе предельное значение лежит между 0 и 1.

### Вероятность поддержания структуры посредством продвижений

Здесь мы ограничимся обсуждением ситуации, когда градации образуют иерархию и продвижение осуществляется только в ближайшую верхнюю градацию. В этом случае соответствующие разностные уравнения запишутся так:

$$n_j(T+1) = n_j(T) - n_{j,j+1} + n_{j-1,j} - W_j + f_j \quad (j=1, 2, \dots, k), \quad (6.47)$$

где  $n_{01} = 0$ ,  $n_{k,k+1} = 0$ . Условие, что структура  $n$  сохраняема, заключается в том, что потоки людей должны удовлетворять неравенствам

$$0 \leq n_{j,j+1} \leq n_j - W_j \quad (j=1, 2, \dots, k-1). \quad (6.48)$$

Правая часть неравенства получается вследствие предположения, что потоки продвижения происходят после потоков увольнений; она отражает лишь тот факт, что численность людей, получивших повышение, не может быть больше, чем число мест в градации. Полагая  $n_j(T+1) = n_j(T) = n_j$  в (6.47), находим

$$n_{j,j+1} = \sum_{i=1}^j f_i - \sum_{i=1}^j W_i \quad (j=1, 2, \dots, k-1). \quad (6.49)$$

Из этого следует, что

$$P_M = Pr \left\{ 0 \leq \sum_{i=1}^j f_i - \sum_{i=1}^j W_i \leq n_j - W_j; \quad j=1, 2, \dots, k-1 \right\}. \quad (6.50)$$

Снова можно точно вычислить  $P_M$ , так как распределения используемых случайных величин полностью описываются марковской моделью. На практике эти вычисления громоздки, но можно воспользоваться аппроксимацией нормальным распределением, как было описано выше. Сосредоточим наше внимание на ситуации, когда все поступления осуществляются только на нижнюю градацию. В этом случае

$$f_1 = \sum_{i=1}^k W_i, \quad f_2 = f_3 = \dots = f_k = 0,$$

и, значит,

$$\begin{aligned} P_M &= Pr \left\{ 0 \leq \sum_{i=1}^k W_i - \sum_{i=1}^j W_i \leq n_j - W_j; \quad j=1, 2, \dots, k-1 \right\} = \\ &= Pr \left\{ \sum_{i=j}^k W_i \leq n_j; \quad j=1, 2, \dots, k-1 \right\}. \end{aligned} \quad (6.51)$$

Последний шаг непосредственно следует из тривиальности левой части неравенства. Так как  $W_i$  независимы и распределены по биномиальному закону, их суммы приблизительно нормальны с

$$E(Z_j) = \sum_{i=j}^k n_i w_i, \quad \text{var}(Z_j) = \sum_{i=j}^k n_i w_i (1 - w_i),$$

$$\text{corr}(Z_j, Z_h) = \left\{ \frac{\sum_{i=h}^k n_i w_i (1 - w_i)}{\sum_{i=j}^k n_i w_i (1 - w_i)} \right\}^{\frac{1}{2}}, \quad h > j,$$

где  $Z_j = \sum_{i=j}^k W_i$ . При нормальной аппроксимации в непрерывном случае получаем

$$\begin{aligned} P_M &\doteq Pr \left\{ u_j \geq \left( -n_j + \sum_{i=j}^k n_i w_i - \frac{1}{2} \right) \left/ \left( \sum_{i=j}^k n_i w_i (1 - w_i) \right)^{\frac{1}{2}} \right. ; \right. \\ &\quad \left. j=1, 2, \dots, k-1 \right\}, \end{aligned} \quad (6.52)$$

где  $u$  — переменные, распределенные по стандартному нормальному закону, имеющие те же коэффициенты корреляции, что и  $Z$ .

Если  $\mathbf{r} \neq (1, 0, \dots, 0)$ , то можно применить аналогичный подход, но область, в которой следует искать вид многомерного нормального распределения, более сложная. Важно различать, как нанятые распределяются по градациям — в заданных пропорциях или с заданными вероятностями. Пока нет теории, соответствующей последнему случаю, но предыдущий рассмотрен в работе Бартоломью (1979), где приведен пример, показывающий зависимость  $P_M$  от  $\mathbf{n}$ . Как и при управлении наймом, граница области  $M$  делит все структуры на те, для которых  $P_M \rightarrow 1$ , и на те, для которых  $P_M \rightarrow 0$ .

## Стратегии поддержания структур посредством найма

Дальнейший анализ дает только частное решение проблемы поддержания. Считая, что неравенства (6.43) выполняются, выбираем

$$f_j = n_j - \sum_{i=1}^k n_{ij} \quad (j=1, 2, \dots, k-1),$$

но на вопрос, что делать, если некоторые из  $f_j$  отрицательны, еще нужно ответить. Мы опишем два подхода к решению этой задачи.

При первом подходе мы откажемся от попытки сохранить структуру точно. Вместо этого только потребуем, чтобы ожидаемый вектор уходов был равен  $\mathbf{n}$  в любой момент времени. Для этого нужно взять

$$\mathbf{f} = \mathbf{n} (\mathbf{I} - \mathbf{P}), \quad (6.53)$$

равенство можно проверить подстановкой  $\mathbf{n}(0) = \mathbf{n}$  и  $R(T - \tau)\mathbf{p}_0 = \mathbf{f}$  в (3.2). (При данном подходе игнорируют тот факт, что  $\mathbf{f}$  должно состоять из целых величин. Это можно исправить, так случайно округляя компоненты, чтобы (6.53) было математическим ожиданием  $\mathbf{f}$ .) Поскольку структура более не является постоянной, полезно найти способ для оценивания, насколько возможны отклонения структур от  $\mathbf{n}$ . Это определяется с помощью дисперсий и ковариаций, которые следуют непосредственно из результатов, приведенных в гл. 3. Сейчас слагаемое  $n_{0j}(T+1)$  в (3.38) постоянно, и поэтому все другие слагаемые, возникающие вследствие отклонений от него, исключаются. Поэтому из результата (3.40) исключается  $\mu_0(T+1)$ . В рассматриваемом случае полезно выделить дисперсии и ковариации и сформировать из них квадратичную матрицу. Разностное уравнение для ковариационной матрицы  $\mathbf{V}(T)$  имеет вид

$$\mathbf{V}(T+1) = \mathbf{P}'\mathbf{V}(T)\mathbf{P} + [\bar{\mathbf{n}}(T)\mathbf{P}]_d - \mathbf{P}'[\bar{\mathbf{n}}(T)]_d\mathbf{P}, \quad (6.54)$$

где  $\mathbf{V}(0) = \mathbf{0}$ .

Преимущество рассматриваемой стратегии в том, что она может быть применена раньше, чем получены значения величин увольнения, но это также означает, что общий размер системы будет меняться. Если число уволившихся известно до того, как принято решение о новом найме, то одна из стратегий, обеспечивающих постоянные размеры

системы, заключается в том, чтобы сохранить общий размер и распределять принятых пропорционально увеличившимся. В этом случае

$$\mathbf{f} = \left( \sum_{i=1}^k W_i \right) \mathbf{r},$$

где  $\mathbf{r} = \mathbf{n} (\mathbf{I} - \mathbf{P}) / \mathbf{n}\mathbf{w}'$ . Можно показать, что такая стратегия является несмещенной в том смысле, что  $\bar{\mathbf{n}}(T) = \mathbf{n}$  для всех  $T$ , и, кроме того,

$$\mathbf{V}(T+1) = \mathbf{Q}' \mathbf{V}(T) \mathbf{Q} + [\bar{\mathbf{n}}(T) \mathbf{Q}]_d - \mathbf{Q}' [\bar{\mathbf{n}}(T)]_d \mathbf{Q} - \\ - \bar{\mathbf{n}}(T) \mathbf{w}' \{[\mathbf{r}]_d - \mathbf{r}' \mathbf{r}\}. \quad (6.55)$$

(см. работу Бартоломью, 1975). Можно ожидать, что такая стратегия отражает более жесткое управление с целью сохранения структуры.

Очевидная ограниченность «постоянной» стратегии состоит в том, что она не учитывает текущее состояние системы. В некоторых случаях можно было бы весьма точно поддерживать целевую структуру, но при постоянной стратегии это не получается. В такой ситуации естественно проанализировать адаптивные стратегии вдоль траекторий, которые приводили к достижению структур в детерминистской внешней среде. Наш второй подход заключается в движении вдоль этих траекторий. На каждом шаге мы стремимся находиться как можно ближе к целевой структуре. Пусть  $N_j = \sum_{i=1}^k n_{ij}$  — общая численность  $j$ -й градации после того, как осуществлены увольнения и продвижения. Очевидно, что структура может быть сохранена точно, если  $n_j \geq N_j$  для всех  $j$ . Когда не выполняется, по крайней мере, одно из неравенств, требуется некоторая модификация. Если мы принимаем  $f_j$  людей в  $j$ -ю градацию, то ее общая численность будет  $N_j + f_j$ . Тогда разумно выбирать такие значения  $f_j$ , чтобы минимизировать

$$\Phi = \sum_{j=1}^k (n_j - N_j - f_j)^2 \quad (6.56)$$

при условии  $\mathbf{f} \geq \mathbf{0}$ . Решение этой задачи очевидно, если взять

$$f_j = \max(0, n_j - N_j).$$

Такое решение означает, что каждая градация, численность которой ниже целевой, должна быть увеличена до необходимого уровня, численности градаций, большие, чем целевые, остаются без изменений. Рассматриваемая стратегия имеет один очевидный и серьезный недостаток. Она приводит к тому, что общие размеры системы обычно больше, чем требуется, и никогда не бывают меньше. Таким образом, она смещена в сторону роста. Можно гарантировать, что общие размеры системы останутся без изменения, если дополнить (6.56) ограничением

$$\sum_{i=1}^k f_j = \sum_{i=1}^k (n_j - N_j).$$

Задача минимизации, по существу, та же, что возникла для функции расстояния  $D_2$  в разделе 6.3; в ней все абсолютные величины заменены пропорциями.

Теперь остается сравнить постоянную и адаптивную стратегии. Процесс, определяемый двумя адаптивными стратегиями, описывается как марковская цепь, в которой множество состояний — все возможные структуры. Следовательно, в принципе могут быть получены такие величины, как средние и дисперсии. На практике число состояний системы реальных размеров так велико, что данный подход не применим. Сравнение основано на имитации. В табл. 6.3 содержатся средние значения векторов состояний, полученные при проведении 10 000 испытаний при использовании тех же матриц переходных вероятностей, что и в табл. 6.2. Когда структуры находятся в середине области  $\mathcal{M}$  и сохраняются при одинаковых наймах на все уровни, смещения в обоих случаях незначительны. Они существенно больше для структур, которые поддерживаются наймом только на нижний уровень. Однако ущерб становится больше выгоды, если сравнивать близость с целевой структурой, как сравнивалась ранее близость с постоянными стратегиями. Результаты этого сравнения представлены в табл. 6.4, где приводятся также средние квадраты ошибок ( $= \sum_{T=1}^{10000} (n_i(T) - n_i)^2 / 10000$  для каждого  $i$ ). Хотя нет ни одной равномерно лучшей стратегии среди всех рассматриваемых, ясно, что адаптивные стратегии предпочтительнее, особенно если целевая структура находится недалеко от центра области  $\mathcal{M}$ . Даже в крайних точках они достаточно хороши, в частности когда числа скрывают незначительность дисперсий, означающую, что вариация структур мала; большие средние квадраты ошибки почти везде возникают из-за смещений.

### Стратегии поддержания структур при управлении продвижением

Исследования, аналогичные только что приведенным, можно продолжить для управления продвижением. Снова возникает неопределенность при переходе от детерминистской постоянной стратегии к стохастической. Мы показали в (6.20), что вероятности поддержания, сохраняющие  $\mathbf{p}^*$ , должны удовлетворять уравнению

$$n_i^* p_{i,i+1} = \sum_{j=1}^i r_j + \sum_{h=1}^k n_h^* w_h - \sum_{j=1}^i n_j^* w_j \quad (j = 1, 2, \dots, k-1). \quad (6.57)$$

Правая часть показывает численность людей, продвинутых из  $i$ -й градации. Первый вариант ( $F_1$ ) постоянной стратегии заключается в том, что это число работников продвигается в каждый момент времени. Тогда получаем

$$n_{i,i+1} = \dot{n}_i p_{i,i+1} \quad (i = 1, 2, \dots, k-1). \quad (6.58)$$

Уравнение (6.57) может также рассматриваться, как будто оно дает долю людей из  $i$ -й градации, которые должны быть продвинуты. Поскольку должна быть сохранена точная структура, не имеет значения, рассматриваем мы относительные численности градации или абсолют-

**Т а б л и ц а 6.3. Средние структуры, поддерживаемые при реализации адаптивных стратегий в серии из 10 000 испытаний.**

**Матрица переходных вероятностей**

$$P = \begin{pmatrix} 0,7 & 0,2 & 0 \\ 0 & 0,8 & 0,1 \\ 0 & 0 & 0,9 \end{pmatrix}$$

	Сохраняется при найме на нижний уровень			Сохраняется при равном найме на все уровни		
	8	8	8	2	5	11
Целевая структура						
Произвольный размер	8,00	9,66	10,67	2,00	5,28	11,45
Постоянный размер	6,16	8,21	9,63	1,75	4,95	11,30
Целевая структура	36	36	36	12	30	66
Произвольный размер	36,00	39,25	41,52	12,00	30,09	66,18
Постоянный размер	32,16	36,54	39,30	11,88	30,00	66,12

**Т а б л и ц а 6.4. Средние квадраты ошибок численностей градаций при постоянных и адаптивных стратегиях в серии из 10 000 испытаний.**

**Матрица переходных вероятностей та же, что и в табл. 6.3**

		Сохраняется при найме на нижний уровень			Сохраняется при равном найме на каждый уровень		
		8	8	8	2	5	11
Целевая структура							
Постоянная стратегия	Произвольный размер	3,83	6,36	7,62	1,32	3,70	8,19
	Постоянный размер	5,30	5,07	5,10	1,25	2,97	3,68
Адаптивная стратегия	Произвольный размер	0,00	6,27	12,26	0,00	0,43	0,90
	Постоянный размер	5,42	2,38	5,97	0,31	0,59	0,71
Целевая структура		36	36	36	12	30	66
Постоянная стратегия	Произвольный размер	15,61	27,05	33,18	5,57	18,32	47,98
	Постоянный размер	24,29	24,37	24,45	6,37	16,74	19,10
Адаптивная стратегия	Произвольный размер	0,00	23,18	50,92	0,00	0,19	0,50
	Постоянный размер	23,66	10,81	25,06	0,17	0,30	0,46

ные Однако как только действительная структура отдалается от  $n^*$ , появляется отличие. В соответствии со второй интерпретацией численность людей, продвинутых из  $i$ -й градации, равна:

$$n_{i,i+1} = n_i p_{i,i+1} \quad (i = 1, 2, \dots, k-1), \quad (6.59)$$

где  $n$  — текущая структура. Назовем эту стратегию  $F_2$ . (Так как  $n_{i,i+1}$  должно быть целым, требуется применить процедуру округления.) На этом этапе возникает другая сложность. Если контролируемые потоки осуществляются прежде, чем сделаны все продвижения, то, возможно, не окажется достаточного числа тех, кого нужно продвигать. Это происходит всякий раз, когда  $n_{i,i+1}$ , определяемое (6.58) и (6.59), превышает  $n_i - W_i$ . Естественно изменить алгоритм так, чтобы продвигать всех, кто есть. Это, конечно, вносит смещение в стратегию, и по этой причине должны быть для последующих вычислений заданы средние значения численностей состояний. Можно было бы считать, с другой стороны, что вначале делаются все продвижения, в этом случае такой проблемы не возникает, хотя затем мы должны решить, может ли индивидуум, выбранный для продвижения, покинуть систему.

Адаптивная стратегия при управлении продвижением может быть сформулирована так, чтобы выбирать число продвигаемых при минимизации расстояния между имеющейся и желаемой структурами. Если, например, текущая структура —  $n$ , а целевая структура —  $n^*$ , то численность  $i$ -й градации после того, как все потоки осуществились, равна:

$$n_i - W_i + f_i - n_{i,i+1} + n_{i-1,i},$$

где  $n_i$ ,  $W_i$  и  $f_i$  заданы, а  $n_{i,i+1}$  и  $n_{i-1,i}$  — неотрицательные переменные, значения которых должны быть определены. Мы предлагаем сделать это, минимизируя

$$\Phi = \sum_{i=1}^k (n_i^* - n_i + W_i - f_i + n_{i,i+1} - n_{i-1,i})^2$$

при условии

$$n_{0,1} = n_{k,k+1} = 0, \quad 0 \leq n_{i,i+1} \leq n_i - W_i \quad (i = 1, 2, \dots, k-1).$$

Перед нами, таким образом, записанная в явном виде задача квадратичного программирования, которая может быть решена численным методом.

Для того чтобы продемонстрировать характеристики этих стратегий, рассмотрим результаты моделирования системы, представленные в табл. 6.5. Приведены три случая, относящиеся к системе численностью в 90 человек, в которой  $w = (0,1; 0,1; 0,1)$  и  $r = (1, 0, 0)$ . Первые две структуры настолько близки к вершинам области поддержания, насколько позволяют ограничения целочисленности; третья структура близка к середине области. Как видим, адаптивная стратегия заметно лучше всех, а  $F_2$  лучше, чем  $F_1$ . Эти выводы подтверждаются другими вычислениями при различных значениях  $w$  и  $r$ .

Т а б л и ц а 6.5. Характеристики стратегий по результатам 20 000 испытаний при  $w=(0,1; 0,1; 0,1)$  и  $g=(1, 0, 0)$

Стратегия		n*								
		(9, 9, 72)			(9, 81, 0)			(36, 32, 22)		
Средняя численность гра- дации	$F_1$	12,6	8,8	68,6	12,8	77,2	0	36,1	32,0	21,9
	$F_2$	9,4	9,1	71,6	9,4	80,6	0	36,0	32,0	22,0
	Адаптивная	10,3	9,1	70,5	10,1	79,9	0	36,0	32,0	22,0
Средний квадрат оши- бки	$F_1$	31,8	2,9	34,6	33,6	33,6	0	27,6	18,0	11,3
	$F_2$	7,9	7,2	13,0	7,9	7,9	0	11,5	7,2	9,6
	Адаптивная	4,8	0,9	6,2	4,4	4,4	0	0	0	0

### 6.5. ДОСТИЖИМОСТЬ В СЛУЧАЙНЫХ СРЕДАХ

В предыдущем разделе мы в неявном виде уже касались достижимости, и нам мало остается, что добавить. Различные стратегии, предлагаемые для поддержания структуры, могут быть применены для любой начальной точки, из которой они позволяют достичь окрестности целевой структуры. Следует напомнить, что адаптивные стратегии первоначально были сформулированы как особые случаи  $T$ -шаговых стратегий для исследования управления на конечном интервале времени в детерминистской внешней среде. Хотя можно сформулировать стохастические варианты общей задачи управления с временными ограничениями и без них, это едва ли представляется целесообразным, так как одношаговые стратегии более просты и эффективны.

Мы не будем больше сравнивать характеристики различных стратегий по средним значениям и средним квадратам ошибок. Далее будем интересоваться скоростью и направлением движения к целевой структуре. На рис. 6.4 и 6.5 показано, как можно изобразить скорости и направления. Они соответствуют случаю управления наймом в трехуровневой системе при

$$P = \begin{pmatrix} 0,7 & 0,2 & 0 \\ 0 & 0,8 & 0,1 \\ 0 & 0 & 0,9 \end{pmatrix}.$$

Задача состоит в том, чтобы достичь целевую структуру (6, 15, 33) из точки (54, 0, 0).

На рис. 6.4 представлены шесть вариантов действия постоянных стратегий в системе «заданного размера», а на рис. 6.5 аналогичные результаты приведены для адаптивной стратегии, при которой минимизируется квадратичная функция расстояния. Как и можно было ожидать, адаптивная стратегия дает более прямой путь к цели и различные варианты пути ближе друг к другу. Очевидно, что при адаптивной стратегии осуществляется более жесткое управление.



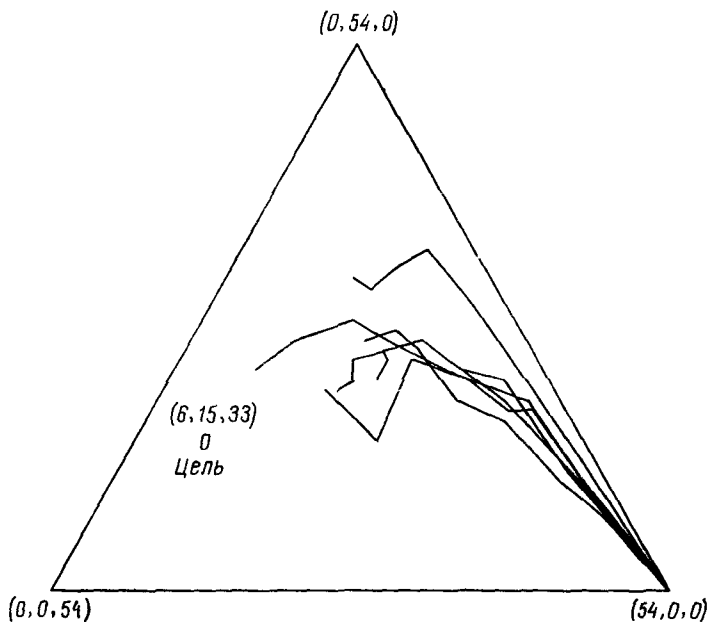


Рис. 6.4. Шесть попыток имитировать достижение цели (6, 15, 33) из точки (54, 0, 0) при постоянной стратегии

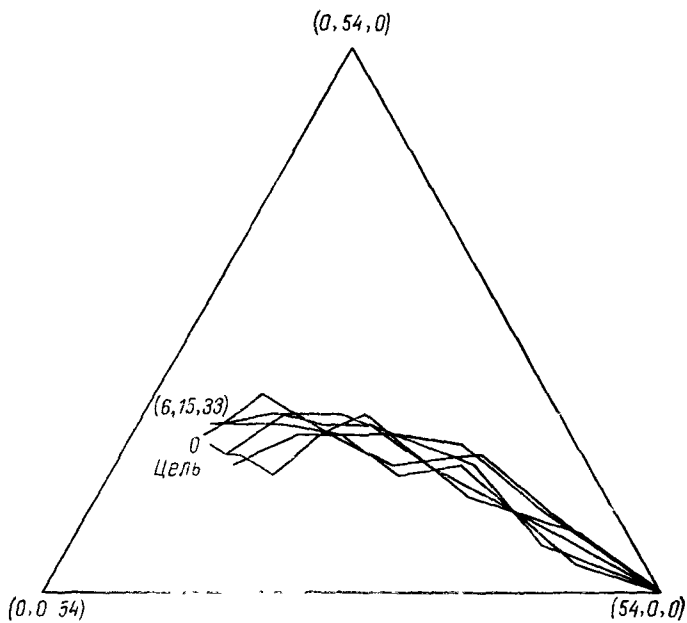


Рис. 6.5. Шесть попыток имитировать достижение цели (6, 15, 33) из точки (54, 0, 0) при адаптивной стратегии

Теория управления — довольно широкое понятие, но существующие работы относятся к техническим приложениям, где связь между переменными выражается с помощью дифференциальных уравнений. Поэтому очень мало литературы с описанием теории, применимой к нашим задачам. Две основные работы по стохастическому управлению — это работы Кушнера (1967, 1971). Санон и его соавторы (1970) излагают теорию управления таким же образом, как в настоящей главе, но в более общем виде. Фактически детерминистские аспекты могли бы рассматриваться как приложение их теории, хотя результаты, приведенные здесь, получены раньше, чем вышла их книга. Мы применили кое-что из их терминологии. В исследованиях в области общественных наук Тинтнер и Сенгупта (1972) обсуждали некоторые задачи управления и математического программирования применительно к стохастическому подходу в экономике. Использование методов математического программирования для управления марковскими цепями изучалось Кушнером и Клейнманом (1971).

Исследование управления иерархическими системами первоначально было приведено в небольшом разделе в конце третьей главы первого издания настоящей книги. Оно было последовательно развито в работах Бартоломью и Форбеса (1971 с), Дэвиса (1973), Вайды (1975), а также других авторов и было представлено целой главой во втором издании. В настоящем издании, чтобы отвести место для новых результатов в теории управления в стохастических средах, опущена большая часть материалов по достижимости, которая имеет меньшее значение для практики. Управление в стохастических средах рассматривается в статьях Бартоломью (1977b, 1979).

Развитие теории управления в детерминистской среде шло в нескольких направлениях. Дэвис (1975) продолжал изучение  $n$ -шагового поддержания. Структура называется поддерживаемой за  $n$  шагов, если существует последовательность таких значений управляющих параметров, при которых система возвращается к целевой структуре через каждые  $n$  шагов. Поддержание, рассмотренное здесь, — одношаговое. Область  $n$ -шагового поддержания ( $n > 1$ ) содержит в себе область при  $n = 1$ , поэтому в этом более широком смысле может быть сохранено большее множество структур. Гринольд и Маршалл (1977) также обсуждали проблемы поддержания и достижимости. Вайда (1975 и 1979) детально обсудил развитие задач управления, описанных в настоящей главе. Он ввел понятие более слабого вида поддержания, согласно которому может быть поставлена задача обеспечить сохранность общей численности только некоторых градаций. У Бартоломью и Форбеса (1979) есть примеры планирования кадров.

Наше краткое изложение теории поддержания при непрерывном времени основано на работе Хассани (1980). Не все задачи с непрерывным временем могут быть сведены соответствующим образом к дискретной модели. При непрерывном времени управление может осуществляться через дискретные промежутки времени или в произвольные моменты. Ясно, что здесь требуются дальнейшие исследования.

Для вычисления  $P_M$  при нормальной аппроксимации требуется программа вычисления объемов под поверхностью многомерного нормального распределения. Мы считаем, что метод Милтона (1972) вполне удовлетворительный, хотя, может быть, и трудоемкий.

Наши исследования по управлению в стохастических средах представлены в сжатом виде, и вычисления носят чисто иллюстративный характер. Более детально эти проблемы изучаются в работах Бартоломью (1977b, 1979), где численные примеры иллюстрируют и другие политики. В частности, обсуждаются стратегии управления, при которых не нужно знать контролируемые потоки. Дополнительные результаты представлены в графическом виде в работе Бартоломью (1976 c). Точное выражение для ковариационной матрицы системы фиксированного размера вместе с некоторыми другими примерами приведено в работе Бартоломью (1975).

Сейчас ведутся работы по применению методов математического программирования для определения оптимальных стратегий в решении задач планирования образования и кадров. Если потоки людей пропорциональны численностям состояний, то эта задача почти та же, что сформулированная здесь для детерминистского случая. При изучении структур и вычислительных аспектов о выпуклости известно больше. Роу и его соавторы (1970) сформулировали задачу управления, в которой требуется минимизировать дисконтированную функцию стоимости на заданном горизонте планирования. Чарнс, Купер, Нихаус и другие разработали теорию, основанную на методах целевого программирования применительно к вложенным марковским цепям. Их подход аналогичен нашему при решении задачи управления с временными ограничениями планирования, но в их модели учитываются и различные бюджетные ограничения. Во всех этих работах внешняя среда детерминистская. Ранняя статья, которая легла в основу данной главы, повторяется в работе Бартоломью (1976b) и в сборнике ранних статей Чарнса с соавторами (1970). Морган (1970) также сформулировал задачу управления как задачу линейного программирования, в которой переменными являются число нанятых и число продвижений, а также определяются затраты на найм новых работников, на сокращение штатов, на излишки штатных единиц. В следующей статье Морган (1971) рассмотрел вопросы управления возрастной структурой, в которой вместо ступеней брались возрастные группы. Он предложил стратегию, аналогичную той, что обсуждалась в разделе 6.3.

Мелман (1980) предлагает метод решения детерминистской задачи управления методом динамического программирования. В его модели разрешается, чтобы управляющими переменными были вектор поступления и матрица переходов, но, кроме того требуется, чтобы их значения были как можно ближе к предпочитаемым величинам.

Статистические основания для этой главы минимальны. С математической точки зрения теория математического программирования и элементы теории выпуклых множеств достаточны для уяснения материала главы. Сведения по этим разделам математики известны.

### 7.1. ПРЕДПОСЫЛКИ

Во многих социальных приложениях продолжительность некоторой деятельности и размер некоторой системы представляют самостоятельный интерес. Так, например, продолжительность человеческой жизни, время пребывания в браке или интервалы между последовательными рожденьями детей в семье — это сфера интересов демографов. Продолжительность жизни пациента после лечения, скажем, заболевания раком, часто — мера эффективности лечения. Продолжительность забастовки, время существования фирмы, время, которое индивидуум работает на одном месте, проживает по определенному адресу, отсутствует на работе, является безработным — все это область социальных и экономических приложений. Аналогично уровень благосостояния и доходы являются ключевыми переменными при изучении социальной дифференциации, а размеры фирм и городов — это характеристики, интересующие экономистов и географов. Все эти самые разнообразные величины имеют две общие черты, позволяющие применить для их изучения метод стохастического моделирования. Во-первых, они случайны по своей природе. Во-вторых, практический опыт показывает, что частотные распределения многих из этих переменных могут быть представлены довольно простыми кривыми, стабильными во времени. Такая регулярность довольно нетипична для общественных наук, и когда она обнаруживается, весьма желательно попытаться узнать, отчего это происходит. Такие исследования необходимы не только для получения некоторых сведений о скрытых мотивах человеческого поведения, но также и для того, чтобы дать фирмам основы статистического анализа.

Широкое распространение различных видов частотных распределений может быть обусловлено несколькими причинами. Например, нормальные распределения часто встречаются потому, что значения многих переменных — результат влияния большого числа независимых переменных.<sup>5</sup> Тогда в соответствии с центральной предельной теоремой их распределения аппроксимируются нормальными. Аналогично распределение экстремальных значений имеет форму, которая в довольно широких пределах не зависит от распределений самих случайных величин. Аналогично этому хорошо известны семейства рас-

пределений, замкнутые по отношению к операциям свертки или при смешивании. Это означает, например, что если взять смесь двух членов семейства, то результирующее распределение также будет принадлежать к тому же семейству. Третье объяснение постоянства вида распределений следует из того, что социальные процессы рассматриваются как стохастические процессы. В предыдущих главах часто отмечалось, что процессы сходятся к устойчивым состояниям, в которых, например, структура групп достигает стабильного, динамического равновесия. Если бы группы в таких процессах состояли из численностей категорий (т. е. исходных подгрупп), то распределение в устойчивом состоянии можно было бы интерпретировать как распределение их частот. Тогда устойчивость определенного вида распределений можно было бы объяснить предположением, что все исходные процессы достигли устойчивых состояний.

Мы рассмотрели модели, относящиеся к каждому из перечисленных случаев. Сначала в разделе 7.2 будут представлены модели распределений времени пребывания в состоянии, относящиеся к ряду приложений. Для описания большинства из них используются стохастические процессы, которые уже встречались в других разделах. Раздел 7.3 переходный, он посвящен логарифмически нормальному распределению для описания как продолжительности пребывания в состоянии, так и размеров системы. В разделе 7.4 обсуждаются модели распределения размеров системы. Продолжительность пребывания в состоянии может, конечно, рассматриваться как размер временного интервала (см. раздел 7.3). Все эти модели могут подвергаться критике вследствие того, что они слишком упрощают жизненные явления. В разделе «Дополнения» освещаются различные направления развития теории с целью дальнейшего приближения к практике. Мы снова хотели бы подчеркнуть, что упрощение — не обязательно недостаток. Выделяя сначала количественные характеристики процесса, мы изучаем суть явления и строим основную модель, а ее в дальнейшем можно уточнить.

## 7.2. МОДЕЛИ ВРЕМЕНИ ПРЕБЫВАНИЯ В СОСТОЯНИИ

Время пребывания в состоянии изучалось не только в общественных науках. Наиболее известные и развитые приложения — теория надежности и технические испытания. Существует много общего между работами в этих областях и тем, о чем пойдет речь в настоящей главе. Однако различие между социальными и техническими приложениями весьма существенно. При технических испытаниях центральную роль играет экспоненциальное распределение, так как вероятность отказа многих типов оборудования почти не зависит от срока службы. Там, где не подходит экспоненциальное распределение, часто используются распределение Вейбулла и гамма-распределение. Ни одно из этих распределений не играет существенной роли в приложениях, с которыми мы будем иметь дело, хотя в следующем разделе мы обсудим экспоненциальное распределение. Другими словами, типы моделей, удовлетворительные при описании времени жизни техниче-

ских элементов, редко подходят для описания времени пребывания в состояниях социальных систем. Мы начнем с моделирования продолжительности занятости человека (полного срока службы, ПСС), но осветим и другие приложения.

Нам потребуются следующие термины и обозначения. Пусть  $T$  — продолжительность пребывания в интересующем нас состоянии; мы не различаем при обозначении, является эта величина случайной или ее реализацией;  $T$  — непрерывная величина. Функция распределения  $T$  — главный объект изучения. Сформируем три эквивалентные формы ее представления.

а) *Функция дожития*. Это вероятность того, что индивидuum пробыдет в данном состоянии интервал времени  $T$ . Обозначим ее  $G(T)$ . Функция дожития в сумме с функцией распределения  $F(T)$  дает 1, но  $G(T)$  более подходит для практических целей.

б) *Плотность функции полного срока службы (ПСС)*. Плотность функции  $F(T)$  связана с функцией дожития следующим соотношением:

$$f(T) = -\frac{dG(T)}{dT}.$$

в) *Интенсивность (или скорость) уходов (или потерь)*. Она обозначается  $\lambda(T)$  и определяется следующим образом:

$Pr \{ \text{уход на интервале } (T, T + \delta T) \mid \text{в момент времени } T \text{ находился в системе} \} = \lambda(T) \delta T$ .

Эта функция в различных приложениях называется по-разному. В теории надежности и технических испытаниях она известна как опасность отказов, функция риска, повозрастная интенсивность выхода из строя или, более просто, интенсивность выхода элементов из строя. При страховании ее называют силой или коэффициентом смертности, а в кадровых приложениях образно определяют как силу или стремление к отделению или склонность к уходу из системы. В гл. 4 и 5 она определена как интенсивность уходов, и здесь мы будем употреблять ту же терминологию. Соотношения между  $\lambda(T)$ ,  $G(T)$  и  $f(T)$  легко выводятся:

$$\begin{aligned} f(T) \delta(T) &= Pr \{ \text{уход за } (T, T + \delta T) \} \\ &= Pr \{ \text{прожить до } T \} Pr \{ \text{уйти за } (T, T + \delta T) \mid \\ &\quad \text{жить до } T \} \\ &= G(T) \lambda(T) \delta T. \end{aligned}$$

Следовательно,

$$\lambda(T) = f(T)/G(T) = -\frac{d \ln G(T)}{dT} \quad (7.1)$$

и, наоборот,

$$G(T) = \exp \left[ -\int_0^T \lambda(x) dx \right]. \quad (7.2)$$

В приложениях к описаниям увольнений или потерь рабочей силы функция  $\lambda(T)$  четко отражает процесс уходов, позволяя по конкретным срокам работы определить моменты, когда интенсивности уходов особенно велики или малы.

На практике иногда проще получить оценку либо для одной, либо для другой функции, но мы можем легко перейти от одного вида к другому с помощью приведенных выше соотношений. В большинстве случаев при построении моделей основное внимание уделяют функциям  $f(T)$  и  $G(T)$ , но интерпретировать модель лучше, используя  $\lambda(T)$ . Данные для оценки этих функций редко представлены в виде простой случайной выборки. Чаще всего они состоят из цензурированных нижней границей сроков работы когорты или результатов переписи, полученных за короткие интервалы времени, относительно численностей групп и потоков между ними.

### Экспоненциальное и смешанное экспоненциальное распределения

Рис и его соавторы (1950) в своей работе привели несколько эмпирических распределений ПСС, полученных при исследовании Глазнерской металлургической компании. Они заметили, что некоторые из распределений можно представить в виде J-образных кривых гиперболической формы, но они не предложили какой-либо теоретической модели для объяснения этого явления. Силкок (1954) пересмотрел литературу, в которой были данные о текучести, и предложил две модели, соответствующие распределениям ПСС. Его первая модель была основана на замечании Риса и его соавторов (1950) о том, что они обнаружили регулярность в текучести, которая является характерной особенностью фирмы и не зависит от экономических и социальных факторов, действующих вне фирмы. Силкок свел это к тому, что принял для простоты постоянными интенсивности уходов, и показал, что это в свою очередь ведет к экспоненциальному распределению ПСС. Весьма сомнительно, что интерпретация Силкока оправдана, так как интенсивность уходов зависит от действующих внутри фирмы факторов, включая стаж работы индивидуума. Тем не менее экспоненциальный случай позволяет получить пробный вариант описания процесса уходов и прокладывает дорогу более точным моделям. Экспоненциальное распределение Силкок применил для описания нескольких распределений ПСС и каждый раз согласие было слабым. В табл. 7.1 приведены два примера, из которых видно, что распределение наблюдаемых значений всегда более асимметрично, чем соответствующее экспоненциальное распределение. По наблюдениям в течение 21 месяца невозможно узнать, какой вид будет у оставшейся части последовательности, но из приведенных данных ясно, что гипотеза об экспоненциальности несправедлива. Опыт Силкока подтверждается почти всеми последующими исследованиями.

Вторая модель, предложенная Силкоком (1954), обобщает первую. В ней также остается упрощающее предположение о постоянстве интенсивности уходов индивидуумов, но предполагается, что она своя для каждого индивидуума в популяции, из которой осуществляются

Таблица 7.1. Наблюдаемые и аппроксимированные распределения полного срока службы для двух фирм

Полный срок службы, месяцев	Гласнерская металлургическая компания				Фирма «Бибби и сыновья»			
	Реальное число ушедших	Экспоненциальная модель	Модель типа XI	Модель смешанного экспоненциального распределения	Реальное число ушедших	Экспоненциальная модель	Модель типа XI	Модель смешанного экспоненциального распределения
До 3	242	160,2	242,0	242,0*	182	103,9	195,4	182,0*
3	152	138,9	150,3	152,0*	103	86,8	87,5	103,0*
6	104	120,4	103,8	101,4	60	72,4	51,8	60,7
9	73	104,5	76,5	72,7	29	60,5	35,0	38,0
12	52	90,6	59,2	55,8	31	50,5	25,6	25,5
15	47	78,5	47,4	45,7	23	42,1	19,7	18,6
18	49	68,1	38,8	39,2	10	35,2	15,8	14,7
21 и больше	487	444,8	488,0	497,2	191	177,6	198,2	186,5
Всего	1206	1206,0	1206,0	1206,0	629	629,0	629,0	629,0

\* Цифры, отмеченные звездочкой, хорошо согласуются с реальными данными, так как в этих точках вычислялись коэффициенты моделей. Подробнее см. работу Бартоломью (1959).

уходы. Это очень плодотворная гипотеза. Люди отличаются друг от друга во всех аспектах поведения, и было бы удивительным, если бы их склонность к уходу с работы была исключением. Обозначим постоянную интенсивность уходов через  $\lambda$ , тогда для любого индивидуума плотность распределения будет следующей:

$$f(T) = \lambda e^{-\lambda T} \quad (\lambda > 0; T \geq 0).$$

Предположим теперь, что  $\lambda$  — случайная величина с функцией распределения  $H(\lambda)$ . Тогда распределение ПСС для выборки из всех занятых будет иметь плотность

$$f(T) = \int_0^{\infty} \lambda e^{-\lambda T} dH(\lambda) \quad (T \geq 0). \quad (7.3)$$

Это распределение всегда более скошенное, чем экспоненциальное с постоянным средним значением, и, следовательно, оно обладает той же особенностью, что и исходные данные. Частичное объяснение этого можно найти, если рассмотреть отношение

$$r = f(T) / \mu^{-1} e^{-T/\mu}, \quad (7.4)$$

где

$$\mu = \int_0^{\infty} T f(T) dT = \int_0^{\infty} \frac{1}{\lambda} dH(\lambda) = E\left(\frac{1}{\lambda}\right).$$



Для малых  $T$

$$r \rightarrow E(\lambda) E\left(\frac{1}{\lambda}\right) \geq 1,$$

так как арифметическое среднее всегда, по крайней мере, меньше, чем геометрическое среднее. Равенство имеет место только в вырожденном случае, когда  $H(\lambda)$  — функция, сосредоточенная в одной точке. Для больших  $T$  запишем:

$$r = \mu \int_0^{\infty} \lambda e^{-(\lambda-1/\mu)T} dH(\lambda).$$

При  $\lambda > \mu^{-1}$  значение интеграла стремится к нулю, при  $\lambda < \mu^{-1}$  оно стремится к бесконечности. Как только распределение  $H(\lambda)$  таково, что оно дает отличное от 0 или 1 значение вероятностей тех  $\lambda$ , которые больше  $\mu^{-1}$ , то  $r \rightarrow \infty$ . Получена другая граница рассматриваемого случая, если только  $H(\lambda)$  не вырождается в точку.

Для того чтобы (7.3) согласовать с данными, нужно взять конкретное  $H(\lambda)$ . Силкок (1954) выбирает для нее гамма-распределение с плотностью

$$\frac{dH(\lambda)}{d\lambda} = \frac{c^\nu}{\Gamma(\nu)} \lambda^{\nu-1} e^{-c\lambda} \quad (\nu > 0; c > 0; \lambda \geq 0)$$

и показывает, что

$$f(T) = \frac{\nu}{c} \left(1 + \frac{T}{c}\right)^{-(\nu+1)} \quad (T \geq 0). \quad (7.5)$$

Соответствующие выражения для функции дожития и интенсивности потерь имеют вид

$$G(T) = (1 + T/c)^{-\nu} \quad \text{и} \quad \lambda(T) = \frac{\nu}{c} (1 + T/c)^{-1}. \quad (7.6)$$

Это J-образное распределение из семейства Пирсона, которое классифицируется как тип XI. Функция интенсивности потерь имеет очень простой вид, она показывает монотонное уменьшение склонности к уходу в зависимости от увеличения времени пребывания в состоянии. С первого взгляда может показаться парадоксальным, что уменьшение  $\lambda(T)$  получилось при принятии постоянной интенсивности уходов для индивидуумов. Однако в группе индивидуумов, которую мы рассматриваем, не все имеют одну и ту же интенсивность ухода. Чем дольше индивидуумы группы работали в организации, тем более вероятно, что у них будет меньшее значение  $\lambda$ . Таким образом, получаем соответствующее уменьшение  $\lambda$  с течением времени. Когда Силкок (1954) описал имеющиеся данные этим распределением, он получил гораздо лучшее совпадение. Некоторые результаты вычислений Силкока приведены в табл. 7.1 (модель типа XI).

Гамма-распределение  $\lambda$ , использованное выше, достаточно гибко, так как оно зависит от двух параметров. Его главное преимущество при применении заключается в возможности проводить анализ с по-

мощью математических методов. Однако частный вид функции, который рассматривался здесь, не является определяющим. Точное выражение для  $f(T)$  можно получить еще, используя сопряженное\* гамма-распределение. Представляет интерес случай

$$\frac{dH(\lambda)}{d\lambda} = \frac{c^\rho}{\Gamma(\rho)\Gamma(1-\rho)} \lambda^{-1} (\lambda - c)^{-\rho} \quad (\lambda \geq c; 0 < \rho < 1), \quad (7.7)$$

когда  $f(T)$  сама становится J-образным гамма-распределением. Для некоторых случаев удобнее положить, что  $\lambda$  принимает только два значения, скажем,  $\lambda_1$  и  $\lambda_2$ . Если соответствующие вероятности равны  $p$  и  $1 - p$ , то распределение ПСС будет иметь вид

$$f(T) = p\lambda_1 e^{-\lambda_1 T} + (1-p)\lambda_2 e^{-\lambda_2 T} \quad (0 < p < 1, \lambda_1, \lambda_2 > 0; T \geq 0). \quad (7.8)$$

Эта двучленная смесь экспоненциальных распределений рассмотрена в гл. 4. Она была применена Бартоломью (1959) для исходных данных Силкока и позволила получить результаты, приведенные в табл. 7.1. Разницу между смесью экспоненциальных распределений и распределением типа XI в этих двух примерах трудно обнаружить, несмотря на радикальное отличие вида функции  $H(\lambda)$  в этих двух случаях.

Дальнейшие исследования не позволяют заключить, что «смешанные» модели, предложенные Силкоком (1954), дают верное описание процесса уходов. Существуют другие модели, которые дают лучшее согласование с исходными данными, и, по крайней мере, одна из них позволяет получить тот же вид функции  $f(T)$ . Если модель отражает хоть в какой-то степени истинное положение вещей, то ее уже можно применять для управления персоналом. Так, согласно модели существует естественная разница между индивидуумами, и способ управления процессом увольнений заключается в том, что все работники классифицируются по группам. Если мы могли определить людей с меньшими или большими значениями  $\lambda$ , то заработную плату можно было бы уменьшить или увеличить в соответствии с этой классификацией. В модели предполагается, что у нас нет эффективных методов классификации людей по их склонности к уходу из системы. Это задача психологов, но успех их методов должен быть, конечно, проверен статистически.

Ясно, что двучленная смесь экспоненциальных распределений может быть заменена распределением с тремя и более членами, но число параметров при этом увеличивается очень быстро, что затрудняет применение, так как для определения параметров требуется большое число данных. В следующем разделе и в гл. 8 мы встретимся с распределением вида

$$f(T) = \sum_{i=1}^k p_i \lambda_i e^{-\lambda_i T} \quad \left( T \geq 0; \sum_{i=1}^k p_i = 1 \right). \quad (7.9)$$

\* В оригинале использован термин «inverse», что в переводе означает «обратное», но в данном случае лучше применить ранее встречавшийся в нашей литературе термин «сопряженное». Тогда смесь гамма-распределений с сопряженными весами (7.7) дает снова гамма-распределение. — *Примеч. ред.*

Если  $p_i$  неотрицательны, то оно принадлежит к классу распределений (7.3), но это не обязательно. Когда модели разного рода приводят к одному и тому же распределению ПСС с положительными  $p_i$ , то не существует статистических методов их различия, использующих исходные данные только о сроках службы.

## Сетевые модели

Интересный класс моделей процесса уходов из системы может быть получен при использовании марковского процесса, описывающего переходы индивидуумов из одного внутреннего (гипотетического) состояния в другое перед тем, как покинуть систему. Хербст (1963) первым предложил этот подход с применением так называемой модели процесса принятия решения. Аналогичные модели, хотя и с различными интерпретациями, были рассмотрены Хоемом (1971) для непрерывного времени, а Янгом (1971) и Файхтингером (1971) для дискретного времени. Весьма схожая модель в приложении к описанию процесса выживания после лечения заболевания раком обсуждалась в гл. 5, но там не изучалось время жизни. Особенность модели Хербста состоит в том, что в ней промежуточные состояния скрыты, т. е. непосредственно их наблюдать нельзя. Хербст убедительно показывает,

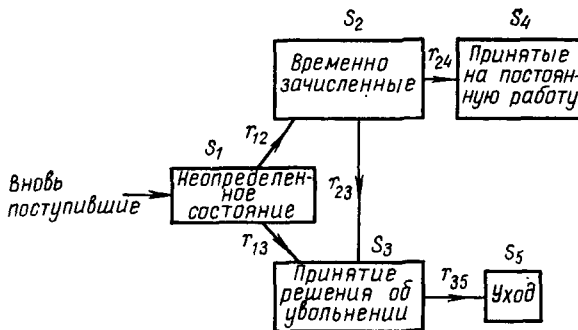


Рис. 7.1. Возможные состояния марковского процесса для модели, описывающей процесс уходов

почему состояния, которые им определены в модели, соответствуют тому, что происходит в действительности, но об их существовании можно судить только косвенно по форме наблюдаемого распределения ПСС.

Модель легко объяснить, если обратиться к рис. 7.1. Каждый индивидуум по предположению поступает в систему в непредрешенное состояние, а затем проходит путь по сети до одного из поглощающих состояний  $S_4$  и  $S_5$ . Все, что мы можем наблюдать на практике, это время, затраченное на достижение состояния  $S_5$ , обо всех остальных качествах системы можно судить только по виду распределения ПСС. Движение по сети характеризуется множеством интенсивностей переходов, обозначенных на рисунке через  $r$ . Таким образом, модель

системы — это марковский процесс с непрерывным временем (см. гл. 4) с матрицей интенсивностей переходов.

$$\mathbf{R} = \begin{pmatrix} r_{11} & r_{12} & r_{13} & 0 & 0 \\ 0 & r_{22} & r_{23} & r_{24} & 0 \\ 0 & 0 & r_{33} & 0 & r_{35} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad (7.10)$$

где  $r_{11} = -r_{12} - r_{13}$ ,  $r_{22} = -r_{23} - r_{24}$ ,  $r_{33} = -r_{35}$ . Так как матрица треугольного вида, ее собственные значения равны ее диагональным элементам.  $\bar{n}_i(T)$  из (4.16) можно интерпретировать как вероятности пребывания в  $i$ -м состоянии в момент времени  $T$ , если в начальный момент времени один индивидуум находится в первом состоянии. Тогда функция дожития определяется выражением

$$1 - \bar{n}_5(T) = G(T) = d_{51} + c_{53} e^{r_{11}T} + c_{54} e^{r_{22}T} + c_{55} e^{r_{33}T}. \quad (7.11)$$

Хербст поставил вопрос, дает ли предлагаемая модель адекватное описание переходам в организационной системе. Далее он сравнил (7.11) с данными Хедберга (1961) о распределении времени пребывания по двум фирмам. В фирме  $A$  уход в единицу времени больше, чем в фирме  $B$ . Необходимо оценить семь параметров, и это очень много, если исходные данные не представляют собой большую совокупность. К счастью, Хербст располагал очень большой выборкой, и по точкам были получены оценки для параметров. Оцененные функции имели вид:

Фирма  $A$ :

$$G(T) = 0,1493 + 0,4544e^{-0,4270T} + 0,4194e^{-0,0966T} - \\ - 0,0231e^{-10,15T}. \quad (7.12)$$

Фирма  $B$ :

$$G(T) = 0,4700 + 0,1274e^{-0,3680T} + 0,4025e^{-0,1142T}. \quad (7.13)$$

Для фирмы  $B$  —  $r_{33}$  практически равен бесконечности. Наблюдаемые и полученные по модели значения  $G(T)$  приведены в табл. 7.2. Сходство для такой большой выборки замечательное, и отсюда следует, что теория Хербста достаточно удачна. Распределения могут быть получены и как смесь экспоненциальных распределений. (Правда, в этом случае имеется отрицательный коэффициент при описании функции дожития для фирмы  $A$ , но его влияние незначительно.) Заметим, что первое слагаемое в одном и другом выражениях не зависит от  $T$ , поэтому можно считать, что функция дожития порождается экспоненциальным распределением с бесконечным средним значением. Оно отражает долю людей, которые в конечном счете достигают категории окончательного постоянства.

Очевидно, что задача статистической оценки параметров требует дальнейших исследований. Свердруп (1965) и Хоем (1971) рассмотрели случай, когда исходные данные представляют собой значения переходов индивидуумов. Если применить эти результаты к приведенно-

Т а б л и ц а 7.2. Наблюдаемые значения и оценки по выражению  $100 \{1 - \bar{p}_5 (T)\}$  (данные Хедберга)

Месяц	Фирма А 7628 поступающих		Фирма В 968 поступающих	
	действительные значения	теоретические значения	действительные значения	теоретические значения
0	100,00	100,00	100,0	100,0
1	82,66	82,66	91,7	91,7
2	68,85	68,85	85,1	85,1
3	58,95	58,94	79,8	79,8
4	52,10	51,66	75,6	75,4
5	46,61	46,17	71,7	71,8
6	42,32	41,92	67,7	68,7
7	38,73	38,53	64,9	66,1
8	35,74	35,77	63,0	63,8
9	33,52	33,46	61,3	61,9
10	31,78	31,56	59,5	60,2
11	30,19	29,81	58,5	58,7
12	28,46	28,33	57,2	57,4
13	27,20	27,03	55,9	56,2
14	25,85	25,87	55,4	55,2
15	24,79	24,83	54,8	54,3
18	22,25	22,30	53,1	52,2
21	20,36	20,43	51,4	50,7
24	19,06	19,04	49,7	49,6
27	18,01	18,00	48,8	48,9
30	17,31	17,23	48,2	48,3
36	16,45	16,22	47,7	47,7
42	15,77	15,65	47,4	47,3
48	15,35	15,33	47,2	47,2
54	15,18	15,16	47,1	47,1
60	15,05	15,06	—	—
66	15,00	15,00	—	—

му примеру, то мы увидим, что совпадение настолько хорошее, что более эффективные процедуры были бы излишни.

Описанный выше метод получения распределения ПСС весьма общий. Он применим, когда  $\mathbf{P}$  — матрица любой размерности.

До сих пор наше внимание было сосредоточено на моделях, которые описывают системы с единственным выходом. Имеется много приложений, когда выходов несколько, например, когда увольняющиеся классифицируются по причинам увольнения. В терминах страхового дела в проблему вписываются убыли (смерть) по разным причинам. Если нас интересует общее время пребывания в системе независимо от причин ухода, то требуется только определить все поглощающие, или конечные, состояния, не различая их. С другой стороны, если можно классифицировать увольняющихся по причинам увольнения, то мы получим распределения длительности пребывания в системе по соответствующим группам.

Чтобы проиллюстрировать предлагаемый метод, предположим, что состояние 4 в модели Хербста называется «объявленные лишними из-за сокращения штатов» вместо «принятые на постоянную работу». Тог-

да мы могли бы рассмотреть два условных распределения времени пребывания в системе для тех, кого увольняют по сокращению штатов, и тех, кто уходит по собственному желанию. Для тех, кто уходит по собственному желанию, существует простое выражение условной вероятности, позволяющее получить следующую функцию дожития:

$$G_{\alpha}(T) = \left\{ G(T) - \frac{r_{12} r_{24}}{r_{11} r_{22}} \right\} / \left( 1 - \frac{r_{12} r_{24}}{r_{11} r_{22}} \right), \quad (7.14)$$

где  $(1 - r_{12} r_{24} / r_{11} r_{22})$  — вероятность уходов в состоянии 5.

### Модели случайного блуждания

Если рассматривать время как случайные шаги в направлении достижения поглощающего экрана, то мы получим другой тип модели. Идея модели была впервые высказана Ланкастером (1972), который предложил базовую модель для описания продолжительности забастовки. Мы возьмем за основу это приложение, ссылаясь, если потребуется, на другие.

Забастовки начинаются как последнее средство, когда конфликтующие стороны не могут прийти к соглашению. Вначале допустим, что возможности проведения забастовки ограничены некоторой «величиной  $d$ ». Можно считать, что эта величина — сумма денег, которая должна быть выплачена, когда конфликт прекращается. Однако предмет конфликта может быть и более сложным. По мере того, как забастовка развивается, расхождение между позициями сторон меняется, так как начинают влиять новые факторы; затруднения и потери, обусловленные забастовкой, приводят к ее концу. Модель Ланкастера — описание пути к прекращению конфликта, использующее понятия теории вероятностей.

Прогресс в переговорах может быть представлен графически, как это показано на рис. 7.2. Точки на графике — это промежуточные шаги в переговорах. Когда путь достигает горизонтальной линии, конфликт прекращается, и тогда продолжительность забастовки выражается расстоянием от начальной точки оси времени до проекции точки пересечения. Ланкастер предложил считать путь переговоров винеровским процессом. Это означает, что при переходе от точки  $T$  к  $T + \delta T$  приращение по оси ординат — случайная величина, распределенная по нормальному закону со средним  $\mu \delta T$  и дисперсией  $\sigma^2 \delta T$  независимо от предыдущих приращений. Свойства этого процесса хорошо известны (см., например, работу Кокса и Миллера (1965)). Время до поглощения конечно (с вероятностью 1) только тогда, когда  $\mu \geq 0$ , т. е. тогда, когда процесс имеет неотрицательный «снос», и плотность функции распределения времени до поглощения представляет собой выражение

$$f(T) = \frac{d}{\sqrt{2\pi} \sigma T^{\frac{3}{2}}} \exp \left[ -\frac{1}{2} \left( \frac{d - T\mu}{\sigma \sqrt{T}} \right)^2 \right] \quad (7.15)$$

$$(d, \sigma, \mu > 0).$$

Это распределение известно как сопряженное\* к гауссовскому, его свойства описаны, например, в книге Чхикара и Фольке (1978). Если  $\mu < 0$ , то время до поглощения не обязательно конечно, т. е. вероятность того, что  $T$  — бесконечность, неотрицательна и равна  $1 - \exp \times (2d/\sigma^2)$ . Распределение описывается тогда как усеченное. Хотя в (7.15) появляются три параметра, распределение, сопряженное к гауссовскому, — распределение с двумя параметрами, так как его функ-

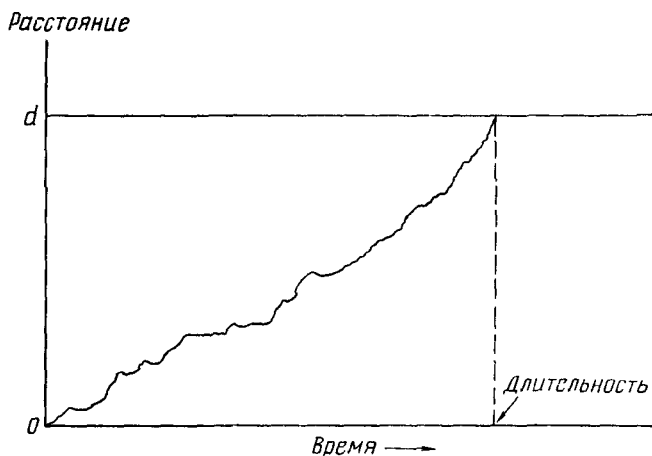


Рис. 7.2. Длительность случайного блуждания до поглощения на границе

ция плотности может быть выражена через  $d/\sigma$  и  $\mu/\sigma$ . Ланкастер (1972) применил его для описания длительности забастовок в Англии и как показано в табл. 7.3, добился значительного успеха.

Итон и Уайтмор (1977) предложили ту же модель для описания времени пребывания больных шизофренией в психиатрической больнице, хотя Ван Корфф (1979) показал, что двучленная смесь экспоненциальных распределений лучше согласуется с исходными данными. В этом приложении  $d$  — число «выздоровевших», т. е. пациентов, которых можно выписывать. Согласно модели процесс лечения идет к окончанию как случайный процесс, т. е. представляет собой винеровский процесс. Отрицательный снос объясняется тем, что болезни не всегда поддаются лечению, и может произойти смерть пациента или принято решение о постоянной госпитализации. Модель применима также для описания процесса увольнений. В этом случае можно представить, что недовольство накапливается и достигает критического уровня, когда работник заявляет о своем уходе. Уайтмор (1979) говорил о приводящей к такого же рода финалу эрозии «преданности» организации. Он применил свою модель для описания данных Силкока (1954), включая данные Бибби, представленные в табл. 7.1.

\* См. примечание на с. 218 — Примеч. ред.

Т а б л и ц а 7.3. Согласие распределения, сопряженного к гауссовскому, с продолжительностью забастовок, начавшихся в 1965 г. в британской промышленности (Ланкастер, 1972)

Продолжительность, дней	Все забастовки на транспорте		Все забастовки в металлообрабатывающей промышленности	
	наблюдаемые значения	по модели	наблюдаемые значения	по модели
2	34	34,0	43	47,1
3	19	18,2	37	30,8
4	10	11,3	21	21,0
5	8	7,7	19	15,3
6	6	5,6	11	11,6
7	5	4,2	8	9,2
8	2	3,3	8	7,4
9	3	2,6	9	6,1
10	2	2,1	3	5,1
11—15	6	6,5	16	16,7
16—20	4	3,1	4	9,1
21—25	4*	4,3*	4	5,6
26—30	—	—	3	3,7
31—40	—	—	3	4,3
41—50	—	—	5	2,3
>50	—	—	5	3,7
<b>Всего</b>	<b>103</b>	<b>102,9</b>	<b>199</b>	<b>199,0</b>

\* Цифры, отмеченные звездочкой, относятся ко всем забастовкам продолжительностью более 20 дней. Параметры оценивались по методу максимального правдоподобия. При моделировании не учитывались забастовки продолжительностью менее 1 дня (они не регистрируются); здесь приведено усеченное распределение (см. более детальное описание в работе Ланкастера (1972)).

Согласование было блестящим, даже лучше, чем при различных смесях экспоненциальных моделей.

Винеровский процесс — один из многих процессов случайного блуждания, которые могут быть использованы при непрерывном и дискретном времени. Как показано в работе Уайтмора (1979), можно обобщать метод смеси распределений, полагая, что параметры — случайные величины.

### 7.3. ЛОГНОРМАЛЬНАЯ МОДЕЛЬ РАЗМЕРОВ СИСТЕМЫ И ПРОДОЛЖИТЕЛЬНОСТИ ПРЕБЫВАНИЯ В СОСТОЯНИИ

Если логарифм положительной случайной величины  $X$  имеет нормальное распределение, то говорят, что величина  $X$  распределена по логарифмически-нормальному (логнормальному) закону. Это распределение положительно асимметрично в той же степени, в какой оно зависит от стандартного отклонения логарифма  $X$ . С помощью логнормального распределения успешно описываются многие распределения продолжительности пребывания в состоянии и размеров системы. Логарифмирование дает существенные статистические



преимущества, так как, имея дело с логарифмами переменных, мы применяем все методы теории нормальных распределений. С практической точки зрения важно понять, какие процессы приводят к логнормальному распределению. В работе Айтчисона и Брауна (1957) дано полное описание распределения с его свойствами, а также рассмотрены некоторые приложения.

Видимо, Лане и Андро (1955) были первыми, кто открыл, что распределения ПСС могут быть успешно смоделированы логнормальными распределениями. С тех пор их результат неоднократно был подтвержден во многих странах с разной точностью. Это распределение обладает важной особенностью — ненулевой модой, и во многих приложениях имеет преимущество по сравнению со смесью экспоненциальных распределений. Хотя, когда данные грубо обработаны, наличие моды не очевидно, при лучшей группировке она легко выявляется, и предпочтительнее иметь модель, обладающую этой особенностью. Некоторые распределения размеров системы также лучше описывать в виде логнормальных распределений. Айтчисон и Браун (1957) рассмотрели эти распределения применительно к описанию распределений доходов и привели множество примеров. В работе Стейндла (1965) исследуется пример из экономики с распределением заводов обрабатывающей промышленности по числу работающих и фирм по объему годового оборота. В естественных науках логнормальное распределение используется для описания распределения размеров частей при дроблении. Хотя с точки зрения общественных наук непосредственного интереса эти результаты не представляют, рассмотренный механизм позволяет провести аналогию с некоторыми видами распределений социальных процессов.

Как уже отмечалось, широкое распространение нормального распределения объясняется действием центральной предельной теоремы, когда суммируются случайные величины. Если переменная распределена по логнормальному закону, то следует задать вопрос, действительно ли это следствие того, что она представляет собой произведение случайных величин, а значит, их логарифм равен сумме. Это приводит нас к необходимости исследовать, возможна ли такая интерпретация в различных приложениях. Так как все описанные здесь модели имеют одну и ту же структуру, удобно начать с изложения наиболее существенных результатов в абстрактном виде.

Рассмотрим последовательность случайных величин  $T_1, T_2, T_3, \dots$ , образованных следующим образом.  $T_1$  задано, а другие величины ряда образуются, если взять

$$T_{j+1} = T_j u_{j+1}, \quad j = 1, 2, 3, \dots, \quad (7.16)$$

где  $u_2, u_3$  — случайные величины, распределенные по одному и тому же закону для любых  $T$ . Иными словами, из выражения (7.16) вытекает, что относительное изменение  $T$  при переходе от  $j$  к  $j+1$ ,  $(T_{j+1} - T_j)/T_j$ , распределено по закону, не зависящему от значения  $T_j$ . В соответствии с (7.16)

$$T_{j+1} = T_1 \prod_{i=2}^{j+1} u_i, \quad j \geq 1,$$

или

$$\log T_{j+1} = \log T_1 + \sum_{i=2}^{j+1} \log u_i. \quad (7.17)$$

Предположим, что распределение величин  $u$  таково, что для их суммы справедлива центральная предельная теорема, поэтому при увеличении  $j$  величина  $\log T_{j+1}$  асимптотически стремится к нормальному закону распределения, и, следовательно, величина  $T_{j+1}$  — к логнормальному. Это полностью справедливо, если  $u$  независимы и имеют одинаковые распределения, но во многих случаях такие требования могут быть ослаблены без нарушения асимптотической нормальности.

Сейчас нам нужно рассмотреть, как эту общую модель можно интерпретировать в различных приложениях. Мы не говорим, что модель близка к действительности всегда, но она отражает особенность социальных процессов, для которых характерно, что в предельном случае они описываются с использованием логнормального закона распределения. Первая версия модели связана с многократным дроблением на части. Предположим, что крупный землевладелец делит имущество среди наследников. Далее эти наследники делят то же имущество среди своих наследников, и так из поколения в поколение. Какое распределение будет иметь размер состояния через несколько поколений? В среднем состояние становится, конечно, меньше, но нам интересен характер распределения. Если  $T_j$  — случайная величина размера состояния в  $j$ -м поколении, то  $u_{j+1}$  — доля состояния отца, которая достается случайно выбранному наследнику следующего поколения. Если пропорции деления наследства в некоторой мере случайны и существует требуемая независимость между поколениями, то на основании изложенного можно ожидать, что распределение будет асимптотически-логнормальным. Аналогичную аргументацию можно применить при рассмотрении повторяющегося деления любой количественно выраженной величины. Например, общество при изучении делится на все более мелкие группы, численность которых в пределе имеет приближенно логнормальное распределение. Как отмечено Айтчисоном и Брауном (1957), эти идеи приводят к разработке теории группирования, в соответствии с которой группы получают путем последовательного деления. Например, при изучении контингента рабочих они вначале подразделяются по уровню квалификации, затем по полу, затем по уровню физического труда и т. д. Распределение численности рассматриваемых групп может быть асимптотически-логнормальным.

Вторую версию модели иногда используют для описания распределения по уровню благосостояния. Пусть, например, уровень благосостояния некоторого индивидуума в начальный момент времени равен величине  $T_1$ , он изменяется через интервалы времени, помеченные  $j = 1, 2, 3, \dots$ . За любой интервал уровень благосостояния может измениться случайным образом. Очень вероятно, что чем выше начальный уровень, тем больше изменение. Если допустить, что относительное изменение распределено по одному и тому же закону, когда

уровень благосостояния индивидуума и изменения независимы, то в общей модели величины  $u$  независимы и одинаково распределены. В этом случае модель обладает нежелательным свойством. Как следует из (7.17), дисперсия  $\log T_{j+1}$  пропорциональна величине  $(j - 1)$ , отсюда распределение по уровню благосостояния обладает увеличивающимся со временем разбросом. Айтчисон и Браун (1957) приводят некоторые примеры, иллюстрирующие такое положение, хотя оно в основном, по-видимому, не имеет места. На практике оказывается, что действуют другие факторы, особенно это не выполняется в предельном случае, когда под их воздействием дисперсия не увеличивается. Это подтверждает замечание Стейндла (1965, с. 11) по поводу того, что согласование логнормальной модели с практикой удовлетворительно в средней области, но слабое в предельном случае. Аналогичные рассуждения могут быть и при описании увеличения размеров города, рост которого по схеме размножения и гибели (или аналогичной) определяется изменениями, пропорциональными размерам в данный момент времени.

Третья интерпретация модели может быть дана при предположении, что изменения происходят не последовательно, а более или менее одновременно. В соответствии с этим, например, поступления индивидуумов определяются большим числом независимых факторов, действие которых умножается. Так, рост цен на предметы потребления повышает доходы всех производителей товаров в одинаковой пропорции, а взимание налогов при одинаковом проценте уменьшает их в той же пропорции. Если таким образом действует большое число факторов, то существуют условия для того, чтобы ситуация описывалась по логнормальному закону.

Применение логнормальной модели для описания полного срока службы может быть аналогично, но это сопряжено с некоторым риском. Рассмотрим последовательность мест работы какого-либо индивидуума и обозначим через  $T_j$  время работы на  $j$ -м месте. Тогда линейное уравнение (7.16) будет просто описанием того, как текущее состояние определяет поведение в будущем. Это подразумевает наличие связи между временами пребывания, так как продолжительное пребывание на одном месте предрасполагает к работе продолжительное время на другом и т. д. Этим можно объяснить логнормальный закон пребывания на многих местах работы, но нельзя дать объяснения логнормальному времени пребывания на первой работе. Чтобы рассмотреть этот случай, необходимо обратиться к третьей интерпретации распределения размеров. Предположим, что  $T_j$  — время, которое индивидуум, поступая на работу, собирается проработать на данном месте. В первый период времени на данном рабочем месте он подвержен влиянию многих факторов. Некоторые из них благотворные и склоняют его остаться здесь дольше, чем он предполагал. Другие факторы отрицательные, они обуславливают стремление индивидуума уйти. Таким образом, первоначальные намерения работника меняются под действием различных факторов. Если выразить влияние каждого фактора на ожидаемое время службы в виде случайной величины  $u$ , то  $T_j$  представляет собой действительную продолжительность работы инди-

видуума на данном месте под влиянием  $j$  факторов. Если принять то же предположение о случайности  $u$ , то действительное время работы на одном месте будет распределено по логнормальному закону.

Согласно этой версии модели индивидуальные отличия будут отражаться на величине  $T_1$ . Это влияние можно получить, пытаясь изменить средние значения  $u$  так, чтобы увеличить (или уменьшить, если это необходимо) действительное время службы.

В разделе 7.2 мы показали, на каких основаниях в модель вводится учет различия между индивидуумами. Тот же подход может применяться для описания многих распределений размеров систем. Для того чтобы реализовать его здесь, положим, что  $\ln T \sim N(\omega, \sigma^2)$ . По крайней мере, в приложениях к описанию увольнений есть эмпирические основания считать, что скорее всего неоднородность приводит к вариации  $\omega$ . Если мы допустим, что функция распределения равна  $H(\omega)$ , то плотность функции смеси логнормальных распределений равна:

$$f(T) = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}\sigma T} \exp\left[-\frac{1}{2}\left(\frac{\ln T - \omega}{\sigma}\right)^2\right] dH(\omega). \quad (7.18)$$

Заметим, что область значений  $\omega$  — вся вещественная ось. Если положить  $\omega \sim N(\mu, \tau^2)$ , то легко показать, что

$$\ln T \sim N(\mu, \sigma^2 + \tau^2). \quad (7.19)$$

Другими словами, распределение имеет тот же вид, но оно более асимметрично, чем первоначальное, так как имеет большую дисперсию  $\ln T$ . Таким образом, если смесь распределений нормальна (и многие характеристики людей включены здесь же), то логнормальный вид распределения сохраняет неоднородность шкалы параметров.

Подводя итоги, можно сделать вывод, что аргументация и модели, полученные в настоящем разделе, дают основания полагать, что логнормальное распределение будет постоянно появляться. Такому постоянству пока нет полного объяснения. Например, если даже первая версия модели правильна, то мы не можем показать, почему соотношение между  $T_{j+1}$  и  $T_j$  просто линейно, как зафиксировано в (7.16).

#### 7.4. МОДЕЛИ РАЗМЕРОВ СИСТЕМ

Обычно приписывают Парето открытие того факта, что логарифм доли людей с доходами, большими, чем  $x$ , часто примерно пропорционален  $-\log x$ , когда  $x$  велико. Распределения, обладающие этим свойством, часто называют распределениями Парето, а константа пропорциональности известна под названием параметра Парето. С практической точки зрения очень легко сравнить распределения доходов графическими методами. Зипф (1949) обнаружил, что той же особенностью обладают другие распределения, описывающие размеры систем, включая, например, размеры городов в США и в других западных странах. В этом смысле широко известна закономерность Зипфа. Зипф сам пытался доказать, что эта закономерность — следствие то-

го, что он назвал принципом наименьших усилий, который, по его мнению, лежит в основе поведения людей. Другие исследователи, начиная с Симона (1955), пытались построить стохастическую модель роста, которая привела бы к распределениям устойчивых состояний, обладающих свойством Парето. Получается, что большой класс процессов сводится к распределениям, имеющим форму Парето, и в этом разделе мы опишем некоторые из самых простых случаев. Эти модели в вероятностных терминах описывают, как размеры систем меняются во времени. Время и размеры могут быть непрерывными или дискретными. В этом разделе мы будем полагать, что размер системы — дискретная переменная, принимающая значения 1, 2, 3, ... . Время будет считаться непрерывным в первых двух моделях и дискретным в третьей. Симон (1955) вывел закон в несколько более общем виде, чем он описан выше, и показал, что для многих распределений размеров систем поведение функции вероятности при больших  $i$  имеет вид

$$p_i \propto i^{-\alpha} \beta^i, \quad (7.20)$$

где  $\beta$  так близок к единице, что  $\beta^i$  существенно влияет на  $p_i$  только тогда, когда  $i$  очень велико и  $\alpha > 0$ .

### Модель Шоррокса

Эта модель основана на марковском процессе с непрерывным временем, она эквивалентна процессу размножения и гибели с иммиграцией, рассматриваемому в большинстве учебников по теории случайных процессов. Поскольку нет ограничений сверху на размеры системы (по крайней мере, теоретически), процесс имеет бесчисленное множество состояний, в то время как до сих пор мы рассматривали процессы с конечным числом состояний. Ключевым моментом задачи является вопрос, существует ли устойчивое состояние, так как область распределения может быть бесконечной. Мы будем рассматривать эти случаи последовательно, но для полного обоснования этапов необходимо сослаться на основные учебники, перечисленные в разделе «Дополнения» в гл. 5. Когда для описания модели роста используется марковский процесс,  $r_{ij}\delta T$  представляет собой вероятность того, что индивидуум передвигается из категории размера  $i$  в категорию размера  $j$  за период времени  $(T, T + \delta T)$ . Тогда вероятности состояний  $p_i(T)$  — распределение численности системы в момент времени  $T$ . Нас будут интересовать лишь некоторые виды этого распределения.

Шоррокс (1975) предложил следующую модель распределения уровней благосостояния:

$$\left. \begin{aligned} r_{i, i+1} &= (\nu + \lambda i), \quad i = 1, 2, \dots \\ r_{i, i-1} &= \mu i, \quad i = 2, 3, \dots \\ r_{ij} &= 0, \quad \text{когда } i \neq j \end{aligned} \right\} \quad (7.21)$$

По определению эта модель относится к описанию замкнутой популяции, но ее можно приспособить и для описания открытых популяций, если положить, что тот, кто покидает систему, заменяется новым,

получающим уровень благосостояния выбывшего. Вероятности распределения уровней для таких процессов удовлетворяют уравнениям (см. (4.4)):

$$\begin{aligned} \frac{dp_i(T)}{dT} &= -\{v + (\lambda + \mu) i\} p_i(T) + \{v + \lambda(i-1)\} p_{i-1}(T) + \\ &\quad + \mu(i+1) p_{i+1}(T), \quad i > 1, \\ \frac{dp_1(T)}{dT} &= -v p_1(T) + \mu p_2(T). \end{aligned} \quad (7.22)$$

Если существует устойчивое распределение, оно может быть найдено из (7.22) после замены производной нулем и при исключении  $T$ . При  $\lambda < \mu$  последовательно решая уравнения и используя тот факт, что  $\sum_{i=1}^{\infty} p_i = 1$ , находим

$$p_i = \frac{\Gamma(i+\rho-1)}{\Gamma(\rho)\Gamma(i)} (1-\beta)^\rho \beta^{i-1}, \quad i = 1, 2, \dots, \quad (7.23)$$

где  $\rho = v/\lambda$  и  $\beta = \lambda/\mu$ . Устойчивое состояние существует, если  $\lambda < \mu$ . Это означает, что интенсивность «размножения»  $\lambda$  должна быть меньше, чем интенсивность «гибели». Уравнения (7.23) определяют отрицательное биномиальное распределение, которое можно получить, когда  $i$  велико, на основании того факта, что

$$\Gamma(x)/\Gamma(x+a) \sim x^{-a}, \quad \text{когда } x \rightarrow \infty.$$

Таким образом, находим

$$p_i \sim \text{постоянная величина} \times i^{\rho-1} \beta^i,$$

что соответствует (7.20) при  $\alpha = 1 - \rho$ . Для того чтобы получить распределение, согласующееся с (7.20), интенсивности размножения и гибели должны быть примерно равными (полагаем  $\beta = 1$ ) и  $0 < \rho < 1$ .

Сейчас необходимо установить основание выбора интенсивностей в (7.21). Ограничения продвижения в смежные группы заключаются в том, что рост осуществляется непрерывно, это для многих приложений вполне обоснованное предположение. Если взять  $v = 0$ , то интенсивности переходов согласуются с законом «эффекта пропорциональности», так как, скажем, чем выше уровень благосостояния в настоящий момент, тем больше увеличение. Из этого, в свою очередь, следует, что в среднем изменения для тех, кто находится в группе  $i$ , за время  $(T, T + \delta T)$  пропорциональны  $i$ . Видимо, такая версия модели больше всего подходит для описания размеров городов, так как при отсутствии миграции можно ожидать изменения (положительные или отрицательные), пропорциональные размерам.

Если интенсивности размножения и гибели примерно равны, то проявляется закономерность Зипфа,  $\alpha = 1$ . Если разрешается миграция из внешней среды, то  $v > 0$ , тогда параметр Парето меньше единицы. Применительно к описанию уровней благосостояния положительное значение  $v$  соответствует тому, что уровень благосостояния повышается при поступлениях из некоторых источников, независимых от имеющегося уровня. В модели может быть  $\lambda = 0$ . Это означает, что

рост не определяется настоящим размером. В этом случае решение (7.22) сводится к распределению Пуассона:

$$p_i = \frac{\rho^{i-1}}{(i-1)!} e^{-\rho}, \quad i = 1, 2, \dots \quad (7.24)$$

Дальнейший анализ проводится при предположении об однородности, которое заключается в том, что для всех членов популяции справедлив один и тот же закон роста. Ранее предполагалось, что должны разрешаться индивидуальные различия. Чтобы реализовать это, нужно считать один или несколько параметров случайными величинами. Однако построенная модель уже и так достаточно многозначна, и сейчас мы это покажем. Предположим, что изменения, сформулированные в (7.21), применимы к любому данному индивидууму с  $\lambda = 0$ , а индивидуальные различия непрерывны и их плотность выражается гамма-распределением

$$f(\rho) = c^\nu \rho^{\nu-1} e^{-c\rho} / \Gamma(\nu), \quad \rho \geq 0.$$

Тогда

$$\begin{aligned} p_i &= c^\nu \int_0^\infty \rho^{i+\nu-2} e^{-\rho(1+c)} d\rho / \Gamma(\nu) (i-1)! = \\ &= \frac{\Gamma(\nu+i-1)}{\Gamma(\nu) \Gamma(i)} \left( \frac{c}{1+c} \right)^\nu \left( \frac{1}{1+c} \right)^{i-1}, \quad i = 1, 2, \dots, \end{aligned} \quad (7.25)$$

что соответствует (7.23). Следовательно, распространенность отрицательного биномиального распределения может быть результатом или того, что система однородна с  $\lambda > 0$ , или того, что имеется смесь систем с  $\lambda = 0$ .

Хотя отрицательное биномиальное распределение принадлежит к семейству распределений, удовлетворяющих (7.20), значение параметра Парето для него не может быть больше единицы. На практике в распределениях они имеют большие значения, обычно от 1 до 2. Следовательно, нам необходимо рассмотреть вторую модель, для которой не нужно вводить такое ограничение.

### Модель Юла—Симона

Это также марковская модель, но с интенсивностями переходов, равными

$$\begin{aligned} r_{i, i+1} &= (\nu + \lambda i) \quad (i = 1, 2, \dots), \quad r_{i, 1} = \mu \quad (i = 2, 3, \dots), \\ r_{ij} &= 0 \quad \text{всегда, когда } i \neq j. \end{aligned} \quad (7.26)$$

В смысле учета роста эта модель идентична (7.21). Однако она позволяет описывать открытые системы заданного размера, в которых при уходах индивидуумы замещаются новыми, поступающими на нижний уровень. Выбор постоянных интенсивностей ухода  $\mu$  означает, что вероятности ухода одинаковы для любого уровня. Очевидно, что  $\lambda$  должно быть неотрицательно, поэтому численность системы пропор-

циональна ее размерам и вероятности их переходов — постоянные величины. Симон (1955) предложил вариант модели с дискретным временем и обратил внимание на то, что Юл также сделал такое предположение, чтобы описать распределение числа типов заводов, имеющих  $i$  видов.

Из этой модели неясно, как определять размеры городов, поскольку, хотя скорость возрастания и можно аппроксимировать с помощью линейного роста интенсивности, города полностью не исчезают. Симон (1955) дал другую, более предпочтительную интерпретацию модели. Он показал, что то же самое распределение получается как квазиустойчивое состояние увеличивающейся численности населения городов, которое обеспечивается за счет новых городов минимального размера.

Легко показать, что уравнения устойчивого состояния для модели (7.26) будут такими:

$$\begin{aligned} - (v + i\lambda + \mu) p_i + (v + (i - 1)\lambda) p_{i-1} &= 0, \quad i = 2, 3, \dots, \\ p_1 &= \mu / (v + \lambda + \mu). \end{aligned}$$

Таким образом,

$$p_i = \frac{\prod_{j=1}^{i-1} (\rho + j)}{i \prod_{j=1}^i (\rho + \eta + j)} \eta, \quad i = 2, 3, \dots, \quad (7.27)$$

$$p_1 = \eta / (1 + \rho + \eta),$$

где  $\eta = \mu/\lambda$ . Симон (1955) рассмотрел только особый случай, когда  $v = \rho = 0$ , для которого

$$p_i = \eta B(i, \eta + 1), \quad i = 1, 2, \dots \quad (7.28)$$

Для больших значений  $i$  можно выразить (7.27) в виде гамма-распределения и использовать такую же аппроксимацию, которая была рассмотрена ранее. Таким образом, находим

$$p_i \sim \text{const} \times i^{-(1+\eta)}, \quad (7.29)$$

так что  $(1 + \eta)$  — параметр Парето, в качестве которого можно взять любое число больше единицы.

Кроме описанного выше приложения, Симон (1955) рассмотрел распределение частоты употребления слов и числа статей различных авторов, опубликованных в научных журналах за несколько лет. Результаты относительно статей и заметок в журнале «Эконометрия» за 20-летний период приведены в табл. 7.4. Естественно, модель слишком проста, чтобы полно отразить ситуацию, но стоит отметить, как много можно сделать с ее помощью. Оценка для значения  $\eta = 0,69$  означает, что автор, вероятно, в 69% случаев прекращает писать новые статьи.

Так же, как в модели Шоррокса, распределение (7.27) может быть получено с помощью смеси нескольких распределений. Как и ранее, положим  $\lambda = 0$ , чтобы рост больше не зависел от численности. В этом



случае распределение (7.27) сводится к геометрическому распределению:

$$p_i = \left( \frac{\nu}{\nu + \mu} \right)^{i-1} \left( 1 - \frac{\nu}{\nu + \mu} \right), \quad i = 1, 2, \dots \quad (7.30)$$

Пусть теперь  $x = \nu / (\nu + \mu)$  имеет бета-распределение с функцией плотности

$$f(x) \propto x^{\rho-1} (1-x)^{\eta-1}.$$

Тогда

$$p_i = \frac{B(i + \rho - 1, \eta + 1)}{B(\rho, \eta)}, \quad i = 1, 2, \dots, \quad (7.31)$$

что представляет собой несколько иную форму записи (7.27). Это приводит к той же дилемме, что и ранее. Хорошее согласие модели с данными из табл. 7.4 может означать, что чем больше статей написано автором, тем, вероятно, он напишет их еще больше. Это может означать также, что авторы имеют различные склонности к сотрудничеству в журналах, а написанное ранее не влияет на то, что они пишут в настоящее время.

**Таблица 7.4. Число авторов, написавших  $i$  статей в журнал «Эконометрика» в течение 20-летнего периода\***

	1	2	3	4	5	6	7	8	9	10	11 или более	Всего
Действительные значения	436	107	61	40	14	23	6	11	1	0	22	721
Значения, вычисленные по (7.28)	453	119	51	27	16	11	7	5	4	3	25	721

\* Из статьи Симона (1955).

Очевидно, что, используя описание непрерывного марковского процесса, можно получить множество других моделей этого типа. Предоставив это читателю, мы перейдем к моделям с дискретным временем.

### Модели размеров с дискретным временем

В принципе, чтобы перейти к моделям с дискретным временем, нужно произвести только одно изменение: заменить интенсивности переходов вероятностями переходов. Однако в этом имеются некоторые сложности, и их необходимо рассмотреть. Если мы хотим, чтобы выполнялся закон пропорциональности (размерам численностей), мы не можем больше ограничиться переходами в смежные группы. В этом случае возможны изменения только за счет введения различных пропорций текущих значений численностей, поэтому одни и те же пропорции изменений не могут определяться одними и теми же вероятностями. В любом случае может появиться необходимость допустить пере-

ходы в несколько групп, если дискретный интервал времени не выбран подходящим образом. Даже если разрешить переходы между несмежными градациями, построение матрицы переходных вероятностей, которая позволяет получить те же пропорции изменений численности при тех же значениях вероятностей, невозможно при небольшом числе параметров. Однако эти трудности можно обойти, полагая, что относительные размеры — непрерывная переменная, а состояния цепи образованы группировкой. Если мы сформируем группы таким образом, что границы (а следовательно, и средние точки) образуют геометрическую прогрессию, то переход из  $i$ -й в  $(i + h)$ -ю группу влечет за собой одинаковые относительные изменения численностей для всех  $i$  (в среднем). Таким образом, вся матрица переходных вероятностей может быть определена через элементы одной строки. Некоторые переделки должны иметь место для нижнего ряда, так как, например, невозможно перейти ниже из самой нижней группы. На практике число групп будет конечным: в наивысшей группе будут все члены системы, находящиеся выше определенного уровня. Однако из теоретических соображений положим, что имеется неограниченное число групп. Несмотря на «концевые эффекты», можно увидеть, наблюдая эмпирическую матрицу переходных вероятностей, правдоподобна ли она при выполнении закона пропорциональности. Ниже дается матрица оценок, приводящая к распределению доходов, которая получена Шорроксом (1976); она хорошо согласуется с моделью;

$$\begin{pmatrix} 0,64 & 0,29 & 0,04 & 0,03 & 0,00 \\ 0,14 & 0,56 & 0,26 & 0,03 & 0,01 \\ 0,02 & 0,22 & 0,54 & 0,21 & 0,01 \\ 0,01 & 0,04 & 0,27 & 0,54 & 0,14 \\ 0,00 & 0,01 & 0,05 & 0,27 & 0,67 \end{pmatrix}.$$

Чампернаун (1953, 1973) с помощью марковских цепей с дискретным временем получил ряд моделей для доходов, в которых учитывается влияние закона пропорциональности. В простейшей из них индивидуумы могут двигаться вверх на одну группу, оставаться в своей или двигаться вниз на  $r$  групп. Если в результате такого движения получается, что они должны опуститься ниже самой низшей группы, то по предположению они переходят в самую низшую группу. Так как матрица переходных вероятностей не зависит от текущего состояния, ее можно записать в следующем виде:

$$\left. \begin{aligned} i > r, p_{ij} &= p_{-(i-j)}, j = i-r, i-r+1, \dots, i+1, \\ &\bullet p_{ij} = 0 \text{ в других случаях;} \\ i \leq r, p_{ij} &= p_{-(i-j)}, j = 2, 3, \dots, i+1, \\ p_{i1} &= \sum_{h=i-1}^r p_{-h}, p_{ij} = 0 \text{ в других случаях.} \end{aligned} \right\} (7.32)$$

Если устойчивая структура  $q$  существует, то она удовлетворяет уравнению  $q = qP$ , в котором элементы  $P$  определяются (7.32). Если выписать их полностью, то получим

$$q_i = \sum_{j=i-1}^{i+r} q_j p_{i-j}, \quad i = 2, 3, \dots \quad (7.33)$$

Первым уравнением при  $i = 1$  можно пока пренебречь. Рассмотрим возможное решение в виде  $q_i = Ax^i$ . Если такое решение существует, то подстановка его в (7.33) показывает, что  $x$  должно удовлетворять уравнению

$$x = p_1 + xp_0 + x^2 p_{-1} + \dots + x^{r+1} p_{-r}. \quad (7.34)$$

Обозначим правую часть (7.34) через  $\Phi(x)$ . Для того чтобы  $\{q_i\}$  была распределением вероятностей, необходимо, чтобы выполнялось неравенство  $0 < x < 1$ . Имеются ли подходящие корни — можно исследовать графически, получая пересечения  $y = x$  и  $y = \Phi(x)$ . Многочлен  $\Phi(x)$  монотонно возрастает (так как все его коэффициенты неотрицательны), начиная с  $y = p_1$  при  $x = 0$ . Есть точка пересечения с  $y = x$  при  $x = 1$ , так как  $\sum_{h=-r}^1 p_h = 1$ . Для того чтобы существовали корни на отрезке  $(0, 1)$ , тангенс угла наклона  $\Phi(x)$  при  $x = 1$  должен быть строго больше единицы. Тогда

$$p_0 + 2p_{-1} + \dots + (r+1)p_{-r} = 1 - p_{-1} + 2p_{-2} + \dots + rp_{-r} = 1 + \mu > 1,$$

где  $\mu$  определяется как среднее значение числа шагов, сделанных вниз. Таким образом, условие, что процесс приближается к устойчивому состоянию, заключается в том, что  $\mu \geq 0$ . Если  $\mu > 0$ , то распределение в устойчивом состоянии представляет собой геометрическую прогрессию, так как на отрезке  $(0, 1)$  существует по крайней мере один корень, такой, что

$$q_i = (1-x)x^{i-1}, \quad i = 1, 2, \dots, \quad (7.35)$$

где  $x$  — корень функции  $\Phi(x) = x$  на отрезке  $(0, 1)$ . Вспомним, что  $i$  измеряется по логарифмической шкале. Нижняя граница для  $i$ -й группы будет иметь вид  $a^{i-1}$  по арифметической шкале, поэтому если обозначить через  $S$  размеры, превышающие некоторую величину, то вероятность записывается как

$$Pr \{S \geq S_0\} \propto S_0^{-\alpha}. \quad (7.36)$$

Таким образом, распределение имеет форму Парето для всех  $S$ , а не только для больших  $S$ . Как отметил Чампернаун (1973), предположения, принятые при построении модели применительно к доходам, сильно ее огрубляют, но он мог бы показать, что вид распределения вышних членов ряда — это распределение Парето при различных обобщениях, когда выполняется закон пропорциональности.

Тот же вид распределения по закону геометрической прогрессии получается и при дискретном варианте модели Юла—Симона. В этом случае мы имеем

$$\begin{aligned} p_{11} &= 1 - p_1, \quad p_{ii} = p_0, \quad i > 1, \\ p_{i, i+1} &= p_1, \quad p_{i1} = 1 - p_0 - p_1 \quad \text{для всех } i. \end{aligned} \quad (7.37)$$

Для данной ситуации устойчивое распределение всегда существует и

$$q_i = \frac{(1-p_0-p_1)}{(1-p_1)} \left( \frac{p_0}{1-p_1} \right)^{i-1}, \quad i = 1, 2, \dots \quad (7,38)$$

Можно получить общий вариант для обеих рассмотренных выше моделей, если ввести нижние ограничения размера — минимум дохода или пороговый размер, который должен быть превышен, чтобы индивидуум мог попасть в популяцию. Это не влияет на вид распределений в области больших размеров.

Труднее ответить на вопрос, почему действие закона пропорциональности не приводит к логнормальному закону распределения численностей в этих моделях, что можно было бы предположить, если следовать логике развития модели из раздела 7.3. В модели Юла—Симона изменение добавки дохода пропорционально ему, поэтому он зависит от  $i$ , отсюда условия применимости центральной предельной теоремы не выполняются. То же самое можно сказать и о модели Чампернауна.

Хотя область применения моделей размеров, которые были здесь рассмотрены, весьма ограничена, мы должны были достаточно подробно их обсудить, так как часто встречаются положительно скошенные распределения типа логнормального или распределения Парето. Здесь нет ничего удивительного, поскольку это явление тесно связано с законом пропорциональности.

## 7.5. ДОПОЛНЕНИЯ

При построении моделей, обсуждавшихся в этой главе, использовалась теория случайных процессов, изложенная в работах, на которые даются ссылки в «Дополнениях» гл. 3 и 5. Кроме уже знакомого материала, здесь рассматривались счетные бесконечные пространства состояний, а в случае винеровских процессов — непрерывные множества состояний. Оба эти вопроса в достаточной степени разработаны. Вывод, что линейные функции распределенных по различным законам случайных величин являются асимптотически нормально распределенными, был получен в соответствии с теорией вероятностей, например, в работе Феллера (1966, теорема 3, ст. 256). Исследование логарифмически-нормального распределения провели Айтчисон и Браун (1957). Логнормальное распределение иногда называют распределением (или законом) Гибрата по предложению экономистов, которые изучали распределения доходов или уровней благосостояния.

Мы обсудили большой класс моделей процессов уходов, не рассматривая все возможности даже тех из них, которые были нами сформулированы. Несмотря на это, некоторые из распределений, широко применяемых в теории испытаний и теории надежности, не обсуждались. Среди них гамма-распределение, различные распределения экстремальных значений, распределение Вейбулла. Эти распределения дают возможность точнее находить вероятности больших отклонений, которые получаются только для очень малых значений параметров их формы. Мы уже отмечали, что J-образное гамма-распределение может

быть получено как смесь экспоненциальных распределений, а у уни-  
модального гамма-распределения монотонно *возрастающая* интенсив-  
ность отказов.

Распределение Вейбулла, примененное Хорватом (1968) для описания  
распределения длительностей войн и забастовок, хорошо согласуется  
с данными. Хорват предложил модель, основанную на экстремальных  
значениях, аналогичную оригинальной модели Вейбулла для сопро-  
тивления материалов. При развитии конфликта, будь то война или  
забастовка, существует множество препятствий на пути к достижению  
соглашения. Время до достижения соглашения по каждому из пунк-  
тов может рассматриваться как случайная величина, и конфликт по  
данному вопросу считается исчерпанным, как только разрешен пер-  
вый пункт разногласий. Таким образом, согласно теории «слабых то-  
чек» выделяются все возможные потенциальные точки прорыва и «ус-  
пех» достигается в одной из них. Если существует много таких точек  
и время до достижения соглашения по каждому из пунктов — случай-  
ные величины с одинаковым распределением, то согласно теории эк-  
стремальных значений соответствующее предельное распределение  
представляет собой распределение Вейбулла. Привлекательность этой  
модели заключается в ее общности, так как она не зависит от вида со-  
ставляющих распределений. Однако это спорное достоинство для дру-  
гих приложений. Во многих процедурах выработки соглашения нет  
оснований предполагать, что конфликт будет прекращен, когда одно  
из препятствий будет преодолено. Может быть, необходимо достичь  
соглашения по нескольким или, возможно, по всем пунктам. Даже  
если допустить, что модель реалистична с этой точки зрения, весьма  
сомнительно, чтобы время достижения соглашения по любому пункту  
имело одинаковые распределения или даже в качестве хорошей ап-  
проксимации было бы равно их предельным значениям. Тем не менее  
модель согласуется с данными и может быть принята.

Другая форма модели длительности войн была предложена Вейс-  
сом (1963). Его первая модель связывает возможность окончания кон-  
фликта в данный момент времени с числом погибших к этому моменту  
времени. Во второй модели возможность окончания войны является  
функцией как времени, так и числа погибших. Эти модели более об-  
щего характера, чем те, что обсуждались в настоящей главе, так как  
в них интенсивности переходов представляют собой функции не толь-  
ко времени, но и других параметров (хотя в этом случае число смертей  
— монотонная функция времени). Такое обобщение — шаг к большему  
учету реальности, но это и дополнительные сложности, оказываю-  
щиеся серьезным препятствием на пути к решению. Трудно отдать пре-  
имущество рассмотренным простейшим моделям или более сложным.  
Видимо, усложнение полезно тогда, когда появляется возможность  
дополнительно получить более детальные исходные данные.

Двучленная смесь экспоненциальных распределений была связа-  
на с распределением длительности пребывания без работы, при этом  
экспериментальным путем были выделены две популяции, одна из ко-  
торых определялась как «постоянное ядро» безработных, а другая —  
как «временные» безработные (см. IMS Monitor, т. 2, 1973, с. 29—32).

Непрерывные смеси распределений аналогичного вида описаны в работе Ланкастера и Никелла (1980). Уже отмечалось использование смесей экспоненциальных распределений Ван Корффом (1979) для описания длительности пребывания в больнице. В следующей главе будут рассмотрены многошаговые процессы восстановления. Мы определим распределение состояния системы и покажем, что оно также имеет вид смеси экспоненциальных распределений.

Сопряженное распределение Гаусса обладает рядом достоинств, и Уайтмор (1979) отметил, что оно сходно по форме с логнормальным, если  $d |\mu|/\sigma^2$  равно 10 или больше, а  $\mu > 0$ . Свойства этого распределения и полная библиография приведены в работе Чхикора и Фолькса (1978). Уайтмор (1976) пересмотрел перечень приложений этих моделей и, кроме названных здесь, включил в него распределение времени реализации ценных бумаг заданной стоимости. Оно основано на том, что изменение стоимости ценных бумаг на бирже представляет собой случайный процесс (см. работу Кутнера, 1964). Краткий обзор моделей, описывающих дискретный случайный процесс с произвольными распределениями, приведен во втором издании настоящей книги (см. гл. 6, с. 201).

Применение смесей экспоненциальных и логарифмических распределений, которые сводятся к распределениям ПСС, обсуждается в книге Бартоломью и Форбеса (1979). В ней даются ссылки (см. с. 75) на неопубликованную работу Д. Кронина, посвященную логарифмически-логистическому распределению. Логистическое распределение по форме аналогично нормальному распределению, поэтому те же выводы можно сделать относительно логарифмически-логистического и логарифмически-нормального распределений. Логарифмически-логистическое распределение может быть получено как смесь распределений Вейбулла, которые, в свою очередь, получают из теории экстремальных значений. Интересная особенность логарифмически-логистического распределения становится очевидной, если рассмотреть функцию дожития следующего вида:

$$G(x) = \{1 + Ax^\alpha\}^{-1},$$
$$G(x) \sim 1/Ax^\alpha,$$

если  $x$  велико, то большие отклонения распределены по закону Парето. Таким образом, различные смеси экспоненциальных распределений, логарифмически-нормальные, сопряженные гауссовские и логарифмически-логистические распределения приводят к распределениям продолжительности пребывания в состоянии, близким по форме. Преимущества этого заключаются в том, что мы рассматриваем соответствующий вид распределения для решения статистических задач и выбираем то, которое наиболее подходит для исследуемого объекта. Если, с другой стороны, нам недостаточно понятна суть социального процесса, то с позиции практики невозможно различить модели. Все, что мы можем исследовать, это то, что можно получить, изучая продолжительность пребывания в состоянии. Для дальнейших шагов необходимо изучать другие аспекты процесса.

Основные исследования распределений размеров проведены Чампернауном (1979), Стейндлом (1965), Ирием и Симоном (1977), а также Зипфом (1949), хотя Зипф не рассматривал стохастические модели. Клифф и соавторы (1975, гл. 2) рассмотрели несколько моделей, включая модель Зипфа, для исследования распределений размеров в приложении к географии. Кендалл (1961) привел несколько дополнительных примеров распределения Юла-Симона. Последние работы по распределению доходов и уровней благосостояния — это работы Лилларда и Уиллиса (1978), а также Пестьо и Поссена (1979).

Другой подход к выводу закономерности Зипфа был применен в работах Хилла (1970, 1974), Хилла и Вудруфа (1975), Чена (1980). Здесь в начале рассматривалась классическая задача размещения, сводящаяся к задаче заполнения ящиков шарами. Уолд и Уйтл (1957) предложили развитие этой модели, основанное на процессах размножения и гибели, в применении к описанию распределения уровней благосостояний. Наряду с другими предложениями они считали, что наследники делят имущество на равные части. Обзор применения теории вероятностей для описания распределений размеров предприятий с множеством численных примеров и ссылок на литературу приведен в работе Коллинза (1973). Шоррокс (1975) показал ограниченность рассмотрения только устойчивых состояний и в своей статье рассмотрел также переходный режим.

Одна из возникающих проблем стохастического моделирования заключается в том, что две или более моделей практически неразличимы. Это явление встречалось несколько раз в предыдущих главах и особенно сильно оно проявляется для моделей длительности пребывания и размеров. Ланкастер и Никелл (1980) перечислили эти трудности в связи с рассмотрением длительности пребывания. Но эта проблема не новая. Она была поставлена в связи с изучением причин несчастных случаев, где, если рассматривать только частоту распределения числа несчастных случаев, невозможно определить, действительно ли относительное увеличение вероятности связано с уже происшедшими несчастными случаями или оно определяется индивидуальными изменениями. С формальной точки зрения эта дилемма аналогична той, на которую мы обратили внимание при рассмотрении модели Шоррокса и Симона для распределения численностей. Для исключения двусмысленности нам необходимы различные модели, показывающие, какого рода данные требуются для исследования.

### 8.1. ВВЕДЕНИЕ

В настоящей главе мы вернемся к рассмотрению динамики социальных систем, в которых индивидуумы перемещаются между группами, выбранными по некоторому признаку. Здесь, однако, мы предположим, что размеры этих групп заданы, что является полной противоположностью допущениям, принятым при изучении марковских моделей. Там были заданы вероятности переходов, а размеры групп представляли собой случайные величины. Здесь размеры групп определены для каждого момента времени, а потоки индивидуумов — интересующие нас случайные величины. Похожая задача уже решалась, когда в гл. 3 и 5 рассматривались марковские модели при заданных размерах системы. Теперь мы пойдем дальше и примем установленными размеры всех групп.

Существует несколько областей применения такого рода моделей. Одна из них — планирование кадров, когда группы определяются, например, конкретными квалификационными уровнями, а их размер устанавливается бюджетом или количеством работы для каждого уровня. Найм и продвижение в этом случае могут происходить только тогда, когда возникают вакансии из-за уходов или развития системы. При изучении территориальной мобильности заданный жилищный фонд в различных местностях — это ограничения на перемещения, которые показывают, что модель с заданными размерами подходит больше. Уайт (1970а) применил модели этого типа для изучения перемещений духовных лиц некоторых больших американских религиозных групп. Все обязанности классифицировались по уровню престижности или статусу, и индивидуумы переходили от выполнения одних обязанностей к другим по мере образования вакансий. Уайта не интересовали вопросы планирования кадров, он занимался только изучением прикладных аспектов социальной мобильности. Позднее Стьюман (1975) применил метод Уайта для изучения кадров государственной полиции.

Элемент случайности в таких процессах принципиально обуславливается механизмом уходов; индивидуумы покидают системы или замещают образовавшиеся вакансии, что обуславливает последовательности внутренних перемещений. Кроме того, случайность появляется вследствие того, как заполняются вакансии. Раздел теории ве-



роятностей, в котором рассматриваются эти модели, называется теорией восстановления. Она появилась в связи с изучением изменений численности человеческих популяций, обуславливаемых рождением и смертью. Сравнительно недавно основной сферой приложения этой теории стали области техники, в частности теория надежности. Аналогия в восстановлении компонентов ЭВМ или авиалайнеров, а также в замещении служащих или в занятии жилищ очевидна. Основная случайная переменная в этих приложениях — продолжительность периода времени, в течение которого живет элемент. В гл. 7 мы подготовили почву для изучения этого вопроса, рассмотрев модели продолжительности пребывания. В настоящей главе мы рассмотрим некоторые из этих моделей при более общем подходе к системам с заданными размерами.

Некоторые ожидают, что исследование этих моделей мы будем проводить по такой же схеме, по которой рассматривались марковские модели. Там мы делили все модели на два подкласса: замкнутые (открытые) системы и модели с непрерывным (дискретным) временем. Однако очевидно, что системы с фиксированными размерами групп обязательно должны быть открытыми, иначе не будет вакансий. Можно разделить модели на модели с дискретным и непрерывным временем. Для вычислительных работ в области планирования кадров и использования трудовых ресурсов первый подход дает много преимуществ. Он был применен в работе Бартоломью и Форбеса (1979). Здесь же более естественно применение моделей с непрерывным временем. Это следует из анализа, проведенного в гл. 7, и определяется относительно большей разработкой стандартного аппарата теории восстановления.

## 8.2. СИСТЕМЫ, СОСТОЯЩИЕ ИЗ ОДНОЙ ГРУППЫ

Мы начнем с простейшего вида систем восстановления, в которых нет внутреннего деления на группы. Наша цель — рассмотреть, как поступления, уходы и возрастное распределение зависят от распределения времени жизни элементов. Это распределение считается заданным. Для того чтобы выразить основные идеи, обратимся к наиболее простой и довольно нереалистичной ситуации. Пусть новый дом занят жильцом в течение периода времени  $t_1$ . Вместо него сразу же поселяется другой, который живет там период времени  $t_2$  и т. д. Таким образом, моменты времени, когда происходят изменения, будут  $t_1, t_1 + t_2, t_1 + t_2 + t_3 + \dots$ . Если предполагается, что величины  $t$  независимы и одинаково распределены, то мы имеем процесс восстановления. Теория таких процессов позволяет нам определить, например, число восстановлений в течение какого-либо интервала времени. На практике мы часто наблюдаем такие системы в единственной точке, и тогда нас может интересовать, как долго элемент занимает данное положение и какой промежуток времени пройдет до момента, когда он передвинется. Назовем время пребывания «возрастом» элемента, и покажем, как найти его распределение. Очень редко нас интересуется единственный дом или место работы, а чаще

всего — население региона или численность служащих фирмы. Обобщенное поведение таких систем можно найти, «складывая» эти процессы.

Основная функция, от которой все зависит, называется функцией восстановления. Ее обозначают  $H(T)$ , и она представляет ожидаемое число восстановлений на интервале  $(0, T)$ . Известно, что ее производная  $h(T)$  — плотность восстановлений;  $h(T) \delta T$  — ожидаемое число восстановлений за период времени  $(T, T + \delta T)$ , когда  $\delta T$  — малая величина. При этих условиях вероятность двух или более восстановлений за этот интервал времени пренебрежимо мала, поэтому  $h(T) \delta T$  можно считать вероятностью восстановления на интервале  $(T, T + \delta T)$ . Мы будем рассматривать непрерывные распределения времени жизни, для которых  $h(T)$  существует, и нам удобнее иметь дело с  $h(T)$ , а не с  $H(T)$ .

### Уравнение восстановления

Рассмотрим событие, заключающееся в том, что осуществляется перемещение за интервал времени  $(T, T + \delta T)$ . Вероятность этого события  $h(T) \delta T$ . Оно может рассматриваться как результат осуществления следующих взаимно исключающих и исчерпывающих событий; тогда сумма их вероятностей должна быть равна  $h(T) \delta T$ .

<i>Событие</i>	<i>Вероятность</i>
а) первое восстановление за период $(T, T + \delta T)$	$f(T) \delta T$
б) первое восстановление за период $(x, x + \delta x)$ , а какое-либо другое за период $(T, T + \delta T)$	$f(x) \delta x h(T-x) \delta T$

Через  $f(t)$  обозначена функция плотности времени жизни индивидуума. При записи второй вероятности мы использовали факт, что фиксируется момент времени, когда начинается процесс восстановления для нового индивидуума. Таким образом, вероятность того, что восстановление произойдет после момента времени  $(T-x)$ , равна  $h(T-x) \delta T$ . В случае б) мы представляем интервал  $(0, T)$  разбитым на участки  $\delta x$  точками, поэтому для каждого участка вероятности подсчитываются. Общая вероятность может быть получена, если проинтегрировать по  $x$ , тогда

$$f(T) \delta T + \delta T \int_0^T f(x) h(T-x) dx.$$

Приравнивая это выражение к  $h(T) \delta T$  и сокращая на  $\delta T$ , получаем уравнение восстановления

$$h(T) = f(T) + \int_0^T f(x) h(T-x) dx. \quad (8.1a)$$

При очевидной замене переменных находим эквивалентное уравнение

$$h(T) = f(T) + \int_0^T h(x) f(T-x) dx. \quad (8.16)$$

Это уравнение иногда можно решить с помощью преобразования Лапласа для обеих частей, поскольку интеграл в правой части будет в преобразованном виде представлен как произведение изображений. Таким образом,

$$h^*(s) = f^*(s) + f^*(s) h^*(s),$$

следовательно,

$$h^*(s) = f^*(s) / \{1 - f^*(s)\}. \quad (8.2)$$

Для того чтобы применять этот метод, нужно найти  $f^*(s)$  и сделать обратные преобразования правой части уравнения (8.2).

Если время жизни распределено по экспоненциальному закону с параметром  $\lambda$ , решение находится довольно просто. В этом случае  $f^*(s) = \lambda / (\lambda + s)$ , откуда получаем  $h^*(s) = \lambda / s$  и  $h(T) = \lambda = \mu^{-1}$ , где  $\mu$  — среднее время жизни. Интуитивно понятно, что если люди в среднем остаются в каком-либо состоянии 4 года, то можно ожидать, что в течение года 25% их общего числа покинет это состояние. Однако этот результат справедлив, вообще говоря, только для экспоненциального распределения, а мы уже видели, что многие распределения длительности пребывания имеют далеко не такой вид. В этом случае интуиция — плохой помощник, и необходим более глубокий анализ. Замечание «вообще говоря» весьма важно, так как и при более общих условиях справедливо, что

$$\lim_{T \rightarrow \infty} h(T) = \mu^{-1}. \quad (8.3)$$

Таким образом, интенсивность перемещений стремится к определенному значению при достаточно больших  $T$ . Остается исследовать, насколько велико должно быть значение  $T$ , чтобы (8.3) было хорошей аппроксимацией, и как ведет себя система на более коротком интервале времени. Ниже мы проведем это исследование применительно к задаче планирования кадров, когда время жизни — полный срок службы (ПСС).

**Решение уравнения восстановления, когда распределение полного срока службы — смесь экспоненциальных распределений**

Когда распределение полного срока службы имеет вид

$$f(T) = p\lambda_1 e^{-\lambda_1 T} + (1-p)\lambda_2 e^{-\lambda_2 T},$$

изображение Лапласа выглядит так:

$$f^*(s) = p \left( \frac{\lambda_1}{\lambda_1 + s} \right) + (1-p) \left( \frac{\lambda_2}{\lambda_2 + s} \right). \quad (8.4)$$

Если подставить это выражение в (8.2) и применить обратные преобразования Лапласа к  $h^*(s)$ , то, как показано в работе Бартоломью (1959),

$$h(T) = \mu^{-1} + \{p\lambda_1 + (1-p)\lambda_2 - \mu^{-1}\} \exp\{- (p\lambda_2 + (1-p)\lambda_1 T)\}, \quad (8.5)$$

где  $\mu = p/\lambda_1 + (1-p)/\lambda_2$  — среднее функции  $f(T)$ . Из этой формулы следует, что  $h(T)$  в пределе приближается к экспоненциальной кривой. Далее, так как  $p\lambda_1 + (1-p)\lambda_2 \geq \mu^{-1}$ , число вновь принимаемых в систему всегда будет больше, чем ожидалось бы по предсказаниям теории устойчивого состояния. Ожидаемую численность для необходимого поступления за интервал времени можно найти, интегрируя  $h(T)$ . Это положение иллюстрируют численные значения, приведенные в табл. 8.1, в ней используются смесь экспоненциальных распределений из табл. 7.1.

Таблица 8.1. Значения относительного числа поступлений и увольнений по кварталам для новых фирм, где распределение полного срока службы — смесь экспоненциальных распределений, %

Квартал	Пополнение при использовании смеси экспоненциальных распределений ПСС	
	Гласнерская металлургическая компания	Фирма «Бибби и сыновья» (мужчины)
1	22,6	34,6
2	18,2	28,6
3	15,3	24,2
4	13,4	20,9
5	12,1	18,5
6	11,3	16,8
7	10,8	15,5
8	10,4	14,6
	69,5	108,3
	44,6	65,4

Мы не приводим результаты вычислений в моменты времени, большие, чем двухлетний интервал, так как, без сомнения, смесь экспоненциальных распределений хорошо согласуется в области больших уклонений.

Данные табл. 8.1 показывают достаточно быстрое уменьшение числа вновь поступающих. Это характерно для новых организаций. В приведенных примерах значения приблизительно в два раза меньше во втором году, чем в первом. Эти результаты, очевидно, определяются как планируемым наймом, так и увольнениями. Предполагается, что первоначальное большое число увольнений можно уменьшить, тщательно отбирая работников, т. е. бракуя тех, у кого ожидается короткий период службы, и что распределение полного срока службы не меняется во времени. Это предположение спорно, что особенно вероятно тогда, когда речь идет о новой организации.

Табл. 8.1 имеет большое значение для выяснения текучести кадров (текучесть кадров определяется как общее число покинувших систему за рассматриваемый период, деленное на среднюю числен-

ность работающих за тот же период). Изменение интенсивностей увольнений необязательно показывает изменение условий, определяющих увольнения. Оно может просто отражать изменения в структуре, определяющие сроки службы. Отсюда следует, что бесполезно, например, сравнивать текучесть кадров на новых и старых предприятиях.

### Приближенное решение уравнения восстановления

Мы видели, что некоторые эмпирические распределения ПСС могут вполне удовлетворительно заменяться смесями экспоненциальных распределений, по крайней мере, на интервале от 0 до 21 месяца. Однако, как было указано, чем длиннее период, тем менее подходит логнормальное распределение, потому что оно дает большую вероятность больших уклонений. То же самое справедливо для распределения типа XI из модели Силкока. К сожалению, невозможно получить простое точное выражение  $h(T)$  ни для одного из этих распределений. Следовательно, для изучения вида  $h(T)$  на длительном интервале времени нам необходимо получить приближенное решение уравнения восстановления.

Аппроксимация, которую мы будем использовать, получена Бартоломью (1963b). Она предполагается в том случае, когда распределение полного срока службы очень сильно асимметрично и имеет вид

$$h^0(T) = f(T) + F^2(T) \left| \int_0^T G(x) dx, \right. \quad (8.6)$$

где  $G(T) = 1 - F(T)$ . Эта аппроксимация, когда существует точное решение уравнения восстановления, обладает следующими свойствами:

- а) если  $f(T) = \lambda e^{-\lambda T}$ , то  $h^0(T) = h(T) = \lambda$  для любых  $T$ ;
- б)  $\lim_{T \rightarrow \infty} h^0(T) = \lim_{T \rightarrow \infty} h(T) = \mu^{-1}$ ;
- в)  $h^0(0) = h(0) = f(0)$ ;

$$\text{г) } \left. \frac{d^i h^0(T)}{dT^i} \right|_{T=0} = \left. \frac{d^i h(T)}{dT^i} \right|_{T=0} \quad (i = 1, 2).$$

Из этих свойств следует, что аппроксимация вполне удовлетворительна, когда  $f(T)$  близко к экспоненциальному, и всегда удовлетворительна при  $T = 0$  и больших значениях  $T$ . При применении метода производных из работы Бартоломью (1963b) также предполагается, что аппроксимация будет удовлетворительна, если функция  $f(T)$  имеет сравнительно большую вероятность больших уклонений. Вычисления, произведенные Бутлером (1970), подтверждают это для логнормального распределения. Еще одно полезное свойство аппроксимации заключается в том, что  $h^0(T)$  — верхнее ограничение для  $h(T)$ , если интенсивность увольнения, связанная с  $f(T)$ , не возрастает

ет. Такие распределения хорошо изучены в теории надежности, где они известны как распределения с уменьшающейся вероятностью выхода из строя (см., например, книгу Барлоу и соавторов (1972, гл. 5)).

Наиболее простую форму имеет аппроксимация, когда мы используем модель Силкока. В этом случае

$$f(T) = \frac{\nu}{c} \left(1 + \frac{T}{c}\right)^{-(\nu+1)}, \quad G(T) = \left(1 + \frac{T}{c}\right)^{-\nu}$$

и

$$\int_0^T G(x) dx = \mu \left\{ 1 - \left(1 + \frac{T}{c}\right)^{-\nu+1} \right\}.$$

Следовательно, из (8.5)

$$\left. \begin{aligned} h^0(T) &= \frac{\nu}{c} \left(1 + \frac{T}{c}\right)^{-(\nu+1)} + \frac{\left\{ 1 - \left(1 + \frac{T}{c}\right)^{-\nu} \right\}^2}{\mu \left\{ 1 - \left(1 + \frac{T}{c}\right)^{-\nu+1} \right\}} \\ h^0(T) &\sim \frac{1}{\mu} \left\{ 1 - \left(\frac{c}{T}\right)^{\nu-1} \right\}, \text{ если } \nu > 1 \end{aligned} \right\} \quad (8.7)$$

Когда  $\nu$  значительно меньше единицы, приближение к устойчивому состоянию очень медленное. Если  $0 < \nu < 1$ , то среднее равно бесконечности и  $h(T)$  стремится к нулю как  $T^{\nu-1}$ . В случае когда  $\nu = 1$ , она стремится к нулю как  $(\log T)^{-1}$ . Согласно вычислениям, выполненным Силкоком (1954), для восьми распределений в шести случаях из восьми значения  $\nu$  находятся между 0,5 и 1. Следовательно, можно ожидать, что интенсивности увольнений уменьшаются, вероятно, медленнее для всех периодов времени, интересных с практической точки зрения.

Аппроксимация логнормальным распределением полного срока службы также имеет простой вид. Таким образом, получаем

$$\left. \begin{aligned} f(T) &= \frac{1}{\sqrt{2\pi} \sigma T} \exp \left\{ -\frac{1}{2} \left( \frac{\ln T - \omega}{\sigma} \right)^2 \right\}, \\ F(T) &= \Phi \left( \frac{\ln(T) - \omega}{\sigma} \right) \\ \text{и} \\ \int_0^T G(x) dx &= T \left\{ 1 - \Phi \left( \frac{\ln T - \omega}{\sigma} \right) \right\} + \mu \Phi \left( \frac{\ln T - \omega}{\sigma} - \sigma \right) \end{aligned} \right\} \quad (8.8)$$

где  $\mu = e^{\frac{\omega+1}{2}\sigma^2}$ , а  $\Phi(\cdot)$  — стандартная функция нормального распределения. Следовательно, аппроксимация может быть рассчитана с помощью стандартных таблиц интеграла вероятностей. Из (8.8) вытекает, что устойчивости аппроксимации не будет до тех пор, пока

величина  $X = (\ln T - \omega)/\sigma$  не станет настолько большой, что  $\Phi(X)$  будет близко к единице. Чтобы исследовать это утверждение более детально, рассмотрим вид  $h^0(T)$  для больших  $T$ . Если  $X$  велико, то мы можем записать:

$$\Phi(X) \sim 1 - \frac{1}{\sqrt{2\pi} X} e^{-\frac{1}{2} X^2},$$

тогда обычные преобразования дают

$$h^0(T) \sim \mu^{-1} \left\{ 1 - \frac{1}{\sqrt{2\pi}} \frac{\sigma}{X(X-\sigma)} e^{-\frac{1}{2}(X-\sigma)^2} \right\}^{-1}. \quad (8.9)$$

Если, например,  $\sigma = 2$  и  $X = 4$ , то  $h^0(T)|_{X=4} = \frac{1,02}{\mu}$ .

Если  $X = 4$ , то  $T = \mu e^\sigma = 400\mu$ . Это означает, что должно пройти время, равное 400 средним значениям полного срока службы, чтобы устойчивая величина была приближена с 2%-ной точностью. Понимая, что логнормальное распределение не будет отражать реальной ситуации через 40 или 45 лет, можно считать, что предельное поведение не представляет практического интереса.

Переходный режим модели восстановления иллюстрируется данными из табл. 8.2, при этом предполагается, что распределение ПСС логнормально. Вычисления производились для типичного случая, когда  $\omega = 0$ , и  $\sigma = 2$ . Если единица времени — один год, то  $\omega = 0$  соответствует медиане (половине жизни) в один год. Для этого распределения  $\mu = e^2 = 7,389$ , потому что  $h(\infty) = 0,1353$ .

Таблица 8.2. Аппроксимация плотности восстановления для логнормального распределения полного срока службы с параметрами  $\omega = 0$  и  $\sigma = 2$

$\ln(T)$	0	1	2	3	4	5	10	$\infty$
$T$	1	2,71	7,39	20,1	54,6	148,4	22,026	$\infty$
$h^0(T)$	0,574	0,424	0,318	0,244	0,194	0,164	0,125	0,135

Данные таблицы свидетельствуют, что  $h^0(T)$  стремится к предельно малому значению крайне медленно. Даже через 20 лет значение плотности восстановления почти в два раза больше значения в установившемся режиме. Тот факт, что  $h(T)$  для соответствующей смеси экспоненциальных распределений достигает устойчивого состояния за несколько лет, показывает, что при долгосрочном прогнозировании необходима аккуратность в аппроксимации распределения для больших уклонов.

## Распределение численностей нанятых

В теории восстановления существуют стандартные методы нахождения распределения численности нанятых в систему, так же как и распределения соответствующих значений их средних. Начнем с рассмотрения организационной системы численностью  $N = 1$ . Тогда ожидаемое число вновь поступающих в систему за период времени  $(0, 1)$  равно:

$$\bar{n}(T) = H(T) = \int_0^T h(x) dx.$$

Известно (см. работу Кокса, 1962, с. 40), что при больших  $T$  величина  $n(T)$  распределена асимптотически-нормально со средним  $T\mu^{-1}$  и дисперсией  $V^2 T\mu^{-1}$ , где  $V^2$  — квадрат коэффициента вариации распределения полного срока службы. Для логнормального распределения  $V^2 = e^{\sigma^2} - 1$ , которое при  $\sigma = 2$  равно 53,6. Таким образом, долговременный прогноз числа вновь поступающих в высшей степени неточен. Однако, как мы видели в случае оценки средних значений, приближение к предельному значению осуществляется так медленно, что асимптотические результаты практически бесполезны. То же относится к дисперсии, поэтому нам необходимы точные результаты, когда мы имеем дело с такого рода распределениями полного срока службы.

Точное значение дисперсии  $n(T)$  может быть найдено из следующего уравнения:

$$E\{n^2(T)\} = \bar{n}(T) + 2 \int_0^T \bar{n}(T-x) h(x) dx \quad (8.10)$$

(см. работу Парзена, 1962, с. 179). Так как у нас есть методы определения точно или приближенно  $\bar{n}(T)$  и  $h(T)$  для любого распределения полного времени пребывания, мы можем вычислить  $E\{n^2(T)\}$  и, следовательно, дисперсию  $n(T)$ . Проиллюстрируем эти вычисления на примере смеси экспоненциальных распределений, найти для которых  $n(T)$  легче. Если положить

$$a = p\lambda_1 + (1-p)\lambda_2 - \mu^{-1},$$

$$b = p\lambda_2 + (1-p)\lambda_1,$$

то

$$\left. \begin{aligned} h(T) &= \mu^{-1} + ae^{-bT} \\ \bar{n}(T) &= T\mu^{-1} + \frac{a}{b}(1 - e^{-bT}) \end{aligned} \right\} \quad (8.11)$$



Осуществляя необходимые подстановки в (8.10) и вычитая  $\bar{n}^2(T)$ , находим следующее выражение для дисперсии:

$$\text{var}\{n(T)\} = \frac{T}{\mu} \left\{ 1 + \frac{2a}{b} \right\} + \frac{a}{b} \left\{ 1 - \frac{4}{\mu b} + \frac{a}{b} \right\} + \frac{2a}{b} \left\{ \frac{1}{\mu} - \frac{a}{b} \right\} T e^{-bT} - \frac{a}{b} \left\{ 1 - \frac{4}{\mu b} \right\} e^{-bT} - \frac{a^2}{b^2} e^{-2bT}. \quad (8.12)$$

Соответствующие результаты для организационных систем численностью  $N$  легко получить, умножая средние значения и дисперсии, полученные выше, на  $N$ . Так как  $n(T)$  для больших организационных систем может рассматриваться как сумма систем, состоящих из одного индивидуума, в соответствии с центральной предельной теоремой она асимптотически-нормальна. Некоторые численные значения для средних и среднеквадратичных отклонений приведены в табл. 8.3. Чтобы произвести вычисления, рассматривались две гипотетические фирмы, в которых работает по 1000 сотрудников. Мы предполагаем, что в обоих случаях полный срок службы распределен как смесь экспоненциальных распределений. В первом случае мы использовали значения параметров кривой, согласующиеся с данными фирмы Глазиерской металлургической компании; во втором случае были взяты более скошенные распределения из семейства смесей экспоненциальных распределений, согласующиеся с данными Силкока (1954) по объединенной сталелитейной компании.

Таблица 8.3. Средние и среднеквадратичные отклонения числа нанятых в систему в различные интервалы времени для двух смесей экспоненциальных распределений полного срока службы, когда  $N=1000$

		Временные интервалы, месяцев			
		0-3	0-6	0-12	0-24
$p=0,6513$ $\lambda_1=0,2684$ $\lambda_2=2,4228$	Среднее	226	408	695	1142
	Среднеквадратичное отклонение	16,2	21,8	29,0	39,2
	Аппроксимация среднеквадратичного отклонения	18,4	22,6	29,2	39,2
$p=0,5377$ $\lambda_1=0,2187$ $\lambda_2=4,8940$	Среднее	458	738	1072	1508
	Среднеквадратичное отклонение	27,0	34,3	41,8	52,5
	Аппроксимация среднеквадратичного отклонения	32,9	36,3	42,4	52,6

Строки «аппроксимация среднеквадратичного отклонения» были получены, когда использовались только первые два слагаемых для вычисления дисперсии в выражении (8.12). Если аппроксимированные значения превышают истинные, то разница становится пренебрежимо малой спустя короткий период времени. Из таблицы ясно также, что

предсказание даже на короткие периоды времени может быть иногда вполне удовлетворительным.

Далее рассмотрим способы определения общего числа замещений за период времени  $(0, T)$ . Чаще нам бывает необходимо знать количество замещений за некоторый относительно короткий период времени  $(T_1, T_2)$ , находящийся на некотором расстоянии от начальной точки. Если это расстояние по оси времени таково, что система достигла устойчивого состояния, то, как известно, численности сменивших место работы на непересекающихся интервалах — независимые переменные, распределенные по закону Пуассона. Этот результат приведен в книге Хинчина (1963, гл. 5). Когда система, в которой осуществляются замещения, медленно приближается к устойчивому состоянию, весьма желательно иметь результаты по характеру распределений числа замещений в переходный период. Предельный результат, относящийся к этому случаю, был получен Григелионисом (1964). Он был повторен Бутлером (1970), а адекватность аппроксимации была исследована Бартоломью и Бутлером (1971).

Рассматриваем число замещений в системе численностью  $N$  индивидуумов за интервал времени  $(T_1, T_2)$  при  $T_2 \rightarrow T_1$  и  $N \rightarrow \infty$ , при постоянном среднем числе увольнений. Пусть интервал  $T_2 - T_1 = \delta T$  так мал, что вероятность двух или более замещений пренебрежимо мала. Тогда

$$\left. \begin{aligned} Pr \{ \text{одно замещение на } i\text{-ю должность} \} &= h(T_1) \delta T + O(\delta T^2), \\ Pr \{ \text{нет ни одного замещения на } i\text{-ю должность} \} &= \\ &= 1 - h(T_1) \delta T + O(\delta T^2). \end{aligned} \right\} \quad (8.13)$$

Таким образом, ожидаемое число замещений за интервал времени  $(T_1, T_1 + \delta T)$  для всей системы равно  $Nh(T_1) \delta T$ . Для того чтобы эта величина оставалась постоянной при  $\delta T \rightarrow 0$ ,  $N \rightarrow \infty$ , необходимо, чтобы  $\delta T = K/N$  для некоторого постоянного значения  $K$ . Таким образом, общее число замещений — сумма переменных Бернулли, вероятности которых не равны нулю и стремятся к нулю, когда значения переменных стремятся к бесконечности. При этих предположениях предельное распределение — это распределение Пуассона со средним  $h(T) K$  (см. работу Феллера, 1968, с. 282). Попутно заметим, что этот результат справедлив, когда вероятность замещения, соответствующая  $i$ -й должности, зависит от  $i$ . (Результат Феллера получен для более общего случая.) Это означает, что для различных должностей могут быть различные полные сроки службы или что начальное время для потоков работы может быть различным.

Практическое значение предельных теорем определяется тем, насколько точно осуществляется аппроксимация до достижения предельных значений. Бартоломью и Бутлер (1971) исследовали адекватность пуассоновской аппроксимации как теоретически, так и с помощью имитационного моделирования, используя смеси экспоненциальных распределений полного срока службы. Они произвели точные расчеты вероятностей (8.13) и смоделировали небольшую систему,  $N = 33$ . Их заключение коротко сводится к тому, что аппроксимация достаточ-

но точно только в том случае, когда интервал ( $T_2 - T_1$ ) меньше одной десятой средней величины полного срока службы. Практически средний ПСС обычно составляет несколько лет, поэтому аппроксимация обоснованна, если интервал прогнозирования не больше нескольких месяцев.

Бутлер (1971) проверил также адекватность теории на основе данных по служащим сталелитейной промышленности. Он обнаружил, что потоки переходов для малых однородных групп за относительно короткие периоды времени приблизительно распределены по закону Пуассона. Однако существовало несколько больших потоков, которые не соответствовали гипотезе распределения по закону Пуассона. Видимо, они возникали из-за переходов групп индивидуумов из одних отраслей в другие. В среднем согласование с теорией было обнадеживающим. Полученные ранее распределения Пуассона не означают, что численности уволенных за непересекающиеся интервалы времени независимы. В общем случае они могут быть зависимыми. Бартоломью и Бутлер (1971) исследовали также и этот вопрос. Из их расчетов следует, что коэффициент корреляции между количествами замещений за смежные интервалы времени пренебрежимо мал, если эти интервалы достаточно коротки, что требуется и для адекватности аппроксимации по закону Пуассона.

Несмотря на множество оговорок, при которых получена пуассоновская аппроксимация, она весьма полезна для практических целей и свидетельствует о том, что среднеквадратическая ошибка аппроксимации может быть вычислена достаточно просто, если извлечь квадратный корень из средних значений. Последнее необходимо сделать в любом случае, поэтому трудоемкость определения ошибки измерения минимальна.

## Распределение по стажу

Полезный способ изучения прошлого организационной системы и определения возможных проблем в будущем заключается в рассмотрении распределения по стажу. Как уже отмечалось, стаж в принятом здесь смысле представляет собой время пребывания в системе одного человека, и мы будем оперировать этим понятием вместо ранее используемого определения продолжительности работы, чтобы избежать путаницы с полным сроком службы в одной системе. Определение ненормальности характеристик распределения по стажу предполагает, что мы знаем, что считать нормальным. Теория восстановления позволяет определить вид распределений по стажу при различных предположениях о прошлом. Чтобы продемонстрировать метод, обратимся к организационной системе постоянного размера, в которой все места в начальный момент времени заняты, и будем действовать при допущениях для модели восстановления.

Некоторые из членов организационной системы заняли свой пост при ее основании, и в момент времени  $T$ , очевидно, имели «стаж»  $T$ . Таким образом, вероятность того, что некоторый случайным образом

выбранный индивидиум имеет стаж  $T$ , — это доля тех, кто прослужит до момента времени  $T$ . Она равна:

$$Pr \{x = T\} = G(T), \quad (8.14)$$

где  $x$  означает стаж. Для значений  $x < T$  имеем:

$$Pr \{ \text{индивидиум в момент времени } T \text{ имеет стаж в интервале } (x, x + \delta x) \} = a(x|T) \delta x = Pr \{ \text{он поступил в систему в интервале времени } (T-x, T-x+\delta x) \text{ и прослужил } x \text{ лет} \} = Pr \{ \text{он поступил в интервал времени } (T-x, T-x+\delta x) \} \cdot Pr \{ \text{он прослужил } x \text{ лет} \},$$

поскольку время службы считается независимым от времени поступления. Таким образом,

$$a(x|T) \delta x = h(T-x) \delta x G(x) \quad (0 \leq x < T)$$

или

$$a(x|T) = h(T-x) G(x). \quad (8.15)$$

Распределение по стажу зависит от распределения полного срока службы непосредственно через функцию  $G(x)$  и опосредованно через плотность восстановления. Мы видели, что для обычных распределений полного срока службы  $h(T)$  — убывающая функция почти по всей области определения. Следовательно,  $h(T-x)$  — возрастающая функция  $x$  при фиксированном  $T$ . Функция дожития  $G(x)$  всегда невозрастающая по  $x$ . Отсюда произведение двух функций может возрастать или убывать, но если  $T$  велико, а  $x$  мало, то  $h(T-x)$  изменяется очень медленно, так что функция  $a(x|T)$  будет подобна функции  $G(x)$  в начальный период. В пределе при  $T \rightarrow \infty$  такое соответствие выполняется для всех  $x$ . При выполнении этих условий  $G(T) \rightarrow 0$  с тем, чтобы разница в вероятностях при  $x = T$  стала пренебрежима малой. Кроме того, для заданного  $x \lim_{T \rightarrow \infty} h(T-x) = \mu^{-1}$ . Следовательно,

$$a(x|\infty) = \mu^{-1} G(x) \quad (0 \leq x < \infty). \quad (8.16)$$

Таким образом, распределение по стажу в установившемся режиме пропорционально функции дожития, поэтому оно является неубывающей функцией стажа.

Полученный результат наводит на мысль об оценке функции дожития, когда отсутствует информация о покинувших систему. Это требует, конечно, выполнения допущения об установившемся режиме, которое маловероятно, чтобы выполнялось достаточно точно. Тем не менее в том случае, когда нет данных о потоках, можно оценить функцию дожития. А это лучше, чем ничего. Если, с другой стороны, можно оценить  $G(x)$  по данным о потоках, то полученную оценку можно сравнить с распределением по стажу и посмотреть, насколько система далека от установившегося режима.

Распределение по стажу может быть вычислено при более сложных допущениях о прошлом системы, и, таким образом, по форме распределения по стажу можно судить об истории системы. Например, вид-

но, как в системе поддерживаемого размера со временем получается распределение по стажу, у которого с увеличением стажа уменьшается функция плотности. Наличие «горба» в наблюдениях какой-либо возрастной группы объясняет, что произошло в прошлом и стало причиной увеличения численности группы.

## Развивающиеся системы

Полученные результаты можно легко обобщить, чтобы рассмотреть развивающуюся систему. Для этого необходимо только объединить отдельные процессы, используя в качестве слагаемых их значения в тех точках, где происходит восстановление. Итак, предположим, что начальное состояние системы  $N(0)$  и она увеличивается на величину  $M(T_i)$  к моменту времени  $T_i$  ( $i = 1, 2, \dots$ ). Тогда ожидаемое число замещений в интервале времени  $(T, T + \delta T)$  определяется выражением

$$N(0)h(T)\delta T + \sum M(T_i)h(T - T_i), \quad (8.17)$$

причем суммирование осуществляется по всем  $i$ , для которых  $T_i < T$ . Если нам необходимо представить увеличение системы как непрерывный процесс, то величина  $M(T)\delta T$  будет означать ее рост в интервале времени  $(T, T + \delta T)$ , и соответствующее выражение для интенсивностей замещений будет таким:

$$R(T) = N(0)h(T) + \int_0^T M(x)h(T-x)dx, \quad (8.18)$$

где  $R(T)\delta T$  означает ожидаемое число замещений в интервале  $(T, T + \delta T)$ . Это естественное развитие результатов, полученных в предыдущих главах. Мы видим, что результат также зависит от плотности восстановления.

Проиллюстрируем применение (8.18), показав, как оценки для текучести кадров зависят от возраста системы. Из-за этого они не могут служить характеристиками интенсивности увольнений, что мы обсудим в разделе 8.3. Оценки текучести кадров на определенном интервале вычисляются как отношение числа увольнений к средним размерам системы, т. е.

$$\omega(T_1, T_2) = \int_{T_1}^{T_2} R(T) dT \left/ \left\{ \frac{1}{(T_2 - T_1)} \int_{T_1}^{T_2} N(T) dT \right\} \right., \quad (8.19)$$

где  $N(T) = \int_0^T M(x)dx + N(0)$  — общая численность системы к моменту времени  $T$ .

Для того чтобы проиллюстрировать влияние развития системы на оценку текучести кадров, рассмотрим два примера. Допустим, что распределение полного срока службы — смесь экспоненциальных распределений, тогда мы уже знаем, что

$$h(T) = \mu^{-1} + ae^{-bT},$$

где

$$a = p\lambda_1 + (1 - p)\lambda_2 - \mu^{-1}, \quad b = p\lambda_2 + (1 - p)\lambda_1$$

и

$$\mu = p/\lambda_1 + (1 - p)/\lambda_2.$$

В нашем первом примере предположим, что численность системы линейно возрастает по закону

$$M(T) = M,$$

где  $M$  — скорость роста. Тогда функция  $R(T)$  может быть получена интегрированием непосредственно из (8.18). Подставляя результат в (8.19), находим

$$\omega(T_1, T_2) = \frac{T_2 - T_1}{\mu} + \frac{2a}{b(T_1 + T_2)} \left\{ T_2 - T_1 - \left( \frac{e^{-bT_1} - e^{-bT_2}}{b} \right) \right\}. \quad (8.20)$$

В табл. 8.4 приведены некоторые результаты вычислений по формуле (8.20) для трех видов распределений полного срока службы из семейства смесей экспоненциальных распределений. Значения параметров выбраны таким образом, чтобы распределения согласовывались с распределением данных по Гласнерской металлургической компании, фирме «Бибби и сыновья» и объединенной сталелитейной компании.

**Таблица 8.4. Увольнения по кварталам для групп, которые увеличиваются с постоянной скоростью, %**

Квартал	Гласнерская металлургическая компания	Фирма «Бибби и сыновья»	Объединенная сталелитейная компания
1	23,5	35,8	49,6
2	21,3	32,9	40,5
3	19,5	30,2	33,5
4	18,0	28,0	28,6
5	16,8	26,2	25,2
6	15,9	24,6	22,7
7	15,1	23,3	20,9
8	14,5	22,2	19,4

Если сравнить эти значения с данными табл. 8.1, то мы увидим, что текучесть кадров еще уменьшается, но не так быстро, как раньше. Из (8.20) видно, что  $\omega(T_1, T_2)$  не зависит от  $M$ , следовательно, число увольнений приближается к некоторому устойчивому значению независимо от скорости развития системы.

В предыдущем примере применялась непрерывная функция роста. Во втором примере примем, что  $N(T)$  изменяется скачком от первоначального значения  $N(0)$  к новому значению  $N(0) + N$  в момент

времени  $T = T_J$ . Очевидно, что в этом случае среднее число увольнений равно:

$$\bar{n}(T) = N(0) \int_0^T h(x) dx \quad (T \leq T_J),$$

$$\bar{n}(T) = N(0) \int_0^T h(x) dx + N \int_0^{T-T_J} h_2(x) dx \quad (T > T_J).$$

Некоторые значения текучести кадров для такого вида развития приведены в табл. 8.5.

**Таблица 8.5. Увольнения по кварталам для фирмы, увеличивающейся от  $N(0)$  до  $N(0) + N$  в момент времени  $T = T_J$ , когда предполагается, что применяется распределение полного срока службы объединенной сталелитейной компании, %**

Квартал	$T_J = 1/2$ года		$T_J = 1$ год	
	$N = \frac{1}{2} N(0)$	$N = N(0)$	$N = \frac{1}{2} N(0)$	$N = N(0)$
1	45,8	45,8	45,8	45,8
2	28,0	28,0	28,0	28,0
3	27,9	32,4	19,0	19,0
4	18,9	21,2	14,4	14,4
5	14,4	15,6	23,3	29,0
6	12,1	12,7	16,7	19,5
7	11,0	11,3	13,3	14,7
8	10,4	10,6	11,5	12,3

Вычисления произведены для двух значений как  $T_J$ , так и  $N$  при предположении, что распределение полного срока службы аналогично распределению в объединенной сталелитейной компании.

Влияние резкого увеличения размеров системы заключается в задержке уменьшения текучести кадров. При значительном увеличении размеров временно увеличивается число увольнений, но этот эффект непродолжителен и через год он почти перестанет сказываться.

### Сокращающиеся системы

Мы рассмотрели развивающиеся системы, которые получены суммированием отдельных процессов, но нельзя произвести расчеты в обратном направлении, так как будущее поведение зависит от процессов, которые не восстанавливаются. Мы не будем разрабатывать теорию сокращающихся систем, а вместо этого рассмотрим основополагающий случай, когда нужно определить максимальную скорость сокращения системы. Очевидно, что она достигается при полном прекращении найма. Задача заключается в том, чтобы определить ожидаемый размер системы как функцию времени, когда

наим полностью прекращен. Примем такой момент за начало отсчета времени. Пусть в этот момент численность системы равна  $N(0)$ . Через интервал времени  $T$  ожидаемый размер системы будет равен:

$$N(T) = N(0) \{1 - D(T)\},$$

где  $D(T)$  — функция распределения времени, по истечении которого выбранный в момент времени  $T = 0$  случайным образом индивидуум уволится из системы. Это так называемое распределение остаточного времени обслуживания. Для индивидуума, имеющего в момент времени  $T = 0$  стаж  $x$ , функция распределения имеет вид

$$D(T|x) = \{F(T+x) - F(x)\} / \{1 - F(x)\}, \quad (8.21)$$

где  $F(\cdot)$  — функция распределения полного стажа работы. Если первоначальная плотность функции распределения стажа  $a(x)$ , то

$$D(T) = \int_0^{\infty} a(x) D(T|x) dx. \quad (8.22)$$

Функция распределения стажа может быть получена, как было описано ранее, или оценена эмпирически. Пока будем считать, что это какая-то произвольная функция. Обычно принято выражать желательную скорость сокращения системы через относительную. В случае непрерывного времени обозначим численность системы через

$$N(T) = N(0) e^{-\alpha T} \quad (0 < \alpha < \infty). \quad (8.23)$$

Возникает важный практический вопрос, насколько велико должно быть значение  $\alpha$ , чтобы не было необходимости сокращения штатов. Другими словами, нужно найти наибольшее значение  $\alpha$  для

$$N(0) e^{-\alpha T} > N(0) \{1 - D(T)\} \quad (8.24)$$

при всех  $T$ .

Когда распределение полного срока службы экспоненциальное,  $D(T) = e^{-\lambda T}$ , неравенство выполняется, если  $\alpha \leq \lambda$ . Таким образом, в сокращении штатов нет необходимости, если скорость убывания системы меньше, чем текучесть кадров. С первого взгляда этот результат очевиден, и можно предположить, что его можно распространить на распределение полного срока службы любого вида. Чтобы показать, что это заблуждение, рассмотрим в качестве распределения полного срока службы смесь экспоненциальных распределений. Подставляя функцию плотности в (8.21), а затем в (8.22), находим

$$\begin{aligned} 1 - D(T) &= e^{-\lambda_1 T} \int_0^{\infty} a(x) \frac{\rho e^{-\lambda_1 x}}{\rho e^{-\lambda_1 x} + (1-\rho) e^{-\lambda_2 x}} dx + \\ &+ e^{-\lambda_2 T} \int_0^{\infty} \frac{a(x) (1-\rho) e^{-\lambda_2 x}}{\rho e^{-\lambda_1 x} + (1-\rho) e^{-\lambda_2 x}} dx, \\ &= A e^{-\lambda_1 T} + B e^{-\lambda_2 T}, \text{ скажем, где } A + B = 1, A, B \geq 0. \end{aligned}$$



В этом случае неравенство (8.24) выполняется для любого  $T$ , если  $\alpha \leq \min(\lambda_1, \lambda_2)$ . Мы уже получали текучесть кадров в установившемся режиме. Она равна  $\{p/\lambda_1 + (1-p)/\lambda_2\}^{-1}$ , что всегда меньше  $\min(\lambda_1, \lambda_2)$ . Следовательно, невозможно уменьшить систему со скоростью больше той, которая определяется текучестью кадров. Для некоторых распределений полного срока службы таких, как тип XI, функция  $\{1-D(T)\}$  уменьшается медленнее, чем любая экспонента, поэтому при заданной скорости убывания системы рано или поздно непременно потребуются сокращение штатов.

### 8.3. ИЗМЕРЕНИЕ ОБОРОТА ТРУДОВЫХ РЕСУРСОВ

Мы уже отмечали применительно к другим случаям роль стохастических моделей в создании адекватных мер социальных явлений. Один из простейших примеров этого — задача измерения оборота трудовых ресурсов. Оборот трудовых ресурсов определяется как численность поступающих и увольняющихся. Поскольку в настоящей главе мы пользуемся моделями восстановления, в которых поступление бывает только тогда, когда есть увольнения, найм у нас эквивалентен увольнению. Будем считать эти понятия аналогичными и предположим, что все измерения сводятся к измерению только потоков увольнений. Иногда говорят об измерении стабильности, которая, естественно, по смыслу противоположна обороту кадров. Но эти два понятия эквивалентны: высокая стабильность означает низкую интенсивность увольнений и наоборот.

Измерения оборота трудовых ресурсов необходимы по двум причинам. Во-первых, это важная составная часть моделей, описывающих кадровые системы. Например, ранее в этой главе мы связывали оценки текучести кадров с распределением индивидуумов системы по стажу. Во-вторых, оборот служит показателем морального климата в организации или показателем эффективности использования кадров. Именно с этих позиций мы будем рассматривать оборот в данном разделе. Высокий оборот кадров обычно связывают с неблагоприятным моральным климатом, что обуславливает низкую эффективность деятельности организации. Существует обширная литература, посвященная тому, как измерять и интерпретировать оборот кадров. Хотя было предложено множество мер, все они оказались противоречивыми, и в этом вопросе все еще существует неясность. Некоторые из работ были перечислены в гл. 7 в связи с обсуждением моделей процессов выбытия. К ним можно добавить статьи Боуэя (1969), Бибби (1970), Хаймана (1970), Форбеса (1971b), Ван дер Мерве и Миллера (1971), а также Клоуэса (1972). Здесь мы хотим внести ясность в введущие споры, используя результаты стохастического анализа. Измерение, конечно, не может автоматически стать хорошим только потому, что оно основано на вероятностной модели. Неподходящая модель приводит к ошибочным выводам, как показано у Бартоломью (1971) на примере работы Стойкова (1971).

Вероятно, наиболее ранней мерой и определенно одной из наиболее частых мер оборота кадров является оценка, определяемая как

$$I = \frac{\text{число покинувших систему за данный интервал времени}}{\text{среднее число работающих за этот интервал}}. \quad (8.25)$$

Неадекватность такого показателя без труда определяется на основе анализа, проведенного в разделе 8.2, где этот показатель использовался, чтобы показать, как влияет на уход возраст системы. Мера, которая так сильно зависит от распределения по стажу, не может адекватно отражать другие показатели до тех пор, пока система не достигла установившегося режима. Таким образом, разница в оценках интенсивностей для двух фирм может просто отражать тот факт, что одна из фирм более «молодая» с точки зрения распределения служащих по возрасту. Лане и Андрю (1955) приводят замечательную иллюстрацию для ситуации такого рода по данным для двух фирм. В одной из них большее число увольнений, хотя в ней и большее ожидаемое время службы.

В демографии и при страховании аналогичная ситуация возникает при измерении коэффициента смертности, где «западня» при использовании грубых коэффициентов хорошо видна. Например, на благоприятных для здоровья приморских курортах высокая смертность не потому, что их положительное влияние сомнительно, а потому что они привлекательны как место жительства для непропорционально большого количества пожилых людей. Демографы и работники страхования, сталкивающиеся с подобными проблемами, пользуются стандартизованными коэффициентами смертности, позволяющими оценивать, какова была бы смертность, если бы изучаемое население имело стандартное распределение по возрасту. То же самое мы имеем, когда измеряем оборот кадров. Предположим, мы получили стандартное распределение по стажу с функцией плотности  $s(x)$ . Тогда ожидаемое число тех, кто покинет систему за время  $(T, T + \delta T)$  из группы по стажу  $(x, x + \delta x)$ , равно:

$$N(T) s(x) \delta x \lambda(x) \delta T,$$

где  $N(T)$  — численность системы в момент времени  $T$ , а  $\lambda(x)$  — стремление к уходу, одинаковое для всех индивидуумов системы со стажем  $x$ . Таким образом, ожидаемое общее число тех, кто покинет систему за интервал времени  $(T_1, T_2)$ , равно:

$$\int_{T_1}^{T_2} N(T) dT \int_0^{\infty} s(x) \lambda(x) dx, \quad (8.26)$$

где опять  $N(T)$  считается непрерывной переменной. Среднее число индивидуумов, находящихся в системе в период времени  $(T_1, T_2)$ , равно:

$$\int_{T_1}^{T_2} N(T) dT / (T_2 - T_1),$$

поэтому оценка коэффициента текучести для заданной функции  $s(x)$  равна:

$$I = (T_2 - T_1) \int_0^{\infty} s(x) \lambda(x) dx. \quad (8.27)$$

Это выражение показывает, как оценка коэффициента текучести зависит от структуры по стажу и стремления к уходу. Используя стандартную форму распределения по стажу, мы исключаем изменения, определяемые ею, и получаем меру, зависящую только от стремления к уходу.

Наиболее естественная форма  $s(x)$  — устойчивое распределение по стажу, которое было получено для системы постоянного размера в виде (8.16). Это показывает, что оценка для интенсивности уходов должна быть взята в установившемся режиме. Осуществляя перестановку, получаем

$$I = (T_2 - T_1) \int_0^{\infty} \frac{1}{\mu} G(x) \lambda(x) dx = \frac{T_2 - T_1}{\mu} \int_0^{\infty} f(x) dx = \frac{T_2 - T_1}{\mu}. \quad (8.28)$$

Таким образом, стандартизованный коэффициент текучести эквивалентен среднему значению полного срока службы, т. е. мере, предложенной Лане и Андрию (1955).

Предыдущий анализ нужен был для того, чтобы подчеркнуть очевидный вывод, а именно: мера увольнений должна зависеть только от склонности к уходу. Она выражается через  $\lambda(x)$  или через эквивалентные функции  $f(x)$  и  $G(x)$ . «Лучший» способ описания процесса увольнений заключается в оценке одной или нескольких перечисленных функций. В настоящее время, когда ЭВМ снабжены графическими терминалами, графики этих функций легко получить на экране дисплея или изобразить на бумаге. Такую диаграмму понять и точно интерпретировать легче, чем какой-либо коэффициент. Однако можно исследовать, как содержащаяся в  $\lambda(x)$  информация выразить обобщенно через одно число. Это можно сделать через среднее значение полного срока службы, о чем уже говорилось в связи с рассмотрением стандартизованного коэффициента смертности. Видимо, медиана имеет свои привлекательные свойства, кроме того, ее легко вычислять и интерпретировать. Также простой мерой является доля тех, кто прослужит, как можно ожидать, некоторый заданный период времени.

В гл. 7 нам не удалось найти параметра распределения, с помощью которого можно было бы описать данные по полному сроку службы. Следовательно, любая попытка выразить одним числом всю информацию, заключенную в  $\lambda(x)$ , приведет к неадекватности. Например, для двух распределений будут одинаковые доли тех, кто работает, скажем, один год, но они вполне могут различаться в других точках. Меньший риск потерять важную информацию, если использовать два показателя, каждый из которых выражает различную особенность распределения. Два параметра  $\omega$  и  $\sigma$  эмпирического логнормального распределения содержат наиболее важную информацию о распределении полного сро-

ка службы, хотя их не легко интерпретировать. Возможно, наиболее простой парой мер являются значения  $G(x_1)$  и  $G(x_2)$  для двух соответствующим образом выбранных величин длительности службы  $x_1$  и  $x_2$  ( $x_1$  может быть моментом окончания обучения, а  $x_2$  — среднее время, когда индивидуум готов для продвижения).

В настоящем анализе, так же как и в гл. 7, считалось, что уходы зависят только от длительности пребывания. В действительности  $\lambda(x)$  является функцией, зависящей от многих других факторов, таких, как пол, место жительства, уровень квалификации и т. п. Следовательно, очевидное различие двух фирм по указанным мерам может стать результатом различного соотношения между числом мужчин и числом женщин или определяться другими факторами. Необходимо подчеркнуть, что показатели, основанные на измерении длительности службы, можно правильно применять только в том случае, когда группы однородны относительно других факторов, влияющих на процесс уходов. При сравнении разнородных групп показатели должны быть стандартизованы. Это легко сделать так же, как при исследовании стажа.

#### 8.4. МОДЕЛИ МНОГОУРОВНЕВЫХ СИСТЕМ

Если в системе имеется несколько градаций, то нет необходимости заполнять вакансии вновь поступающими. Вместо этого может быть перемещен индивидуум, занимающий некоторый другой уровень, и тогда образуется новая вакансия и т. д. В этом разделе мы рассмотрим только системы, в которых имеется простая иерархия ступеней, такая, что вакансии или заполняются нанятыми, или замещаются индивидуумами, находящимися на ближайшей расположенной ниже ступени. Насколько это возможно, мы сохраним обозначения и терминологию из предыдущих глав. Так как численности на ступенях считаются постоянными, обозначим их через  $n_1, n_2, \dots, n_k$ ,  $\sum_{i=1}^k n_i = N$  и опустим аргумент «время». Постоянство численностей на ступенях определяет некоторые ограничения на потоки, которые имеются во всех моделях. Таким образом, рассмотрим произвольный интервал времени единичной длины и определим случайные величины, представляющие потоки следующим образом:

$$\left. \begin{aligned} R_i & \text{— число вновь поступивших на } i\text{-ю} \\ & \text{ступень,} \\ W_i & \text{— число уволившихся с } i\text{-й ступени,} \\ P_i & \text{— число тех, кто перемещен с } i\text{-й} \\ & \text{ступени } (i = 1, 2, \dots, k - 1). \end{aligned} \right\} (i = 1, 2, \dots, k)$$

Так как размеры каждой ступени заданы, общий входной поток должен быть сбалансирован с выходным потоком. Это запишется как

$$P_{i-1} + R_i = P_i + W_i \quad (i = 1, 2, \dots, k), \quad (8.29)$$

где  $P_0 = P_k = 0$ . Отсюда следует, что средние этих случайных величин должны удовлетворять тем же уравнениям. Далее проанализируем изменение процесса, рассматривая достаточно малый интервал времени. Для этих целей определим следующие функции:

$n_i \omega_i (T) \delta T$  — среднее число увольнений с  $i$ -й ступени за интервал времени  $(T, T + \delta T)$ ;

$n_i h_i (T) \delta T$  — ожидаемое число продвижений с  $i$ -й ступени за интервал времени  $(T, T + \delta T)$ ;

$R_i (T) \delta T$  — ожидаемое число вновь нанятых на  $i$ -ю ступень за интервал времени  $(T, T + \delta T)$ .

При таких обозначениях можно считать, что  $\omega_i (T)$  и  $h_i (T)$  — интенсивности, которые выражают склонность к уходу из системы или склонность к продвижению в момент времени  $T$ . Непосредственно из (8.29) имеем

$$n_{i-1} h_{i-1} (T) + R_i (T) = n_i h_i (T) + n_i \omega_i (T) \quad (i = 1, 2, \dots, k), \quad (8.30)$$

где  $h_0 (T) = h_k (T) = 0$ .

### Модель с заданными интенсивностями уходов

Любая модель определяется двумя факторами: механизмом уходов и способом заполнения вакансий. В нашей первой модели мы полагаем, что каждый индивидуум имел постоянную склонность к уходу из системы, зависящую только от ступени. Положим, таким образом,

$$\omega_i (T) = \lambda_i \quad (i = 1, 2, \dots, k) \quad (8.31)$$

и будем считать, что поведение одних индивидуумов не зависит от поведения других. Эти предположения аналогичны тем, которые были сделаны в гл. 5 для марковских моделей с непрерывным временем. Когда есть уходы с  $i$ -й ступени ( $i > 1$ ), то предполагается, что вакансии с вероятностью  $s_i$  заполняется тем, кто находится на ступеньку ниже; в другом случае она замещается вновь нанятым. Такое предположение подтверждается практикой. Все вакансии ступени 1 (самой начальной) заполняются вновь поступившими. Предположение о характере увольнений и (8.30) позволяют записать

$$R_i (T) = n_i h_i (T) - n_{i-1} h_{i-1} + n_i \lambda_i. \quad (8.32)$$

В соответствии со вторым допущением

$$R_i (T) = (1 - s_i) \{n_i h_i (T) + n_i \lambda_i\}. \quad (8.33)$$

Исключая  $R_i (T)$  из (8.32) и (8.33), получаем следующую рекурсивную формулу для интенсивностей продвижений:

$$n_{i-1} h_{i-1} (T) = s_i n_i h_i (T) + s_i n_i \lambda_i,$$

откуда

$$n_i h_i (T) = \sum_{j=i+1}^k n_j \lambda_j \prod_{r=i+1}^j s_r. \quad (8.34)$$

В этой модели ни  $h_i (T)$ , ни  $R_i (T)$  не зависят от  $T$ . Проиллюстрируем применение этих формул на примере двух частных случаев.

Случай 1.  $\lambda_i = \lambda$  ( $i = 1, 2, \dots, k$ ).

Этот случай эквивалентен одноуровневой модели с экспоненциальным распределением времени жизни, что относится как ко входному, так и выходному потокам. Тогда для интенсивностей продвижений записываем

$$h_i = \frac{\lambda \sum_{j=i+1}^k n_j}{n_i} \prod_{r=i+1}^j s_r \quad (i = 1, 2, \dots, k-1). \quad (8.35)$$

Из этого выражения видно влияние численностей на ступенях на каждую интенсивность продвижений. Интенсивности изменяются прямо пропорционально общей численности на более высоких ступенях и обратно пропорционально рассматриваемой ступени. Кроме того, интенсивность пропорциональна каждой вероятности  $s_r$ , относящейся к более высокой ступени. Если найм возможен лишь на самую низкую ступень, то  $s_1 = 0$ ,  $s_i = 1$  ( $i > 1$ ) и

$$h_i = \lambda (N - N_i)/n_i \quad (i = 1, 2, \dots, k-1), \quad (8.36)$$

где  $N_i = \sum_{j=1}^i n_j$ .

Допустим, что возникает вопрос, какой должна быть структура ступеней, чтобы интенсивности переходов были равными. Из уравнения (8.36) следует, что

$$n_i \propto (N - N_i) \quad (i = 1, 2, \dots, k-1),$$

а из этого вытекает

$$n_{k-1} = cn_k, \quad n_i = n_{i+1} (1 + c) \quad (i = 1, 2, \dots, k-1). \quad (8.37)$$

Начиная с высшей ступени, их численности образуют геометрическую прогрессию. Аналогичный результат был получен в гл. 5 для систем с заданным поступлением и случайными размерами ступеней. Там мы показали, что если интенсивности продвижений равны, то средние численности на ступенях были членами геометрической прогрессии. Этот результат дает основания считать, что марковские модели и модели восстановления эквивалентны, если их интерпретировать детерминистским образом, т. е. когда случайные величины заменяются их математическими ожиданиями.

Случай 2.  $s_i = s$  ( $i = 2, 3, \dots, k$ ).

В этом случае имеем

$$n_i h_i = \sum_{j=i+1}^k n_j \lambda_j s^{j-i}. \quad (8.38)$$

Допустим еще, что  $\lambda_j = \lambda s^{-j}$ , тогда

$$h_i = \lambda (N - N_i)/n_i s^i.$$

Если бы  $n_i \propto (N - N_i)$ , то интенсивности продвижений увеличивались бы по мере продвижения по иерархии как члены геометрической прогрессии. Это непосредственное следствие того, что на более высоких

ступенях иерархии уходы имеют большую величину. Если, с другой стороны,  $\lambda_i = \lambda$ , то

$$h_i = \lambda \sum_{j=i+1}^k n_j s^j / n_i s^i .$$

Для таких систем, чтобы обеспечить равные интенсивности продвижений, необходимо выполнение следующих отношений, характеризующих их структуры:

$$n_i s^i \propto \sum_{j=i+1}^k n_j s^j .$$

Это снова приводит к геометрической структуре, хотя множитель прогрессии равен  $s(1+c)$ , а не  $(1+c)$ . Другими словами, структура такова, что увеличение численности с повышением ступени иерархии происходит быстрее, чем в случае, когда нет набора на верхние ступени.

### Распределение потоков

Теперь мы в состоянии показать, что все потоки имеют пуассоновское распределение, если только все интенсивности увольнений постоянны. Если для каждого из  $n_i$  индивидуумов  $i$ -й ступени интенсивность ухода равна  $\lambda_i$ , то число уходов из этой градации распределено по закону Пуассона\* со средним  $n_i \lambda_i$ . Для того чтобы получить распределение оставшихся потоков, необходимо воспользоваться сделанным ранее предположением о способе заполнения вакансий. Согласно ему условное распределение  $P_i$  для данных  $P_i + W_i$  — биномиальное распределение с параметрами  $P_i + W_i$  и  $s_i$ . Чтобы полностью найти распределение, нам понадобится следующая лемма.

*Лемма.* Если случайная величина  $X$  распределена по закону Пуассона с параметром  $\mu$  и условное распределение  $Y$  при заданной величине  $X$  биномиально с параметрами  $(X, p)$ , то величина  $Y$  распределена по закону Пуассона с параметрами  $(p\mu)$ .

Начнем с  $k$ -й градации, для которой  $P_{k-1} + R_k = W_k$ . В соответствии с леммой  $X = W_k$ ,  $p = s_k$ . Мы получаем, что  $P_{k-1}$  распределено по закону Пуассона с параметром  $(n_k \lambda_k s_k)$  и аналогично  $R_k$  с параметром  $\{\lambda_k (1 - s_k)\}$ . Переходя к следующей  $(k-1)$ -й градации, имеем  $P_{k-2} + R_{k-1} = W_{k-1} + P_{k-1}$ . Полагая  $X = W_{k-1} + P_{k-1}$ ,  $Y = P_{k-2}$  и отмечая, что сумма  $W_{k-1} + P_{k-1}$  будет распределена по закону Пуассона как сумма двух независимых переменных, распределенных по закону Пуассона, согласно лемме получим, что  $P_{k-2}$  распределено по закону Пуассона с параметром  $\{s_{k-1} (n_k \lambda_k s_k + n_{k-1} \lambda_{k-1})\}$ . Рассматривая сверху вниз все ступени иерархии аналогичным способом, мы определим, что потоки поступлений и продвижений распределены по закону Пуассона. Ограничения (8.29) показывают, что потоки продвижений и найма являются зависимыми.

\* Начиная с этого места и до конца раздела все рассуждения автора носят приближенный характер, так как распределение было бы пуассоновским, если бы число уходов не было ограничено  $n_i$ . — *Примеч. ред.*

Обоснованность наших доводов не зависит от того, на каком временном интервале рассматриваются потоки и каков возраст системы. Это прямое следствие допущения о постоянстве интенсивностей уходов. В общем случае, если интенсивности уходов зависят от стажа индивидуума или уровня иерархии, оба эти фактора должны учитываться, и вывод сделать гораздо труднее.

### Длительность пребывания

Другой важный аспект систем восстановления — длительность пребывания индивидуумов как на отдельных ступенях иерархии, так и в системе в целом. Как уже отмечалось, последнюю величину можно сравнить с другими распределениями длительностей пребывания, которые можно получить с помощью моделей из гл. 7.

Индивидуум с  $i$ -й ступени уходит по двум причинам. Одна — увольнение с постоянной интенсивностью  $\lambda_i$ , а другая — продвижение. Вначале мы рассмотрим, как определить функцию дожития индивидуума, попавшего на  $i$ -ю ступень, если на него действует только одна из причин. Так как причины по сделанному ранее предположению независимы друг от друга, функция дожития индивидуума при действии обеих причин получается как произведение двух функций дожития, найденных для каждой из причин порознь. Если нет продвижения, то, как следует из предположения о постоянстве интенсивностей уходов, функция дожития будет равна  $e^{-\lambda_i \tau}$ , где  $\tau$  выражает длительность пребывания в градации.

Время ожидания продвижения при отсутствии уходов зависит от того, как выбирают тех, кого необходимо продвигать. Рассмотрим два правила продвижения. Первое заключается в случайном выборе тех, кто может быть повышен; второе предполагает, что продвигается тот, кто дольше пробыл на ступени. Если выбор осуществляется случайным образом, мы определим вначале функцию дожития при условии, что на интервале  $(0, \tau)$  было  $P_i$  продвижений. При продвижении вероятность того, что выбран какой-либо конкретный индивидуум, всегда равна  $1/n_i$ . Следовательно, вероятность того, что конкретный индивидуум не продвинут, когда осуществлено  $P_i$  продвижений, равна

$$\left(1 - \frac{1}{n_i}\right)^{P_i}.$$

Безусловную вероятность можно определить на основании того факта, что случайная величина  $P_i$  распределена по закону Пуассона со средним  $n_i h_i \tau$ . Таким образом, вероятность пробыть на  $i$ -й ступени до момента времени  $\tau$  при условии, что нет уходов, равна:

$$Pr = e^{-n_i h_i \tau} \sum_{P_i=0}^{\infty} \left(1 - \frac{1}{n_i}\right)^{P_i} \frac{(n_i h_i \tau)^{P_i}}{P_i!} = e^{-h_i \tau}. \quad (8.39)$$



Перемножая две функции дожития, получаем, что вероятность пробыть на  $i$ -й ступени до момента времени  $\tau$  равна:

$$Pr = e^{-(h_i + \lambda_i) \tau}. \quad (8.40)$$

Это означает, что длительность пребывания на  $i$ -й ступени — случайная величина, распределенная по экспоненциальному закону со средним  $(h_i + \lambda_i)^{-1}$ .

Допустим теперь, что продвигается индивидуум, который дольше всех пробыл на ступени. Поступая на  $i$ -ю ступень (не важно, каким образом), индивидуум впереди себя имеет в очереди  $(n_i - 1)$  человек. Чтобы продвинуться, он должен вначале ждать, пока все они уйдут со ступени, а затем ждать, когда появится новая вакансия. Вакансии образуются на  $i+1$  ступени как пуассоновский процесс с параметром интенсивности  $n_i h_i$ , кроме того, каждый индивидуум может уйти с интенсивностью  $\lambda_i$ . Таким образом, процесс, когда  $n_i - 1$  человек покидают  $i$ -ю ступень, — это простейший процесс гибели с интенсивностью гибели, когда остается  $m$  индивидуумов, равной:

$$\begin{aligned} \mu_0 &= 0, \\ \mu_m &= n_i h_i + \lambda_i m \quad (m = 1, 2, \dots, n_i - 1). \end{aligned} \quad (8.41)$$

Используя теорию процессов гибели (см., например, работу Балея (1964) или Феллера (1968)), можно найти вероятность того, что к моменту времени  $\tau$  остается  $m$  человек. Обозначим ее через  $P_m(\tau)$ . Вероятность  $P_0(\tau)$  равна дополнению дожития, так как, если в момент времени  $\tau$  никого не остается, период времени для всех, кто покидает систему, меньше  $\tau$ . В нашем случае дифференциально-разностные уравнения для вероятности  $P_m(\tau)$  записываются так:

$$\begin{aligned} P'_m(\tau) &= -(m \lambda_i + n_i h_i) P_m(\tau) + \overline{(m+1) \lambda_i + n_i h_i} P_{m+1}(\tau) \\ &\quad (m = 1, 2, \dots, n_i - 2), \\ P'_0(\tau) &= (\lambda_i + n_i h_i) P_1(\tau), \\ P'_{n_i-1}(\tau) &= -\{(n_i - 1) \lambda_i + n_i h_i\} P_{n_i-1}(\tau) \end{aligned} \quad (8.42)$$

с начальными условиями  $P_{n_i-1}(0) = 1$ ,  $P_m(0) = 0$  ( $m < n_i - 1$ ). Будем решать эти уравнения, используя преобразования Лапласа для этих вероятностей. Чтобы произвести преобразование Лапласа, следует умножить обе части на  $e^{-st}$  и проинтегрировать от 0 до бесконечности. Известно также, что изображение производной  $P'_m(\tau)$  равно  $sP_m^*(s) - P_m(0)$ , если учесть начальные данные. Отсюда следует, что

$$\begin{aligned} \{s + m \lambda_i + n_i h_i\} P_m^*(s) &= \overline{\{m+1\} \lambda_i + n_i h_i} P_{m+1}^*(s) \quad (m = \\ &\quad = 1, 2, \dots, n_i - 2), \\ sP_0^*(s) &= (\lambda_i + n_i h_i) P_1^*(s), \quad sP_{n_i-1}^*(s) = \{(n_i - 1) \lambda_i + n_i h_i\} P_{n_i-1}^*(s). \end{aligned} \quad (8.43)$$

Начав с  $P_{n_i-1}^*(s)$ , из этих уравнений получаем:

$$P_0^*(s) = \frac{1}{s} \prod_{j=1}^{n_i-1} \left[ \frac{j\lambda_i + n_i h_i}{j\lambda_i + n_i h_i + s} \right]. \quad (8.44)$$

Это изображение функции распределения времени ожидания. Изображение функции плотности —  $sP_0^*(s)$ . После того как  $n_i - 1$  индивидуумов покинуло ступень, нужно ждать следующую вакансию. Это время ожидания распределено по экспоненциальному закону с параметром  $n_i h_i$ . Таким образом, сумма двух времен ожидания имеет изображение

$$sP_0^*(s) \left( \frac{n_i h_i}{n_i h_i + s} \right).$$

Объединяя это выражение с (8.44), получаем, что при отсутствии уходов стаж работы имеет изображение

$$\prod_{j=0}^{n_i-1} \frac{j\lambda_i + n_i h_i}{j\lambda_i + n_i h_i + s}. \quad (8.45)$$

Последнее представляет собой изображение распределения Эрланга, которое может быть получено путем разложения на простые дроби и применения обратного преобразования Лапласа почленно. Однако, предполагая, что численность градаций достаточно большая, покажем вырождение функции плотности, что означает постоянство в пределе времени ожидания. Для этого возьмем натуральный логарифм (8.45) и допустим, что  $n_i$  велико, тогда

$$-\sum_{j=0}^{n_i-1} \ln \left[ 1 + \frac{s}{j\lambda_i + n_i h_i} \right] \sim -\frac{s}{\lambda_i} \sum_{j=0}^{n_i-1} \frac{1}{j + n_i h_i / \lambda_i}. \quad (8.46)$$

Последняя сумма просто выражается функцией

$$\Phi(x) = \sum_{i=1}^{\infty} \frac{x}{i(i+x)}$$

следующим образом:

$$-\frac{s}{\lambda_i} \{ \Phi(n_i - 1 + n_i h_i / \lambda_i) - \Phi(-1 + n_i h_i / \lambda_i) \}.$$

При больших  $x$   $\Phi(x) \sim \ln x + 0,5772 \dots$  и, следовательно, (8.46) в пределе равно

$$-\frac{s}{\lambda_i} \ln \left[ \frac{h_i + \lambda_i}{h_i} \right].$$

Таким образом, в пределе распределение времени ожидания до продвижения имеет изображение Лапласа

$$\exp \left\{ -\frac{s}{\lambda_i} \ln \left[ \frac{h_i + \lambda_i}{h_i} \right] \right\}. \quad (8.47)$$

Очевидно, что это изображение распределения, в котором вся вероятность сконцентрирована в одной точке

$$\tau_i = \frac{1}{\lambda_i} \ln \left[ \frac{h_i + \lambda_i}{h_i} \right], \quad (8.48)$$

где  $h_i$  определяется по (8.34). Следовательно, при соблюдении правила продвижения по стажу работы длительность пребывания для тех, кто не покидает систему, в пределе постоянна. Таким образом, множество значений  $\{\tau_i\}$  представляет типовую карьеру тех, кто остается в фирме.

Но каждый индивидuum может убыть со ступени  $i$ , на которой он находится, по двум причинам, поэтому его функция дожития имеет вид  $e^{-\lambda_i \tau}$  при  $0 \leq \tau < \tau_i$  и 0 при  $\tau \geq \tau_i$ .

Предыдущие результаты могут быть использованы для определения распределения длительности службы для индивидуума, поступающего в систему на нижнюю ступень иерархии. (Для тех, кто поступает на более высокие ступени, расчеты аналогичные, но с меньшим числом градаций.) Пусть  $G(T)$  — функция дожития для индивидуума, поступающего на нижнюю ступень иерархии, тогда можно записать:

$$G(T) = \sum_{i=1}^k Pr \{ \text{служить до момента времени } T \text{ и в этот момент находиться на } i\text{-й ступени} \}. \quad (8.49)$$

Для того чтобы находиться в момент времени  $T$  на  $i$ -й ступени, необходимо продвинуться  $i - 1$  раз и не уйти из системы. Допустим, что эти продвижения происходят в моменты времени  $T_1, T_2, \dots, T_{i-1}$ , и вначале найдем условные вероятности того, что индивидuum не уйдет из системы до этих моментов времени. На интервале времени от  $T_j$  до  $T_{j+1}$  интенсивность ухода индивидуума равна  $\lambda_j$ . Следовательно, вероятность того, что он прослужит до момента времени  $T$  и не уйдет, равна:

$$Pr = \exp \left\{ - \sum_{j=0}^{i-1} \lambda_j (T_{j+1} - T_j) \right\}, \quad (8.50)$$

где  $T_0 = 0, T_i = T$ . Чтобы найти безусловную вероятность, нам необходимо определить плотность функции распределения  $T_1, T_2, \dots, \dots, T_{i-1}$  и вероятность того, что нет продвижения в интервале времени  $T - T_{i-1}$ . Таким образом,

$$b_1(T_1) b_2(T_2 - T_1) \dots b_{i-1}(T_{i-1} - T_{i-2}) \{1 - B_i(T - T_i)\}, \quad (8.51)$$

где  $b_i(\cdot)$  — плотность функции длительности пребывания на  $i$ -й ступени (уже найденная) и  $B_i(\cdot)$  — ее функция распределения. Требуемая вероятность получается умножением (8.50) и (8.51) и интегрированием по области  $0 \leq T_1 \leq T_2 \leq \dots \leq T_i \leq T$ . Полученный интеграл имеет вид свертки нескольких функций, поэтому все относительно сложные вычисления могут быть легко осуществлены, если

использовать преобразования Лапласа. Действительно, если интеграл имеет вид

$$\iint \dots \int a_1(x_1) a_2(x_2 - x_1) \dots a_{h-1}(X - x_h) dx_1 \dots dx_h, \\ 0 \leq x_1 \leq x_2 \dots x_h \leq X$$

то его изображение будет следующим:

$$a_1^*(s) a_2^*(s) \dots a_i^*(s).$$

В нашем случае

$$a_j(u) = b_j(u) e^{-\lambda_j u} \quad (j = 1, 2, \dots, i-1), \\ a_i(u) = \{1 - B_i(u)\} e^{-\lambda_i u},$$

поэтому требуемыми изображениями будут

$$a_j^*(s) = b_j^*(s + \lambda_j) \quad (j = 1, 2, \dots, i-1), \\ a_i^*(s) = \{1 - B_i^*(s + \lambda_i)\} / (s + \lambda_i).$$

И наконец, изображение  $G(T)$  имеет вид

$$G^*(s) = \sum_{i=1}^k \left\{ \frac{1 - b_i^*(s + \lambda_i)}{s + \lambda_i} \right\} \prod_{j=1}^{i-1} b_j^*(s + \lambda_j). \quad (8.52)$$

При использовании правила продвижения путем случайного отбора получаем из (8.39):

$$b_j^*(s + \lambda_j) = h_j / (\lambda_j + h_j + s) \quad \text{с } h_k = 0.$$

Подставляя это в (8.52) и используя тот факт, что

$$f^*(s) = 1 - sG^*(s),$$

находим

$$f^*(s) = \frac{\lambda_1}{\lambda_1 + h_1 + s} + \sum_{i=2}^k \left( \frac{\lambda_i}{\lambda_i + h_i + s} \right) \prod_{j=1}^{i-1} \left( \frac{h_j}{\lambda_j + h_j + s} \right) \quad (h_k = 0). \quad (8.53)$$

В правильности соотношения (8.53) можно убедиться, сведя изображение к  $\lambda / (\lambda + s)$ , если  $\lambda_j = \lambda$  для всех  $j$ , и к 1, если  $s = 0$ . Правую часть (8.53) можно записать в виде суммы простых дробей и, делая обратное преобразование Лапласа почленно, получим плотность функции распределения в следующем виде:

$$f(T) = \sum_{i=1}^k p_i (\lambda_i + h_i) e^{-(\lambda_i + h_i)T}, \quad (8.54)$$

$$\text{где } h_k = 0 \text{ и } \sum_{i=1}^k p_i = 1.$$

Это знакомая нам смесь экспоненциальных распределений, хотя, как мы увидим ниже, не все  $p_i$  должны быть положительными.

В простейшем случае, когда  $k = 2$ ,

$$p_1 = (\lambda_1 - \lambda_2) / (\lambda_1 - \lambda_2 + h_1) = n_1 (\lambda_1 - \lambda_2) / \{n_1 (\lambda_1 - \lambda_2) + n_2 \lambda_2\},$$

для которого выполняется  $0 \leq \rho_1 \leq 1$ , если  $\lambda_1 \geq \lambda_2$ . Таким образом, всякий раз, когда на первой ступени интенсивность уходов больше, чем на второй, мы будем иметь смесь экспоненциальных распределений, аналогичную тем, которые встречались ранее. Если описывать этим распределением реальные данные, типичные значения параметров получаются примерно такими:  $\rho_1 = \frac{1}{2}$  и  $(\lambda_1 + h_1)/\lambda_2 = 10$ . Для нахождения таких значений в данной модели необходимо, чтобы  $\lambda_1/\lambda_2 = 5,5$  и  $n_2/n_1 = 4,5$ . Если бы описанная модель отражала истинное положение дел, то мы не могли бы отделить ее от других по наблюдениям только общей продолжительности службы. Для того чтобы сделать это, необходимо после выделения двух ступеней проверить, являются ли интенсивности уходов в них постоянными величинами, и, если это так, посмотреть, соответствуют ли оцененные значения тем, которые предсказаны распределениями, согласующимися с экспериментальными данными по общим срокам службы.

При  $k > 2$  распределения полного срока службы можно взять в различных формах. Для того чтобы проиллюстрировать некоторые возможности, рассмотрим четыре примера для случая  $k = 3$ :

$$а) n_1 = n_2 = n_3, \lambda_j = \lambda/j \quad (j = 1, 2, 3),$$

$$f(T) = \frac{\lambda}{324} (209 e^{-11\lambda T/6} + 75 e^{-5\lambda T/6} + 40 e^{-\lambda T/3});$$

$$б) n_1 = n_2 = n_3, \lambda_j = \lambda/(4-j) \quad (j = 1, 2, 3),$$

$$f(T) = \frac{\lambda}{60} (209 e^{-11\lambda T/6} - 405 e^{-3\lambda T/2} + 216 e^{-\lambda T});$$

$$в) n_1 = 2n_2 = 3n_3, \lambda_j = \lambda/j \quad (j = 1, 2, 3),$$

$$f(T) = \frac{\lambda}{18} (14 e^{-4\lambda T/3} + 3 e^{-2\lambda T/3} + e^{-\lambda T/3});$$

$$г) n_1 = 2n_2 = 4n_3, \lambda_j = \lambda/(4-j) \quad (j = 1, 2, 3),$$

$$f(T) = \frac{\lambda}{6} (20 e^{-4\lambda T/3} - (18 - 9T) e^{-\lambda T}).$$

Слагаемое, включающее переменную  $T$  в случае (г), появляется вследствие того, что в этом примере  $\lambda_2 + h_2 = \lambda_3$ , поэтому обратное преобразование Лапласа, приводящее к (8.54), сделать нельзя. Чтобы такое затруднение преодолеть, следует перейти к пределу  $f(T)$  при  $\lambda_3 \rightarrow \lambda_2 + h_2$ , как и было сделано в соотношении (г).

В тех случаях, когда интенсивности уходов уменьшаются по мере продвижения вверх по иерархической лестнице, мы имеем смесь экспоненциальных распределений с положительными коэффициентами. Последнее показывает, кроме того, что полученные модели могут объяснить распределения полных сроков службы. Когда интенсивности уходов увеличиваются по мере продвижения вверх, распределение полного срока службы унимодально и менее асимметрично.

Аналогичным образом рассмотренный метод можно применять, когда правило продвижения — продвижение по стажу работы. Для

больших численностей на ступенях аппроксимация, представленная в виде (8.48), дает результат немедленно без преобразований Лапласа. Это следует из того, что время ожидания продвижения в пределе — постоянная величина и что опасность уходов для полного срока службы определяется соотношениями

$$\lambda(T) = \lambda_1, \quad 0 \leq T < \tau_1,$$

$$\lambda(T) = \lambda_i \text{ для } \sum_{j=1}^{i-1} \tau_j \leq T < \sum_{j=1}^i \tau_j \quad (i = 2, 3, \dots, k).$$

Тогда функцию плотности находят подстановкой в

$$f(T) = \lambda(T) \exp \left\{ - \int_0^T \lambda(x) dx \right\}, \quad T \geq 0.$$

Кривая состоит из последовательности экспоненциальных сегментов. Если интенсивности  $\lambda$  образуют убывающую последовательность, то по сравнению с экспоненциальной функция плотности будет асимметричнее, т. е. полученная функция будет больше экспоненциальной на концах области определения. В этом отношении она по форме аналогична смеси экспоненциальных распределений, которая получается для правила случайного продвижения. Следовательно, вообще говоря, правило продвижения не играет определяющей роли для формы распределения индивидуумов по ступеням. Но оно, конечно, сильно влияет на ожидаемую карьеру индивидуумов.

### Распределение по стажу

Как и при исследовании одноуровневой системы, мы можем найти распределение индивидуумов по стажу в любой момент времени. Это распределение будет зависеть от времени, но в точках, близких к устойчивому состоянию, оно легко может быть выражено через найденные уже распределения полного срока службы. Рассмотрим вычисления для обоих случаев.

Обозначим через  $a_i(t)$  равновесную функцию плотности распределения вероятностей стажа работы индивидуумов, поступивших в систему на  $i$ -ю ступень;  $a(t)$  — соответствующая плотность для всех индивидуумов системы. Ясно, что

$$Na(t) = \sum_{i=1}^k n_i a_i(t). \quad (8.55)$$

Когда действует правило продвижения по стажу, решение задачи получается почти сразу же. В предположении о достаточно больших численностях ступеней продвижения из  $i$ -й в  $i + 1$ -ю градацию происходят в моменты времени

$$t_i = \sum_{j=1}^i \tau_j \quad (i = 1, 2, \dots, k-1).$$

Следовательно, все занятые на ступени  $i$  имеют стаж от  $t_i$  до  $t_{i+1}$  и

$$a_i(t) = a(t) \int_{t_i}^{t_{i+1}} a(x) dx, \quad t_i \leq t < t_{i+1},$$

$$a_i(t) = Na(t)/n_i \quad (i = 1, 2, \dots, k), \quad (8.56)$$

где  $a(t) = \mu^{-1} G(t)$ ,  $t_0 = 0$  и  $t_k = \infty$ .

Чтобы решать задачу, когда действует правило продвижения случайным образом, заметим, что

$$a_i(t) \propto a(t) Pr \{ \text{индивидуум со стажем } t \text{ находится на } i\text{-й ступени} \}. \quad (8.57)$$

Индивидуум будет на  $i$ -й ступени в момент времени  $t$  только в том случае, если он продвинулся  $i - 1$  раз на интервале времени  $(0, t)$ . Число продвижений на интервале  $(0, t)$  — простейший процесс рождения с интенсивностью рождения  $h_j$ , если уже было  $j$  продвижений. Вероятность события, состоящего в том, что на интервале  $(0, t)$  будет  $i - 1$  продвижений, может быть выписана (см. работу Бартлетта, 1955, раздел 3.2). В наших обозначениях эта вероятность равна:

$$P(i-1|t) = \prod_{j=1}^{i-1} h_j \sum_{j=1}^i e^{-h_j t} \prod_{\substack{r=1 \\ r \neq j}}^i \frac{1}{(h_r - h_i)} \quad (i = 1, 2, \dots, k-1). \quad (8.58)$$

Ниже в нашем примере мы рассмотрим ситуацию с равными  $h$ . В этом случае число продвижений имеет распределение

$$P(i-1|t) = \frac{(ht)^{i-1}}{(i-1)!} e^{-ht} \quad (i = 1, 2, \dots, k-2),$$

$$P(k-1|t) = 1 - \sum_{j=0}^{k-2} P(j|t). \quad (8.59)$$

Возвращаясь к (8.57), имеем

$$a_i(t) = a(t) P(i-1|t) \int_0^\infty a(x) P(i-1|x) dx \quad (i = 1, 2, \dots, k-1),$$

$$a_k(t) = \frac{1}{n_k} \left\{ Na(t) - \sum_{j=1}^{k-1} n_j a_j(t) \right\}. \quad (8.60)$$

Для примера положим, что в системе  $\lambda_i = \lambda$  для всех  $i$  и

$$n_{k-1} = n_k, \quad n_i = n_{k-1} 2^{k-i-1} \quad (i = 1, 2, \dots, k-2) \quad (8.61)$$

при больших  $n$ . Эта система обладает тем свойством, что все интенсивности увольнений и продвижений равны  $\lambda$ . Время до продвижения с  $i$ -й ступени из (8.48) равно:

$$t_i = \sum_{j=1}^i \tau_j = \frac{i}{\lambda} \ln 2 = \frac{0,6932i}{\lambda}.$$

Таким образом, промежутки времени между двумя последовательными продвижениями увеличиваются по закону арифметической прогрессии.

Легко вычислить средний стаж индивидуумов на каждой ступени:

$$\left. \begin{aligned} \xi_i &= \frac{\lambda N}{n_i} \int_{((i-1)\ln 2)/\lambda}^{(i \ln 2)/\lambda} t e^{-\lambda t} dt = \frac{1}{\lambda} \{1 + (i-2) \ln 2\} \\ (i &= 1, 2, \dots, k-1) \\ \xi_k &= \frac{1}{\lambda} \{1 + (k-1) \ln 2\} \end{aligned} \right\} \quad (8.62)$$

Средняя продолжительность пребывания индивидуума на  $i$ -й ступени равна:

$$\begin{aligned} \xi_i - t_{i-1} &= \frac{1 - \ln 2}{\lambda} = \frac{0,3068}{\lambda} \quad (i = 1, 2, \dots, k-1; \quad t_0 = 0), \\ \xi_k - t_{k-1} &= \frac{1}{\lambda}. \end{aligned}$$

Эта система обладает очень интересным свойством: средний стаж деятельности на ступени одинаков для всех ступеней, кроме наивысшей. Когда действует правило продвижения случайным образом, мы подставляем (8.59) в (8.60) с  $a(t) = \lambda e^{-\lambda t}$  и находим

$$a_i(t) = \frac{2\lambda}{(i-1)!} (2\lambda t)^{i-1} e^{-2\lambda t} \quad (i = 1, 2, \dots, k-1).$$

Следовательно,

$$\left. \begin{aligned} \xi_i &= i/2\lambda \quad (i = 1, 2, \dots, k-1) \\ \xi_k &= (k+1)/2\lambda. \end{aligned} \right\} \quad (8.63)$$

Сравнивая результаты, выражаемые данным уравнением, с (8.62) видим, что  $\xi_2$  одно и то же в обоих случаях. Если действует правило продвижения по стажу работы, то для  $i > 2$  средний стаж больше, кроме случая, когда  $i = 1$ , где он меньше. При правиле продвижения случайным образом средний стаж на каждой ступени равен:

$$\left. \begin{aligned} \xi_i - \xi_{i-1} &= 1/2\lambda \quad (i = 1, 2, \dots, k-1; \quad \xi_0 = 0) \\ \xi_k - \xi_{k-1} &= 1/\lambda \end{aligned} \right\} \quad (8.64)$$

Как и в предыдущем примере, находим, что при правиле продвижения случайным образом средний стаж больше на каждой ступени.

Мы показали, что гибкая политика продвижений, когда она не так жестко определяется очередностью, имеет существенные практические преимущества.



## 8.5. РАНГОВЫЕ СИСТЕМЫ С БОЛЕЕ ОБЩИМ МЕХАНИЗМОМ УХОДОВ

Анализ последнего раздела допускал, что интенсивности ухода были зависимы от ступеней, но не от стажа работы или очередности. Обе последние зависимости можно ввести в модели, но, чтобы получить точные результаты, более жесткими становятся условия относительно равновесия или аппроксимации «большими численностями на ступенях». Мы не будем даже развивать эту тему, хотя несколько коснемся ее в разделе «Дополнения». А сейчас проиллюстрируем с помощью одного примера, как на полученные результаты влияет невыполнение допущения о постоянстве склонности к уходу.

Мы рассматриваем случай простой иерархии с наймом только на самую нижнюю ступень. Склонность к уходу будет считаться функцией продолжительности службы, а не пребывания на ступени или очередности. Таким образом, она может быть описана с помощью распределения полного срока службы или опасности ухода. Ограничим обсуждение случаем, когда все вакансии занимаются путем продвижения тех, кто был первым в очереди на ближайшей нижней ступени. Из этого правила продвижения по стажу вытекает простое следствие, которое позволяет найти аппроксимацию для средних интенсивностей потоков. Если это правило выполняется, то каждый индивидуум на  $i + 1$ -й ступени должен служить, по крайней мере, столько же, сколько любой индивидуум на  $i$ -й ступени ( $i = 1, 2, \dots, k - 1$ ). Это значит, что для любого момента времени  $T$  будет, например, последний срок  $t_i(T)$ , по истечении которого индивидуумы должны быть продвинуты с  $i$ -й на  $i + 1$ -ю ступень. Здесь, конечно, возможны колебания, но для ступеней с большой численностью их дисперсия должна быть небольшой. (Результаты последнего раздела иллюстрируют эту особенность для рассмотренной здесь модели.) Отсюда доля занятых на ступени  $i$  и ниже будет такой же, как доля занятых со стажем, меньшим  $t_i(T)$ . Приравнявая эти пропорции, получаем

$$\int_0^{t_i(T)} a(x|T) dx = N_i/N \quad (i = 1, 2, \dots, k - 1), \quad (8.65)$$

где  $a(x|T)$  — функция плотности распределения вероятностей по стажу. Если  $a(x|T)$  имеет вероятность в точке  $x = T$  ту же, что и в (8.15), требуется небольшое изменение (8.65). Читатель может проверить, что при  $a(x|T) = e^{-\lambda x}$  (8.65) согласуется с (8.48).

Для системы, находящейся в устойчивом режиме,  $a(x|T) = \mu^{-1}G(x)$ , тогда (8.65) имеет вид

$$\mu^{-1} \int_0^{t_i} G(x) dx = N_i/N \quad (i = 1, 2, \dots, k - 1). \quad (8.66)$$

Из этого уравнения можно найти значения стажа, при которых осуществляются продвижения. Для того чтобы найти интенсивности увольнений в устойчивом режиме, заметим, что

$$\sum_{j=1}^i n_j \omega_j = \frac{N}{\mu} \int_0^{t_i} f(x) dx = \frac{N}{\mu} E(t_i) \quad (i=1, 2, \dots, k-1). \quad (8.67)$$

Эти уравнения определяют интенсивности увольнений; остальные интенсивности получаются из (8.30) при  $R_i(T) = 0$  для  $i > 1$ . Для исследования влияния сделанных обобщений рассмотрим функцию

$$f(T) = \frac{\nu}{c} (1 + T/c)^{-(\nu+1)},$$

для которой

$$\mu = c/(\nu - 1), \quad G(T) = (1 + T/c)^{-\nu}$$

и

$$\int_T^{\infty} G(x) dx = \mu (1 + T/c)^{-\nu+1}.$$

Подставляя эти выражения в (8.67) и используя (8.66), чтобы исключить  $t_i$ , находим:

$$\sum_{j=0}^i n_j \omega_j \sim \frac{N}{\mu} \left\{ 1 - \left( \frac{N - N_i}{N} \right)^{\nu/(\nu-1)} \right\} \quad (i=1, 2, \dots, k-1). \quad (8.68)$$

Устремляя  $\nu$  к бесконечности в уравнении (8.68), получаем решение для экспоненциального распределения. Чтобы проиллюстрировать теоретические положения, приведем в табл. 8.6 для  $k=2$  некоторые значения  $\omega_1$  для распределения типа XI полного срока службы при  $\mu=1$ . Дисперсия распределения приведена в последней строке таблицы; ее не существует для  $\nu \leq 2$ :

Т а б л и ц а 8.6. Значения  $\omega_1$ , когда  $k=2$  для распределения типа XI полного срока службы при различных параметрах  $\nu$

Структура	$\nu$				
	В пределе $\nu \rightarrow 1$	2	3	11	$\infty$
$n_1 = 2n_2$	1,50	1,33	1,21	1,05	1,00
$n_1 = n_2$	2,00	1,50	1,29	1,07	1,00
$n_1 = \frac{1}{2} n_2$	3,00	1,67	1,37	1,08	1,00
Дисперсия $\{\nu/(\nu-2)\}$	—	—	3,00	1,22	1,00

Эти вычисления показывают, что влияние увеличения дисперсии распределения полного срока службы заключается в увеличении интенсивностей уходов с нижних ступеней относительно системы в целом. Это означает снижение  $w_2$  и, следовательно, уменьшение продвижений между первой и второй ступенями. Величина таких изменений зависит от структуры организации и будет больше, когда численность на высших ступенях больше, чем на нижних. С другой стороны, можно сделать вывод, что вероятность продвижений в устойчивом режиме увеличится, если дисперсия полного срока службы уменьшится. В этом случае вывод противоположен тому, что сделан в разделе 5.4. Итак, мы нашли, что уменьшение дисперсии полного срока службы должно привести к уменьшению относительных размеров самой нижней ступени.

## 8.6. ЦЕПОЧКИ ВАКАНСИЙ

Уайт (1970) предложил новый и эффективный метод изучения мобильности в системах с заданной численностью, в котором моделируются потоки вакансий, а не людей. Когда индивидуум перемещается из  $A$  в  $B$ , вакансия передается из  $B$  в  $A$ . Таким образом, результаты одних перемещений обратны другим. Понятие «вакансия» не применялось в моделях с использованием марковских цепей, так как численности групп не фиксировались. Мы говорим о возникновении вакансии только тогда, когда продвижение освобождает место и его необходимо заполнить. Уход из системы образует вакансию, а она приводит к последовательному освобождению мест до тех пор, пока не поступит кто-то извне. Длина и вид последовательности продвижений могут пролить свет на то, как функционирует система. Мы используем понятие *цепочки вакансий* относительно множества продвижений, которые начинаются с уходом индивидуума, они заканчиваются, когда поступает новый работник.

В кадровых системах, где продвижения происходят через единичные интервалы времени, а поступления ограничиваются несколькими уровнями, длины цепочек различаются уровнями градаций, из которых осуществляются уходы или в которые осуществляются поступления. Такой случай представляет ограниченный интерес. Для более сложных систем, рассмотренных Уайтом, изучение цепочки вакансий открывает большие возможности.

Любая модель, сводимая к модели потоков вакансий, требует выполнения двух условий. Во-первых, цепочка должна быть начата, и это можно сделать, определив механизм уходов. Во-вторых, мы должны определить, как «двигается» вакансия. Уайт (1970) предположил, что вакансии «перемещаются» независимо как частицы в марковской цепи. Мы уже рассматривали случай с таким допущением в предыдущей главе, где предполагали, что вакансия в иерархической системе заполняется из внешней среды с заданной вероятностью. Сейчас мы расширим это допущение, разрешив вакансии «перемещаться» в любое состояние и вводя переходные вероятности  $s_{ij}$  ( $i, j = 1, 2, \dots, k$ ) продвижений из  $i$  в  $j$ . Так как вакансия может «покинуть» систему, вве-

дем величину  $s_{i, h+1}$ , означающую вероятность покинуть  $i$ -ю ступень. Таким образом, процесс прохождения вакансии по системе может быть описан как марковская цепь с поглощающим состоянием, для которой имеется матрица переходов

$$\begin{array}{c|c} \mathbf{S} & \mathbf{s}'_{k+1} \\ \hline \mathbf{0} & \mathbf{1}' \end{array},$$

где  $\mathbf{S}$  — матрица с элементами  $\{s_{ij}\}$ ,  $\mathbf{s}_{k+1}$  — вектор-строка с элементами  $\{s_{i, h+1}\}$  и  $\mathbf{1} = (1, 1, \dots, 1)$ . Для того чтобы изучить «пути» отдельных вакансий, нет необходимости знать, как они образовались. Теоретические положения, приведенные в гл. 3 применительно к когортам индивидуумов, могут быть непосредственно использованы и для вакансий. На основе полученных там результатов найдем распределения длины цепочки вакансий. Пусть  $s_i(T)$  — вероятность того, что вакансия, образовавшаяся в группе  $i$ , «покинет» систему через  $T$  шагов, и пусть  $\mathbf{s}(T)$  — вектор вероятностей для разных  $T$ . Тогда

$$\begin{aligned} s_i(T) &= \sum_{r=1}^k Pr \{ \text{переход из } i \text{ в } r \text{ за } T-1 \text{ шагов} \} Pr \{ \text{уйти из } r \} = \\ &= \sum_{r=1}^k s_{ir}^{(T-1)} s_{r, h+1}, \end{aligned}$$

откуда

$$\mathbf{s}(T) = \mathbf{S}^{T-1} \mathbf{s}_{h+1} = \mathbf{S}^{T-1} (\mathbf{I} - \mathbf{S}) \mathbf{1}'. \quad (8.69)$$

Последнее равенство получается из-за того, что

$$\sum_{j=1}^k s_{ij} + s_{i, h+1} = 1 \text{ для всех } i.$$

Уайт сравнил распределения длин наблюдаемых цепочек с теоретическими распределениями (8.69). Вектор средней длины цепочки определяется как

$$\mu = \sum_{T=1}^{\infty} T \mathbf{s}(T) = (\mathbf{I} - \mathbf{S})^{-1} \mathbf{1}'. \quad (8.70)$$

Такой вывод противоположен результату (3.30). В соответствии с ним средние определяются суммами по строкам фундаментальной матрицы.

Эти результаты можно применить к системам с дискретным и непрерывным временем. Так как вакансии заполняются мгновенно, «реальное» время в анализе не учитывается. Введенная выше величина  $T$  просто отражает число шагов вакансии. Обобщение теории для случая, когда вакансии мгновенно не заполняются, для систем с дискретным временем содержится в книге Бартоломью и Форбеса (1979).

## 8.7. ДОПОЛНЕНИЯ

Мы в большой степени опирались на теорию восстановления, приведенную в большинстве учебников по теории случайных процессов например, в книгах Бартлетта (1955), Феллера (1968), Кокса и Мил-

лера (1965). Использовались также обзорная статья Смита (1958) и более поздняя статья о марковских процессах восстановления Синлара (1975), где рассмотрены особые случаи процессов восстановления. Может быть, наиболее доступно теория и прикладные вопросы изложены в монографии Кокса (1962).

Основной математический аппарат, применяемый здесь, — преобразования Лапласа. Книга Уиддера (1946) — наиболее исчерпывающая по этому вопросу, но имеется много работ по применению математики в технике с таблицами изображений. К выводу, что  $\lim_{T \rightarrow \infty} h(T) = \mu^{-1}$ , можно прийти посредством преобразований Лапласа с применением известных Тауберовых теорем.

Модели восстановления, описывающие социальные процессы, исследуются в данной главе совершенно не так, как в предыдущих изданиях. Там модели вводились в более общем виде. Результаты, приведенные здесь, проще и они представляют собой частные случаи при постоянных интенсивностях уходов, больших численностях градаций или в устойчивых режимах, когда  $T \rightarrow \infty$ . При этом мы стремились найти как можно больше параллелей между моделями восстановления и марковскими моделями, описанными в предыдущих главах. Так, в большинстве случаев допускалась зависимость интенсивностей уходов от номера градации, но не допускалась зависимость от стажа или очередности. Тем самым мы надеялись существенно упростить многие результаты, чтобы сделать их более явными и стимулировать их большее применение. Читатели, которым необходимо обстоятельное изложение, могут обратиться к оригинальным статьям.

Точность аппроксимации функцией  $h^0(T)$  плотности восстановления была исследована Бартоломью (1963) для дискретного времени и Бутлером (1970) для непрерывного времени. Эти исследования показали, что аппроксимация очень хороша для распределений времени жизни, более асимметричных, чем экспоненциальное. Она стремится к верхней границе  $h(T)$ , но условия, когда это выполняется, неизвестны. Уравнение восстановления легко может быть обобщено на случай, когда распределение времени жизни — функция возраста системы. Аппроксимация  $h^0(T)$  также может быть записана в более общем виде. Доказано, что она удовлетворительна. Это особенно полезный результат, так как из него следует, что в данном случае не существует более общих методов решения уравнения восстановления.

Точное исследование  $k$ -уровневой системы, когда интенсивности уходов — функции стажа, содержится в работе Бартоломью (1963). В ней показано, что метод аппроксимации из раздела 8.5, достаточно точен для практического использования при описании малых систем. Можно также продемонстрировать, что чем больше система приближается к устойчивому режиму, тем больше интенсивности продвижений и поступлений стремятся достичь предельных значений, независимых от  $T$ . Следовательно, в устойчивом режиме они становятся, по крайней мере в среднем, похожими на основные модели, рассмотренные в этой главе, где мы предполагали, что интенсивности уходов из градаций постоянны. Допущения для вывода распределений по стажу и полного срока службы в разделе 8.4 заключаются в том, что потоки

представляют собой пуассоновские процессы. Это в пределе верно для больших  $k$ -уровневых систем, что можно показать, если оперировать аргументами, аналогичными тем, что использовались в разделе 8.2 для одноуровневой системы.

Исследования, подобные приведенным здесь, могут быть выполнены для дискретного времени. Незначительные сложности возникают при попытке упорядочить некоторые потоки и при замене правила продвижения по стажу работы таким, когда допускается, чтобы было несколько индивидуумов с одинаковым приоритетом. С точки зрения практики планирования использования трудовых ресурсов модели восстановления с дискретным временем имеют ряд преимуществ вычислительного характера. Исследование с использованием потоков индивидуумов и вакансий содержится в работе Бартоломью и Форбеса (1979). В ней рассмотрен также вариант модели, где могут быть задержки при заполнении вакансий.

Модели восстановления нашли применение во многих приложениях в планировании трудовых ресурсов, но обычно в более общем виде, чем они обсуждались здесь. Эти модели хорошо сводятся к вычислительному алгоритму и поддаются численному анализу. Многие из них известны под названием КЕНТ-моделей. КЕНТ-модели, предложенные Хоупсом, обсуждаются в книге Бартоломью (1976), а Уишартом — в книге Смита (1976). Применение теории процессов восстановления с  $k$  уровнями в описании замещений в промышленности освещено в работе Бартоломью (1963), а в планировании трудовых ресурсов — в работе Робинсона (1974). Идея цепочек вакансий предложена Уайтом (1971) для описания сдачи домов в наем.

## ● БИБЛИОГРАФИЯ

- Adelman I. G. (1958). A stochastic analysis of the size distribution of firms. *J. Amer. Statist. Ass.* 53. 893—904.
- Aitchison J. and Brown J. A. C. (1957). *The Lognormal Distribution*. Cambridge University Press.
- Anderson R. B. W. (1974). A Markov chain model of medical speciality choice. *J. Math. Sociology.* 3. 259—274.
- Anderson T. W. (1954). Probability models for analyzing time changes in attitudes. In P. F. Lazarsfeld (1954). 17—66.
- Armitage P. H., Phillips C. M. and Davies J. (1970). Towards a model of the upper secondary school system (with discussion). *J. R. Statist. Soc.* A133. 166—205.
- Armitage P. H., Smith C. S. and Alper P. (1969). *Decision Models for Educational Planning*. Allen Lane. The Penguin Press. London.
- Bailey N. T. J. (1964). *The Elements of Stochastic Processes with Applications to the Natural Sciences*. John Wiley. New York.
- Barlow R. E., Bartholomew D. J., Bremner J. M. and Brunk H. D. (1972). *Statistical Inference under Order Restrictions*. John Wiley. Chichester.
- Bartholomew D. J. (1959). Note on the measurement and prediction of labour turnover. *J. R. Statist. Soc.* A122. 232—239.

- Bartholomew D. J. (1963a). A multistage renewal process. *J. R. Statist. Soc.* B25, 150—168.
- Bartholomew D. J. (1963b). An approximate solution of the integral equation of renewal theory. *J. R. Statist. Soc.* B25, 432—441.‡
- Bartholomew D. J. (1963c). Two-stage replacement strategies. *Operat. Res. Quart.* 14, 71—87.
- Bartholomew D. J. (1969). Renewal theory models for manpower systems. In N. A. B. Wilson (1969), 120—128.
- Bartholomew D. J. (1971). The statistical approach to manpower planning. *Statistician.* 20, 3—26.
- Bartholomew D. J. (1972). The effect of changes in quits and hires on the length of service composition of employed workers: a comment on Stoikov's paper. *Brit. J. Indust. Rel.* 10, 130—133.
- Bartholomew D. J. (1975). A stochastic control problem in the social sciences. *Bull. Int. Statist. Inst.* 46, 670—680.
- Bartholomew D. J. (1976 a). Continuous time diffusion models with random duration of interest. *J. Math. Sociology.* 4, 187—199.
- Bartholomew D. J. (1976b). *Manpower Planning*, Penguin Modern Management Readings. Penguin Books, Harmondsworth, Middlesex, England.
- Bartholomew D. J. (1976c). Statistical problems of prediction and control in manpower planning. *Math. Scientist.* 1, 133—144.
- Bartholomew D. J. (1977a). The analysis of data arising from stochastic processes. Chapter 5 in C. A. O'Muircheartaigh and C. Payne (Eds.), *The Analysis of Survey Data*, Vol. II, Model Fitting. John Wiley, Chichester, 145—174.
- Bartholomew D. J. (1977b). Maintaining a grade or age structure in a stochastic environment. *Adv. Appl. Prob.* 9, 1—17.
- Bartholomew D. J. (1979). The control of a grade structure in a stochastic environment using promotion control. *Adv. Appl. Prob.* 11, 603—615.
- Bartholomew D. J. (1981). *Guidebook for Social Scientists*. John Wiley, Chichester.
- Bartholomew D. J. and Butler A. D. (1971). The distribution of the number of leavers for an organization of fixed size. In A. R. Smith (1971), 417—426.
- Bartholomew D. J. and Forbes A. F. (1979). *Statistical Techniques for Manpower Planning*. John Wiley, Chichester.
- Bartholomew D. J. and Smith A. R. (Eds.) (1971). *Manpower and Management Science*. English Universities Press, London and D. C. Heath and Co. Lexington, Mass.
- Bartlett M. S. (1955). *An Introduction to Stochastic Processes*. Cambridge University Press, London.
- Bartlett M. S. (1956). Deterministic and stochastic models for recurrent epidemics. *Proc. Third Berkeley Symp. Math. Statist. Prob.* 4, 81—109.
- Bell E. J. (1974). Markov analysis of land use change — an application of stochastic processes to remotely sensed data. *Socio-Econ. Plan. Sci.* 8, 311—316.
- Berry B. J. L. (1971). Monitoring trends forecasting change and evaluating goal achievement in the urban environment. In M. Chisholm, A. E. Frey and P. Haggett (1971), 93—117.
- Bibby J. (1970). A model to control for the biasing effects of differential wastage. *Brit. J. Indust. Rel.* 8, 418—420.
- Bibby J. (1975). Methods of measuring mobility. *Quality and Quantity.* 9, 107—136.
- Blumen I., Kogan M. and McCarthy P. J. (1955). *The Industrial Mobility of Labour as a Probability Process*. Cornell University Press, Ithaca, New York.
- Boag J. (1949). Maximum likelihood estimates of the proportion of patients cured by cancer therapy. *J. R. Statist. Soc.* B11, 15—53.
- Boudon R. (1973). *Mathematical Structures of Social Mobility*. Elsevier, Amsterdam, London and New York.
- Bowey A. M. (1969). Labour stability curves and labour stability. *Brit. J. Indust. Rel.* 7, 69—84.

- Box G. E. P. and Jenkins G. M. (1971). *Times Series Analysis Forecasting and Control*. Holden-Day. San Francisco. Cambridge. London and Amsterdam. Русский перевод: Бокс Г. Е., Дженкинс Г. М. Анализ временных рядов. Прогноз и управление. — М.: Мир, 1974.
- Britney R. R. (1975). Forecasting educational enrolments: comparison of a Markov chain and circuitless flow network model. *Socio-Econ. Plan. Sci.* 9. 53—60.
- Brown L. A. (1970). On the use of Markov chains in movement research. *Econ. Geog.* 46. 393—403.
- Bush R. R. and Mosteller C. F. (1955). *Stochastic Models for Learning*. John Wiley. New York.
- Butler A. D. (1970). *Renewal theory applied to manpower systems*. Ph. D. Thesis. University of Kent.
- Butler A. D. (1971). An analysis of flows in a manpower system. *The Statistician*. 20. 69—84.
- Canon M. D., Cullum C. D. and Polack E. (1970). *Theory of Optimal Control and Mathematical Programming*. McGraw-Hill. New York.
- Champernowne D. G. (1953). A model of income distribution. *Econ. J.* 63. 318—351.
- Champernowne D. G. (1973). *The Distribution of Income Between Persons*. Cambridge University Press.
- Charnes A., Cooper W. W., Niehaus R. J. and Sholtz D. (1970). A model for civilian manpower management and planning in the U. S. Navy. In A. R. Smith (1971). 247—263.
- Chen W. C. (1980). On the weak form of Zipf's law. *J. Appl. Prob.* 17. 611—622.
- Chikara R. S. and Folks J. L. (1978). The inverse Gaussian distribution and its statistical application — a review. *J. R. Statist. Soc.* 40. 263—289.
- Chung K. L. (1967). *Markov Chains with Stationary Transition Probabilities*. (2nd ed.) Springer-Verlag. Berlin. Русский перевод: Чжуи Кай-лай. Однородные цепи Маркова. — М.: Мир, 1964.
- Cinlar E. (1975). Markov renewal theory: a survey. *Man. Sci.* 21. 727—752.
- Clark W. A. V. (1965). Markov chain analysis in geography: an application to the movement of rental housing areas. *Ann. Assoc. of Amer. Geog.* 55. 351—359.
- Cliff A. D., Haggett P., Ord J. K., Bassett K. and Davies R. (1975). *Elements of Spatial Structure: A Qualitative Approach*. Cambridge University Press.
- Clough D. J. and McReynolds W. P. (1966) State transition model of an educational system incorporating a constraint theory of supply and demand. *Ontario J. Educational Research*. 9. 1—18.
- Cloves G. A. (1972). A dynamic model for the analysis of labour turnover. *J. R. Statist. Soc.* A135. 242—256.
- Coleman J. S. (1961). *The Adolescent Society*. The Free Press of Glencoe. New York.
- Coleman J. S. (1964a) *Models of Change and Response Uncertainty*. Prentice-Hall. Englewood Cliffs.
- Coleman J. S. (1964b). *Introduction to Mathematical Sociology*. The Free Press of Glencoe and Collier-Macmillan. London.
- Collins L. (1973). Industrial size distribution and stochastic processes. *Progress in Geography*. 5. 121—165.
- Collins N. R. and Preston L. E. (1961). The size structure of the largest industrial firms. 1907—1958. *Amer. Econ. Rev.* 51. 986—1011.
- Conlisk J. (1976). Interactive Markov chains. *J. Math. Sociology*. 4. 157—185.
- Conlisk J. (1978). A stability theorem for an interactive Markov chain. *J. Math. Sociology*. 6. 163—168.
- Cootner P. H. (1964). *The Random Character of Stock Market Prices*. M. I. T. Press. Cambridge. Mass.
- Cox D. R. (1962). *Renewal Theory*. Methuen. London. Русский перевод: Кокс Д. Р., Смит В. Л. Теория восстановления. — М.: Сов.радио, 1967.



- Cox D. R. and Miller H. D. (1965). *The Theory of Stochastic Processes*. Methuen, London.
- Creedy J. (1979). The analysis of labour market flows using a continuous time model. Private communication.
- Cyert R. M., Davidson H. J. and Thompson G. L. (1962). Estimation of the allowance for doubtful accounts by Markov chains. *Man. Sci.* 8. 287—303.
- Davies G. S. (1973). Structural control in a graded manpower system. *Man. Sci.* 20. 76—84.
- Davies G. S. (1975). Maintainability of structures in Markov chains in models under recruitment control. *J. Appl. Prob.* 12. 376—382.
- Davies R., Johnson D., Farrow S. (1975). Planning patient care with a Markov model. *Operat. Res. Quart.* 26. 599—607.
- Dent W. T. (1967). Application of Markov analysis to international wool flows. *Rev. of Economics and Statistics.* 49. 613—616.
- Drewett J. R. (1969). A stochastic model of the land conversion process. *Reg. Studies.* 3. 269—280.
- Dryden M. (1969). Share price movements: a Markovian approach. *J. Finance.* 24. 49—60.
- Eaton W. W. and Whitmore G. A. (1977). Length of stay as a stochastic process: a general approach and application to hospitalization for schizophrenia. *J. Math. Sociology.* 5. 273—292.
- Fenehey G. (1973). Two models for multi-regional population dynamics. *Environment and Planning.* 5. 31—43.
- Feichtinger G. (1971). *Stochastische Modelle Demographischer Prozesse*. Springer-Verlag, Berlin.
- Feichtinger G. (1972). Stochastic decrement models of demography. *Biom. Zeit.* 14. 106—125. Русский перевод: Файхтингер Г. Стохастические модели выбытия в статистике населения. — В кн.: Демографические модели. М.: Статистика, 1977.
- Feichtinger G. (1973). Markovian models for some demographic processes. *Statistische Hefte.* 14. 310—334.
- Feichtinger G. (1976). On the generalization of stable age distributions to Gani-type person-flow models. *Adv. Appl. Prob.* 8. 433—445.
- Feller W. (1966). *An Introduction to Probability Theory and its Applications*. Vol. II. John Wiley, New York. Русский перевод: Феллер В. Введение в теорию вероятностей и ее приложение. Т. 2. — М.: Мир, 1967.
- Feller W. (1968). *An Introduction to Probability Theory and its Applications*. Vol. I. (3rd edn.). John Wiley, New York. Русский перевод: Феллер В. Введение в теорию вероятностей и ее приложение. Т. 1. — М.: Мир, 1964.
- Fielitz B. D. and Bhargava T. H. (1973). The behaviour of stock-price relatives — a Markovian analysis. *Operat. Res.* 21. 1183—1199.
- Fix E. and Neyman J. (1951). A simple stochastic model of recovery, relapse, death and loss of patients. *Human Biology.* 23. 205—241.
- Foley J. D. (1967). A Markovian model of the University of Michigan executive system. *Communications of the A. C. M.* 10. 584—588.
- Forbes A. F. (1971a). Markov chain models for manpower systems. In D. J. Bartholomew and A. R. Smith (1971). 93—113.
- Forbes A. F. (1971b). Non-parametric methods of estimating the survivor function. *Statistician.* 20. 27—52.
- Forbes A. F. (1971c). Promotions and recruitment policies for the control of quasi-stationary hierarchical systems. In A. R. Smith (1971). 401—414.
- Frazer R. A., Duncan W. J. and Collar A. R. (1946). *Elementary Matrices*. Cambridge University Press.
- Gani J. (1963). Formulae for projecting enrolments and degrees awarded in universities. *J. R. Statist. Soc. A.* 126. 400—409.
- Гантмахер Ф. П. (1967). *Теория матриц*. — М.: Наука.
- Gilbert G. (1973). Semi-Markov processes and mobility: a note. *J. Math. Sociology.* 3. 139—145.
- Ginsberg R. B. (1971). Semi-Markov processes and mobility. *J. Math. Sociology.* 1. 233—262.

- Ginsberg R. B. (1978a). Probability models of residence histories: analysis of times between moves, in W. A. V. Clark and E. G. Moore (Eds). *Population Mobility and Residential Change*. Northwestern studies in Geography. No. 24. Evanston, Illinois.
- Ginsberg R. B. (1978b). The relationship between timing of moves and choice of destination in stochastic models of migration. *Environment and Planning*. A10. 667—679.
- Ginsberg R. B. (1978c). Timing and duration effects in residential histories and other longitudinal data, II: studies of duration effects in Norway 1965—1971. Fels Discussion Paper No. 121. School of Public and Urban Policy. University of Pennsylvania.
- Ginsberg R. B. (1979a). *Stochastic Models of Migration: Sweden 1961—1975*. North-Holland. New York and Amsterdam.
- Ginsberg R. B. (1979b). Tests of stochastic models of timing in mobility histories: comparison of information derived from different observation plans. *Environment and Planning*. A11. 1387—1404.
- Ginsberg R. (1979c). Timing and duration effects in residential histories and other longitudinal data, I: stochastic and statistical models. *Regional Science and Urban Economics*. 9. 311—331.
- Glass D. V. (Ed.). (1954). *Social Mobility in Britain* Routledge and Kegan Paul. London.
- Григелионис Б. И. (1964). Предельные теоремы для сумм процессов восстановления. — В кн.: *Кибернетику на службу коммунизму*. Т. 2. М.: Энергия.
- Grinold R. C. and Marshall K. T. (1977). *Manpower Planning Models*. North-Holland. New York.
- Hajnal J. (1976). On products of non-negative matrices. *Math. Proc. Camb. Phil. Soc.* 79. 521—530.
- Hart P. E. and Prais S. J. (1956). The analysis of business concentration. *J. R. Statist. Soc.* A119. 150—191.
- Hassani H. (1980). Markov renewal models for manpower systems. Ph. D. Thesis. University of London.
- Hawes L. C. and Foley J. M. (1973). A Markov analysis of interview communication. *Speech Monographs*. 40. 208—219.
- Hawkes A. G. (1969). An approach to the analysis of electoral swing. *J. R. Statist. Soc.* A132. 68—79.
- Hedberg M. (1961). The turnover of labour in industry, an actuarial study. *Acta Sociologica*. 5. 129—143.
- Hedge B. J., Everitt B. S. and Firth C. (1978). The role of gaze in dialogue. *Acta Psychologica*. 42. 453—475.
- Henry N. W., McGinnis R. and Tegtmeyer H. W. (1971). A finite model of mobility. *J. Math. Sociology*. 1. 107—116.
- Herbst P. G. (1963). Organizational commitment: a decision model. *Acta Sociologica*. 7. 34—45.
- Hill B. M. (1970). Zipf's law and prior distributions for the composition of a population. *J. Amer. Statist. Ass.* 65. 1220—1232.
- Hill B. M. (1974). The rank-frequency form of Zipf's law. *J. Amer. Statist. Ass.* 69. 1017—1026.
- Hill B. M. and Woodroffe M. (1975). Stronger forms of Zipf's law. *J. Amer. Statist. Ass.* 70. 212—219.
- Хичин А. Я. (1963). Математические методы в теории массового обслуживания. — В кн.: *Работы по математической теории массового обслуживания*. М.: Гос. изд. физ.-мат. лит.
- Hodgk R. W. (1966). Occupational mobility as a probability process. *Demography*. 3. 19—34.
- Hoem J. M. (1971). Point estimation of forces of transition in demographic models. *J. R. Statist. Soc.* 33. 275—289.
- Holland P. W. and Leinhardt S. (1977). A dynamic model for social networks. *J. Math. Sociology*. 5. 5—20.
- Horvath W. J. (1968). A statistical model for the duration of wars and strikes. *Behavioural Science* 13. 18—28.

- Hyman R. (1970). Economic motivation and labour stability. *Brit. J. Indust. Rel.* 8. 159—178.
- Ijiri Y. and Simon H. A. (1977). *Skew Distributions and the Sizes of Business Firms*. Elsevier. Amsterdam.
- Isaacson D. L. and Marsden R. W. (1976). *Markov Chains Theory and Applications*. John Wiley, New York.
- Johnstone J. N. and Philip H. (1973). The application of a Markov chain in educational planning. *Socio-econ. Plan. Sci.* 7. 283—294.
- Joseph G. (1974). Inter-regional population distribution and growth in Britain— a projection exercise. *Scottish. J. Political Economy.* 21. 159—170.
- Kamat A. R. (1968a). Estimating wastage in a course of education. *Sankhya.* B30.5—12.
- Kamat A. R. (1968b). Mathematical schemes for describing progress in a course of education. *Sankhya.* B30.13—24.
- Kamat A. R. (1968c). A stochastic model for progress in a course of education. *Sankhya.* B30. 25—32.
- Kao E. P. C. (1974). Modelling the movement of coronary patients within a hospital by semi-Markov processes. *Operat. Res.* 22. 683—698.
- Karlin S. and Taylor H. (1975). *A First Course in Stochastic Processes*. Academic Press, New York.
- Katz L. and Procter C. H. (1959). The concepts of configurations of interpersonal relations in a group as a time-dependent stochastic process. *Psychometrika.* 24. 317—327.
- Кемени J. G. and Снелл L. (1962). *Mathematical Models in the Social Sciences*. Ginn and Co. Boston. Русский перевод: Кемени Дж., Снелл Дж. Кибернетическое моделирование. Некоторые приложения.— М.: Сов. радио, 1972.
- Кемени J. G. and Shell L. (1976). *Finite Markov Chains*. Springer-Verlag, Berlin (1st Edn. van Nostrand, 1960. Princeton, N. J). Русский перевод: Кемени Дж., Снелл Дж. Конечные цепи Маркова.— М.: Наука, 1970.
- Kendall M. G. (1961). Natural law in the social sciences. *J. R. Statist. Soc.* A124. 1—19.
- Kingman J. F. C. (1962). The imbedding problem for finite Markov chains. *Zeit. Wahrscheinlichkeitstheorie.* 1. 14—24.
- Krishnan P. (Ed.). (1977a). *Mathematical Models of Sociology*. Sociological Review Monograph 24. University of Keele. Keele. Staffs. U. K.
- Krishnan P. (1977b). A stochastic process model of conjugal history. In *Mathematical Models of Sociology* (P. Krishnan Ed.) Sociological Review Monograph 24. University of Keele. Keele Staffs. U. K. 117—125.
- Kruskal W. (Ed.). (1970). *Mathematical Sciences and Social Sciences*. Prentice-Hall, Englewood Cliffs, N. J.
- Kushner H. (1967). *Stochastic Stability and Control*. Academic Press, New York.
- Kushner H. (1971). *Introduction to Stochastic Control*. Holt, Rinehart and Winston, New York.
- Kushner H. J. and Kleinman A. J. (1971). Mathematical programming and the control of Markov chains. *Int. J. Control.* 13. 801—820.
- Lancaster A. (1972). A stochastic model for the duration of a strike. *J. R. Statist. Soc.* A135. 257—271.
- Lancaster A. and Nickell S. J. (1980). The analysis of re-employment probabilities for the unemployed. *J. R. Statist. Soc.* A143. 141—165.
- Lancaster P. (1969). *Theory of Matrices*. Academic Press, New York. Русский перевод: Ланкастер П. Теория матриц.— М.: Наука, 1978.
- Land K. C. (1969). Duration of residence and prospective migration. *Demography.* 6. 133—140.
- Lane K. F. and Andrew J. E. (1955). A method of labour turnover analysis. *J. R. Statist. Soc.* A118. 296—323.
- Ledermann W. (Editor-in-Chief) (1980). *Handbook of Applicable Mathematics* (6 vols.). John Wiley, Chichester.

- Lillard L. A. and Willis R. J. (1978). Dynamic aspects of earning mobility. *Econometrica*. 46. 985—1012.
- Long L. H. (1970). On measuring geographic mobility. *J. Amer. Statist. Ass.* 65. 1195—1203.
- Lu K. H. (1968). A Markov chain analysis of caries process with consideration for the effect of restoration. *Archs. Oral. Biol.* 13. 1119—1132.
- McCall J. J. (1971). A Markovian model of income dynamics. *J. Amer. Statist. Ass.* 66. 439—447.
- McClean S. I. (1980). A semi-Markov model for a multi-grade population with Poisson recruitment. *J. Appl. Prob.* 17. 846—852.
- McFarland D. (1970). Inter-generational social mobility as a Markov process: including a time-stationary Markovian model that explains observed declines in mobility rates over time. *Amer. Sociol. Rev.* 35. 463—476.
- McGinnis R. (1968). A stochastic model of social mobility. *Amer. Sociol. Rev.* 33. 712—721.
- Marshall A. W. and Goldhamer H. (1955). An application of Markov-processes to the study of the epidemiology of mental disease. *J. Amer. Statist. Ass.* 50. 99—129.
- Matras J. (1960a). Comparison of intergenerational occupational mobility patterns: an application of the formal theory of social mobility. *Population Studies*. 14. 163—169.
- Matras J. (1960b). Differential fertility, intergenerational occupational mobility and change in the occupational distribution: some elementary inter-relationships. *Population Studies*. 15. 187—197.
- Matras J. (1967). Social mobility and social structure: some insights from the linear model. *Amer. Sociol. Rev.* 32. 608—614.
- Mayhew B. H. (1972). Growth and decay of structure in interaction. *Comparative Group Studies*. May. 131—160.
- Mayhew B. H. and Gray L. N. (1971). The structure of dominance relations in triadic interaction systems. *Comparative Group Studies*. May. 161—190.
- Mehlman A. (1977a). A note on the limiting behaviour of discrete-time Markovian manpower models with inhomogeneous independent Poisson input. *J. Appl. Prob.* 14. 611—613.
- Mehlman A. (1977b). Markovian manpower models in continuous time. *J. Appl. Prob.* 14. 249—259.
- Mehlman A. (1979). Semi-Markovian manpower models in continuous time. *J. Appl. Prob.* 16. 416—422.
- Mehlman A. (1980). An approach to optimal recruitment and transition strategies for manpower systems using dynamic programming. *J. Operat. Res. Soc.* 31. 1009—1015.
- Menges G. and Elstermann G. (1971). Capacity models in university management. In D. J. Bartholomew and A. R. Smith (1971). 207—221.
- Meredith J. (1973). A Markovian analysis of a geriatric ward. *Man. Sci.* 19. 604—612.
- Miller G. A. (1952). Finite Markov processes in psychology. *Psychometrika*. 17. 149—167.
- Miller W. L. (1972). Measures of electoral change using aggregate data. *J.R. Statist. Soc. A* 135. 122—142.
- Milton R. C. (1972). Computer evaluation of the multivariate normal integral. *Technometrics*. 14. 881—889.
- Mores G. (1975). On Johnstone and Philp's Markov chain educational planning model. *Socio-econ. Plan. Sci.* 9. 89—91.
- Moran P. A. P. (1968). *An Introduction to Probability Theory*. Clarendon Press. Oxford.
- Morgan R. W. (1970). Manpower planning in the Royal Air Force: an exercise in linear programming. In A. R. Smith (1971). 317—325.
- Morgan R. W. (1971). The use of a steady state model to obtain the recruitment, retirement and promotion policies of an expanding organization. In D. J. Bartholomew and A. R. Smith (1971). 283—291.
- Morrison D. G. (1966). Testing brand-switching models. *J. Marketing Res.* 3. 401—409.

- Morrison D. G., Massy W. F. and Silverman F. N. (1971). The effect of non-homogeneous populations on Markov steady state probabilities. *J. Amer. Statist. Ass.* 66. 268—274.
- Morrison P. (1967). Duration of residence and prospective migration: the evaluation of a stochastic model. *Demography*. 4. 533—561.
- Moya-Angeler J. (1976). A model with shortage of places for educational and manpower systems. *Omega*. 4. 719—730.
- Myers G. C., McGinnis R. and Masnick G. (1967). The duration of residence approach to a dynamic stochastic model of internal migration: a test of the axiom of cumulative inertia. *Eugenics Quarterly*. 14. 121—126.
- Navarro V. (1969). Planning personal health services: a Markovian model. *Medical Care*. 7. 242—249.
- Parzen E. (1962). *Stochastic Processes*. Holden-Day. San Francisco.
- Pestieau P. and Possen U. M. (1979). A Model of wealth distribution. *Econometrica*. 47. 761—772.
- Plewisi I. (1980). Using longitudinal data to model teachers ratings of classroom behaviour as a dynamic process. Private communication.
- Pollard J. H. (1966). On the use of the direct matrix product in analysing certain stochastic population models. *Biometrika*. 53. 397—415.
- Pollard J. H. (1967). A note on certain discrete time stochastic population models with Poisson immigration. *J. Appl. Prob.* 4. 209—213.
- Prais S. J. (1955a). Measuring social mobility. *J. R. Statist. Soc.* A118. 56—66.
- Prais S. J. (1955b). The strike-proneness of large plants in Britain. *J. R. Statist. Soc.* A141. 368—384.
- Preston L. E. and Bell E. J. (1961). The statistical analysis of industry structure: an application to food industries. *J. Amer. Statist. Ass.* 56. 925—932.
- Pyke R. (1961a). Markov renewal processes: definition and preliminary properties. *Ann. Math. Statist.* 32. 1231—1242.
- Pyke R. (1961b). Markov renewal processes with finitely many states. *Ann. Math. Statist.* 32. 1243—1259.
- Quetelet A. (1835). *Essai de Physique Sociale* (2 vols). Bachalier, Paris (2nd Edn. 1869. Brussels).
- Quetelet A. (1849). *Letters on the Theory of Probabilities as Applied to the Moral and Political Sciences* (translated from the French by O. G. Downes). C. and E. Layton. London.
- Rees P. H. and Wilson A. G. (1973). Accounts and models for spatial demographic analysis: I. aggregate population. *Environment and Planning*. 5. 61—90.
- Rice A. K., Hill J. M. M. and Trist E. L. (1950). The representation of labour turnover as a social process. *Human Relations*. 3. 349—381.
- Robinson D. (1974). Two stage replacement strategies and their application to manpower planning. *Man. Sci.* 21. 199—208.
- Rogers A. (1968). *Matrix Analysis of Inter-regional Population Growth and Distribution*. University of California Press. Berkeley, California.
- Rogoff N. (1953). *Recent Trends in Occupational Mobility*. Free Press. Glencoe. Illinois.
- Rosenbaum S. (1971). A report on the use of statistics in social science research (with discussion). *J. R. Statist. Soc.* A134. 534—610.
- Rowe S. M., Wagner W. G. and Watersby G. B. (1970). A control theory solution to optimal faculty staffing. Paper P-11. Research Program in University Administration University of California.
- Sales P. (1971). The validity of the Markov chain model for a branch of the Civil Service. *Statistician* 20. 85—110.
- Salkin M. S., Lianos T. P. and Paris Q. (1975). Population predictions for the western United States: a Markov chain approach. *J. Reg. Sci.* 15. 53—60.
- Schinnaar A. P. and Stewman S. (1978). A class of Markov models of social mobility with duration memory patterns. *J. Math. Sociology*. 6. 61—86.
- Seal H. L. (1945). The mathematics of a population composed of  $k$  stationary strata each recruited from the stratum below and supported at the lowest level by a uniform annual number of entrants. *Biometrika*. 33. 226—230.

- Seneta E. (1973). *Non-negative Matrices*. George Allen & Unwin, London.
- Shorrocks A. F. (1975). On stochastic models of size distributions *Rev Econ. Studies*. 42. 631—641.
- Shorrocks A. F. (1976). Income mobility and the Markov assumption. *Econ. J.* 86. 566—578.
- Shorrocks A. F. (1978). The measurement of mobility. *Econometrica*. 46. 1013—1024.
- Silcock H. (1954). The phenomenon of labour turnover. *J. R. Statist. Soc. A117*. 429—440.
- Simon H. A. (1955). On a class of skew distribution functions. *Biometrika*. 42. 425—440.
- Singer B. and Spilerman S. (1976a). Some methodological issues in the analysis of longitudinal surveys. *Ann. of Econ. and Social Measurement*. 5. 447—474.
- Singer B. and Spilerman S. (1976b). The representation of social processes by Markov models. *Amer. J. Sociol.* 82. 1—54.
- Singer B. and Spilerman S. (1977a). Fitting stochastic models to longitudinal survey data—some examples in the social sciences. *Bull. Int. Statist. Inst.* 47. Book 3. 283—300.
- Singer B. and Spilerman S. (1977b). Trance inequalities for mixtures of Markov chains. *Adv. Appl. Prob.* 9. 747—764.
- Singer B. and Spilerman S. (1979). Clustering on the main diagonal in mobility matrices. In *Sociological Methodology (1979)*. K. Schuessler (Ed.). Jossey — Bass, San Francisco.
- Slivka R. T. and Cannavale F. (1973). An analytical model of the passage of defendants through a court system. *J. Res. in Crime and Delinquency*. June. 132—140.
- Smith A. R. (Ed.) (1976). *Monpower Planning in the Civil Service*. Civil Service Studies No. 3. H. M. S. O. London.
- Smith W. L. (1958). Renewal theory and its ramifications. *J. R. Statist. Soc. B20*. 243—302. Русский перевод: Кокс Д. Р., Смит В. Л. Теория восстановления. — М.: Сов. радио, 1967.
- Sommers P. M. and Conlisk J. (1979). Eigenvalue immobility measures for Markov chains. *J. Math. Sociology*. 6. 169—234.
- Sorenson A. B. and Hallinan M. T. (1977). A stochastic model for change in group structure. In *Mathematical Models of Sociology* (Ed. P. Krishnan). Sociological Review Monograph 24. University of Keele. Keele. Staffs. U. K. 143—166.
- Spilerman S. (1972a). The analysis of mobility processes by the introduction of independent variables into a Markov chain. *Amer. Sociol. Rev.* 37. 277—294.
- Spilerman S. (1972b). Extensions of the mover-stayer model. *Amer. J. Sociol.* 78. 599—626.
- Staff P. J. and Vagholkar M. K. (1971). Stationary distributions of open Markov processes in discrete time with application to hospital planning. *J. Appl. Prob.* 8. 668—680.
- Stafford J. (1977). Urban growth as an absorbing Markov process. In *Mathematical Models of Sociology*. (Ed. P. Krishnan). Sociological Review Monograph 24. University of Keele. Keele Staffs. U. K. 135—142.
- Steindl J. (1965). *Random Processes and the Growth of Firms*. Griffin, London.
- Stewman S. (1975a). An application of the job vacancy chain model to a Civil Service internal labour market. *J. Math. Sociology*. 4. 37—59.
- Stewman S. (1975b). Two Markov models of system occupational mobility: underlying conceptualizations and empirical tests. *Amer. Sociol. Rev.* 40. 298—321.
- Stewman S. (1976). Markov models of occupational mobility theoretical development and empirical support. Part 1 Careers. Part 2. Continuously operative job systems. *J. Math. Sociology*. 4. 201—278.
- Stoikov V. (1971). The effect of changes in quits and hires on the length-of-service composition of employed workers. *Brit. J. Indust. Rel.* 9. 225—233.

- Sto ne R. (1972). A Markovian educational model and other examples linking social behaviour to the economy (with discussion). *J. R. Statist. Soc.* A135. 511—543.
- Sv alagosta K. (1959). *Prestige, Class and Mobility*. Heinemann. London.
- Sve r d r u p E. (1965). Estimates and test procedures in connexion with stochastic models of deaths, recoveries and transfers between different states of health. *Skand. Aktuarietidskrift.* 46. 184—211.
- Sy kes Z. M. (1969). Some stochastic versions of the matrix model for population dynamics. *J. Amer. Statist. Ass.* 64. 111—130.
- Tar ver J. D. and Gur ley W. R. (1965). A stochastic analysis of geographic mobility and population projections of the census divisions in the United States. *Demography.* 2. 134—139.
- Th om p s on M. E. (1980). Estimation of mobility parameters from longitudinal data over a fixed time interval using semi-Markov models. Private communication.
- Th on st ad T. (1969). *Education and Manpower: Theoretical Models and Empirical Applications*. Oliver and Boyd. Edinburgh and London.
- T in t n e r G. and S e n g u p t a J. K. (1972). *Stochastic Economics: Stochastic Processes. Control and Programming*. Academic Press. New York.
- T u m a N. B., H a n n a n M. T. and G r o e n v e l d L. P. (1979). Dynamic analysis of event histories. *Amer. J. Sociol.* 84. 820—854.
- V a j d a S. (1947). The stratified semi-stationary population. *Biometrika.* 34. 243—254.
- V a j d a S. (1948). Introduction to a mathematical theory of a graded stationary population. *Bull. de l'Ass. Actuaire. Suisses.* 48. 251—273.
- V a j d a S. (1975). Mathematical aspects of manpower planning. *Operat. Res. Quart.* 26. 527—542.
- V a j d a S. (1978). *Mathematics of Manpower Planning*, John Wiley, Chichester.
- V a l l i a n t R. and M i l k o v i c h G. T. (1977). Comparison of semi-Markov and Markov models in a personnel forecasting application. *Decision Sciences.* 8. 465—477.
- V a n d e r M e r w e R. and M i l l e r S. (1971). The measurement of labour turnover. *Human Relations.* 24. 233—253.
- V a n K o r f f M. (1979). A statistical model of the duration of mental hospitalization: the mixed exponential distribution. *J. Math. Sociology.* 6. 169—175.
- W e i s s G. H. and Z e l e n M. (1965). A semi-Markov model for clinical trials. *J. Appl. Prob.* 2. 269—285.
- W e i s s H. K. (1963). Stochastic models for the duration and magnitude of a «deadly quarrel». *Operat. Res.* 11. 101—121.
- W h i t e H. C. (1963). Cause and effect in social mobility tables. *Behavioural Science.* 8. 14—27.
- W h i t e H. C. (1969). Control and evolution of aggregate personnel: flows of men and jobs. *Admin. Sci. Quart.* 14. 4—11.
- W h i t e H. C. (1970a). *Chains of Opportunity*. Harvard University Press. Cambridge, Mass.
- W h i t e H. C. (1970b). Matching, vacancies and mobility. *J. Pol. Econ.* 78. 97—105.
- W h i t e H. C. (1970c). Stayers and movers. *Amer. J. Sociol.* 76. 307—324.
- W h i t e H. C. (1971). Multipliers, vacancy chains and filtering in housing. *J. Amer. Inst. Planners.* 37. 88—94.
- W h i t m o r e G. A. (1976). Management applications of the inverse Gaussian distribution. *Omega.* 4. 215—223.
- W h i t m o r e G. A. (1979). An inverse Gaussian model for labour turnover. *J. R. Statist. Soc.* A142. 468—478.
- W i d d e r D. V. (1946). *The Laplace Transform*. Princeton University Press. Princeton, N. J.
- W o l d H. O. A. and W h i t t l e P. (1957). A model explaining the Pareto distribution of wealth. *Econometrica.* 25. 591—595.
- W y n n H. P. and S a l e s P. (1973a). A simple model for projecting means and variances of population grade sizes. In *Stochastic Analysis of National Manpower Problems*. Research Report. University of Kent.

- Wynn H. P. and Sales P. (1973b). The mover-stayer model and the 1963 labour mobility survey. In *Stochastic Analysis of National Manpower Problems*. Research Report. University of Kent.
- Yang M. C. K. and Hirsch C. J. (1973). The use of a semi-Markov model for describing sleep patterns. *Biometrics*. 29. 667—676.
- Young A. (1971). Demographic and ecological models for manpower planning. In D. J. Bartholomew and B. R. Morris (Eds.) (1971). 75—97.
- Young A. and Almond G. (1961). Predicting distributions of staff. *Comp. J.* 3. 246—250.
- Young A. and Vassiliou P. C. G. (1974). A non-linear model on the promotion of staff. *J. R. Statist. Soc. A137*. 584—595.
- Zahl S. (1955). A Markov process model for follow-up studies. *Human. Biology*. 27. 90—120.
- Zipf G. K. (1949). *Human Behaviour and the Principle of Least Effort*. Addison-Wesley Reading, Mass.

## ● ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

- Адаптивная стратегия управления (adaptive strategies for control) 197, 204, 205, 206, 207, 208, 209
- Аппроксимация (approximation) 200, 202, 211
- Асимметрия распределения (skewness of distribution) 215, 216, 223, 245, 269, 270, 277  
— логнормального (skewness of lognormal density) 228
- Базисное решение задачи линейного программирования (basic solution of linear programming problem) 189, 191
- Безработица (unemployment) 23, 212  
— длительность (duration) 237
- Биномиальная случайная величина (Бернулли) (Bernoulli, binomial variable) 27, 199, 202, 250, 263
- Биномиальное распределение (binomial distribution) 27, 199, 202, 263
- Благосостояния уровень (wealth) 230, 236  
— распределение (wealth distribution) 226, 236, 239
- Блуждание случайное (random walk) 222—224, 238
- Большие отклонения (upper tail) 244, 245, 247
- Вакансии (vacancy) 27, 261, 263, 265, 266, 273, 275, 276, 278
- Вейбулла распределения (Weibull distribution) 213, 236, 237  
— смесь (mixture of) 238
- Вероятность перехода за малый промежуток времени (infinitesimal transition probability) 105, 125, 206  
— поддержания, сохранения (probability maintainable) 198, 201  
— продвижения (promotion probability) 201, 205
- Верхняя треугольная матрица (upper-triangular matrix) 73, 81, 83, 180, 183, 185, 191
- Вершиной вверх, структура групп (top-heavy grade structure) 173, 178, 185, 197, 200  
— вниз, — (bottom-heavy grade structure) 180, 185
- Винера процесс (Wiener process) 222, 224, 236
- Вложенная марковская цепь (embedded Markov chain) 106, 124, 171
- Возрастная структура (age distribution structure) 49, 193, 199, 241, 251, 252, 258, 259, 270—272  
— поддерживаемая (maintainable) 180, 200, 203—207  
— установившаяся (steady state) 252
- Восстановления, теория (renewal theory) 103, 180, 241, 276  
— марковские процессы (Markov renewal processes) 169, 277  
— плотность (renewal density) 242, 247, 252, 253, 277  
— уравнение (equation) 242, 245—247, 277  
— функция (function) 242  
— система (system) 242  
— многоступенчатый процесс (multystage renewal process) 242, 277, 278



- Время дискретное (discrete time) 25, 29  
 — непрерывное (continuous) 25, 29, 105  
 — пребывания (sojourn time) 106, 108, 109, 113, 124, 126, 171  
 — службы (seniority time, length of service) 125, 158, 160  
 Выживания интенсивность (survival rate) 145  
 Выпуклая оболочка множества точек (convex hull set of points) 177, 178, 211  
 Гамма, распределение (gamma distribution) 121, 213, 217, 218, 231, 236  
 — функция (function) 122, 232  
 Гаусса распределение (Gaussian distribution) 238  
 Геометрическая прогрессия (geometric series) 93, 152, 167, 262  
 Геометрический закон распределения (geometric distribution) 46, 108, 233, 238, 263  
 Геометрическое среднее (geometric mean) 42  
 — структура ступеней (geometric grade structure) 152, 167, 263  
 Гибрата распределение (Gibrat distribution) 236  
 Гипотетические состояния, ступени, градации (hypothetical grades) 165—168, 219  
 Городов распределение по размерам (City size distribution) 228, 230  
 Группа социальная (social class) 29, 30, 31  
 Дисперсия (variance) 36, 38, 39, 45, 46, 103, 120, 127, 200, 205, 222  
 — числа восстановлений (variance of number of renewals) 249  
 — распределения полных сроков службы (variance of CLS distribution) 274, 275  
 Дифференциальные уравнения (differential equations) 107, 136, 158, 210, 214  
 — разностные уравнения (differential difference equations) 265  
 Длительность пребывания (duration, length of service) 46, 55, 57, 118, 214, 217, 243  
 — безработицы (duration of unemployment) 212, 222, 224, 237  
 — войны (—duration of war) 237  
 — пребывания в больнице (—duration of stay in hospital) 238  
 Должития функция (survivor function) 212, 214, 217, 220, 222, 252, 265, 267  
 — —  $T$  лет ( $T$ -year survival function, rate) 142  
 Допустимое решение (feasible solution) 188, 189, 191  
 Достижимость (attainability) 175, 186—197, 208—210  
 Доходов распределение (income distribution) 101, 212, 225, 228, 234, 236, 239  
 — неравенство (inequality of wealth) 212  
 Естественная смертность (natural mortality) 20, 145  
 — — интенсивность (force of natural mortality) 143  
 Желаемая структура (goal, desired structure) 173—175, 207, 209  
 Желаемые численности (goal, desired, target stocks) 68, 186—192  
 Жизни продолжительность (length of human life) 212  
 — — измерение (measure of, survival) 212  
 — — распределение (life distribution) 241, 242, 243, 277  
 Закон (правило) пропорциональности (law of proportionate effect) 23, 230, 233, 234, 236  
 Закрытые, замкнутые системы (closed system) 25, 28, 29, 85, 105—135  
 Запас, численность (stock) 28, 68, 78, 80, 103, 138, 139, 141, 153, 154, 203, 207, 211, 215  
 Зипфа закономерность (Zipf's law) 228, 230, 239  
 Иерархическая система (hierarchical systems) 73, 74, 79, 81, 82, 148—152, 157, 162, 168, 169, 269, 273, 275  
 Измерение (measurement) 20, 21, 39  
 — мобильности (of mobility) 30, 39—41, 102  
 — стабильности, текучести (of stability) 257—260  
 Имитация (simulation) 22, 23, 205, 209  
 Интенсивность выхода из строя по возрастам (age specific failure rate) 277  
 — продвижения (promotion rate) 150, 152, 160—169  
 — уходов, зависящая от времени (time-dependent loss rate) 157—161, 273—276  
 Интенсивность переходов (transition intensity) 27, 106, 108, 109, 112, 126, 127, 132, 134, 135, 136, 142, 143  
 — грубая (crude intensity) 142  
 — поступления (input rate) 152, 161  
 — рождения (birth rate) 170, 230, 271

- смерти, гибели (death rate) 143, 170, 220, 258, 265
- уходов, увольнений (loss wastage) 150, 152, 157—161, 182, 214, 215, 256, 264, 274, 277
- чистая (net) 142, 146
- Испытания технические (на надежность) (industrial life testing) 214, 215
- Кадров планирование (manpower planning) 20, 68, 81, 101, 173, 176, 193, 210, 211, 240, 241, 243, 278
  - текучесть (crude wastage) 245, 258
  - ые приложения (manpower application) 210
  - ые системы (manpower systems) 82, 148, 154, 157, 173, 188, 274
- Квази-стационарные системы (quasi-stationary system) 181
- Квази-устойчивое состояние (quasi-steady state) 232
- Ковариация размеров групп (covariance of grade size) 36—39, 84—86, 97—98
- Ковариационная матрица (variance — covariance matrix) 203, 211
- Когорта (cohort) 67, 71, 80—87, 136, 141, 216, 275
- Корни см. Собственные числа (roots) 26, 110, 111
- Коэффициент вариации (coefficient of variation) 248
  - корреляции (correlation) 45
  - рождаемости (birth rate) 47, 49, 170
  - — зависимый от групп (differential birth rate) 47, 48, 49
- Логистическое распределение (logistic distribution) 238
- Логнормальное (логарифмически-нормальное) распределение (lognormal distribution) 143, 228, 235—238
  - — смесь (mixture of) 228
  - модель размеров и длительности (lognormal model of size and duration) 224—228
- Марковское свойство (Markov property) 25, 33, 34, 55, 82, 103
- Марковская цепь (Markov chain) 20, 30, 31, 34, 116
  - — вложенная (embedded) 124—135
  - — зависящая от времени (time-dependent) 46, 100, 106
  - — поглощающая (absorbing) 57
  - — примитивная (primitive) 61
- Матрица вероятностей переходов (transition probability matrix) 27, 30, 31, 35, 40, 106, 107
- Медиана (median, «half-life») 20, 42
- Мера (measure) 32, 42
- Мобильность (mobility)
  - внутри поколения (intragenerational) 29
  - между поколениями (intergenerational) 29, 40, 42
  - профессиональная (occupational) 29, 39, 126
  - рабочей силы (labour) 29
  - социальная (social) 17, 21, 28, 29, 34, 35
  - структурная (structural) 103
  - чистая (pure) 103
- Модели, выводы (Model solving) 19
  - построение (building) 19, 22, 25, 215
  - проверка (testing) 19, 33
- Модель взаимодействия марковская (interactive Markov chain model) 29, 58—66
  - голосования (voting model) 64, 66
  - движения, зависящая от времени (time-dependent mobility model) 46, 47
  - математическая (mathematical) 17
  - МС («mover-stayer» model) 7, 53, 54, 118—120
  - процесса принятия решений (decision process model) 219
  - процесса уходов (model for leaving process) 215—218
  - сетевая (network model) 219—221
  - Силкока (Silcock's) 215—218, 245
- Найм (recruitment) 18
  - управление (control) 176—182
  - планирование (planning) 244
  - политика (politics) 19, 176
  - стратегия (strategies) 203—205
- Накопленная инерция (cumulative inertia) 55—58, 118, 125
- Немарковские модели, системы, процессы (non-Markovian models, systems, processes) 148, 165
- Неопределенность (competing risk) 126, 141, 171
- Неподвижные общества (immobile society) 42, 49
- Неравенство (inequality)
  - для следов матриц (trace) 53, 54
  - Иенсена (Jensen's) 54
- Нормальное распределение (normal distribution) 27, 200, 202, 222—225, 228, 246
  - — многомерное (multivariate) 200, 210
- Оборот рабочей силы (labour turnover) 257—260

- — — измерение (measurement) 258, 259
- — — индекс (crude index) 259
- Обращаемый (обратимый) процесс Маркова (reversible Markov process) 34, 35, 43, 103
- цепь Маркова (reversible Markov chain) 34, 35, 43, 103
- Оптимальность стратегий управления (optimality of control strategies) 187, 189, 191, 197
- Открытая система (open system) 25, 29, 69, 112, 126, 136—172
- Отрицательное биномиальное распределение (negative binomial distribution) 122, 123, 230, 231
- Оценивание (estimation) 20, 224
- интенсивностей перехода (of transition rate) 220
- переходных вероятностей (of transition probabilities) 104
- функции дожития (—of a survivor function) 220
- Очередей теория (queues theory) 22, 165
- Парето распределение (Pareto distribution) 228, 235, 236, 238
- параметр (Pareto parameter) 228
- Переходные вероятности (transition probabilities) 29, 50, 115, 149, 162, 199
- Период полужизни (half-life) 42, 247, 259
- Поглощающее состояние (absorbing, terminal state) 136, 137, 141, 148, 219, 222, 223, 276
- Поддержание (maintainability) 176—186, 198—207
- на  $n$  шагах ( $n$ -step maintainability) 210
- Поддерживаемые структуры (maintainable structures) 181, 182, 200, 203—207
- Полиномиальное распределение (multinomial distribution) 35, 38, 39, 70, 199, 200
- Полный срок службы, ПСС (completed length of service, CLS) 159, 163, 214, 215, 259, 269
- Понижение (demotion) 93, 157, 174
- интенсивность (demotion rate) 182
- Поток (flow) 28, 38, 101, 157, 173, 175, 190, 198, 202, 211, 263
- Правило (rule) 167, 264, 270, 271
- Предельная структура (limiting structure) 44, 111
- Предельный режим (steady state, limiting behaviour) 154—157
- Преобразование Лапласа (Laplace transform) 27, 128—130, 134, 153, 155, 156, 172, 242, 265—268
- Программы для ЭВМ (computer programs) 25, 38, 73, 98
- Продвижение по стажу работы (promotion by seniority) 19, 264, 269, 271
- случайным образом (at random) 264, 270—272
- Производящая функция (generating function) 27, 87—89, 139, 140
- Процесс размножения и гибели (Birth and death process) 170, 229, 239
- полумарковский (semi-Markov process) 112, 125—135, 161—169, 171
- Пуассона (Poisson process) 27, 120, 124, 125, 139, 140, 171, 264, 278
- — зависящий от времени (time-dependent) 135, 140, 154, 157
- Равновесие (установившееся распределение, устойчивая структура) (equilibrium) 33, 43, 103, 133, 250
- Развитие системы (expanding system) 91—99, 173, 185, 253—256
- — с заданным (установленным) размером (with given, fixed size) 91—99, 153—161
- Разностные уравнения (differential equations) 34, 91, 171, 203
- Распределение (distribution)
- прибыли, дохода (of income) 101, 212, 225, 236, 239
- принятых (recruitment) 70, 186
- размеров (of size) 228—236, 239
- Рост фирмы, предприятия (growth of firm) 23
- линейный (linear) 232, 253
- по закону геометрической прогрессии (geometric) 92—96, 152
- управление (control under growth) 180—185
- Свертка (convolution) 186, 267
- Семейная история (family history) 30, 33
- линия (line) 30, 31, 35, 36, 43, 44, 45, 48, 50
- Склонность, стремление к уходу (propensity to leave, force of separation) 20, 214, 216, 217, 258, 259, 261, 273
- к движению (to move) 173
- След матрицы (trace of matrix) 42, 53, 54, 115, 123, 131
- Смертность естественная (mortality natural) 20, 143, 145
- измерение (measurement of) 143, 145
- стандартизованная (standardized) 258
- Смещение в поддерживаемых структурах (bias in maintaining a structures) 205, 207

- Собственные векторы (eigenvector) 72, 79, 100, 121
- числа, значения (eigenvalues) 26, 41—43, 54, 67, 71, 72, 74, 78, 79, 85, 92, 100, 103, 109, 110, 111, 117, 121, 131, 138, 143, 149, 150, 172, 220
- Совершенно мобильное общество (perfectly mobile society) 38, 40, 49
- Сокращающаяся система (contracting system) 185, 255, 256
- Сокращение штатов (redundancy) 98, 99, 189, 211, 221, 222, 256
- Сопряженное гамма-распределение (inverse gamma distribution) 218
- Сопутствующие, составляющие матрицы (spectral set) 41, 71—76, 79, 109, 110
- Спектральное разложение, представление (spectral representation) 41, 71—77, 92, 109, 110, 122
- Средние квадратические ошибки (standard errors) 38, 249, 252
- Средний квадрат ошибки (mean square errors) 205, 206, 208
- Страхование (actuaries) 106, 125, 143, 145, 147, 214, 221, 258
- Структура групп (class structure) 30, 49, 173, 213
- Супердиагональная матрица (super-diagonal matrix) 179, 183, 185
- Таблицы, матрицы мобильности (mobility matrices, tables) 40, 102, 103, 104**
- Тауберовы теоремы (Tauberian theorems) 130, 277
- Текучесть (turnover, crude wastage) 20, 21, 174, 214, 215, 244, 245, 257—260
- Теория надежности (reliability theory) 106, 213, 214, 236, 241, 246
- Тип XI распределения Пирсона (Pearson type XI distribution) 216, 217, 245, 257, 274
- Трудовых ресурсов движение (labour, manpower mobility) 101, 171
- Убыль по многим причинам (multiple decrements) 125, 171, 221**
- Увольнения, в административном порядке (wastage, dismissing) 18, 175
- вероятность (wastage probability) 74
  - интенсивность (wastage rate) 182, 184, 245, 255, 256, 274, 275, 277
- Управление без временных ограничений (control, free time) 187—192, 197, 198, 208
- персоналом (personnel management) 218
  - с временными ограничениями (fixed time control) 192—198, 208
  - стохастическое (stochastic) 210
- Уравнение в конечных разностях (difference equations) 99, 176, 203
- Установившаяся структура (steady-state; ultimate, limiting structure, distribution) 41, 42, 44, 62, 64, 78, 79, 90, 97, 112, 124, 138, 154—157, 213, 229
- Фиксированная, установленная численность (fixed size) 91—99, 153—157
- Фирмы размер (size of firm) 101, 212
- Форма J-образная (J-shaped curves) 215, 217, 218, 239, 277
- Функция дожития (survivor function) 214, 217, 220, 252, 265, 267
- правдоподобия (likelihood function) 20
  - производящая (generating function) 87, 88, 135, 140
  - распределения вероятностей (distribution function) 26, 214
- Хербста модель (Herbst's model) 221
- Целевая, желаемая структура (goal, target structure) 173, 192, 194, 198, 204, 207, 208
- Центральная предельная теорема (central limit theorem) 199, 212, 226, 239, 249
- Цепочки вакансий (vacancy chains) 27, 275, 276, 278
- Цель Маркова (Markov chain) 20, 30, 31, 55, 60, 61, 64, 81, 94, 105, 115, 121, 173
- Шоррокса модель (Shorrocks model) 229, 230, 239**
- Экспоненциальный закон распределения (exponential distribution)
- плотность (density) 108, 126, 143, 160, 168, 213, 215—219, 243, 245, 262, 266, 274
  - рост (growth) 155, 157
  - скорость (rate) 155, 157
  - случайная величина (variate) 165
  - — — смесь (mixed exponential distribution) 168, 215—219, 269
  - — — ПСС (CLS) 159, 214
- Экстремальных значений распределение (extreme value distribution) 212, 236
- Эрланга распределение (Erlang distribution) 266
- Юла—Симона модель (Yule—Simon model) 231—233**

Предисловие к русскому изданию . . . . .	5
Предисловие . . . . .	13
<b>Г л а в а 1. Стохастическое моделирование . . . . .</b>	<b>15</b>
1.1. Физика социальных явлений . . . . .	15
1.2. Модели и моделирование . . . . .	16
1.3. Некоторые основные вопросы . . . . .	21
1.4. Предпосылки, терминология и обозначения . . . . .	24
<b>Г л а в а 2. Модели мобильности в замкнутых социальных системах . . . . .</b>	<b>28</b>
2.1. Введение . . . . .	28
2.2. Марковские модели социальной мобильности . . . . .	29
Основная модель. Обращение времени в марковской модели. Изменчивость размеров группы.	
2.3. Измерение мобильности . . . . .	39
Измерение зависимости между поколениями. Меры количества перемещений. Другие подходы.	
2.4. Некоторые обобщения марковской модели . . . . .	46
Модели, зависящие от времени. Модель с зависимым от группы коэффициентом рождаемости. Неоднородность переходных вероятностей. Модели накопленной инерции.	
2.5. Марковские модели взаимосвязей . . . . .	58
Модели притяжения и отталкивания. Модель неудовлетворенности. Модель голосования.	
<b>Г л а в а 3. Марковские модели открытых систем . . . . .</b>	<b>67</b>
3.1. Введение . . . . .	67
3.2. Модель системы с заданным входом . . . . .	69
Базовая модель. Спектральное разложение $\bar{p}(T)$ . Иерархическая структура без понижения в должности. Предельное поведение $\bar{p}(T)$ . Применения к когорте. Дисперсии и ковариации размеров групп. Пример с периодической переходной матрицей.	
3.3. Модель развития системы с заданным размером . . . . .	91
Уравнения модели. Точное решение для роста по закону геометрической прогрессии. Предельная структура. Дисперсии и ковариации размеров уровней. Контрактные системы.	
3.4. Дополнения к гл. 2 и 3 . . . . .	100
Планирование системы образования и планирование кадров. Демография. География. Экономика. Планирование здравоохранения.	
<b>Г л а в а 4. Модели с непрерывным временем для замкнутых социальных систем . . . . .</b>	<b>105</b>
4.1. Основы теории . . . . .	105
Основные уравнения. Решение основных уравнений. Предельное поведение.	

4.2.	Примеры . . . . .	112
4.3.	Дискретное наблюдение непрерывного процесса . . . . .	115
	Предельное поведение.	
4.4.	Полумарковские модели . . . . .	125
	Вероятности переходов на фиксированном интервале времени.	
	Примеры. Наблюдение на интервале с произвольной начальной точкой. Предельный режим. Модели, основанные на более общих точечных процессах.	
<b>Глава 5. Модели открытых социальных систем с непрерывным временем . . . . .</b>		<b>136</b>
5.1.	Марковская модель с заданным входом . . . . .	136
	Предельный режим. Распределение численностей.	
5.2.	Практический пример описания системы с заданным входом . . . . .	141
	Выживание после лечения заболевания раком. Изучение когорты. Применение теории. Иерархическая кадровая система.	
5.3.	Модель, описывающая систему установленного размера . . . . .	153
	Предельный режим. Числовой пример.	
5.4.	Системы с установленными поступлениями и уходами, зависящими от продолжительности пребывания . . . . .	157
	Теория. Иллюстрация теоретических результатов.	
5.5.	Полумарковские модели открытых систем . . . . .	161
	Прямой подход. Метод введения гипотетических состояний.	
5.6.	Дополнения к гл. 4 и 5 . . . . .	169
	Теория. Приложения. Разновидности и обобщения.	
<b>Глава 6. Теория управления на базе марковских моделей . . . . .</b>		<b>173</b>
6.1.	Постановка задачи . . . . .	173
	Управляющие переменные	
6.2.	Поддержание состояний в детерминистской внешней среде . . . . .	176
	Управление наймом. Управление наймом при развитии и сокращении системы. Управление посредством продвижений. Поддержание при непрерывном времени.	
6.3.	Достижимость в детерминистской внешней среде . . . . .	186
	Управление наймом без временных ограничений. Управление продвижением без временных ограничений. Управление наймом и продвижением с временным ограничением. Численное сравнение стратегий при управлении наймом.	
6.4.	Поддержание состояний при случайных внешних воздействиях . . . . .	198
	Вероятность поддержания структуры посредством найма. Вероятность поддержания структуры посредством продвижений. Стратегии поддержания структур посредством найма. Стратегии поддержания структур при управлении продвижением.	
6.5.	Достижимость в случайных средах . . . . .	208
6.6.	Дополнения . . . . .	210
<b>Глава 7. Модели продолжительности пребывания в состоянии и размеров системы . . . . .</b>		<b>212</b>
7.1.	Предпосылки . . . . .	212
7.2.	Модели времени пребывания в состоянии . . . . .	213
	Экспоненциальное и смешанное экспоненциальное распределения. Сетевые модели. Модели случайного блуждания.	
7.3.	Логнормальная модель размеров системы и продолжительности пребывания в состоянии . . . . .	224
7.4.	Модели размеров систем . . . . .	228
	Модель Шоррокса. Модель Юла — Симона. Модели размеров с дискретным временем.	
7.5.	Дополнения . . . . .	236

Глава 8. Модели социальных систем с установленными размерами групп . . . . .	240
8.1. Введение . . . . .	240
8.2. Системы, состоящие из одной группы . . . . .	241
Уравнение восстановления. Решение уравнения восстановления, когда распределение полного срока службы — смесь экспоненциальных распределений. Приближенное решение уравнения восстановления. Распределение численностей нанятых. Распределение по стажу. Развивающиеся системы. Сокращающиеся системы.	
8.3. Измерение оборота трудовых ресурсов . . . . .	257
8.4. Модели многоуровневых систем . . . . .	260
Модель с заданными интенсивностями уходов. Распределение потоков. Длительность пребывания. Распределение по стажу.	
8.5. Ранговые системы с более общим механизмом уходов . . . . .	273
8.6. Цепочки вакансий . . . . .	275
8.7. Дополнения . . . . .	276
Библиографии . . . . .	278
Предметный указатель . . . . .	288

26 **Бартоломью Д.** Стохастические модели социальных процессов / Пер. с англ.; Под ред. О. В. Староверова. — М.: Финансы и статистика, 1985. — 295 с., ил. — (Математико-статистические методы за рубежом).

В пер.: 2 р. 80 к. 3700 экз.

Книга английского ученого Д. Бартоломью посвящена применению статистических методов для анализа использования трудовых ресурсов. Рассматриваются модели различных замкнутых и открытых ранговых социальных систем, способы управления структурной организационной системой. Изложение сопровождается разбором большого количества примеров.

Для специалистов-статистиков, экономистов, социологов.

Б  $\frac{070200000-022}{010(01)-85}$  84-85

ББК 22.172  
517.8

**Дэвид Дж. Бартоломью**

## **СТОХАСТИЧЕСКИЕ МОДЕЛИ СОЦИАЛЬНЫХ ПРОЦЕССОВ**

Книга одобрена на заседании редколлегии серии «Математико-статистические методы за рубежом» 31.08.81

Зав. редакцией *А. В. Павлюков*  
Редактор *Е. В. Крестьянинова*  
Мл. редактор *О. А. Ермилина*  
Техн. редакторы *Л. Г. Чельшева, Г. А. Полякова*  
Корректоры *Г. А. Башарина, Т. М. Иванова*  
и *Т. В. Рослякова*  
Худож. редактор *Ю. И. Артюхов*

ИБ № 1571

---

Сдано в набор 31.08.84. Подписано в печать 24.12.84.  
Формат 60×90<sup>1/16</sup>. Бум. тип. № 1. Гарнитура «Литературная».  
Печать высокая. Усл. п. л. 18,5. Усл. кр.-отт. 18,5. Уч.-изд. л.  
19,97. Тираж 3700 экз. Заказ 433. Цена 2 р. 80 к.  
Издательство «Финансы и статистика»,  
101000, Москва, ул. Чернышевского, 7

---

Московская типография № 4 Союзполиграфпрома  
при Государственном комитете СССР  
по делам издательств, полиграфии и книжной торговли  
129041, Москва, Б. Переяславская ул., 46