МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РФ
Федеральное государственное автономное образовательное учреждение высшего образования
Национальный исследовательский
Томский государственный университет

Факультет инновационных технологий
Кафедра информационного обеспечения инновационной деятельности

МАТЕМАТИЧЕСКОЕ И ИМИТАЦИОННОЕ МОДЕЛИРОВАНИЕ ЭКОНОМИЧЕСКИХ ПРОЦЕССОВ

ТОМСК – 2016

Составитель А.А. Мицель

Математическое и имитационное моделирование экономических процессов

Томск: Изд-во ТГУ, 2016. – 193с.

В пособии приведены следующие темы дисциплины «Математическое и имитационное моделирование экономических процессов» — основные понятия математического моделирования в экономике, модели производства, функция полезности, балансовые модели, моделирование финансовых операций, математическое и компьютерное моделирование, сущность метода имитационного моделирования, имитационные модели глобальных систем, метод Монте-Карло и проверка статистических гипотез, моделирование случайных событий, системы массового обслуживания, модели управления запасами.

Учебное пособие предназначено для студентов бакалавриата направления подготовки 09.03.03 – прикладная информатика (профиль – прикладная информатика в информационной сфере). Кроме того, это пособие может быть использовано студентами других смежных экономических специальностей.

СОДЕРЖАНИЕ

Тема 1. Основные понятия математического моделирования в экономике

1.1 Краткий исторический обзор

Экономико-математические методы и модели применяют с целью отыскания наилучшего решения, то есть решения, оптимального в том или ином смысле (максимума или минимума).

Становление математических методов анализа и выработки хозяйственных решений как самостоятельной ветви математики произошло в XVIII веке.

Во Франции Франсуа Кенэ, врач и экономист, предпринял одну из первых попыток экономико-математического моделирования механизма движения финансов. Он построил экономическую таблицу, рассматривающую экономику государства как единую систему. Кенэ применил идею кровообращения человека к кругообороту экономических отношений.

Карл Маркс, используя таблицы Кенэ, ввел алгебраические формулы и мечтал «вывести главные законы кризисов». В работах Маркса впервые сделано математическое формализованное описание процесса расширенного воспроизводства.

В 1838 году французский математик Антуан Курно выпустил книгу «Исследование математических принципов теории богатства». В ней была впервые предложена математическая зависимость спроса и цены товара. Эти величины связаны коэффициентом эластичности, который показывает, как изменяется спрос при росте или снижении цепы на 1%. Функция спроса позволила вскрыть ряд закономерностей. Продавать подороже не всегда выгодно. Все зависит от коэффициента эластичности. Спрос на товары, для которых он больше единицы, при снижении цены растет так быстро, что общая прибыль от продажи увеличивается.

В 1874 году швейцарский экономист Л. Вальрас ввел статистическую модель системы экономического равновесия, затем итальянский экономист В. Парето предложил модель распределения доходов населения.

Конец XIX и начало XX века характеризуется значительной активизацией работ развивающих математические методы решения экономических задач. Одной из первых задач, решенных на основе математического подхода, является «задача о землекопе», сформулированная Фредериком Тейлором в 1885 году. В задаче требовалось определить оптимальную разовую массу подбираемой земли, обеспечивающую максимум объема работ землекопа за день. Если землекоп за раз забирает много земли, то усталость его быстро нарастает. Если брать за раз мало земли, то падает общий объем работ.

Становление современного математического аппарата оптимальных экономических решений началось в 40-е годы, благодаря первым работам Н. Винера, Р. Беллмана, С. Джонсона, Л. В. Канторовича.

В 1938 году перед двадцатипятилетним профессором ЛГУ Канторовичем Л. В. была поставлена задача: как наилучшим образом распределить работу восьми станков фанерного треста при условии, что известна производительность каждого станка по каждому из пяти видов обрабатываемых материалов. В 1939 году выдающийся советский математик и экономист публикует работу «Математические методы организации и планирования производства», в которой впервые формулирует задачу линейного программирования и разрабатывает алгоритм ее решения. Совместно с американским ученым Т. Купмансом в 1975 году Л. Канторович получает Нобелевскую премию за вклад в теорию оптимизации распределения ресурсов.

Исторически общая задача линейного программирования ставится в 1947 году Данцигом Дж. и Вудом М. в департаменте ВВС США. Данцигом предлагается универсальный алгоритм решения задач линейного программирования, названный им симплекс-методом. В 1941 году Хичкок и независимо от него Купсман в 1947 году формулируют транспортную задачу, Стиглер в 1945 году — задачу о диете.

В 50-60-х годах появляются значительные работы в области экономико-математического моделирования и у нас, в том числе: Канторович Л. В. «Экономический расчет наилучшего исследования ресурсов» (1959); Канторович Л. В., Гавурин М. К. «Применение математических методов в вопросах анализа грузопотоков» (1949); работы Новожилова В. В. по оптимальному планированию народного хозяйства. В I960 г. академик Немчинов В. С. при Новосибирском отделении АН СССР создает лабораторию экономико-математического моделирования.

1.2 Математические методы и моделирование экономических процессов

Термин *экономико-математические методы* понимается как обобщающее название комплекса экономических и математических научных дисциплин, объединенных для изучения социально-экономических систем и процессов.

Под *социально-экономической системой* будем понимать сложную вероятностную динамическую систему, охватывающую процессы производства, обмена, распределения и потребления материальных и других благ.

Основным методом исследования систем и процессов является *метод моделирования,* т. е. способ теоретического анализа и практического действия, направленный на разработку и использование моделей. При этом под *моделью* будем понимать образ реального объекта (процесса) в материальной или идеальной форме (т. е. описанный знаковыми средствами на каком-либо языке), отражающий существенные свойства моделируемого объекта (процесса) и замещающий его в ходе исследования и управления. В дальнейшем мы будем говорить только об экономико-математическом моделировании, т. е. об описании знаковыми математическими средствами социально-экономических систем. Практическими задачами экономико-математического моделирования являются:

- анализ экономических объектов и процессов;
- прогнозирование развития экономических процессов;
- выработка управленческих решений на всех уровнях хозяйственной иерархии.

Следует, однако, иметь в виду, что далеко не во всех случаях данные, полученные в результате экономико-математического моделирования, могут использоваться непосредственно как готовые управленческие решения. Они скорее могут быть рассмотрены как «консультирующие» (советующие) средства. Принятие управленческих решений остается за человеком. Таким образом, экономико-математическое моделирование является лишь одним из компонентов (пусть очень важным) в человеко-машинных системах планирования и управления экономическими системами.

Важнейшим понятием при экономико-математическом моделировании, как и при всяком моделировании, является понятие ***адекватности*** *модели,* т. е. соответствия модели моделируемому объекту или процессу по тем свойствам, которые считаются существенными для исследования. Проверка адекватности экономико-математических моделей является весьма серьезной проблемой, тем более, что ее осложняет трудность измерения экономических величин. Однако без такой проверки применение результатов моделирования в управленческих решениях может не только оказаться мало полезным, но и принести существенный вред.

В наше время экономико-математическое моделирование применяют к широкому классу задач, связанному со сложными организационными структурами современной экономики. Наша естественная склонность ставить и решать подобные задачи проявляется в выражениях типа «с наименьшими затратами», «максимальная прибыль», «полная отдача» и т. п. Сюда относятся задачи наиболее эффективного управления предприятием, распределения ресурсов, управления технологическими процессами, создания оптимальных конструкций, управления грузопотоками, персоналом и многие другие.

Эти задачи возникают не только в промышленности, но и в повседневной жизни каждого человека. Например, задача программирования утреннего одевания.*) Следует выбрать такой вариант последовательности одевания — программу, которая позволит выполнить определенные ограничения или общепринятые правила. Если программа включает шесть предметов одежды: ботинки» носки, брюки, рубашку, галстук, пиджак г — то программа — любой порядок, в котором можно надеть эти предметы. Всего в этом случае существует 6! = 720 различных программ. Многие из них недопустимы (носки поверх ботинок, галстук под рубашку) и если их отбросить, все равно остается несколько допустимых программ, которые нужно исследовать. Как же выбрать окончательное, оптимальное решение?

В этой или любой, другой задаче; где необходимо анализировать все возможные варианты решений, и выбрать единственный, оптимальный, имеется некая основная цель (целевая функция, критерий качества), позволяющая сравнивать эффективность этих допустимых вариантов (программ действий). Если мы можем указать целевую функцию, то тем самым можем выбрать и оптимальную программу действий. Если целевая функция связана с затратами времени, то оптимальная программа утреннего одевания: носки, рубашка, брюки, галстук, ботинки, пиджак — минимизирует время на одевание без нарушения общепринятых ограничений. Но может быть выбрана и другая целевая функция, например, минимизация утреннего шума — как можно меньше открывать и закрывать дверцы и шкафчики. Тогда будет и другое оптимальное решение.

Задачи математического программирования существуют только тогда, когда имеется много допустимых решений (по крайней мере, от двух и более). Если допустимое решение единственное, то не возникает никакой проблемы по поиску решения.

Неоптимальное решение этих задач приводит к излишним затратам сырья и времени. Допустим, что при интуитивном распределении людей на работы эффективность их использования по сравнению с оптимальным вариантом, рассчитанным на компьютере, ухудшается всего на 3%. Казалось бы, очень небольшая погрешность, на которую можно и не обратить внимания. Такая погрешность означала бы, например, в гончарном цехе прошлых веков с 30 работниками неполную загрузку в течение рабочего дня всего лишь одного из них. А в наши дни, если принять число занятых в народном хозяйстве 60 млн. человек» такая же погрешность может явиться причиной сокращения числа рабочих мест почти для 2 млн. человек.

1.3 Этапы математического моделирования

Выделим шесть основных этапов ЭММ: постановка экономической проблемы, ее качественный анализ; построение математической модели; математический анализ модели; подготовка исходной информации; численное решение; анализ численных результатов и их применение. Рассмотрим каждый из этапов более подробно.

1. Постановка экономической проблемы и ее качественный анализ. На этом этапе требуется сформулировать сущность проблемы, принимаемые предпосылки и допущения. Необходимо выделить важнейшие черты и свойства моделируемого объекта, изучить его структуру и взаимосвязь его элементов, хотя бы предварительно сформулировать гипотезы, объясняющие поведение и развитие объекта.

2. Построение математической модели. Это этап формализации экономической проблемы, т. е. выражения ее в виде конкретных математических зависимостей (функций, уравнений, неравенств и др.). Желательно построить модель, относящуюся к хорошо изученному классу математических задач, что может потребовать некоторого упрощения исходных предпосылок модели, не искажающего основных черт моделируемого объекта. Однако возможна и такая ситуация, когда формализация проблемы приводит к неизвестной ранее математической структуре.

Для некоторых сложных объектов целесообразно строить несколько разноаспектных моделей; при этом каждая модель выделяет лишь некоторые стороны объекта, а другие стороны учитываются агрегированно и приближенно.

3. Математический анализ модели. На этом этапе чисто математическими приемами исследования выявляются общие свойства модели и ее решений. В частности, важным моментом является доказательство существования решения сформулированной задачи. При аналитическом исследовании выясняется, единственно ли решение, какие переменные могут входить в решение, в каких пределах они изменяются, каковы тенденции их изменения и т. д. Однако модели сложных экономических объектов с большим трудом поддаются аналитическому исследованию; в таких случаях переходят к численным методам исследования.

4. Подготовка исходной информации. В экономических задачах это, как правило, наиболее трудоемкий этап моделирования, так как дело не сводится к пассивному сбору данных. Математическое моделирование предъявляет жесткие требования к входной информации; при этом надо принимать во внимание не только принципиальную возможность подготовки информации требуемого качества, но и затраты на подготовку информационных массивов. В процессе подготовки информации используются методы теории вероятностей, теоретической и математической статистики для организации выборочных обследований, оценки достоверности данных и т.д. При системном экономико-математическом моделировании результаты функционирования одних моделей служат исходной информацией для других.

5. Численное решение. Этот этап включает разработку алгоритмов численного решения задачи, подготовку программ на ЭВМ и непосредственное проведение расчетов; при этом значительные трудности вызываются большой размерностью экономических задач. Обычно расчеты на основе экономико-математической модели носят многовариантный характер. Многочисленные модельные эксперименты, изучение поведения модели при различных условиях возможно проводить благодаря высокому быстродействию современных ЭВМ. Численное решение существенно дополняет результаты аналитического исследования, а для многих моделей является единственно возможным..

6. Анализ численных результатов и их применение. На этом этапе прежде всего решается важнейший вопрос о правильности и полноте результатов моделирования и применимости их как в практической деятельности, так и в целях усовершенствования модели. Поэтому в первую очередь должна быть проведена проверка адекватности модели по тем свойствам, которые выбраны в качестве существенных (другими словами, должны быть произведены верификация и валидация модели)[1]. Применение численных результатов моделирования в экономике направлено на решение практических задач (анализ экономических объектов, экономическое прогнозирование развития хозяйственных и социальных процессов, выработка управленческих решений на всех уровнях хозяйственной иерархии).

---

[1] *Верификация* модели — проверка правильности структуры (логики) модели; *валидация* модели — проверка соответствия данных, полученных на основе модели, реальному процессу.

Перечисленные этапы экономико-математического моделирования находятся в тесной взаимосвязи, в частности, могут иметь место возвратные связи этапов. Так, на этапе построения модели может выясниться, что постановка задачи или противоречива, или приводит к слишком сложной математической модели; в этом случае исходная постановка задачи должна быть скорректирована. Наиболее часто необходимость возврата к предшествующим этапам моделирования возникает на этапе подготовки исходной информации. Если необходимая информация отсутствует или затраты на ее подготовку слишком велики, приходится возвращаться к этапам постановки задачи и ее формализации, чтобы приспособиться к доступной исследователю информации.

1.4 Классификация математических моделей

Выделяют следующие признаки классификации, или классификационные рубрики.

По общему целевому назначению экономико-математические модели делятся на *теоретико-аналитические,* используемые при изучении общих свойств и закономерностей экономических процессов, и *прикладные,* применяемые в решении конкретных экономических задач анализа, прогнозирования и управления. Различные типы прикладных экономико-математических моделей как раз и рассматриваются в данном учебном пособии.

По степени агрегирования объектов моделирования модели разделяются на *макроэкономические и микроэкономические.* Хотя между ними и нет четкого разграничения, к первым из них относят модели, отражающие функционирование экономики как единого целого, в то время как микроэкономические модели связаны, как правило, с такими звеньями экономики, как предприятия и фирмы.

По конкретному предназначению, т. е. по цели создания и применения, выделяют *балансовые* модели, выражающие требование соответствия наличия ресурсов и их использования; *трендовые* модели, в которых развитие моделируемой экономической системы отражается через тренд (длительную тенденцию) ее основных показателей; *оптимизационные* модели, предназначенные для выбора наилучшего варианта из определенного числа вариантов производства, распределения или потребления; *имитационные* модели, предназначенные для использования в процессе машинной имитации изучаемых систем или процессов и др.

По типу информации, используемой в модели, экономико-математические модели делятся на *аналитические,* построенные на априорной информации, и *идентифицируемые,* построенные на апостериорной информации.

По учету фактора времени модели подразделяются на *статические,* в которых все зависимости отнесены к одному моменту времени, и *динамические,* описывающие экономические системы в развитии.

По учету фактора неопределенности модели распадаются на *детерминированные,* если в них результаты на выходе однозначно определяются управляющими воздействиями, и *стохастические* (вероятностные), если при задании на входе модели определенной совокупности значений на ее выходе могут получаться различные результаты в зависимости от действия случайного фактора.

Экономико-математические модели могут классифицироваться также по типу математического аппарата, используемого в модели. По этому признаку могут быть выделены *матричные* модели, модели *линейного* и *нелинейного программирования, корреляционно-регрессионные* модели, модели *теории массового обслуживания,* модели *сетевого планирования и управления,* модели *теории игр* и т.д.

Наконец, по типу подхода к изучаемым социально-экономическим системам выделяют *дескриптивные* и *нормативные* модели. При дескриптивном (описательном) подходе получаются модели, предназначенные для описания и объяснения фактически наблюдаемых явлений или для прогноза этих явлений; в качестве примера дескриптивных моделей можно привести названные ранее балансовые и трендовые модели. При нормативном подходе интересуются не тем, каким образом устроена и развивается экономическая система, а как она должна быть устроена и как должна действовать в смысле определенных критериев. В частности, все оптимизационные модели относятся к типу нормативных; другим примером могут служить нормативные модели уровня жизни.

Рассмотрим в качестве примера экономико-математическую модель межотраслевого баланса (ЭММ МОБ). С учетом приведенных выше классификационных рубрик это прикладная, макроэкономическая, дескриптивная, детерминированная, балансовая, матричная модель; при этом существуют как статические, так и динамические ЭММ МОБ.

Вопросы для самопроверки

1. Что понимается под математическими методами в экономике?
2. Что такое социально-экономическая система?
3. Перечислите задачи математического моделирования в экономике.
4. Что такое адекватность модели?
5. Что такое оптимальное решение?
6. Перечислите основные этапы математического моделирования.
7. Перечислите основные признаки классификации математических моделей.

Тема 2. Модели производства

2.1 Производственные функции

2.1.1 Понятие производственной функции одной переменной

Производственная функция - это функция, независимая переменная которой принимает значения объемов затрачиваемого или используемого ресурса (фактора производства), а зависимая переменная - значения объемов выпускаемой продукции

$$y = f(x) \tag{2.1}$$

В формуле (1) x $(x \geq 0)$ и y $(y \geq 0)$ – числовые величины, т.е. $y = f(x)$ есть функция одной переменной x. В связи с этим производственная функция (ПФ) называется одноресурсной или однофакторной ПФ, ее область определения - множество неотрицательных действительных чисел (т.е. $x \geq 0$). Запись $y = f(x)$ означает, что если ресурс затрачивается или используется в количестве x единиц, то продукция выпускается в количестве $y = f(x)$ единиц. ПФ задается с точностью до параметров. Более правильной является символика $y = f(x, a)$, где a – вектор параметров ПФ.

Пример 1. Возьмем ПФ в виде $f(x) = ax^b$, где x - величина затрачиваемого ресурса (например, рабочего времени), $f(x)$ – объем выпускаемой продукции (например, число готовых к отправке холодильников). Величины a и b – параметры ПФ. Здесь a и b - положительные числа и число $b \leq 1$.

График производственной функции $y = ax^b$ изображен на рис. 2.1.

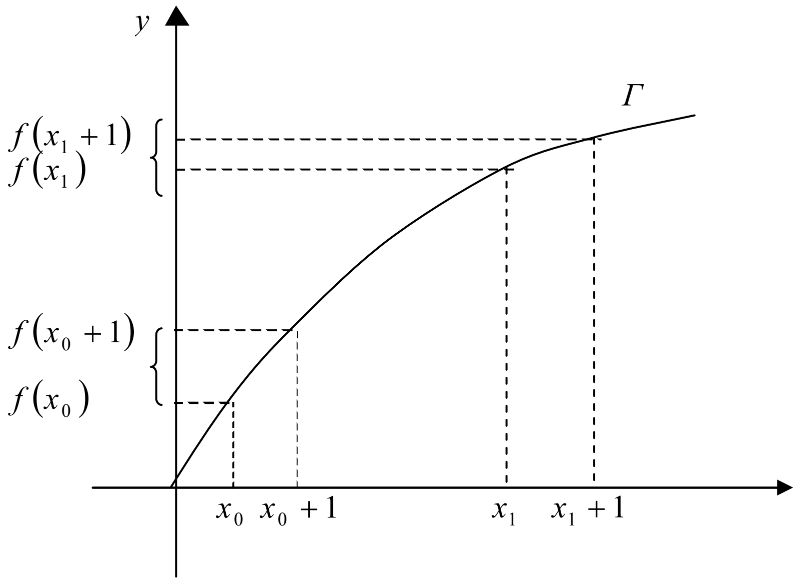


Рис. 2.1

На графике видно, что с ростом величины затрачиваемого ресурса x объем выпуска y растет, однако при этом каждая дополнительная единица ресурса дает все меньший прирост объема у выпускаемой продукции. Отмеченное обстоятельство (рост объема y и уменьшение прироста объема у с ростом величины x) отражает фундаментальное положение экономической теории (хорошо подтверждаемое практикой), называемое законом убывающей эффективности. ПФ $y = ax^b$ является типичным представителем широкого класса однофакторных ПФ.

ПФ могут иметь разные области использования. Принцип "затраты - выпуск" может быть реализован как на микро-, так и на макроэкономическом уровне. Сначала остановимся на микроэкономическом уровне. ПФ $y = ax^b$, рассмотренная выше, может быть использована для описания взаимосвязи между величиной затрачиваемого или используемого ресурса x в течение года на отдельном предприятии (фирме) и годовым выпуском продукции y этого предприятия (фирмы). В роли производственной системы здесь выступает отдельное предприятие (фирма) - имеем микроэкономическую ПФ (МИПФ). На микроэкономическом уровне в роли производственной системы может выступать также отрасль, межотраслевой производственный комплекс. МИПФ строятся и используются в основном для решения задач анализа и планирования, а также задач прогнозирования.

ПФ может быть использована для описания взаимосвязи между годовыми затратами труда в масштабе региона или страны в целом и годовым конечным выпуском продукции (или доходом) этого региона или страны в целом. Здесь в роли производственной системы выступает хозяйственная система региона или страны в целом, т.е. имеем *макроэкономический* уровень и *макроэкономическую* ПФ (МАПФ). МАПФ строятся и активно используются для решения всех трех типов задач (анализа, планирования и прогнозирования).

Точное толкование понятий затрачиваемого (или используемого) ресурса и выпускаемой продукции, а также выбор единиц их измерения зависят от характера и масштаба производственной системы, особенностей решаемых (с помощью ПФ) задач (аналитических, плановых, прогнозных), наличия исходных данных. На микроэкономическом уровне затраты и выпуск могут измеряться как в натуральных, так и в стоимостных единицах (показателях). Годовые затраты труда могут быть измерены в человеко-часах (объем человеко-часов - натуральный показатель) или в рублях выплаченной заработной платы (ее величина - стоимостный показатель); выпуск продукции может быть представлен в штуках или в других натуральных единицах (тоннах, метрах и т.п.) или в виде своей стоимости.

На макроэкономическом уровне затраты и выпуск измеряются, как правило, в стоимостных показателях и представляют собой стоимостные (ценностные) агрегаты, т.е. суммарные величины произведений объемов затрачиваемых (или используемых) ресурсов и выпускаемых продуктов на их цены.

2.1.2 Производственная функция нескольких переменных

Производственная функция нескольких переменных - это функция, независимые переменные $x_1,...,x_n$ которой принимают значения объемов затрачиваемых или используемых ресурсов (число переменных n равно числу ресурсов), а значение функции имеет смысл величин объемов выпуска:

$$y = f(x) = f(x_1,...,x_n) \tag{2.2}$$

В формуле (2) y $(y \geq 0)$ – скалярная, а x – векторная величина, $x_1,...,x_n$ – координаты вектора x, т.е. $f(x_1,...,x_n)$ есть числовая функция нескольких (многих) переменных $x_1,...,x_n$. В связи с этим ПФ $f(x_1,...,x_n)$ называют многоресурсной или многофакторной ПФ. Для многофакторной ПФ также используется символика $f(x_1,...,x_n, a)$, где a – вектор параметров ПФ.

По экономическому смыслу $x_1 \geq 0, ..., x_n \geq 0$, следовательно, областью определения многофакторной ПФ является множество n-мерных векторов x, все координаты $x_1, ..., x_n$ которых неотрицательные числа.

Для отдельного предприятия (фирмы), выпускающего однородный продукт, ПФ может связывать объем выпуска (в натуральном или стоимостном выражении) с затратами рабочего времени по различным видам трудовой деятельности, различных видов сырья, комплектующих изделий, энергии, основного капитала (измеренных обычно в натуральных единицах). ПФ такого типа характеризуют действующую технологию предприятия (фирмы).

При построении ПФ для региона или страны в целом в качестве величины годового выпуска y чаще берут совокупный продукт (доход) региона или страны, исчисляемый обычно в неизменных, а не в текущих ценах. В качестве ресурсов рассматривают основной капитал ($x_1 = K$ – объем используемого в течение года основного капитала), живой труд ($x_2 = L$ – количество единиц затрачиваемого в течение года живого труда), исчисляемые обычно в стоимостном выражении. Таким образом, строят двухфакторную ПФ $y = f(K, L)$. От двухфакторных ПФ переходят к трехфакторным. В качестве третьего фактора иногда вводят объемы используемых природных ресурсов. Кроме того, если ПФ строится по данным временных рядов, то в качестве особого фактора роста производства может быть включен технический прогресс.

ПФ $y = f(x_1, ..., x_n)$ называется **статической**, если ее параметры и ее характеристика f не зависят от времени t, хотя объемы ресурсов могут зависеть от времени ($x_1(t), ..., x_n(t)$).

ПФ называется **динамической**, если:

1) время t фигурирует в качестве самостоятельной переменной величины (как бы самостоятельного фактора производства), влияющего на объем выпускаемой продукции;

2) параметры ПФ и ее характеристика f зависят от времени t.

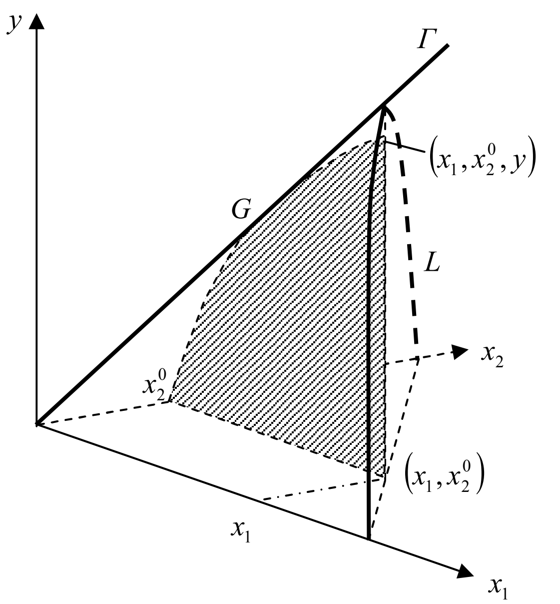


Рис. 2 .2

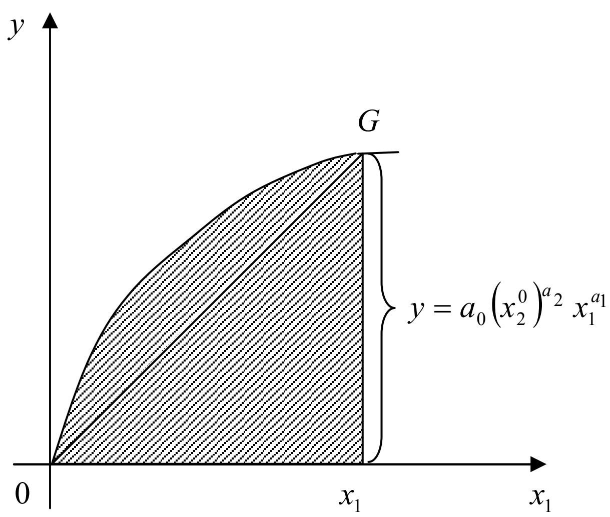


Рис. 2. 3

Пример 2. Для моделирования отдельного региона или страны в целом (т.е. для решения задач на макроэкономическом, а также и на микроэкономическом уровне) часто используется ПФ вида $y = a_0 x_1^{a_1} x_2^{a_2}$, где a_0, a_1, a_2 - параметры ПФ. Это положительные постоянные (часто a_1 и a_2 таковы, что $a_1 + a_2 = 1$). ПФ только что приведенного вида называется ПФ Кобба-Дугласа (ПФКД) по имени двух американских экономистов, предложивших ее использовать в 1929 г. ПФКД активно применяется для решения разнообразных теоретических и прикладных задач благодаря своей структурной простоте. ПФКД принадлежит к классу так называемых мультипликативных ПФ (МПФ). В приложениях ПФКД $x_1 = K$ равно объему используемого основного капитала (объему используемых основных фондов - в отечественной терминологии), $x_2 = L$ – затратам живого труда, тогда ПФКД приобретает вид:

$$y = a_0 K^{a_1} L^{a_2}$$

Графиком ПФ $y = a_0 x_1^{a_1} x_2^{a_2}$ $(a_1 + a_2 = 1)$ в трехмерном пространстве является двумерная поверхность Г, эскиз которой представлен на рис. 2.2. График Г в рассматриваемом случае есть коническая поверхность, направляющей которой является, например, линия *L*, а образующими - лучи, выходящие из точки О. Зафиксируем переменную $x_2 = x_2^0$. Тогда $y = a_0 (x_2^0)^{a_2} x_1^{a_1} = a_0' x_1^{a_1}$ и мы получаем вариант ПФ, аналогичный рассмотренному выше (см. рис. 1 и рис. 2.3). Линия *G* есть пересечение поверхности Г вертикальной плоскостью $x_2 = x_2^0$. На рис. 3 представлен фрагмент рис. .2, относящийся к линии G. Поведение линии G отражает то обстоятельство, что с ростом затрат первого ресурса объем выпуска y растет, но каждая дополнительная единица первого ресурса обеспечивает все меньший прирост выпуска y. Это обстоятельство можно прокомментировать следующим образом. Если число работников и их квалификация остаются неизменными, а число обслуживаемых ими станков (которое уже достаточно велико) увеличивается, например, в два раза, то это естественно не приведет к двойному росту объема выпуска. Отметим, что если $a_1 + a_2 < 1$, то графиком ПФКД является поверхность, которая напоминает выпуклую вверх "горку", крутизна которой падает, если точка (x_1, x_2) перемещается на "северо-восток" по плоскости Ox_1x_2 .

Пример 3. Линейная ПФ (ЛПФ) имеет вид: $y = a_0 + a_1 x_1 + a_2 x_2$ (двухфакторная) и $y = a_0 + a_1 x_1 + \ldots + a_n x_n$ (многофакторная). ЛПФ принадлежит к классу так называемых аддитивных ПФ (АПФ). Переход от мультипликативной ПФ к аддитивной осуществляется с помощью операции логарифмирования. Для двухфакторной мультипликативной ПФ

$$y = a_0 x_1^{a_1} x_2^{a_2}$$

этот переход имеет вид: $\ln y = \ln a_0 + a_1 \ln x_1 + a_2 \ln x_2$. Полагая $\ln y = w;$, $\ln a_0 = a_0';$ $\ln x_1 = v_1;$ $\ln x_2 = v_2$ получаем аддитивную ПФ $w = a_0' + \ a_1 v_1 + a_2 v_2$

Выполняя обратный переход, из аддитивной ПФ получим мультипликативную ПФ.

Если сумма показателей степени в ПФ Кобба- Дугласа $y = a_0 K^{a_1} L^{a_2}$ равна единице $(a_1 + a_2 = 1)$, то ее можно записать в несколько другой форме:

$$\frac{y}{L}=\frac{a_0K^{a_1}L^{a_2}}{L}=\frac{a_0K^{a_1}}{L^{1-a_2}}=\frac{a_0K^{a_1}}{L^{a_1}}=a_0\left(\frac{K}{L}\right)^{a_1}$$

Дроби $\frac{y}{L}=z$ и $\frac{K}{L}=k$ называются соответственно ***производительностью труда и капиталовооруженностью труда***. Используя новые символы, получим

$$z=a_0k^{a_1}$$

т.е. из двухфакторной ПФКД получим формально однофакторную ПФКД. В связи с тем, что $0<a_1<1$, из последней формулы следует, что производительность труда z растет медленнее его капиталовооруженности. Однако этот вывод справедлив для случая статической ПФКД в рамках существующих технологии и ресурсов.

Отметим здесь, что дробь $\frac{y}{K}$ называется ***производительностью капитала*** **или *капиталоотдачей.***

При построении ПФ научно-технический прогресс (НТП) может быть учтен с помощью введения множителя НТП e^{pt}, где параметр p ($p>0$) характеризует темп прироста выпуска под влиянием НТП:

$$y(t)=e^{pt}f(x_1(t),\ x_2(t))\ \ (t=0,1,...,T)$$

Эта ПФ - простейший пример динамической ПФ; она включает нейтральный, то есть не материализованный в одном из факторов, технический прогресс.

2.1.3 Формальные свойства производственных функций

Производственная функция $f(x_1,x_2)$ как формальная конструкция определена в неотрицательном октанте двумерной плоскости. т.е. определена при $x_1\geq 0,\ x_2\geq 0$. ПФ должна удовлетворять ряду свойств:

1. $f(0,0)=0;\ f(0,x_2)=f(x_1,0)=0.$
2. $\forall\ x(1)>x(0)\Rightarrow f(x(1))>f(x(0));\ \frac{\partial f(x)}{\partial x_i}>0;\ i=1,2;\ \ x(k)=(x_1(k),x_2(k));$
3. $\frac{\partial^2 f(x)}{\partial x_i^2}\leq 0;\ \frac{\partial^2 f(x)}{\partial x_1 x_2}\geq 0;\ (i=1,2)$
4. $f(tx_1,tx_2)=t^p f(x_1,x_2)$.
5. Матрица Гессе, составленная из вторых производных производственной функции

$$\Gamma=\begin{pmatrix}\frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 x_2}\\ \frac{\partial^2 f}{\partial x_2 x_1} & \frac{\partial^2 f}{\partial x_2^2}\end{pmatrix}$$

отрицательно определена.

Свойство 1 означает, что без ресурсов (даже при отсутствии хотя бы одного из ресурсов) нет выпуска.

Свойство 2 означает, что с ростом затрат хотя бы одного ресурса объем выпуска растет. Положительность первой частной производной означает, что с ростом затрат одного ресурса при неизменном количестве другого ресурса объем выпуска растет.

Свойство 3 (вторая частная производная ПФ неположительна) означает, что с ростом затрат одного (1-го) ресурса при неизменном количестве другого ресурса величина прироста выпуска на каждую дополнительную единицу 1-го ресурса не растет (закон убывающей эффективности). Неотрицательность второй смешанной производной означает, что при росте одного ресурса предельная эффективность другого ресурса возрастает. Если выполнены условия 3, то график ПФ есть поверхность, расположенная в неотрицательном октанте $x_1 \ge 0,\ x_2 \ge 0,\ y \ge 0$ трехмерного пространства и выпуклая вверх. Вообще геометрический образ ПФ должен прежде всего ассоциироваться с выпуклой горкой, крутизна которой убывает, если точка (x_1 x_2) уходит в плоскости Ox_1x_2 на "северо-восток".

Свойство 4 означает, что ПФ является однородной функцией степени $p > 0$. При $p > 1$ с ростом масштаба производства в t раз (число $t > 1$), т.е. с переходом от вектора x к вектору tx, объем выпуска возрастает в t^p раз, т.е. имеем рост эффективности производства при росте масштаба производства. При $p < 1$ имеем падение эффективности производства при росте масштаба производства. При $p = 1$ имеем постоянную эффективность производства при росте его масштаба (или имеем независимость удельного выпуска от масштаба производства.

Свойство 5 означает, что ПФ является выпуклой вверх функцией.

Для ПФКД $y = a_0 x_1^{a_1} x_2^{a_2}$ $(a_1 + a_2 = 1)$ свойства 1-5 выполняются.

Для ЛПФ $y = a_0 + a_1x_1 + a_2x_2$ $(a_0 > 0,\ a_1 > 0,\ a_2 > 0)$ свойство 1 (при $a_0 \ne 0$) и свойство 4 не выполняются, а матрица Гессе не существует.

Множество точек (линия) ℓ_q уровня $q = f(x_1, x_2)$ ($q > 0$ - действительное число) ПФ $y = f(x_1, x_2)$ называется *изоквантой или линией уровня* ПФ. Иными словами, линия уровня q - это множество точек, в котором ПФ постоянна и равна q.

Различные наборы (v_1, v_2) и (w_1, w_2) затрачиваемых ресурсов, принадлежащие одной и той же изокванте ℓ_q (т.е. $q = f(v_1, v_2) = f(w_1, w_2)$) дают один и тот же объем выпуска q.

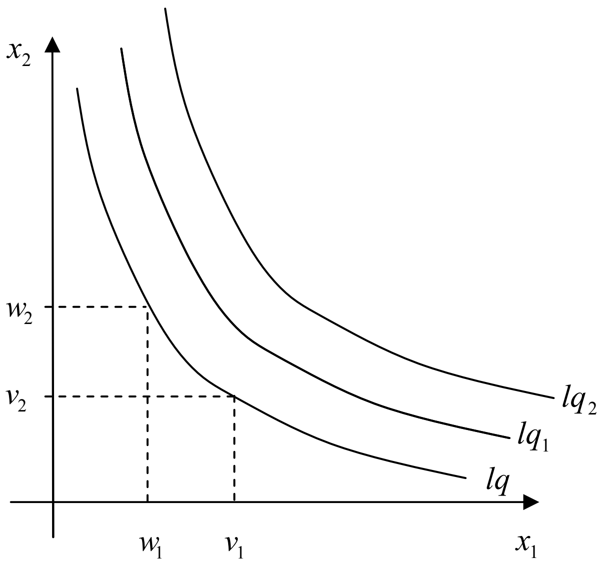


Рис. 2.4

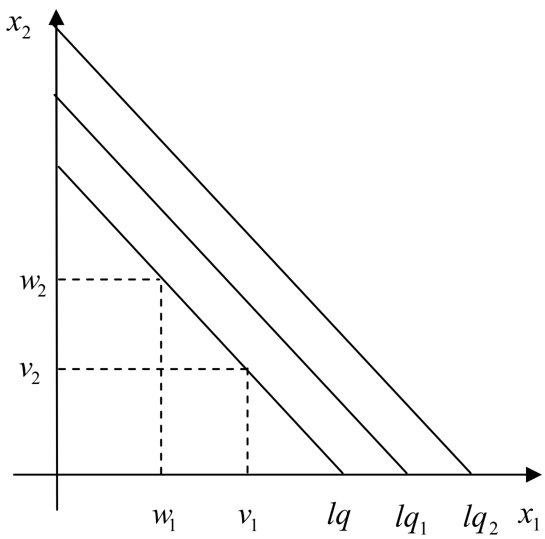


Рис. 2.5

Изокванта есть линия, расположенная в неотрицательном октанте $\{(x_1, x_2) | x_1 \ge 0,\ x_2 \ge 0\}$ двумерной плоскости Ox_1x_2.

Пример 1. На рис. 4 даны изокванты ℓ_{q1}, ℓ_{q2} ПФКД. Отметим, что изокванта ℓ_{q2}, расположенная "северо-восточнее" изованты ℓ_{q1}, соответствует большему объему выпуска (т.е. $q_2 > q_1$). .Если объем используемого основного капитала неограниченно растет (т.е. $x_1 = K \to \infty$), то, как видно на рис. 4, затраты труда неограниченно убывают (т.е. $x_2 = L \to 0$). Аналогично, как видно на рис. 2.4, если $x_2 = L \to \infty$ то $x_1 = K \to 0$. На рис. 5 даны изокванты ℓ_{q1}, ℓ_{q2} $(q_2 > q_1)$ линейной ПФ.

При $n = 2$ для любой ПФ, для которой справедливы все (или часть) свойств 1-4, изокванта (если она не является прямой) есть линия, которая выпукла к точке О.

2.1.4 Характеристики производственной функции

Производительность ресурса. Пусть задана ПФ $y = f(x) = f(x_1, x_2)$. Дробь

$$A_i = \frac{f(x)}{x_i} \quad (i = 1,\ 2) \tag{2.3}$$

называется средней производительностью i-го ресурса (фактора производства) (СПФ) или средним выпуском по i-му ресурсу (фактору производства).

Напомним, что в случае двухфакторной ПФКД $y = a_0 K^{a_1} L^{a_2}$ для средних производительностей $\frac{y}{K}$ и $\frac{y}{L}$ основного капитала и труда были использованы соответственно термины **производительностью капитала** (или **капиталоотдача**) и **производительность труда**. Эти термины используют и применительно к любым двухфакторным ПФ, у которых $x_1 = K,\ x_2 = L$.

Обратные дроби $\frac{K}{y}$ и $\frac{L}{y}$ называются соответственно ***капиталоемкостью и трудоемкостью выпуска***.

Первая частная производная ПФ

$$M_i = \frac{\partial f(x)}{\partial x_i} \ (i = 1, 2) \tag{2.4}$$

называется **предельной (маржинальной) производительностью** i-го ресурса (фактора производства) (ППФ) или предельным выпуском по i-му ресурсу (фактору производства). Обозначим символами Δx_i приращение переменной x_i, а $\Delta_i f(x)$ - соответствующее ей частное приращение ПФ. Здесь $\Delta_1 f(x) = f(x_1 + \Delta x_1, x_2) - f(x_1, x_2)$; $\Delta_2 f(x) = f(x_1, x_2 + \Delta x_2) - f(x_1,\ x_2)$. При малых Δx_i имеем приближенное равенство

$$M_i = \frac{\partial f(x)}{\partial x_i} \approx \frac{\Delta_i f(x)}{\Delta x_i} \quad (i = 1,\ 2)$$

Следовательно, ППФ (приближенно) показывает, на сколько единиц увеличится объем выпуска y, если объем затрат x_i i -го ресурса вырастает на одну единицу при неизменных объемах другого затрачиваемого ресурса.

Пример 1. Для ПФКД $y = a_0 x_1^{a_1} x_2^{a_2}$ найти в явном виде A_1, A_2, M_1, M_2.
Решение задачи. Имеем:

$$A_1 = \frac{y}{x_1} = a_0 x_1^{a_1 - 1} x_2^{a_2}; \; A_2 = \frac{y}{x_2} = a_0 x_1^{a_1} x_2^{a_2 - 1};$$

$$M_1 = \frac{\partial f}{\partial x_1} = a_1 A_1; \; M_2 = \frac{\partial f}{\partial x_2} = a_2 A_2.$$

$$\frac{M_1}{A_1} = a_1 \le 1 \Rightarrow M_1 \le A_1; \; \frac{M_2}{A_2} = a_2 \le 1 \Rightarrow M_2 \le A_2.$$

Последние неравенства означают, что предельные производительности i - го ресурса не больше средней производительности этого ресурса.
Пример 2. Для ЛПФ $y = a_0 + a_1 x_1 + a_2 x_2 \; (a_i > 0, \; i = 1,2,3)$ найти в явном виде A_1, A_2, M_1, M_2.
Решение задачи. Имеем:

$$A_1 = \frac{y}{x_1} = \frac{a_0}{x_1} + a_1 + a_2 \frac{x_2}{x_1}; \; A_2 = \frac{y}{x_2} = \frac{a_0}{x_2} + a_1 \frac{x_1}{x_2} + a_2;$$

$$M_1 = \frac{\partial f}{\partial x_1} = a_1; \; M_2 = \frac{\partial f}{\partial x_2} = a_2.$$

$$\frac{M_1}{A_1} \le 1 \Rightarrow M_1 \le A_1; \; \frac{M_2}{A_2} \le 1 \Rightarrow M_2 \le A_2.$$

Эластичность. Отношение предельной производительности M_i i-го ресурса к его средней производительности A_i

$$E_i = \frac{M_i}{A_i} = \frac{x_i}{f(x)} \frac{\partial f(x)}{\partial x_i}, \quad i = 1, \; 2. \tag{2.5}$$

называется (частной) эластичностью выпуска по i-му ресурсу (по фактору производства) (ЭВФ).
Сумма $E_1 + E_2 = E_x$ называется *эластичностью производства.*

Заменяя $\dfrac{\partial f(x)}{\partial x_i}$ приближенно на $\dfrac{\Delta f(x)}{\Delta x_i}$, получим

$$E_i = \frac{M_i}{A_i} = \frac{\Delta_i f(x)}{f(x)} / \frac{\Delta x_i}{x_i}$$

т.е. E_i приближенно показывает, на сколько процентов увеличится выпуск y, если затраты i-го ресурса увеличатся на один процент при неизменных объемах другого ресурса.

Пример 3. Выписать в явном виде для ПФКД выражения для E_1, E_2 и E.
Решение задачи. Имеем:

$$E_1 = a_1; \quad E_2 = a_2; \quad E_x = E_1 + E_2 = a_1 + a_2.$$

Пример 4. Для ЛПФ $y = a_1x_1 + a_2x_2$ $(a_0 = 0)$ выписать в явном виде выражения для E_1, E_2 и E_x. Решение задачи. Имеем:

$$E_1 = \frac{a_1x_1}{a_1x_1 + a_2x_2}; \; E_2 = \frac{a_2x_2}{a_1x_1 + a_2x_2}; \; E_x = E_1 + E_2 = 1.$$

Предельная норма замены (замещения) ресурсов.

Пусть $y = f(x)$ - ПФ. Предельной нормой R_{ij} замены (замещения) i-го ресурса (фактора производства) j-м (ПНЗФ) называется выражение

$$R_{ij} = -\frac{\partial x_j}{\partial x_i} = \frac{\partial f / \partial x_i}{\partial f / \partial x_j} > 0 \quad (i, j = 1, 2) \tag{2.6}$$

при постоянной y.

Здесь i - номер заменяемого ресурса, j -номер замещающего ресурса. Используется также термин: предельная технологическая норма замены (замещения) i-ого ресурса (фактора производства) j-м ресурсом (фактором производства).

Непосредственно проверяется, что для двухфакторной ПФ справедливо равенство

$$R_{12} = \frac{E_1x_2}{E_2x_1}, \tag{2.7}$$

т.е. (предельная) норма замены первого ресурса вторым равна отношению эластичностей выпуска по первому и второму ресурсам, умноженному на отношение объема второго ресурса к объему первого ресурса. Если $x_1 = K$, $x_2 = L$, то отношение — $\frac{x_1}{x_2} = \frac{K}{L}$ называется капиталовооруженностью труда. В этом случае (предельная) норма замены основного капитала трудом равна отношению эластичностей выпуска по основному капиталу и труду, поделенному на капиталовооруженность труда.

Заменяя бесконечно малые приращения dx_i на конечные Δx_i, можно приближенно записать выражение для предельной нормы замещения ресурсов (для двухфакторной ПФ)

$$R_{12} \approx -\frac{\Delta x_2}{\Delta x_1} \tag{2.8}$$

На основании (6) предельная норма замены ресурсов R_{12} приближенно показывает, на сколько единиц надо увеличить затраты второго ресурса (при неизменном выпуске y), если затраты первого ресурса уменьшатся на одну единицу. См. рис. 2.6, на котором видно, что чем круче касательная к изокванте $l(q)$ в точке (x_1, x_2), тем больше выражение $-\frac{dx_2}{dx_1}$, и следовательно, тем больше норма замены R_{12} первого ресурса вторым.

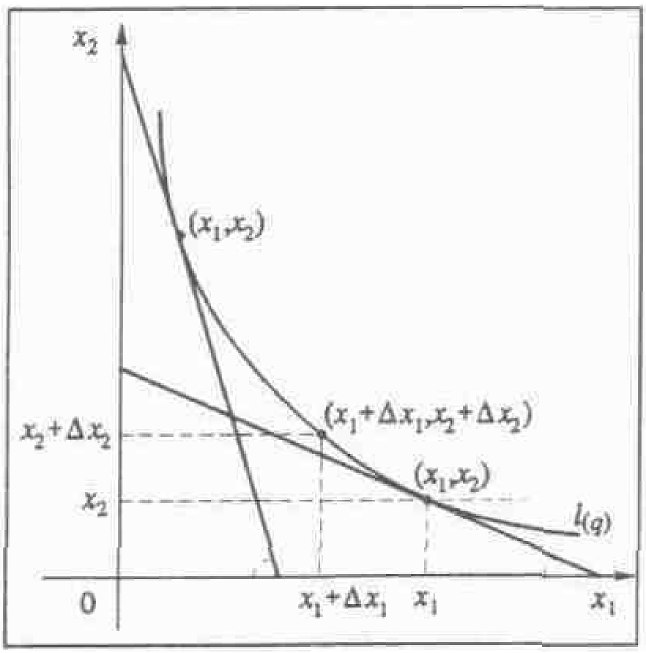


Рис 2.6

Пример 5. Для ПФКД $y = aK^{\alpha}L^{\beta}$ выписать в явном виде выражения R_{12} и R_{21}.

Решение. $R_{12} = \dfrac{\partial y/\partial K}{\partial y/\partial L} = \dfrac{\alpha L}{\beta K}$; $R_{21} = \dfrac{\partial y/\partial L}{\partial y/\partial K} = \dfrac{\beta K}{\alpha L}$.

Эластичность замещения ресурсов

Эластичность замены труда капиталом показывает, на сколько процентов изменится фондовооруженность $k = K/L$ труда при изменении предельной нормы замещения труда капиталом $R_{L,K} = -\dfrac{dK}{dL} = \dfrac{Y'_L}{Y'_K}$ на 1% при неизменном выпуске продукции:

$$E_{L,K} = \frac{d\ln k}{d\ln R_{L,K}}. \quad (2.9)$$

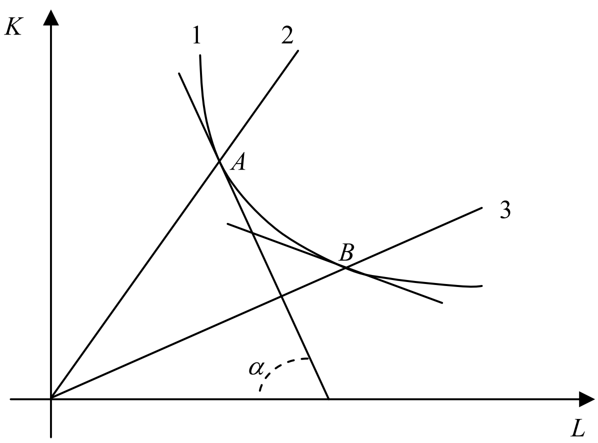


Рис 2.7

На рис. 2.7 показана изокванта (линия уровня) ПФ на плоскости KL (обозначена цифрой 1). Предельная норма замены труда капиталом в т. A равна тангенсу угла наклона касательной, проведенной через т. A ($R_{L,K} = \mathrm{tg}\alpha$). При перемещении из т. A в т. B по изокванте наклон касательной меняется, т.е. меняется величина $R_{L,K}$. При этом меняется и отношение $k = K/L$. Это отношение постоянно вдоль каждой прямой, проходящей через начало координат (вдоль прямых 2 и 3).

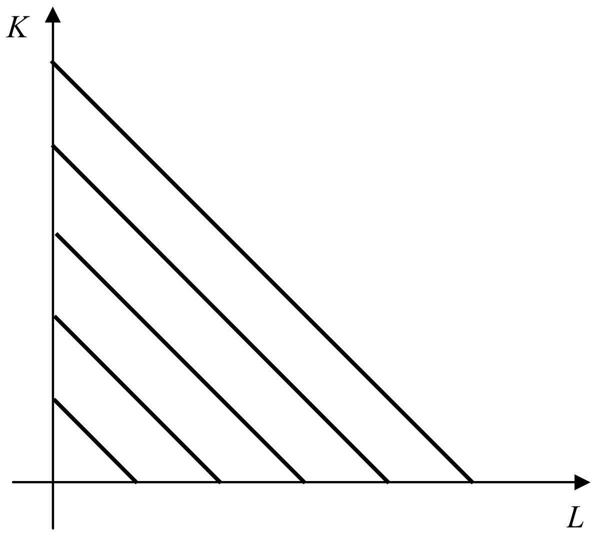


Рис 2.8а

Линейная ПФ $y = a + \alpha K + \beta L$

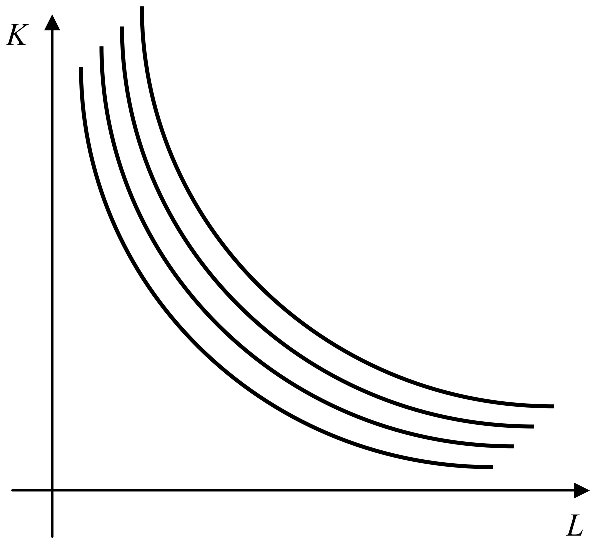


Рис 2.8б

Функция Кобба-Дугласа $y = aK^{\alpha}L^{\beta}$

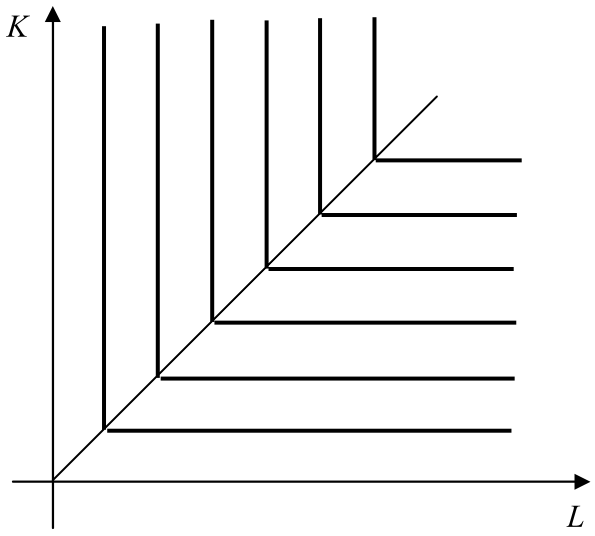


Рис 2.8в

Функция Леонтьева $y = \min(\alpha K, \beta L)$

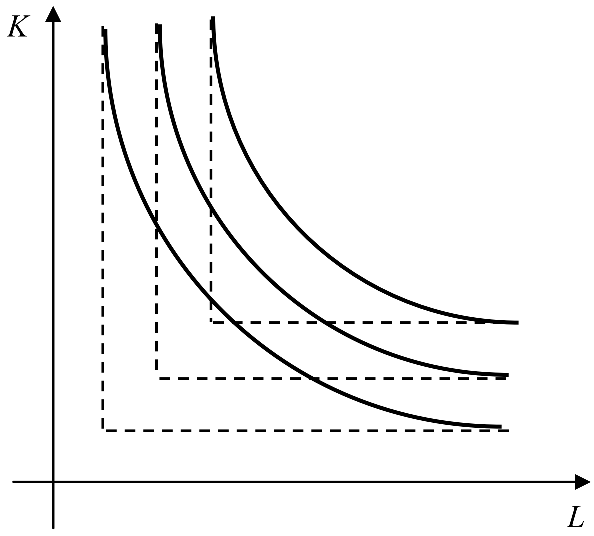


Рис 2.8г

Функция $y = a\left[\delta K^{-\rho} + (1-\delta)L^{-\rho}\right]^{-1/\rho}$

На рис. 2.7 показана изокванта (линия уровня) ПФ на плоскости KL (обозначена цифрой 1). Предельная норма замены труда капиталом в т. A равна тангенсу угла наклона касательной, проведенной через т. A ($R_{L,K} = tg\alpha$). При перемещении из т. A в т. B по изокванте наклон касательной меняется, т.е. меняется величина $R_{L,K}$. При этом меняется и отношение $k = K/L$. Это отношение постоянно вдоль каждой прямой, проходящей через начало координа(вдоль прямых 2 и 3).

На рис. 2.8 изображены линии уровня для линейной ПФ $y = c + aK + bL$, для ПФКД $y = aK^{\alpha}L^{\beta}$, для ПФ Леонтьева $Y = \min(aK, bL)$ и ПФ Солоу (функции с постоянной эластичностью замещения) $Y = a\left[\delta K^{-\rho} + (1-\delta)L^{-\rho}\right]^{-1/\rho}$. Здесь $0 < \delta < 1$, $\rho \ge -1$.

Для функции Кобба-Дугласа $E_{L,K} = 1$. Действительно, имеем $d\ln k = dk/k$, $d\ln R = dR/R$;

$R = -dK/dL = (\beta/\alpha)(Y/a)^{1/\alpha}(1/L^{1+\beta/\alpha}) = (\beta/\alpha)k$; $dR = (\beta/\alpha)dk$.

Отсюда получим $E_{L,K} = 1$.

Рассмотрим линейно однородную ПФ Солоу (CES функцию)

$$y = a\left[\delta K^{-\rho} + (1-\delta)L^{-\rho}\right]^{-1/\rho} \tag{2.10}$$

Можно показать, что при $\rho \to 0$ функция $y = aK^{\delta}L^{1-\delta}$. Действительно имеем

$$y = aL\left[\delta(L/K)^{\rho} + (1-\delta)\right]^{-1/\rho} = aL\left[\delta[(L/K)^{\rho} - 1] + 1\right]^{-1/\rho}.$$

Разложим в ряд выражение в квадратных скобках $(L/K)^{\rho} - 1 \approx \rho\ln(L/K)$. В результате имеем $y = aL\left[\delta\rho\ln(L/K) + 1\right]^{-1/\rho} = aL\left[\rho v + 1\right]^{-1/\rho}$, где $v = \delta\ln(L/K)$. Далее

$$y = aL\left[\rho v + 1\right]^{-1/\rho} = aL\left[[\rho v + 1]^{1/\rho v}\right]^{-v} \to aLe^{-v} = aL(L/K)^{-\delta} = aK^{\delta}L^{1-\delta}$$ ч.т.д.

При $\rho \to \infty$ получаем ПФ Леонтьева. Покажем это. Если $K > L$, то при $\rho \gg 1$ $\delta K^{-\rho} \ll (1-\delta)L^{-\rho}$ и для ПФ получим $y = aL(1-\delta)^{-1/\rho} \to aL$. Если $K < L$, то при $\rho \gg 1$ $\delta K^{-\rho} \gg (1-\delta)L^{-\rho}$ и для ПФ получим $y = aK\delta^{-1/\rho} \to aK$. При $K = L = F$ получим н ч.т.д.

Для ПФ Солоу эластичность замещения труда капиталом равна $E_{L,K} = \dfrac{1}{1+\rho}$.

Отсюда при $\rho \to 0$ получим $E_{L,K} = 1$ (функция Кобба-Дугласа). При $\rho \to -1$ следует $E_{L,K} \to \infty$ (линейная ПФ). При $\rho \to \infty$ получаем $E_{L,K} \to 0$ (ПФ Леонтьева) т.е. Функция Леонтьева имеет нулевую эластичность замещения: ресурсы в ней должны использоваться в заданной пропорции и не могут замещать друг друга.

В качестве примера ПФ CES приведем функцию, полученную Грандбергом А.Г. для экономики СССР за период времени 1960 – 1985 гг.:

$Y = 1.002 \cdot (0.6412K^{-0.81} + 0.3588L^{-0.81})^{-1/0.81}$ - без учета технического прогресса;

$Y = 0.966 \cdot (0.4074K^{-3.03} + 0.3588L^{-3.03})^{-1/3.03} e^{0.0252t}$ - с учетом технического прогресса.

Показатели эластичности замещения ресурсов $E_{L,K} = \dfrac{1}{1+\rho}$ для этих двух функций различны: в первом случае это 0.55, во втором – 0.25. Другими авторами были получены оценки эластичности замещения ресурсов в диапазоне от 0.25 до 0.55. Это говорит о том, что степень заменяемости труда и капитала невысока, во всяком случае гораздо ниже, чем в ПФКД, для которой она равна единице.

Доход. Пусть дана ПФКД и $\alpha + \beta = 1$. Тогда

$$y = \frac{\partial y}{\partial K} K + \frac{\partial y}{\partial L} L . \qquad (2.11)$$

Действительно,

$$\frac{\partial y}{\partial K} K + \frac{\partial y}{\partial L} L = a\alpha K^{\alpha-1} L^{\beta} K + a\beta K^{\alpha} L^{\beta-1} L = aK^{\alpha} L^{\beta} (\alpha + \beta) = y .$$

Если считать, что общество состоит только из работников и предпринимателей, а функция y представлена в стоимостном выражении (доход от продажи продукции), то весь доход (11) распадается на две части, которые можно назвать доходом предпринимателя (предельная фондоотдача, или норма прибыли, умноженная на объем фондов) и доходом работников (предельная производительность труда, умноженная на количество трудовых ресурсов). Аналогичный результат можно получить и для линейной ПФ, у которой $a = 0$.

2.2 Задача производителя

Сформулируем задачу производителя: найти технологию, дающую максимальную прибыль.

Введем некоторые допущения. Пусть производственная функция предприятия имеет вид $y = f(x_1,...,x_n)$, т.е. затраты однозначно определяют выпуск. Пусть p - цена единицы выпускаемой продукции (будем считать ее постоянной).

Тогда задача производителя имеет вид:

$$W(X) = p \cdot f(x_1, x_2,...,x_n) - (\lambda_1 x_1 + \lambda_2 x_2 + ... + \lambda_n x_n) \to \max \qquad (2.12)$$

при $X \geq 0$, где λ_i - цена i -го ресурса.

Необходимое условие максимума имеет вид

$$p\frac{\partial f}{\partial x_i} - \lambda_i = 0, \; i = 1,...,n \qquad (2.13)$$

Система уравнений (2.13) определяет точку экстремума. Матрица вторых производных функции прибыли

$$\Gamma_W = p \cdot \Gamma_f,$$

где Γ_f - матрица вторых производных ПФ. А так как, матрица Гессе Γ_f отрицательно определена, то найденное решение системы (2.13) является точкой максимума функции прибыли.

Итак, соотношения (2.13) дает оптимальное решение задачи производителя - технологию $\{y^*, X^*\}$, где вектор $X^* = \{x_1^*,...,x_n^*\}$ - оптимальное количество перерабатываемых ресурсов, *а* $y^* = f(X^*)$ - оптимальный объем выпуска.

Рассмотрим экономический смысл уравнений (2.13). Преобразуем их к виду

$$p\frac{\partial f}{\partial x_i} dx_i = \lambda_i dx_i$$

Здесь сомножитель левой и правой части dx_i, представляет собой некоторое количество i-го ресурса, дополнительно вовлеченного в производство.

Тогда в левой части имеем стоимость продукта, дополнительно полученного из dx_i единиц i-го ресурса, а в правой - стоимость dx_i единиц i-го ресурса. Получилось равновесие: можно вовлечь в производство дополнительно dx_i единиц i-го ресурса, потратив на его закупку $\lambda_i dx_i$, но выигрыша не будет, так как продукции будет получено ровно на ту же сумму, сколько затратили.

Поэтому наращивание объемов производства идет до тех пор, пока не начнет выполняться соотношение (13).

Пример. Зависимость выручки y за смену от числа столиков S и числа официантов F в небольшом кафе выражается формулой $y = 140 \cdot S^{2/3}F^{1/4}$. Расходы на один столик составляют 50 д.е., зарплата официанта - 100 д.е. за смену. Найти оптимальный размер кафе, т.е. число столиков и число официантов их обслуживающих, исходя из максимума получаемой прибыли

Решение. Составляем функцию прибыли

$$W(S,F) = 140 \cdot S^{2/3}F^{1/4} - 50S - 100F \to \max$$

Берем частные производные по обеим переменным и приравниваем их к нулю

$$\frac{\partial W}{\partial S}=140\cdot\frac{2}{3}\cdot S^{-1/3}F^{1/4}-50=0,$$

$$\frac{\partial W}{\partial F}=140\cdot\frac{1}{4}\cdot S^{2/3}F^{=3/4}-100=0.$$

Решение этой системы $S=12,\ F=2$.

Ответ: 12 столиков и 2 официанта.

2.3 Учет налогов

Рассмотрим влияние различных видов налогов на характер поведения производителя.

1. *Налог с прибыли.* При ставке s налога производитель отчисляет государству часть s прибыли.

Тогда задача производителя примет вид

$$W(X)=[p\cdot f(x_1,x_2,...,x_n)-(\lambda_1x_1+\lambda_2x_2+...+\lambda_nx_n)]\cdot(1-s)\to\max$$

Ясно, что положение точки максимума X^* не изменится, т.к. она будет определяться теми же уравнениями (13); изменится только уровень прибыли

2. *Налог с продаж (акцизный налог).* При акцизном налоге предприятие отчисляет государству некоторую сумму v за каждую проданную единицу продукции.

Тогда задача производителя примет вид

$$W(X)=(p-v)\cdot f(x_1,x_2,...,x_n)-(\lambda_1x_1+\lambda_2x_2+...+\lambda_nx_n)\to\max$$

В этом случае точка максимума будет определяться уже соотношениями, отличными от (13), а именно

$$(p-v)\frac{\partial f}{\partial x_i}=\lambda_i,\ i=1,...,n,$$

т.е. точка максимума сместится влево (будет при меньших значения X).

2.4 Функции спроса на ресурсы

Найденное решение X^* системы уравнений (13) - единственное при сделанных предположениях $X\geq 0;\ p>0;\ \Lambda>0$. Таким образом, для данной производственной функции каждому p и Λ соответствует X^*, т.е. $X^*=\Phi(p,\Lambda)$, или

$$x_i^*=\varphi_i(p,\lambda_1,\lambda_2,...,\lambda_n),\ i=1,...,n \tag{2.14}$$

Эти n функций называются функциями спроса на ресурсы при данных ценах на продукты и ресурсы.

Их смысл: если сложились цены Λ на ресурсы и цена p на выпускаемый продукт, то данный производитель определяет по ним объем перерабатываемых ресурсов и ищет их на рынке. Затем, зная эти объемы, из производственной функции определяются выпуски как функции цен.

Для функции спроса можно строго доказать, что

$$\frac{\partial \varphi_i}{\partial \lambda_i} < 0, \quad i = 1,...,n ,$$

т.е. повышение платы за ресурс всегда приводит к сокращению спроса на этот ресурс.
Также можно доказать, что

$$\frac{\partial \varphi_i}{\partial \lambda_j} = \frac{\partial \varphi_j}{\partial \lambda_i}, \quad i, j = 1,...,n .$$

Это означает, что влияние изменения платы за j-й ресурс на спрос на i-и ресурс точно такое же, как и влияние изменения платы за i-и ресурс на спрос на j-й ресурс.

Назовем ресурсы **взаимодополняемыми**, если $\frac{\partial \varphi_i}{\partial \lambda_j} < 0$ и **взаимозаменяемыми**, если

$$\frac{\partial \varphi_i}{\partial \lambda_j} > 0 .$$

Таким образом, для взаимодополняемых ресурсов повышение цены на один из них приводит к падению спроса на другой (например, бензин и машинное масло), а для взаимозаменяемых ресурсов повышение цены на один из них приводит к увеличению спроса на другой (например, шифер и рубероид).

2.5 Модели ценообразования

Пусть предприятие действует в условиях несовершенной конкуренции. Каким образом выбрать цену на товар, чтобы максимизировать прибыль?
Прибыль от продажи Y единиц продукции запишется в виде

$$W(Y) = R(Y) - C(Y) = P(Y) \cdot Y - (vY + F) \tag{2.15}$$

где $R(Y) = P(Y) \cdot Y$ - доход от продажи Y единиц продукции по цене $P(Y)$, зависящей от сбыта; $C(Y) = (vY + F)$ - издержки производства: F — постоянные издержки, те. издержки, величина которых не изменяется с изменением объема производства (административные расходы, расходы, связанные с арендой, использованием машин и оборудования, капитальным ремонтом и т.п.); v - издержки на одну единицу продукции (затраты на электроэнергию, расходные материалы и т.д.). Тогда условие максимума прибыли имеет вид

$$\frac{dW}{dY} = \frac{d}{dY}\big(P(Y) \cdot Y - (vY + F)\big) = 0$$

или

$$\frac{dP}{dY} Y + P(Y) - v = 0 \tag{2.16}$$

Если использовать понятие эластичности спроса (сбыта) по цене $z = \frac{dY}{dP}\frac{P}{Y}$, то соотношение (16) запишется в виде

$$\frac{1}{z} P(Y) + P(Y) = v$$

или

$$P(Y)=\frac{z}{1+z}v \qquad (2.17)$$

Таким образом, формула (17) представляет собой модель ценообразования на продукцию - назначение оптимальной цены (если известен коэффициент эластичности спроса).

Тогда максимальная прибыль составит (подставим (17) в (15))

$$W^*(Y)=-\frac{vY}{z+1}-F$$

Но реализация модели (17) на практике осложняется тем, что кривая спроса, как правило, не известна.

Рассмотрим две эмпирические модели, которые позволяют установить цену, исходя из степени достоверности рыночной информации, имеющейся в распоряжении руководителя предприятия.

Будем считать, что руководство предприятия располагает достоверными сведениями о своих издержках и предполагает достичь вполне определенного уровня прибыли от продажи продукции (назовем ее *целевой прибылью* W^0).

Также, с разной степенью достоверности известна информация о спросе Q

$$Q_{min}\le Q\le Q_{max} \qquad (2.18)$$

и информация о существующих рыночных ценах на продукты – аналоги

$$P_{min}\le P\le P_{max} \qquad (2.19)$$

1. Если более достоверной является информация о спросе, то, используя формулу (15), по заданному наиболее вероятному значению спроса Q^0 вычисляют размер цены P^0, необходимой для получения целевой прибыли

$$P^0=v+\frac{W^0+F}{Q^0} \qquad (2.20)$$

2. Если сведения о спросе менее достоверны, чем о ценах, то сначала по формуле (15) определяют спрос Q_k для каждого значения цены из интервала (19)

$$Q_k=\frac{W^0+F}{P_k-v} \qquad (2.21)$$

Затем экспертными методами оценивают наиболее вероятный спрос Q^0 и далее по формуле (20) рассчитывают цену P^0, соответствующую спросу Q^0.

Понятно, что модели (20), (21) взаимосвязаны. Их взаимосвязь проявляется не только в том, что они дополняют друг друга, но и в том, что они позволяют при определенных условиях получить значения цен, которые, во-первых, сходны между собой, и, во-вторых, близки к их оптимальным уровням в соответствии с теоретическим правилом (17) определения цен.

Пример. Уличный продавец газет берет их в издательстве по цене 2 д.е. за экземпляр. Объем продажи y связан с назначаемой им ценой p формулой $y=2800-1000p$; издержки по продаже равны 0,1 д.е. на экземпляр. Какое оптимальное количество газет должен брать продавец в издательстве и какова оптимальная цена продажи газеты?

Решение. Функция прибыли запишется в виде

$$W(Y) = yp - 0.1y - 2y = yp - 2.1y$$

Выразим цену газеты через объем продаж:

$$p = \frac{2800 - y}{1000} \qquad (2.22)$$

Тогда функция прибыли имеет вид

$$W(y) = yp - 2.1y = y(2.8 - 0.001y) - 2.1y = -0.001y^2 + 0.7y$$

Возьмем производную по y и приравняем к нулю

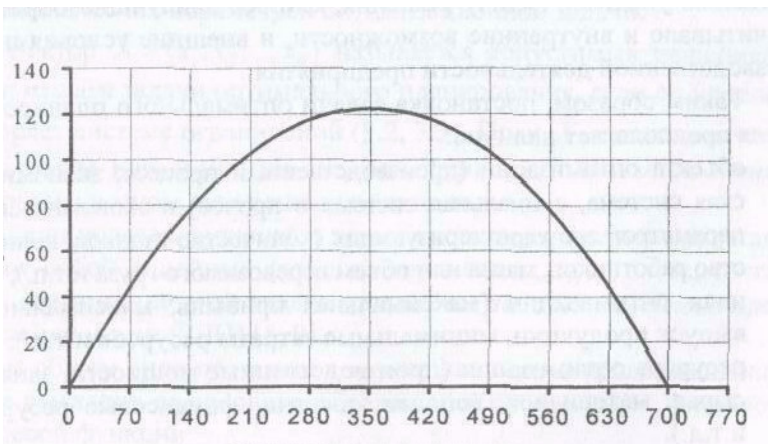
$$W'(y) = -0.002y - 0.7 = 0$$

Отсюда $y^* = 350$.

Из (22) получаем $p^* = 2.45$.

Итак, оптимальная цена продажи газет 2 руб. 45 коп., а оптимальное их количество – 350

экземпляров. Прибыль при этом — 122,5 руб. (рис. 1*)*.



Количество экземпляров газет

Вопросы для самопроверки

1. Дайте определение производственной функции. Как классифицируются производственные функции?
2. Что в статической производственной функции не зависит от времени, а что может зависеть от времени?
3. В чем суть закона убывающей эффективности?
4. Как определяется производительность труда и капиталовооруженность труда? Как связаны эти характеристики в случае линейной ПФ и производственной функции Кобба-Дугласа?
5. Основные свойства производственной функции. Пример ПФ, которая отдельными свойствами не обладают. Пример производственных функций, которые обладают всеми основными свойствами.
6. Приведите характеристики однофакторной производственной функции.
7. Приведите характеристики двухфакторной производственной функции.

8. Понятие изокванты и ее экономический смысл. Построить изокванту линейной ПФ и ПФКД.
9. Как определяется средняя и предельная производительности капитала?
10. Как определяется средняя и предельная производительности труда?
11. Какая существует связь между производительностью труда и капиталовооруженностью труда в случае линейной ПФ и производственной функции Кобба-Дугласа.
12. Связь между средней и предельной производительностью капитала (труда) в общем случае и в случае производственной функции Кобба-Дугласа.
13. Эластичность выпуска по i-му ресурсу $(i = 1, 2)$ и эластичность производства. Экономический смысл.
14. Норма замены одного ресурса другим. Экономический интерпретация этого понятия. Графическая интерпретация.
15. Эластичность замещения ресурсов. Графическая интерпретация.
16. Как меняется предельная норма замены одного ресурса другим при движении по изокванте? Дайте содержательную интерпретацию характеру изменения предельной нормы замены.
17. Поясните смысл ПФ CES. Каковы ее свойства и основные характеристики?
18. Что такое доход предпринимателя и доход работников?
19. В чем состоит экономический смысл задачи производителя?
20. Как влияют налоги на характер производителя?
21. Дайте экономическую интерпретацию функции спроса на ресурсы
22. Какие ресурсы называются взаимодополняемыми, а какие – взаимозаменяемыми?
23. Поясните смысл постоянных и переменных издержек производства
24. Что такое эластичность спроса по цене?
25. Как связана оптимальная цена на продукцию с эластичностью спроса и изержками?

Тема 3. Функция полезности

Основными участниками процесса производства, распределения и потребления благ являются отдельные потребители и производители. Каждый из них имеет свои интересы. Построим теорию поведения потребителя.

3.1. Множество благ

Пусть на рынке благ индивидуальному потребителю предлагается n различных благ. Под *набором благ* будем понимать совокупность неотрицательных чисел $x_1,...,x_n$, где x_j – количество j-го блага ($x_j \geq 0,\ j=1,...,n$). Такой набор неотрицательных чисел можно рассматривать как n-й вектор $x=(x_1,...,x_n)$ благ. Всевозможные наборы благ образуют *множество благ*

$$R_+^n = \left\{ x=(x_1,...,x_n);\ x_j \geq 0,\ j=1,...,n \right\}, \tag{3.1}$$

которое геометрически представляет собой неотрицательный ортант n-го пространства.

Для каждого блага j заданы ограничения

$$x_j^{\min} \leq x_j \leq x_j^{\max},\ j=1,...,n. \tag{3.2}$$

Поэтому, множество наборов благ это ограниченное множество G в n-ом пространстве

$$G = \left\{ x=(x_1,...,x_n);\ x_j^{\min} \leq x_j \leq x_j^{\max},\ j=1,...,n \right\}. \tag{3.3}$$

Множество наборов благ может быть и неограниченным. В этом случае его можно представить как

$$G = \left\{ x=(x_1,...,x_n);\ x_j \geq x_j^{\min},\ j=1,...,n \right\} \tag{3.4}$$

или

$$G = \left\{ x=(x_1,...,x_n);\ x_j \leq x_j^{\max},\ j=1,...,n \right\}, \tag{3.5}$$

при этом неравенства $x_j \leq x_j^{\min}$ или $x_j \leq x_j^{\max}$ могут выполняться не для всех j.

В распоряжении потребителя имеется ограниченное количество денежных средств M (доход), которые он может использовать для приобретения благ. Блага приобретаются по ценам, которые устанавливаются рынком и не зависят от отдельного потребителя. Обозначим их $p_1, p_2,..., p_n$ соответственно для 1-го, 2-гобюююб n-го блага. Вектор $p=(p_1, p_2,..., p_n)$ называют *вектором цен благ*, при этом $p_j \geq 0,\ j=1,...,n$. Тогда стоимость набора благ равна $p^T x = \sum_{j=1}^{n} p_j x_j$.

Стоимость набора благ для потребителя не превосходит его дохода M, т.е.

$$\sum_{j=1}^{n} p_j x_j \leq M. \tag{3.6}$$

Неравенство (3.6) называют *бюджетным ограничением.*

Неравенство (3.6) вместе с неравенством (3.2) определяет множество K *доступных наборов* благ на рынке товаров и услуг. В силу наличия бюджетного неравенства (3.6) множество K всегда будет ограниченным множеством R^n.

Например, в случае множества G, заданного неравенствами вида (3.2), получаем, что множество K доступных благ может быть одним из трех:

1) пустое множество, т.е. при имеющемся доходе M и ценах $p = (p_1, p_2, ..., p_n)$ потребитель не может приобрести на рынке даже минимального набора благ и можно сказать, что в этом случае он живет за чертой бедности;

2) часть n-мерного параллелепипеда G, т.е. потребителю доступны не все наборы благ при имеющемся доходе M и установившихся ценах p на рынке благ;

3) весь n-мерный параллелепипед G, т.е. потребитель настолько богат, что ему доступен любой набор благ при его доходе и действующих ценах.

На рис. 3.1. изображены эти случаи при $n = 2$.

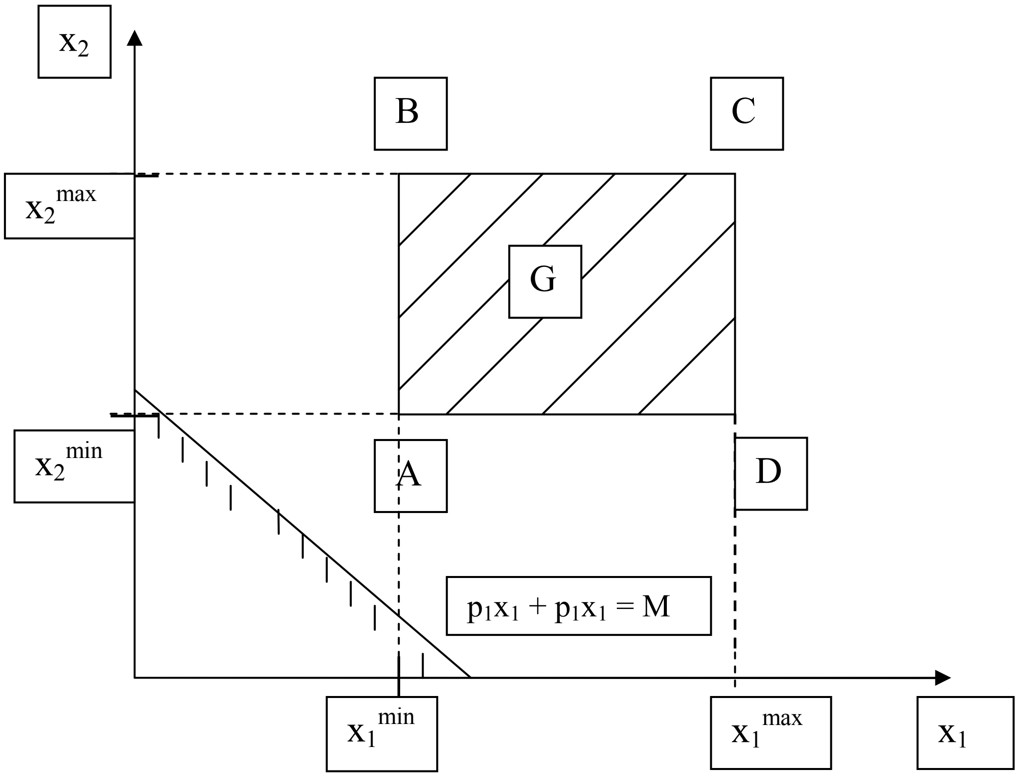


Рис. 3.1а. Пустое множество доступных благ

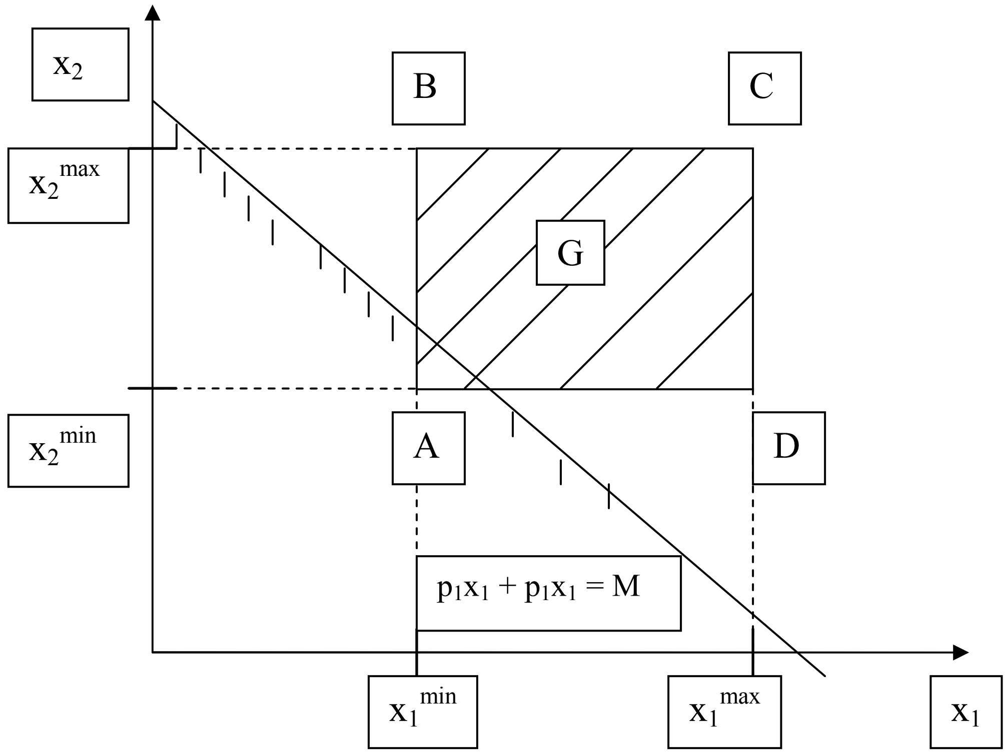


Рис. 3.1б. Частичное множество доступных благ

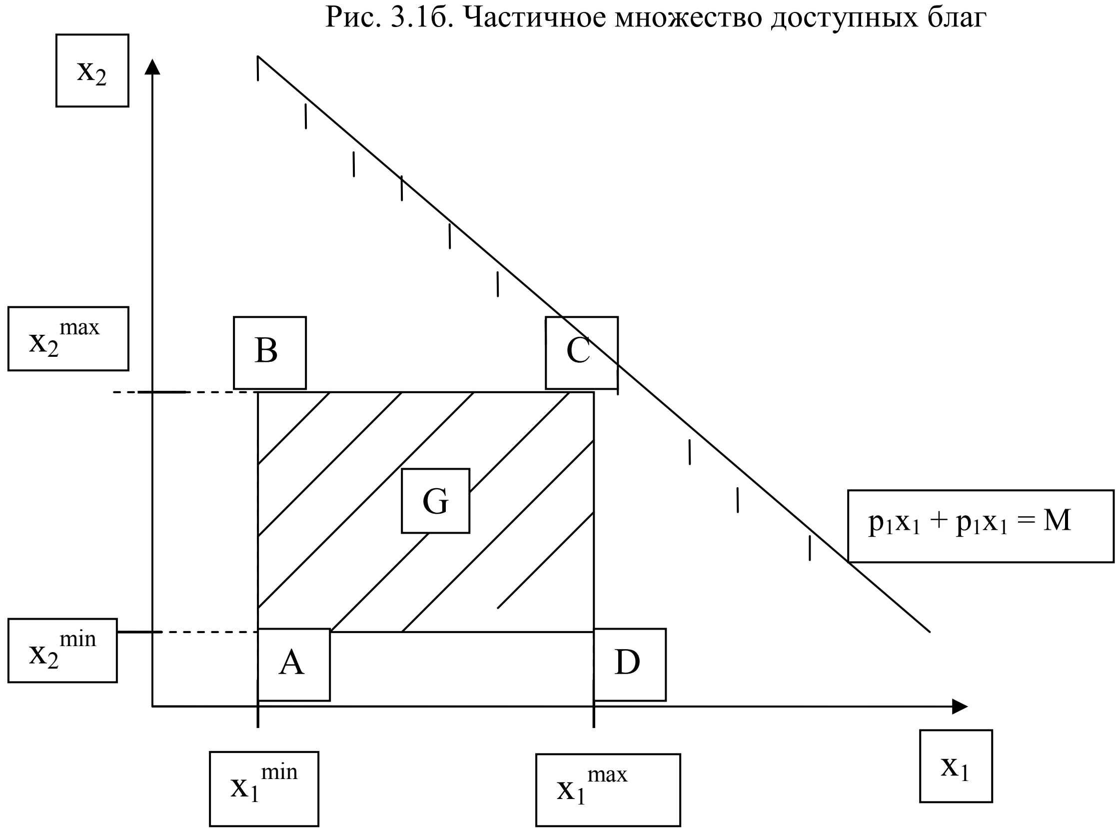


Рис. 3.1в. Полное множество доступных благ

Таким образом, множество K доступных благ является выпуклым множеством, определяемым системой неравенств

$$x_j \geq x_j^{\min}, \; j \in J_1, \; x_j \leq x_j^{\max}, \; j \in J_2, \; \sum_{j=1}^{n} p_j x_j \leq M, \qquad (3.7)$$

где J_1, J_2 – некоторые подмножества множества $J = \{1, 2, ..., n\}$.

3.2. Функция полезности и ее свойства

Рассмотрим вопрос о выборе набора благ. Каждое благо должно удовлетворять ту или иную потребность. Способность удовлетворять ту или иную потребность называют *полезностью блага.*

Потребитель при рассмотрении двух наборов благ $x = (x_1, ..., x_n)$ и $y = (y_1, ..., y_n)$ принимает одно из следующих решений:

а) набор x предпочтительнее (полезнее), чем набор y;

б) набор y предпочтительнее, чем набор x;

в) наборы x и y равнозначны (равноценны, одинаково полезны).

Запись $x > y$ означает, что набор x предпочтительнее набора y; запись $x \sim y$ – набор x и набор y равнозначны; запись $x \geq y$ означает, что набор x не менее предпочтителен, чем набор y.

Введенное понятие отношения предпочтения позволяет сформулировать следующий *принцип выбора потребителя:* потребитель выбирает наиболее предпочтительный набор среди всех доступных ему наборов благ.

Вместо рассуждений на основе заданного отношения предпочтения $\geq$ используем функцию полезности для предпочтения одного набора благ другому.

Функция $u(x) = u(x_1, ..., x_n)$, определенная на неотрицательном ортанте R_+^n (или некотором $G \subset R_+^n$), называется *функцией полезности,* соответствующей отношению предпочтения $\geq$, если $u(x) \geq u(y)$, тогда и только тогда, когда $x \geq y$,

причем если $u(x) = u(y)$, то $x \sim y$ и обратно, если $x \sim y$, то $u(x) = u(y)$.

Функция полезности $u(x)$, по существу, представляет систему предпочтений потребителя. Основное ее свойство в том, потребитель предпочитает выбирать x, а не y, если $u(x) > u(y)$, она упорядочивает наборы по предпочтению их друг другу. Отсюда следует, что потребитель при выборе набора благ стремится максимизировав свою функцию полезности.

В таблице 3.1 приведены четыре типа функции полезности.

Рассмотрим некоторые общие свойства функции полезности.

1) Функция полезности и дважды дифференцируема и строго выпукла.

2) Функция полезности не насыщена. Свойство ненасыщаемости состоит в том, для любых заданных двух наборов $x, y \in R_+^N$ соотношение $x \geq y$ влечет $u(x) \geq u(y)$, а соотношения $x \geq y$ и $x \neq y$ влекут $u(x) > u(y)$. Значит, функция полезности

является возрастающей по любому ее аргументу, т.е. $\dfrac{\partial u(x)}{\partial x_j} > 0,\ \forall j = \overline{1,n},\ x \in R_+^n$.

3) $\lim\limits_{x_j \to 0} \dfrac{\partial u}{\partial x_j} = \infty,\ \lim\limits_{x_j \to \infty} \dfrac{\partial u}{\partial x_j} = 0,\ j = 1,\dots,n$.

Таблица 3.1

Тип функции полезности	Функция полезности	Ограничения
Логарифмичесчкая	$u(x) = \sum\limits_{j=1}^{n} a_j \ln x_j$	$a_j > 0,\ x_j > 1,\ j = 1,\dots,n$
Мультипликативная	$u(x) = a\prod\limits_{j=1}^{n} x_j^{\alpha_j}$	$0 < \alpha_j < 1,\ x_j > 0,$ $j = 1,\dots,n,\ a > 0.$
Аддитивная	$u(x) = \sum\limits_{j=1}^{n} \alpha_j x_j^{\beta_j}$	$0 < \beta_j < 1,\ x_j > 0,\ \alpha_j > 0,$ $j = 1,\dots,n.$
Квадратичная	$u(x) = \sum\limits_{j=1}^{n} a_j x_j + \dfrac{1}{2}\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n} b_{ij} x_i x_j$	$a_j + \sum\limits_{i=1}^{n} b_{ij} x_i > 0,\ j = \overline{1,n}$ $B < 0$ (отрицательно определена)

В ортанте R_+^n уравнению $u(x) = c$, где c – константа, соответствует определенная поверхность равноценных (одинаковой полезности) наборов благ (множество безразличия), и наоборот, каждому множеству безразличия соответствует некоторая поверхность, определяемая уравнением $u(x) = c$. Эти поверхности называют *поверхностями безразличия.* В случае двух благ, т.е. в R_+^2 их называют *кривыми безразличия.*

Для данной функции полезности кривые (поверхности) безразличия обладают следующими свойствами:

1) через каждую точку множества товаров проходит лишь одна кривая (поверхность) безразличия;

2) линии (поверхности) безразличия не пересекаются, причем кривая, лежащая выше и правее другой кривой, представляет собой более предпочтительные наборы товаров;

3) линии (поверхности) безразличия обращены выпуклостью к началу координат (вытекает из строгой вогнутости функции полезности);

4) множество наборов x, для которых $u(x) \ge c$ (предпочтительное множество) является выпуклым множеством.

3.3. Предельная полезность и предельная норма замещения благ

В теории потребительского выбора большую роль играют предельные полезности благ, которые выражают дополнительное удовлетворение от потребления одной дополнительной единицы блага. Математически этот факт описывается частными производными функции полезности. Установим содержательный смысл частных производных функции полезности. Пусть количество j-го блага изменилось на величину Δx_j, а количества остальных благ не изменились. Это вызывает частное приращение функции полезности $\Delta u_j = u\left(x_1, x_2, ..., x_j + \Delta x_j, ..., x_n\right) - u\left(x_1, x_2, ..., x_j, ..., x_n\right)$, причем $\Delta x_j > 0$.

Величина $\dfrac{\Delta u_j}{\Delta x_j} > 0$ указывает на изменение полезности на дополнительную единицу j-го блага. Переходя к пределу, получим $\lim\limits_{\Delta x_j \to 0} \dfrac{\Delta u_j}{\Delta x_j} = \dfrac{\partial u}{\partial x_j} > 0$. Частная производная $\dfrac{\partial u}{\partial x_j}$ называется *предельной полезностью* j *-го блага.*

Пример 3.1. Дана функция полезности $u(x_1, x_2) = a_1 \ln x_1 + a_2 \ln x_2$. Предельная полезность первого блага равна $\dfrac{\partial u}{\partial x_1} = \dfrac{a_1}{x_1}$, а второго – $\dfrac{\partial u}{\partial x_2} = \dfrac{a_2}{x_2}$. ►

Пример 3.2. Для функции полезности $u(x_1, x_2) = a x_1^{\alpha_1} x_2^{\alpha_2}$ будем иметь

$\dfrac{\partial u}{\partial x_1} = a\alpha_1 x_1^{\alpha_1 - 1} x_2^{\alpha_2}$ – для первого блага, $\dfrac{\partial u}{\partial x_2} = a\alpha_2 x_1^{\alpha_1} x_2^{\alpha_2 - 1}$ – для второго блага. ►

Рассмотрим вопрос о взаимозаменяемости благ. Пусть объемы потребляемых благ изменились соответственно на малые величины $\Delta x_1, \Delta x_2, ..., \Delta x_n$. Тогда полным приращением полезности является величина $\Delta u \approx \dfrac{\partial u}{\partial x_1} dx_1 + \dfrac{\partial u}{\partial x_2} dx_2 + ... + \dfrac{\partial u}{\partial x_n} dx_n$ – полный дифференциал функции полезности. Пусть уровень полезности не изменяется, т.е. изменение набора благ происходит так, что сохраняется одна и та же поверхность безразличия $u(x_1, ..., x_n) = c$. Предположим, что количество всех благ, кроме k-го и l-го, которые взаимозаменяемы, не изменяются. Тогда получим

$$\frac{\partial u}{\partial x_k} dx_k + \frac{\partial u}{\partial x_l} dx_l = 0.$$

Отсюда имеем

$$n_{kl} = \frac{dx_k}{dx_l} = \frac{\Delta x_k}{\Delta x_l} = -\frac{\partial u / \partial x_l}{\partial u / \partial x_k}.$$

Величина $n_{kl} = \dfrac{\Delta x_k}{\Delta x_l}$ называется *коэффициентом (нормой) предельной эквивалентной замены благ,* который обратно пропорционален отношению предельных полезностей этих благ, взятому с обратным знаком. Поскольку $\dfrac{\partial u}{\partial x_l} > 0,\ j = \overline{1,n}$, то $n_{kl} < 0$, т.е. увеличение потребления одного блага вызывает уменьшение другого для сохранения одного и того же уровня полезности.

Пример 3.3.

Для функции полезности $u(x_1, x_2) = a_1 \ln x_1 + a_2 \ln x_2$ имеем

$n_{21} = \dfrac{\Delta x_2}{\Delta x_1} = -\dfrac{\alpha_1 / x_1}{\alpha_2 / x_2} = -\dfrac{\alpha_1 x_2}{\alpha_2 x_1}$. ►

Пример 3.4.

Для функции полезности $u(x_1, x_2) = a x_1^{\alpha_1} x_2^{\alpha_2}$ получим $n_{21} = -\dfrac{\alpha_1 x_2}{\alpha_2 x_1}$. ►

Изучение изменения нормы предельной заменяемости одних благ другими играет важную роль при изучении закономерностей потребления: если потребность в определенном благе удовлетворяется незначительно, то относительная полезность этого блага по отношению к другим для сохранения одного и того же уровня полезности высока.

Рассмотрим теперь смысл вторых частных производных функции полезности. Вторые частные производные $\dfrac{\partial^2 u}{\partial x_i \partial x_j}$ характеризуют изменение предельной полезности $\dfrac{\partial u}{\partial x_j}$ блага j при изменении потребления этого же блага. Пусть $\dfrac{\partial^2 u}{\partial x_i \partial x_j} < 0$, т.е. предельная полезность любого блага уменьшается по мере того, как его потребление увеличивается. Это свойство называют *законом Госсена* – законом убывания предельной полезности. Таким образом, предполагаем, что функция полезности дважды дифференцируема, имеет непрерывные частные производные, а матрица Гессе H, образованная из вторых частных производных, является отрицательно определенной (ее главные миноры нечетного порядка отрицательны, а четного положительны для любого набора $x = (x_1, \ldots, x_n)$. Это требование относительно функции полезности означает, что функция полезности строго вогнута.

Пример 3.5.

Для функции полезности $u(x_1, x_2) = a_1 \ln x_1 + a_2 \ln x_2$ матрица Гессе

равна $H = \begin{pmatrix} -\dfrac{a_1}{x_1^2} & 0 \\ 0 & -\dfrac{a_2}{x_2^2} \end{pmatrix}$.

Так как $\Delta_1 = -\dfrac{a_1}{x_1^2} < 0,\ \Delta_2 = \dfrac{a_1 a_2}{x_1^2 x_2^2} > 0$, то матрица Гессе отрицательно определена.

►

Пример 3.6.

Для функции полезности $u(x_1, x_2) = a x_1^{\alpha_1} x_2^{\alpha_2}$ получим

$$H = \begin{pmatrix} a\alpha_1(\alpha_1 - 1)x_1^{\alpha_1 - 2} x_2^{\alpha_2} & a\alpha_1\alpha_2 x_1^{\alpha_1 - 1} x_2^{\alpha_2 - 1} \\ a\alpha_1\alpha_2 x_1^{\alpha_1 - 1} x_2^{\alpha_2 - 1} & a\alpha_2(\alpha_2 - 1)x_1^{\alpha_1} x_2^{\alpha_2 - 2} \end{pmatrix}.$$

$$\Delta_1 = a\alpha_1(\alpha_1 - 1)x_1^{\alpha_1 - 2} x_2^{\alpha_2} < 0,$$
$$\Delta_2 = a^2\alpha_1\alpha_2 x_1^{2(\alpha_1 - 1)} x_2^{2(\alpha_2 - 1)}(1 - \alpha_1 - \alpha_2) > 0,\ (\alpha_1 + \alpha_2) < 1.$$

Таким образом, при выполнении условия $(\alpha_1 + \alpha_2) < 1$ матрица Гессе отрицательно определена. ►

3.4. Оптимальный выбор благ потребителем

3.4.1. Модель задачи оптимального выбора

В предыдущей лекции отмечалось, что предпочтение потребителя на множестве наборов благ выражается целевой функцией $u(x)$. Поэтому математическая модель выбора благ потребителем имеет следующий вид задачи математического программирования:

$$u(x) = u(x_1, ..., x_n) \to \max \tag{3.8}$$

при условиях:

$$\sum_{j=1}^{n} p_j x_j \le M, \tag{3.9}$$

$$x_j \ge 0,\ j = 1, ..., n. \tag{3.10}$$

Задача (3.8) – (3.10) является простейшей моделью, так как здесь допускается, что выбор благ потребителем ограничен только величиной дохода. На самом деле на выбор благ могут оказывать влияние и другие факторы, например недостаточное предложение (дефицитность) некоторых благ. Так, например, если предложение k*-го* блага ограничено числом $x_k^{\max}$, то нужно ввести ограничение $x_k \le x_k^{\max}$.

Таким образом, более сложные модели содержат ряд дополнительных ограничений.

Для $n = 2$ получаем задачу

$$u(x_1, x_2) \to \max \tag{3.11}$$

при ограничениях

$$p_1 x_1 + p_2 x_2 \le M;\; x_1 \ge 0,\; x_2 \ge 0. \tag{3.12}$$

Геометрическая интерпретация модели (3.11) – (3.12) дана на рис. 3.2.

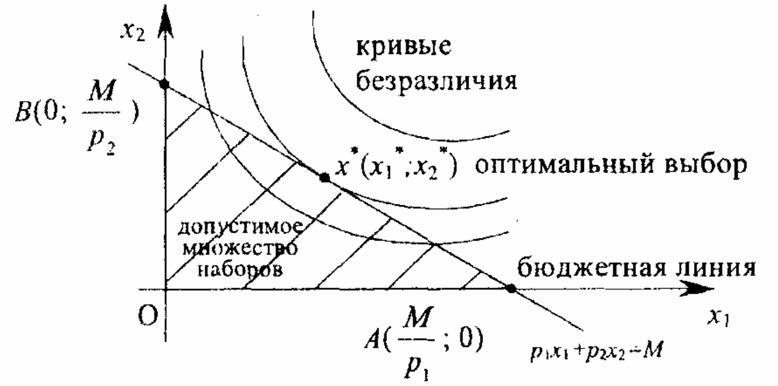


Рис. 3.2. Геометрическая интерпретация модели при $n = 2$.

На рис. 3.2 прямая *AB* соответствует бюджетному ограничению, треугольник OAB – области доступных наборов, а точка $x^*(x_1^*, x_2^*)$ касания кривой безразличия со стороной *AB* треугольника OAB определяет оптимальный набор благ задачи (3.11) – (3.12).

Задача (3.8) –(3.10) является задачей математического программирования и состоит в максимизации строго вогнутой функции при линейном ограничении. Решение такой задачи существует, и оно единственно. Это оптимальное решение называют *точкой равновесия задачи оптимального выбора благ потребителем.*

Необходимым и достаточным условием для решения задачи (3.8) –(3.10) являются условия Куна-Таккера для функции Лагранжа

$$L(x, \lambda) = u(x) + \lambda\left(M - \sum_{j=1}^{n} p_j x_j\right),$$

которые имеют вид

$$\frac{\partial L}{\partial x_j} = \frac{\partial u}{\partial x_j} - \lambda p_j = 0,\; j = \overline{1,n};\; x_j \frac{\partial L}{\partial x_j} = x_j\left(\frac{\partial u}{\partial x_j} - \lambda p_j\right) = 0, \tag{3.13}$$

$$\frac{\partial L}{\partial \lambda} = M - \sum_{j=1}^{n} p_j x_j \ge 0;\; \lambda \frac{\partial L}{\partial \lambda} = \lambda\left(M - \sum_{j=1}^{n} p_j x_j\right) = 0.$$

В (3.13) частные производные и переменные вычислены в оптимальной точке $(x_1^*,...,x_n^*;\lambda^*)$.

Из условий (3.13) следует, что если $x_j^* > 0$, то

$$\frac{\partial u}{\partial x_j} = \lambda^* p_j,\;\; j = 1,...,n. \tag{3.14}$$

Отсюда следует, что предельные полезности пропорциональны ценам соответствующих благ.

Геометрически свойство (3.13) означает, в точке оптимума вектор $p = (p_1,..., p_n)$ бюджетной гиперплоскости (прямой АВ на рис. 3.2) и вектор-градиент функции полезности $grad\left(u(x)\right) = \left(\frac{\partial u}{\partial x_1},...,\frac{\partial u}{\partial x_n}\right)$ коллинеарны, т.е. $grad\left(u(x)\right) = \lambda^* p$. Кроме того, из (3.13) следует, что $\frac{\partial u/\partial x_j}{p_j} = \lambda^* > 0$, так как $\frac{\partial u}{\partial x_j} > 0,\;\; p_j > 0,\;\; j = \overline{1,n}$.

Таким образом, оптимальный множитель Лагранжа λ^* должен быть положительным, а тогда из условия Куна-Таккера $\lambda\left(M - \sum_{j=1}^{n} p_j x_j\right) = 0$ следует, что весь доход используется на приобретение оптимального набора благ, т.е.

$$\sum_{j=1}^{n} p_j x_j = M. \tag{3.14}$$

Полученное оптимальное решение задачи (3.8) – (3.10) зависит от вектора цен p и дохода M, т.е. в общем случае решение задачи может быть записано в виде $x_j^* = x_j^*(p,M),\;\; j = 1,...,n$ и $\lambda^* = \lambda^*(p,M)$ как функций переменных $p_1,..., p_n$ и M.

Из приведенных результатов вытекают следующие следствия, имеющие место при оптимальном выборе благ потребителем.

1. Предельные полезности благ пропорциональны их ценам:

$$\frac{\partial u(x^*)}{\partial x_j} = \lambda^* p_j,\;\; j = 1,...,n.$$

2. . Отношение предельных полезностей двух благ равно отношению их цен:

$$\frac{\partial u(x^*)/\partial x_j}{\partial u(x^*)/\partial x_k} = \frac{p_j}{p_k};\;\; j,k = \overline{1,n},\;\; j \neq k.$$

3. Предельная полезность, приходящаяся на денежную единицу, одинакова для всех приобретаемых благ:

$$\frac{\partial u(x^*)/\partial x_j}{p_j} = \frac{\partial u(x^*)/\partial x_k}{p_k};\ j,k=\overline{1,n},\ \ j \neq k.$$

4. Равные предельные полезности, приходящиеся на денежную единицу, равны множителю λ^* - предельной полезности денежной единицы, которую потребитель расходует для приобретения благ:

$$\frac{\partial u(x^*)/\partial x_j}{p_j} = \lambda^*,\ \ j = 1,\ldots,n.$$

5.Норма замещения:

$$n_{kj} = \frac{\Delta x_k}{\Delta x_j} = -\frac{p_j}{p_k};\ \ j,k = 1,\ldots,n,\ \ k \neq j.$$

6.В оптимальной точке имеет равенство:

$$\frac{\partial u(x^*)}{\partial M} = \lambda^*. \tag{3.15}$$

Действительно, из соотношения $\sum_{j=1}^{n} p_j x_j = M$ следует равенство $\sum_{j=1}^{n} p_j \frac{\partial x_j}{\partial M} = 1$. Используя правило дифференцирования сложной функции и следствие 1, получим

$$\frac{\partial u(x^*)}{\partial M} = \sum_{j=1}^{n} \frac{\partial u(x^*)}{\partial x_j}\frac{dx_j}{dM} = \sum_{j=1}^{n} \lambda^* p_j \frac{dx_j}{dM} = \lambda^* \sum_{j=1}^{n} p_j \frac{dx_j}{dM} = \lambda^*.$$

Таким образом, величина λ^* множителя Лагранжа означает дополнительную полезность, приходящуюся на дополнительную единицу дохода, т.е. предельную полезность денежной единицы дохода потребителя.

3.5. Взаимная задача к задаче оптимального выбора благ потребителем

Мы нашли решение задачи об оптимальном выборе благ в виде

$$x_j^* = x_j^*(p, M),\ \ j = 1,\ldots,n.$$

Если подставить эти значения в функцию полезности, то получим

$$u^* = u^*(p, M),$$

т.е. максимальная полезность зависит в конечном счете от цен на блага и дохода потребителя.

Рассмотрим теперь так называемую взаимную задачу. Зафиксируем значение функции полезности на уровне u_0 и рассмотрим те блага x, для которых $u(x) = u_0$, т.е. те блага, которые дают потребителю один и тот же уровень удовлетворения u_0. Очевидно, что при заданном векторе цен на блага $p=(p_1,...,p_n)$ стоимости $M=\sum_{j=1}^{n} p_j x_j$ таких благ различны. Поставим *взаимную задачу:* какой набор благ, обеспечивающий данный уровень удовлетворения потребностей, самый дешевый. Математически такая взаимная задача формулируется так: найти минимум функции

$$M=\sum_{j=1}^{n} p_j x_j \tag{3.16}$$

при условии

$$u(x_1,...,x_n)=u_0;\; x_j \ge 0,\; j=\overline{1,n} \tag{3.17}$$

Геометрически для $n=2$ задачу (3.16) – (3.17) можно сформулировать следующим образом: для данной кривой безразличия с уравнением $u(x_1,x_2)=u_0$ среди параллельных бюджетных линий найти ту, которая ее касается. Точка касания и будет оптимальным решением (рис. 3.3).

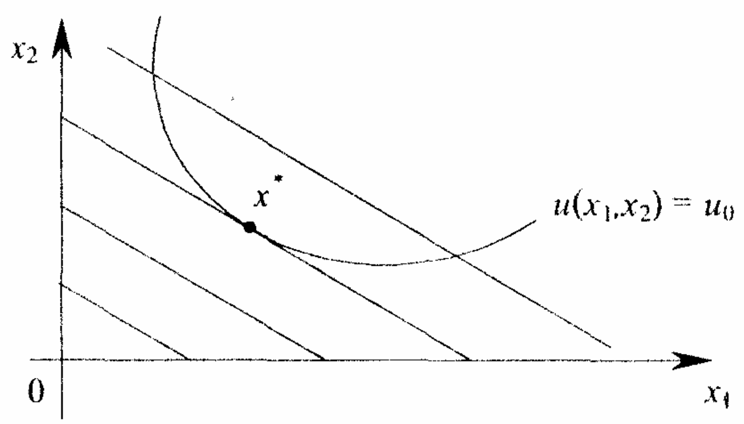


Рис. 3.3 Геометрическая интерпретация модели (3.16) – (3.17)

Задачу (3.16) – (3.17) решим методом множителей Лагранжа. Запишем функцию Лагранжа

$$L(x_1,...,x_n)=\sum_{j=1}^{n} p_j x_j + \mu\left(u_0 - u(x_1,...,x_n)\right).$$

Необходимые условия экстремума имеют вид

$$\frac{\partial L}{\partial x_j}=p_j-\mu\frac{\partial u}{\partial x_j}=0,\; j=1,...,n;$$

$$\frac{\partial L}{\partial \lambda}=u_0-u(x_1,...,x_n)=0,$$

которые перепишем в форме

$$\frac{\partial u}{\partial x_j} = \frac{1}{\mu} p_j, \ \ j = 1,...,n; \quad u(x_1,...,x_n) = u_0. \tag{3.18}$$

Решение системы (3.18) определяет оптимальный набор благ $x^* = (x_1^*,...,x_n^*)$, где

$$x_j^* = x_j^*(p,u_0), \tag{3.19}$$

и множитель μ^*. Минимальные затраты $M^* = \sum_{j=1}^{n} p_j x_j^*$ будут зависеть от величины u_0 и вектора цен p. Меняя уровень потребления u_0, получим $M^* = C(u_0)$, которая называется *функцией затрат потребителя.*

Выясним смысл множителя Лагранжа μ^*. Рассмотрим полный дифференциал функции затрат

$$dM = \sum_{j=1}^{n} p_j dx_j .$$

Так как в точке минимума $x^* = (x_1^*,...,x_n^*)$ справедливы равенства

$p_j = \mu^* \dfrac{\partial u}{\partial x_j}, \ \ j = 1,...,n$, то получаем

$$dM = \sum_{j=1}^{n} \mu^* \frac{\partial u}{\partial x_j} dx_j = \mu^* \sum_{j=1}^{n} \frac{\partial u}{\partial x_j} dx_j = \mu^* du. \tag{3.20}$$

Из (3.20) следует, что в оптимальной точке $\mu^* = \dfrac{dM}{du}$, т.е. множитель Лагранжа μ^* показывает, какие дополнительные затраты необходимо сделать, чтобы уровень удовлетворения (полезность) увеличить на единицу.

Выше было показано (см. (3.15))), что $\lambda^* = \dfrac{\partial u(x^*)}{\partial M}$, т.е. множитель λ^* показывает, какую дополнительную полезность мы получим на дополнительную единицу расходуемого дохода. Значит, по смыслу множители μ^* и λ^* взаимообратны. Установим, как связаны оптимальные решения взаимной задачи и задачи оптимального выбора благ потребителем. Пусть u_0 равно максимальному значению функции полезности u^*, полученному при решении задачи оптимального выбора благ потребителем. Тогда взаимная задача (3.16) – (3.17) примет следующий вид:
найти

$$\min\left(M = \sum_{j=1}^{n} p_j x_j \right) \tag{3.21}$$

при ограничениях

$$u(x_1,...,x_n) = u^*; \ x_j \ge 0, \ \ j = \overline{1,n}. \tag{3.22}$$

Её решение задает минимальный уровень затрат $M = M^*$, обеспечивающий максимальную полезность. Из геометрических соображений следует, что в этом случае оптимальные решения взаимной и исходной задач совпадут. Для $n = 2$ имеем одну и ту же точку касания бюджетной прямой и кривой безразличия, при этом во взаимной задаче задается кривая безразличия с максимальным уровнем удовлетворения потребностей и нужно найти к ней бюджетную прямую среди всевозможных параллельных между собой бюджетных прямых, а в задаче оптимального выбора благ потребителем задается бюджетная прямая уравнением $p_1x_1 + p_2x_2 = M$ и нужно среди всевозможных кривых безразличия найти ту, которая касается данной бюджетной прямой.

В этом случае, как легко видеть, что $\mu^* = \dfrac{1}{\lambda^*}$, а также $M^* = M$, где M – заданный доход потребителя в задаче оптимального выбора благ потребителем (рис. 3.4).

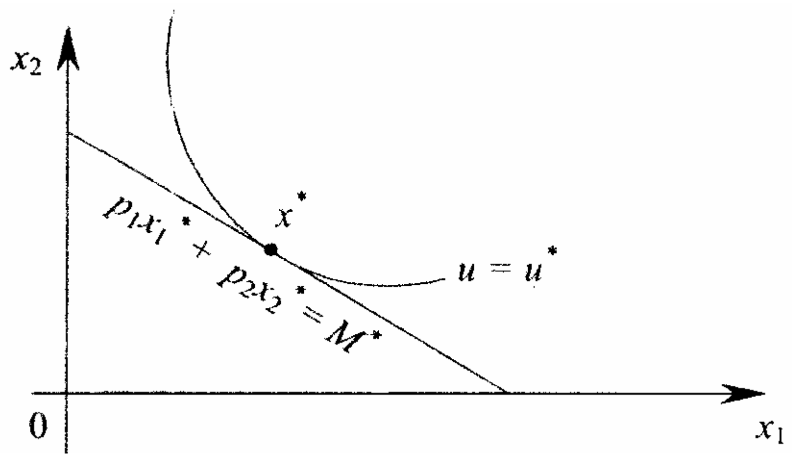


Рис. 3.4 Геометрическая интерпретация модели (3.21) – (3.22)

Пример 3.7. Для мультипликативной функции полезности потребителя
$u(x_1\, x_2) = ax_1^{\alpha_1}x_2^{\alpha_2};\ 0 < \alpha_i < 1,\ i = 1,2;\ a > 0$
найти решение:
а) задачи оптимального выбора благ потребителя;
б) взаимной задачи.

Решение.
а) Задача оптимального выбора благ потребителя имеет вид

$$u(x_1\, x_2) = ax_1^{\alpha_1}x_2^{\alpha_2} \to \max$$

при ограничениях

$$p_1x_1 + p_2x_2 = M;\ x_i \ge 0,\ i = 1,2.$$

Система уравнений (3.13) – (3.14) примет вид

$$\begin{aligned} &a\alpha_1x_1^{\alpha_1-1}x_2^{\alpha_2} = \lambda p_1,\ a\alpha_2x_1^{\alpha_1}x_2^{\alpha_2-1} = \lambda p_2, \\ &p_1x_1 + p_2x_2 = M. \end{aligned} \qquad (3.23)$$

Решая систему (3.23), получим

$$x_1^* = \frac{\alpha_1 M}{(\alpha_1+\alpha_2)p_1},\ x_2^* = \frac{\alpha_2 M}{(\alpha_1+\alpha_2)p_2},$$

$$\lambda^* = a\left(\alpha_1+\alpha_2\right)^{1-\alpha_1-\alpha_2}\left(\frac{\alpha_1}{p_1}\right)^{\alpha_1}\left(\frac{\alpha_2}{p_2}\right)^{\alpha_2} M^{-(1-\alpha_1-\alpha_2)}. \qquad (3.24)$$

Тогда полезность равна

$$u^* = a\left(\frac{\alpha_1}{(\alpha_1+\alpha_2)p_1}\right)^{\alpha_1}\left(\frac{\alpha_2}{(\alpha_1+\alpha_2)p_2}\right)^{\alpha_2} M^{\alpha_1+\alpha_2},$$

а множитель Лагранжа равен

$$\lambda^* = \frac{\alpha_1+\alpha_2}{M}u^*.$$

б) Взаимная задача: найти

$$\min\left(M = p_1x_1 + p_2x_2\right)$$

при условиях

$$ax_1^{\alpha_1}x_2^{\alpha_2} = u_0,\ x_1 \ge 0,\ x_2 \ge 0.$$

Функция Лагранжа этой задачи

$$L = p_1x_1 + p_2x_2 + \mu\left(ax_1^{\alpha_1}x_2^{\alpha_2} - u_0\right).$$

Запишем условия экстремума

$$\begin{aligned} &\frac{\partial L}{\partial x_1} = p_1 - \mu a\alpha_1 x_1^{\alpha_1-1}x_2^{\alpha_2} = 0,\\ &\frac{\partial L}{\partial x_2} = p_2 - \mu a\alpha_2 x_1^{\alpha_1}x_2^{\alpha_2-1} = 0, \qquad (3.25)\\ &ax_1^{\alpha_1}x_2^{\alpha_2} = u_0. \end{aligned}$$

Теперь решаем систему (3.25). В результате получим

$$x_1^* = \left(\frac{u_0}{a}\right)^{\frac{1}{\alpha_1+\alpha_2}}\left(\frac{\alpha_1p_2}{\alpha_2p_1}\right)^{\frac{\alpha_2}{\alpha_1+\alpha_2}},\ x_2^* = \left(\frac{u_0}{a}\right)^{\frac{1}{\alpha_1+\alpha_2}}\left(\frac{\alpha_1p_2}{\alpha_2p_1}\right)^{-\frac{\alpha_1}{\alpha_1+\alpha_2}}. \qquad (3.26)$$

Минимальные затраты и множитель Лагранжа равны

$$M^* = p_1x_1^* + p_2x_2^* = \left(\frac{u_0}{a}\right)^{\frac{1}{\alpha_1+\alpha_2}}\left[p_1\left(\frac{\alpha_1p_2}{\alpha_2p_1}\right)^{\frac{\alpha_2}{\alpha_1+\alpha_2}} + p_2\left(\frac{\alpha_1p_2}{\alpha_2p_1}\right)^{-\frac{\alpha_1}{\alpha_1+\alpha_2}}\right]. \qquad (3.27)$$

$$\mu^* = \frac{1}{a}\left(\frac{u_0}{a}\right)^{\frac{1-\alpha_1-\alpha_2}{\alpha_1+\alpha_2}}\left(\frac{p_1}{\alpha_1}\right)^{\frac{\alpha_1}{\alpha_1+\alpha_2}}\left(\frac{p_2}{\alpha_2}\right)^{\frac{\alpha_2}{\alpha_1+\alpha_2}}. \qquad (3.28)$$

При

$$u_0 = u^* = a\left(\frac{\alpha_1}{(\alpha_1+\alpha_2)p_1}\right)^{\alpha_1}\left(\frac{\alpha_2}{(\alpha_1+\alpha_2)p_2}\right)^{\alpha_2} M^{\alpha_1+\alpha_2} \qquad (3.29)$$

получим

$$x_1^* = \left(\frac{u_0}{a}\right)^{\frac{1}{\alpha_1+\alpha_2}}\left(\frac{\alpha_1 p_2}{\alpha_2 p_1}\right)^{\frac{\alpha_2}{\alpha_1+\alpha_2}} = \frac{\alpha_1 M}{(\alpha_1+\alpha_2)p_1},$$

$$x_2^* = \left(\frac{u_0}{a}\right)^{\frac{1}{\alpha_1+\alpha_2}}\left(\frac{\alpha_1 p_2}{\alpha_2 p_1}\right)^{-\frac{\alpha_1}{\alpha_1+\alpha_2}} = \frac{\alpha_2 M}{(\alpha_1+\alpha_2)p_2}.$$

Отсюда, при $u_0 = u^*$ имеем

$$M^* = p_1 x_1^* + p_2 x_2^* = p_1\frac{\alpha_1 M}{(\alpha_1+\alpha_2)p_1} + p_2\frac{\alpha_2 M}{(\alpha_1+\alpha_2)p_2} = M.$$

Подставим (3.29) в (3.28), получим

$$\mu^* = \frac{1}{a}\left(\frac{1}{\alpha_1+\alpha_2}\right)^{1-\alpha_1-\alpha_2}\left(\frac{p_1}{\alpha_1}\right)^{\alpha_1}\left(\frac{p_2}{\alpha_2}\right)^{\alpha_2} M^{1-\alpha_1-\alpha_2}. \qquad (3.30)$$

Сравнивая (3.30) и (3.24), видим, что $\mu^* = \frac{1}{\lambda^*}$, т.е. значения множителей Лагранжа для рассматриваемых задач при $u_0 = u^*$ взаимообратны. ►

Вопросы для самопроверки

1. Дайте понятие множества благ, ограниченного множества благ и доступных благ
2. Какие возможны множества доступных благ?
3. Дайте понятие функции полезности.
4. Какие типы функции полезности вам известны?
5. Перечислите свойства функции полезности.
6. Что такое поверхность безразличия и кривые безразличия?
7. Перечислите свойства поверхности безразличия.
8. Что такое предельная полезность блага и коэффициент предельной эквивалентной замены благ?
9. В чем состоит задача оптимального выбора благ? Записать математическую модель.
10. Записать решение задачи выбора оптимальных благ.
11. Дайте интерпретацию множителя Лагранжа задачи оптимального выбора благ.
12. Перечислите следствия, вытекающие из оптимального выбора благ?

13. В чем состоит взаимная задача к задаче оптимального выбора благ? Записать математическую модель.
14. Как решить взаимную задачу?
15. Дайте интерпретацию множителя Лагранжа взаимной задачи.
16. Как связаны множители Лагранжа задачи оптимального выбора благ и взаимной задачи.

Тема 4. Балансовые модели

4.1 Балансовый метод. Принципиальная схема межпродуктового баланса

Балансовые модели, как статические, так и динамические, широко применяются при экономико-математическом моделировании экономических систем и процессов. В основе создания этих моделей лежит балансовый метод, т.е. метод взаимного сопоставления имеющихся материальных, трудовых и финансовых ресурсов и потребностей в них. Если описывать экономическую систему в целом, то под балансовой моделью понимается система уравнений, каждое из которых выражает требование баланса между производимым отдельными экономическими объектами количеством продукции и совокупной потребностью в этой продукции. При таком подходе рассматриваемая система состоит из экономических объектов, каждый из которых выпускает некоторый продукт, часть его потребляется другими объектами системы, а другая часть выводится за пределы системы в качестве ее конечного продукта.(см. рис. 4.1). Если вместо понятия п р о д у к т ввести более общее понятие

р е с у р с , то под *балансовой моделью* следует понимать систему уравнений, которые удовлетворяют требованиям соответствия наличия ресурса и его использования. Кроме приведенного выше требования соответствия производства каждого продукта и потребности в нем, можно указать такие примеры балансового соответствия, как соответствие наличия рабочей силы и количества рабочих мест, платежеспособного спроса населения и предложения товаров и услуг и т. д. При этом соответствие понимается либо как равенство, либо менее жестко — как достаточность ресурсов для покрытия потребности и, следовательно, наличие некоторого резерва.

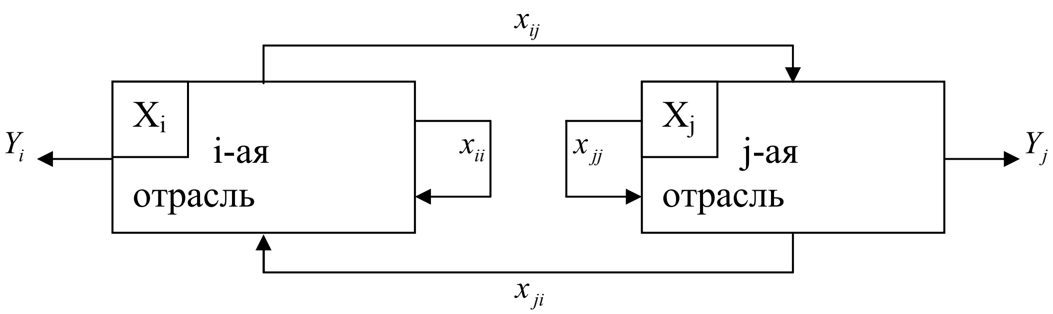


Рис. 4.1. Связь отраслей производства

Важнейшие виды балансовых моделей:

- частные материальные, трудовые и финансовые балансы для народного хозяйства и отдельных отраслей;
- межотраслевые балансы;
- матричные техпромфинпланы предприятий и фирм.

Балансовый метод и создаваемые на его основе балансовые модели служат основным инструментом поддержания пропорций в народном хозяйстве. Балансовые модели на базе отчетных балансов характеризуют сложившиеся пропорции, в них ресурсная часть всегда равна расходной.

Основу информационного обеспечения балансовых моделей в экономике составляет матрица коэффициентов затрат ресурсов по конкретным направлениям их использования. Например, в модели межотраслевого баланса такую роль играет так называемая *технологическая матрица* — таблица межотраслевого баланса, составленная из коэффициентов (нормативов) прямых затрат на производство единицы продукции в натуральном выражении. По многим причинам исходные данные реальных хозяйственных объектов не могут быть использованы в балансовых моделях непосредственно, поэтому подготовка информации для ввода в модель является весьма серьезной проблемой. Так, при построении модели межотраслевого баланса используется специфическое понятие чистой (или технологической) отрасли, т.е. условной отрасли, объединяющей все производство данного продукта независимо от ведомственной (административной) подчиненности и форм собственности предприятий и фирм. Переход от хозяйственных отраслей к чистым отраслям требует специального преобразования реальных данных хозяйственных объектов, например, агрегирования отраслей, исключения внутриотраслевого оборота и др. В этих условиях понятия «межпродуктовый баланс» и «межотраслевой баланс практически идентичны, отличие заключается лишь в единицах измерения элементов баланса.

Принципиальная схема межотраслевого баланса производства и распределения совокупного общественного продукта в стоимостном выражении приведена в табл. 4.1.

Таблица 4.1

Производящие отрасли	Потребляющие отрасли					Конечный продукт	Валовый продукт
	1	2	3	…	n		
1	x_{11}	x_{12}	x_{13}	…	x_{1n}	Y_1	X_1
2	x_{21}	x_{22}	x_{23}	…	x_{2n}	Y_2	X_2
3	x_{31}	x_{32}	x_{33}	…	x_{3n}	Y_3	X_3
…	…	…	…	I	…	…	…
…	…	…	…	…	…	II	…
…	…	…	…	…	…	…	…
n	x_{n1}	x_{n2}	x_{n3}	…	x_{nn}	Y_n	X_n
Итого	$\sum x_{i1}$	$\sum x_{i2}$	$\sum x_{i3}$	…	$\sum x_{in}$	$\sum Y_i$	$\sum X_i$
Амортизация Оплата ттруд Чистый доход	c_1 v_1 m_1	c_2 v_2 m_2	c_3 v_3 m_3	… III …	c_n v_n m_n	IV	
Валовой Продукт	X_1	X_2	X_3	…	X_n		$\sum\limits_{i=1}^{n} X_i = \sum\limits_{j=1}^{n} X_j$

В основу этой схемы положено разделение совокупного продукта на две части: промежуточный и конечный продукт; все народное хозяйство представлено в виде совокупности n отраслей (имеются в виду чистые отрасли), при этом каждая отрасль фигурирует в балансе как производящая и как потребляющая.

Рассмотрим схему МОБ в разрезе его крупных составных частей. Выделяются четыре части, имеющие различное экономическое содержание, они называются квадрантами баланса и на схеме обозначены римскими цифрами.

П е р в ы й к в а д р а н т МОБ — это шахматная таблица межотраслевых материальных связей. Показатели, помещенные на пересечениях строк и столбцов, представляют собой величины межотраслевых потоков продукции и в общем виде обозначаются x_{ij}, где i и j — соответственно номера отраслей производящих и потребляющих. Так, величина x_{32} понимается как стоимость средств производства, произведенных в отрасли с номером 3 и потребленных в качестве материальных затрат в отрасли с номером 2. Таким образом, первый квадрант по форме представляет собой квадратную матрицу порядка n, сумма всех элементов которой равняется годовому фонду возмещения затрат средств производства в материальной сфере.

Во в т о р о м к в а д р а н т е представлена конечная продукция всех отраслей материального производства, при этом под конечной понимается продукция, выходящая из сферы производства в область конечного использования (на потребление и накопление). В табл. 1 этот раздел дан укрупненно в виде одного столбца величин Y_i; в развернутой схеме баланса конечный продукт каждой отрасли показан дифференцированно по направлениям использования: на личное потребление населения, общественное потребление, на накопление, возмещение потерь, экспорт и др. Итак, второй квадрант характеризует отраслевую материальную структуру национального дохода, а в развернутом виде — также распределение национального дохода на фонд накопления и фонд потребления, структуру потребления и накопление по отраслям производства и потребителям.

Т р е т и й к в а д р а н т МОБ также характеризует национальный доход, но со стороны его стоимостного состава как сумму чистой продукции и амортизации; чистая продукция понимается при этом как сумма оплаты труда и чистого дохода отраслей. Сумму амортизации (c_j) и чистой продукции ($v_j + m_j$) некоторой j-й отрасли будем называть условно чистой продукцией этой отрасли и обозначать в дальнейшем Z_j. Таким образом, $Z_j = c_j + v_j + m_j$.

Ч е т в е р т ы й к в а д р а н т баланса находится на пересечении столбцов второго квадранта (конечной продукции) и строк третьего квадранта (условно чистой продукции). Этим определяется содержание квадранта: он отражает конечное распределение и использование национального дохода. В результате перераспределения первоначально созданного национального дохода образуются конечные доходы населения, предприятий, государства. Данные четвертого квадранта важны для отражения в межотраслевой модели баланса доходов и расходов населения, источников финансирования капиталовложений, текущих затрат непроизводственной сферы, для анализа общей структуры конечных доходов по группам потребителей. Более детально составляющие элементы этого квадранта в данном пособии не рассматриваются, однако очень важным является тот факт, что общий итог четвертого квадранта, так же как второго и третьего, должен быть равен созданному за год национальному доходу.

Запишем два важных соотношения. 1) Рассматривая схему баланса по столбцам, можно сделать очевидный вывод, что итог материальных затрат любой потребляющей отрасли и ее условно чистой продукции равен валовой продукции этой отрасли. Данный вывод можно записать в виде соотношения:

$$X_j = \sum_{i=1}^{n} x_{ij} + Z_j;\ j = 1,...,n. \qquad (4.1)$$

Напомним, что величина условно чистой продукции Z_j равна сумме амортизации, оплаты труда и чистого дохода j-и отрасли. Соотношение (4.1) охватывает систему из n уравнений, отражающих стоимостный состав продукции всех отраслей материальной сферы.

2) Рассматривая схему МОБ по строкам для каждой производящей отрасли, можно видеть, что валовая продукция той или иной отрасли равна сумме материальных затрат потребляющих ее продукцию отраслей и конечной продукции данной отрасли:

$$X_i = \sum_{j=1}^{n} x_{ij} + Y_i;\ i = 1,...,n \tag{4.2}$$

Формула (4.2) описывает систему из n уравнений, которые называются уравнениями распределения продукции отраслей материального производства по направлениям использования.

Просуммируем по всем отраслям уравнения (1), в результате получим

$$\sum_{j=1}^{n} X_j = \sum_{j=1}^{n}\sum_{i=1}^{n} x_{ij} + \sum_{j=1}^{n} Z_j$$

Аналогичное суммирование уравнений (2) дает:

$$\sum_{i=1}^{n} X_i = \sum_{i=1}^{n}\sum_{j=1}^{n} x_{ij} + \sum_{j=1}^{n} Y_i$$

Левые части обоих равенств равны, так как представляют собой весь валовой общественный продукт. Первые слагаемые правых частей этих равенств также равны, их величина равна итогу первого квадранта. Следовательно, должно соблюдаться соотношение

$$\sum_{j=1}^{n} Z_j = \sum_{i=1}^{n} Y_i \tag{4.3}$$

Левая часть уравнения (4.3) есть сумма третьего квадранта, а правая часть — итог второго квадранта. В целом же это уравнение показывает, что в межотраслевом балансе соблюдается важнейший принцип единства стоимостного и материального состава национального дохода.

4.2 Экономико-математическая модель межотраслевого баланса

Выше в п. 4.1 было отмечено, что основу информационного обеспечения модели межотраслевого баланса составляет технологическая матрица, содержащая коэффициенты прямых материальных затрат на производство единицы продукции. Эта матрица является также основой экономико-математической модели межотраслевого баланса. Предполагается, что для производства единицы продукции в j-й отрасли требуется определенное количество затрат промежуточной продукции i-й отрасли, равное a_{ij}. Оно не зависит от объема производства в отрасли и является довольно стабильной величиной во времени. Величины a_{ij} называются ***коэффициентами прямых материальных затрат*** и рассчитываются следующим образом:

$$a_{ij} = \frac{x_{ij}}{X_j};\ i, j = \overline{1, n}. \tag{4.4}$$

Определение 1. Коэффициент прямых материальных затрат показывает, какое количество продукции i-й отрасли необходимо, если учитывать только прямые затраты, для производства единицы продукции j-й отрасли.

С учетом формулы (4.4) систему уравнений баланса (2) можно переписать в виде

$$X_i = \sum_{j=1}^{n} a_{ij} X_j + Y_i;\ i = 1, ..., n \tag{4.5}$$

Если ввести в рассмотрение матрицу коэффициентов прямых материальных затрат $A = (a_{ij})$, вектор-столбец валовой продукции X и вектор-столбец конечной продукции Y:

$$X = \begin{pmatrix} X_1 \\ X_2 \\ ... \\ X_n \end{pmatrix}, \qquad Y = \begin{pmatrix} Y_1 \\ Y_2 \\ ... \\ Y_n \end{pmatrix}$$

то система уравнений (4.5) в матричной форме примет вид

$$X = AX + Y \tag{4.6}$$

Система уравнений (5), или в матричной форме (4.6), называется *экономико-математической моделью межотраслевого баланса (моделью Леонтьева, моделью «затраты— выпуск»)*. С помощью этой модели можно выполнять три варианта расчетов:

- Задав в модели величины валовой продукции каждой отрасли (X_i), можно определить объемы конечной продукции каждой отрасли (Y_i):

$$Y = (E - A)X \tag{4.7}$$

- Задав величины конечной продукции всех отраслей (Y_i), можно определить величины валовой продукции каждой отрасли (X_i):

$$X = (E - A)^{-1} Y \tag{4.8}$$

- Для ряда отраслей задав величины валовой продукции, а для всех остальных отраслей задав объемы конечной продукции, можно найти величины конечной продукции первых отраслей и объемы валовой продукции вторых, в этом варианте расчета удобнее пользоваться не матричной формой модели (4.6), а системой линейных уравнений (4.5). В формулах (4.7) и (4.8) E обозначает единичную матрицу n-го порядка, а $(E - A)^{-1}$ обозначает матрицу, обратную к матрице $(E - A)$. Если определитель матрицы $(E - A)$ не равен нулю, то обратная к ней матрица существует. Обозначим эту обратную матрицу через B, тогда систему уравнений в матричной форме (4.8) можно записать в виде

$$X = BY. \tag{4.9}$$

Элементы матрицы B будем обозначать через b_{ij}, тогда из матричного уравнения (4.9) для любой i-й отрасли можно получить следующее соотношение:

$$X_i = \sum_{j=1}^{n} b_{ij} Y_j, \; i = 1,...,n. \qquad (4.10)$$

Из соотношений (4.10) следует, что валовая продукция выступает как взвешенная сумма величин конечной продукции, причем весами являются коэффициенты b_{ij}, которые показывают, сколько всего нужно произвести продукции i-й отрасли для выпуска в сферу конечного использования единицы продукции j-и отрасли. В отличие от коэффициентов прямых затрат a_{ij} коэффициенты b_{ij} называются *коэффициентами полных материальных затрат* и включают в себя как прямые, так и косвенные затраты всех порядков. Если прямые затраты отражают количество средств производства, израсходованных непосредственно при изготовлении данного продукта, то косвенные относятся к предшествующим стадиям производства и входят в производство продукта не прямо, а через другие (промежуточные) средства производства. Более детально этот вопрос рассматривается в п.4.3

Определение 2. Коэффициент полных материальных затрат b_{ij} показывает, какое количество продукции i й отрасли нужно произвести, чтобы с учетом прямых и косвенных затрат этой продукции получить единицу конечной продукции j-й отрасли.

Коэффициенты полных материальных затрат можно применять, когда необходимо определить, как скажется на валовом выпуске некоторой отрасли предполагаемое изменение объемов конечной продукции всех отраслей:

$$\Delta X_i = \sum_{j=1}^{n} b_{ij} \Delta Y_j, \; i = 1,...,n \qquad (4.11)$$

где ΔX_i и ΔY_j — изменения (приросты) величин валовой и конечной продукции соответственно.

4.3 Коэффициенты прямых и полных материальных затрат

Переходя к анализу модели межотраслевого баланса, необходимо прежде всего рассмотреть основные свойства матрицы коэффициентов прямых материальных затрат A. Коэффициенты прямых затрат по определению являются неотрицательными, следовательно, матрица A в целом может быть названа неотрицательной: $A > 0$. Так как процесс воспроизводства нельзя было бы осуществлять, если бы для собственного воспроизводства в отрасли затрачивалось большее количество продукта, чем создавалось, то очевидно, что диагональные элементы матрицы A меньше единицы: $a_{ii} < 1$.

Система уравнений межотраслевого баланса является отражением реальных экономических процессов, в которых содержательный смысл могут иметь лишь неотрицательные значения валовых выпусков; таким образом, вектор валовой продукции состоит из неотрицательных компонентов и называется неотрицательным: $X \geq 0$. Встает вопрос, при каких условиях экономическая система способна обеспечить положительный конечный выпуск по всем отраслям. Ответ на этот вопрос связан с понятием продуктивности матрицы коэффициентов прямых материальных затрат.

Будем называть неотрицательную матрицу A *продуктивной,* если существует такой неотрицательный вектор $X \geq 0$, что

$$X > AX. \tag{4.12}$$

Очевидно, что условие (4.12) означает существование положительного вектора конечной продукции $Y > 0$ для модели межотраслевого баланса (6).

Для того чтобы матрица коэффициентов прямых материальных затрат A была продуктивной, необходимо и достаточно чтобы выполнялось одно из перечисленных ниже условий:

1) матрица $(E - A)$ неотрицательно обратима, т.е. существует обратная матрица $(E - A)^{-1} \geq 0$;
2) наибольшее по модулю собственное значение λ матрицы A, то есть решение характеристического уравнения $|\lambda E - A| = 0$, строго меньше единицы;
3) все главные миноры матрицы $(E - A)$, т.е. определители матриц, образованные элементами первых строк и первых столбцов этой матрицы, порядка от 1 до n, положительны.
4) норма матрицы должна быть строго меньше единицы, т.е. на величина наибольшей из сумм элементов матрицы A *в* каждом столбце должна быть строго меньше единицы

Перейдем к анализу матрицы коэффициентов полных материальных затрат, т.е. матрицы $B = (E - A)^{-1}$. Согласно определению 2 коэффициент этой матрицы показывает, сколько всего нужно произвести продукции i-й отрасли, чтобы получить единицу конечной продукции j-й отрасли.

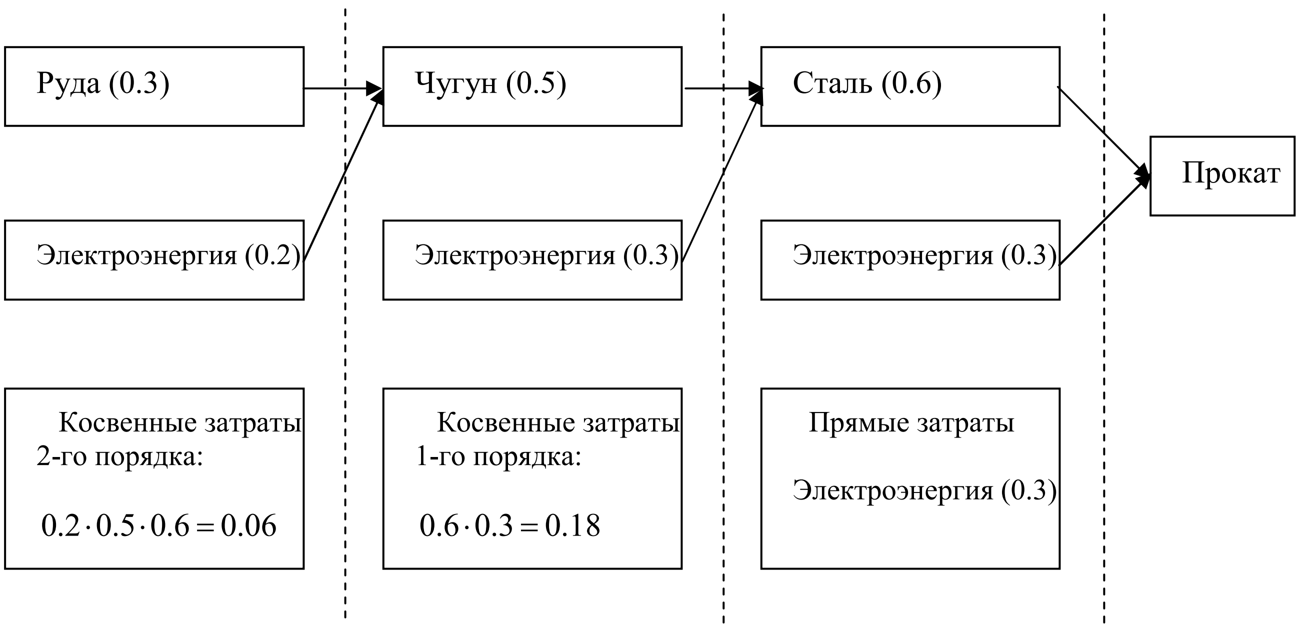


Рис. 4.2. Затраты электроэнергии на выпуск стального проката

Как уже указывалось выше, коэффициент полных материальных затрат включает прямые затраты и косвенные затраты. В отличие от коэффициентов прямых затрат, коэффициенты полных материальных затрат отражают не отраслевые, а народнохозяйственные затраты по производству единицы продукции.

Поясним косвенные затраты на следующем примере. Рассмотрим в качестве примера формирование затрат электроэнергии на выпуск стального проката, при этом ограничимся технологической цепочкой «руда-чугун-сталь-прокат». Затраты электроэнергии при получении проката из стали будут называться прямыми затратами, те же затраты при получении стали из чугуна будут называться косвенными затратами 1-го порядка, а затраты электроэнергии при получении чугуна из руды будут называться косвенными затратами электроэнергии на выпуск стального проката 2-го порядка и т. д. (см. рис 4.2).

Перейдем теперь к вычислительным аспектам решения задач на основе модели межотраслевого баланса. Основной объем расчетов по этой модели связан с вычислением матрицы коэффициентов полных материальных затрат B. Если матрица коэффициентов прямых материальных затрат A задана и является продуктивной, то матрицу B можно находить по формулам обращения матриц, рассматриваемым в курсе вычислительной математики.

Пример 4.1. Для трехотраслевой экономической системы заданы матрица коэффициентов прямых материальных затрат и вектор конечной продукции:

$$A = \begin{pmatrix} 0.3 & 0.1 & 0.4 \\ 0.2 & 0.5 & 0.0 \\ 0.3 & 0.1 & 0.2 \end{pmatrix}; \quad Y = \begin{pmatrix} 200 \\ 100 \\ 300 \end{pmatrix}$$

Найти коэффициенты полных материальных затрат и вектор валовой продукции, заполнить схему межотраслевого материального баланса.

1. Определим матрицу коэффициентов полных материальных затрат.

а) Вычислим матрицу $E - A$.

$$E - A = \begin{pmatrix} 0.7 & -0.1 & -0.4 \\ -0.2 & 0.5 & -0.0 \\ -0.3 & -0.1 & 0.8 \end{pmatrix}.$$

б) Находим обратную матрицу одним из методов обращения матриц (например, методом матричной алгебры)

$$B = (E - A)^{-1} = \begin{pmatrix} 2.041 & 0.612 & 1.020 \\ 0.816 & 2.245 & 0.408 \\ 0.887 & 0.510 & 1.664 \end{pmatrix}$$

2. Найдем величины валовой продукции трех отраслей (вектор X), используя формулу (8'):

$$X = BY = \begin{pmatrix} 2.041 & 0.612 & 1.020 \\ 0.816 & 2.245 & 0.408 \\ 0.867 & 0.510 & 1.664 \end{pmatrix} \cdot \begin{pmatrix} 200 \\ 100 \\ 300 \end{pmatrix} = \begin{pmatrix} 775.3 \\ 510.1 \\ 729.6 \end{pmatrix}$$

3. Для определения элементов первого квадранта материального межотраслевого баланса воспользуемся формулой, вытекающей из формулы (4.4): $x_{ij} = a_{ij} X_j$. Из этой формулы следует, что для получения первого столбца первого квадранта нужно элементы первого столбца заданной матрицы A умножить на величину $X_1 = 775.3$;

элементы второго столбца матрицы A умножить на $X_2 = 510.1$; элементы третьего столбца матрицы A умножить на $X_3 = 729.6$.

Составляющие третьего квадранта (условно чистая продукция) находятся с учетом формулы (1) как разность между объемами валовой продукции и суммами элементов соответствующих столбцов найденного первого квадранта.

Четвертый квадрант в нашем примере состоит из одного показателя и служит, в частности, для контроля правильности расчета: сумма элементов второго квадранта должна в стоимостном материальном балансе совпадать с суммой элементов третьего квадранта. Результаты расчета представлены в табл. 4.2.

Таблица 4.2.. Межотраслевой баланс производства и распределения продукции

Производящие отрасли	**Потребляющие отрасли**				
	1	2	3	*Конечная продукция*	*Валовая продукция*
1	232,6	51,0	291,8	200,0	775,3
2	155,1	255,0	0,0	100,0	510,1
3	232,6	51,0	145,9	300,0	729,6
Условно чистая продукция	155,0	153,1	291,9	600,0	
Валовая продукция	775,3	510,1	729,6		2015,0

4.4 Агрегирование показателей межотраслевого баланса

При моделировании межотраслевых связей важным является вопрос агрегирования нормативных показателей.

Рассмотрим пример. Пусть задана таблица межотраслевых потоков для 4 отраслей.

Таблица 4.3.

Производящая отрасль	Распределение по потребляющим отраслям				*Конечный продукт*	*Валовой продукт*
	1	2	3	4		
1	x_{11}	x_{12}	x_{13}	x_{14}	Y_1	X_1
2	x_{21}	x_{22}	x_{23}	x_{24}	Y_2	X_2
3	x_{31}	x_{32}	x_{33}	x_{34}	Y_3	X_3
4	x_{41}	x_{42}	x_{43}	x_{44}	Y_4	X_4

Определим параметры агрегирования при объединении второй и третьей отраслей. Выделим в табл. 2 отрасли подлежащие агрегированию. Присвоим новой отрасли индекс k и составим другую таблицу, введя в нее отрасль k.

Таблица 4.4.

Производящая отрасль	Распределение по потребляющим отраслям				*Конечный продукт*	*Валовой продукт*
	1	2	3	4		
1	x_{11}	x_{12}	x_{13}	x_{14}	Y_1	X_1
k	x_{k1}	x_{k2}	x_{k3}	x_{k4}	Y_k	X_k
4	x_{41}	x_{42}	x_{43}	x_{44}	Y_4	X_4

Агрегированными окажутся те межотраслевые потоки, которые содержат индекс k.
По таблице межотраслевых потоков x_{ij} построим матрицу коэффициентов прямых затрат A:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix},$$

где $a_{ij} = \dfrac{x_{ij}}{X_j}$, $i, j = \overline{1, n}$.

Далее сформируем оператор агрегирования T. Для этого произведем деформацию единичной матрицы 4 – го порядка (размерность единичной матрицы совпадает с размерностью исходной таблицы МОБ) по следующему правилу: выделим в единичной матрице E те строки, номера которых совпадают с номерами агрегируемых отраслей, и просуммируем их. Результат внесем в k-ую строку матрицы T. Все остальные строки, номера которых в исходной таблице соответствуют номерам отраслей, не подлежащих агрегированию, переписываем в матрицу T без изменения.

Для нашего примера

$$E = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \to T = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Таким образом, матрица T есть результат горизонтальной «деформации» матрицы E.

Построим деформированную весовую матрицу W. Для этого введем веса W_i, означающие вклад валовой продукции исходной i-ой отрасли в валовую продукцию отраслей, представленных в новой агрегированной таблице.

Так, 1 –ая и 4 –ая отрасли в нашем примере не подлежат агрегированию. Следовательно,

$W_1 = 1,\ W_4 = 1,$

$$W_2 = \frac{X_2}{X_2 + X_3},\ W_3 = \frac{X_3}{X_2 + X_3}.$$

Составим весовую матрицу W:

$$W = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & W_2 & 0 & 0 \\ 0 & 0 & W_3 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Деформируем матрицу W по столбцам, объединив второй и третий столбцы. Тогда

$$W^* = \begin{pmatrix} 1 & 0 & 0 \\ 0 & W_2 & 0 \\ 0 & W_3 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

где W^* - весовой оператор агрегирования.

Число строк весового оператора агрегирования W^* совпадает с размерностью исходной таблицы МОБ, число столбцов равно количеству строк деформированной матрицы T.

Агрегированная матрица A^* прямых затрат вычисляется по формуле

$$A^* = T \cdot A \cdot W^*.$$

4.5 Анализ экономических показателей

Рассмотрим применение межотраслевого балансового метода для анализа таких важных экономических показателей, как труд, фонды и цены.

4.5.1 Модель затрат труда

К числу важнейших аналитических возможностей межотраслевого балансового метода относится определение прямых и полных затрат труда на единицу продукции и разработка на этой основе балансовых продуктово-трудовых моделей. Исходной моделью при этом служит отчетный межпродуктовый баланс в **натуральном выражении**. В этом балансе по строкам представлено распределение каждого отдельного продукта на производство других продуктов и конечное потребление (первый и второй квадранты схемы межотраслевого баланса). Отдельной строкой дается распределение затрат живого труда в производстве всех видов продукции; предполагается, что трудовые затраты выражены в единицах труда одинаковой степени сложности.

Обозначим затраты живого труда в производстве j-то продукта через L_j, а объем производства этого продукта (валовой выпуск), как и раньше, через X_j;. Тогда прямые затраты труда на единицу j-го вида продукции *(коэффициент прямой трудоемкости)* можно задать следующей формулой:

$$t_j = \frac{L_j}{X_j}; \; j = \overline{1,n}. \tag{4.13}$$

Введем понятие *полных затрат труда* как суммы прямых затрат живого труда и затрат овеществленного труда, перенесенных на продукт через израсходованные средства производства. Если обозначить величину полных затрат труда на единицу продукции j-го вида через T_j, то произведения вида $a_{ij}T_i$ отражают затраты овеществленного труда, перенесенного на единицу j-го продукта через i -е средство производства; при этом предполагается, что коэффициенты прямых материальных затрат a_{ij} выражены в натуральных единицах. Тогда полные трудовые затраты на единицу j-го вида продукции *(коэффициент полной трудоемкости)* будут равны

$$T_j = \sum_{i=1}^{n} a_{ij} T_i + t_j; \; j = \overline{1,n}. \qquad (4.14)$$

Введем в рассмотрение вектор-строку коэффициентов прямой трудоемкости $t = (t_1, t_2, ..., t_n)$ и вектор-строку коэффициентов полной трудоемкости $T = (T_1, T_2, ..., T_n)$.

Тогда с использованием уже рассматриваемой выше матрицы коэффициентов прямых материальных затрат A (в натуральном выражении) систему уравнений (14) можно переписать в матричном виде:

$$T = TA + t. \qquad (4.15)$$

Из (15) получим получим следующее соотношение для вектора коэффициентов полной трудоемкости:

$$T = t(E - A)^{-1} \qquad (4.16)$$

Матрица $(E - A)^{-1}$ нам уже знакома, это матрица B коэффициентов полных материальных затрат, так что последнее равенство можно переписать в виде

$$T = tB \qquad (4.17)$$

Обозначим через L величину совокупных затрат живого труда по всем видам продукции, которая с учетом формулы (13) будет равна

$$L = \sum_{j=1}^{n} L_j = \sum_{j=1}^{n} t_j X_j = tX \qquad (4.18)$$

Учитывая соотношение $X = BY$ из (18) и (17), приходим к следующему равенству:

$$tX = TY, \qquad (4.19)$$

здесь t и T — вектор-строки коэффициентов прямой и полной трудоемкости, а X и Y — вектор-столбцы валовой и конечной продукции соответственно.

Соотношение (19) представляет собой основное балансовое равенство в теории межотраслевого баланса труда. В данном случае его конкретное экономическое содержание заключается в том, что *стоимость совокупных затрат живого труда равна стоимости конечной продукции, оцененной по полным затратам труда.* С помощью показателей полной трудоемкости более полно и точно, чем при использовании существующих стоимостных показателей, выявляется структура затрат на выпуск различных видов продукции и прежде всего соотношение между затратами живого и овеществленного труда.

На основе коэффициентов прямой и полной трудоемкости могут быть разработаны межотраслевые и межпродуктовые балансы затрат труда и использования трудовых ресурсов. Схематически эти балансы строятся по общему типу матричных моделей, однако все показатели в них (межотраслевые связи, конечный продукт, условно чистая продукция и др.) выражены в трудовых измерителях.

Пример 4.2. Пусть в дополнение к исходным данным примера 4.1 заданы затраты живого труда (трудовые ресурсы) в трех отраслях: $L_1 = 1160$, $L_2 = 460$, $L_3 = 875$

в некоторых единицах измерения трудовых затрат. Требуется определить коэффициенты прямой и полной трудоемкости и составить межотраслевой баланс затрат труда.

1. Воспользовавшись формулой (4.13) и результатами примера 4.1, находим коэффициенты прямой трудоемкости:

$$t_1 = \frac{1160}{775.3} = 1.5;\ t_2 = \frac{460}{510.1} = 0.9;\ t_3 = \frac{875}{729.6} = 1.2;$$

2. По формуле (4.17), в которой в качестве матрицы B берется матрица коэффициента полных материальных затрат, найденная в примере 4.1, находим коэффициенты полной трудоемкости:

$$T = (1.5, 0.9, 1.2) \cdot \begin{pmatrix} 2.041 & 0.612 & 1.020 \\ 0.816 & 2.245 & 0.408 \\ 0.887 & 0.510 & 1.664 \end{pmatrix} = (4.84, 3.55, 3.92).$$

3. В соответствии с формулой $L_i = t_i \cdot X_i = \sum\limits_{j=1}^{n} x_{ij} \cdot t_i + Y_i \cdot t_i$ умножим первую, вторую и третью строки первого и второго квадрантов межотраслевого материального баланса, построенного в примере 1, на соответствующие коэффициенты прямой трудоемкости,

$$xt_{ij} = x_{ij} \cdot t_i \qquad Yt_i = Y_i \cdot t_i \tag{4.20}$$

получим схему межотраслевого баланса труда (в трудовых измерителях) (табл. 1).

Таблица 4.5. Межотраслевой баланс затрат труда

Производящие отрасли	**Потребляющие отрасли**				
	Межотраслевые затраты овеществленного			*Затраты труда на конечную продукцию*	*Затраты труда в отраслях (трудовые ресурсы)*
	1	2	3		
1	348,9	76,5	437,7	300,0	1163,0 459,1
2	139,6	229,5	0,0	90,0	875,5
3	279,1	61,2	175,1	360,0	

Незначительные расхождения между данными таблицы и исходными данными вызваны погрешностями округления при вычислениях.

4.5.2 Модель фондоемкости продукции

Развитие основной модели межотраслевого баланса достигается также путем включения в нее показателей фондоемкости продукции. В простейшем случае модель дополняется отдельной строкой, в которой указаны в стоимостном выражении объемы производственных фондов Φ_j, занятые в каждой j-й отрасли. На основании этих данных и объемов валовой продукции всех отраслей определяются *коэффициенты прямой фондоемкости (капиталоемкости)* продукции j-й отрасли:

$$f_j = \frac{\Phi_j}{X_j};\ j = 1,...,n. \tag{4.21}$$

Коэффициент прямой фондоемкости показывает величину производственных фондов, непосредственно занятых в производстве данной отрасли, в расчете на единицу ее валовой продукции. В отличие от этого показателя коэффициент полной фондоемкости F_j отражает объем фондов, необходимых во всех отраслях для выпуска единицы конечной продукции j-й отрасли. Если a_{ij} — коэффициент прямых материальных затрат, то для коэффициента полной фондоемкости справедливо равенство, аналогичное равенству (4.14) для коэффициента полной трудоемкости:

$$F_j = \sum_{i=1}^{n} a_{ij} F_i + f_j; \; j = \overline{1, n} \tag{4.22}$$

Если ввести в рассмотрение вектор-строку коэффициентов прямой фондоемкости $f = (f_1, f_2, ..., f_n)$ и вектор-строку коэффициентов полной фондоемкости $F = (F_1, F_2, ..., F_n)$, то систему уравнений (4.22) можно переписать в матричной форме:

$$F = FA + f, \tag{4.23}$$

откуда с помощью преобразований, аналогичных применяемым выше для коэффициентов трудоемкости, можно получить матричное соотношение

$$F = fB, \tag{4.24}$$

где $B = (E - A)^{-1}$ — матрица коэффициентов полных материальных затрат.

Для составления межотраслевого баланса фондоемкости продукции вычисляют в соответствии с формулой $\Phi_i = f_i \cdot X_i = \sum_{j=1}^{n} f_i \cdot x_{ij} + f_i \cdot Y_i$ продуктово – фондовую матрицу xf и стоимость фондов на конечную продукцию Yf

$$xf_{ij} = x_{ij} \cdot f_i \qquad Yf_i = Y_i \cdot f_i \tag{4.25}$$

Пусть Φ совокупный объем фондов по всем отраслям

$$\Phi = \sum_{j=1}^{n} \Phi_j \tag{4.26}$$

Подставляя (4.21) в (4.26), получим $\Phi = \sum_{j=1}^{n} f_j X_j = f \cdot X = f \cdot B \cdot Y = F \cdot Y$ или

$$f \cdot X = F \cdot Y \tag{4.27}$$

Соотношение (4.27) представляет собой основное балансовое равенство в межотраслевом балансе фондов. Его экономическое содержание заключается в том, что стоимость прямых фондозатрат валовой продукции равна стоимости конечной продукции, оцененная по полным фондозатратам.

Для более глубокого анализа необходимо дифференцировать фонды на основные и оборотные, а в пределах основных — на здания, сооружения, производственное оборудование, транспортные средства и т.д.

Пусть в целом все производственные фонды разделены на m групп. Тогда характеристика занятых в народном хозяйстве фондов задается матрицей показателей Φ_{kj}, отражающих объем фондов k-ой группы, занятых в j-й отрасли:

$$\Phi = \begin{pmatrix} \Phi_{11} & \Phi_{12} & \dots & \Phi_{1n} \\ \Phi_{21} & \Phi_{22} & \dots & \Phi_{2n} \\ \dots & \dots & \dots & \dots \\ \Phi_{m1} & \Phi_{m2} & \dots & \Phi_{mn} \end{pmatrix}$$

Коэффициенты прямой фондоемкости также образуют матрицу размерности $m \cdot n$, элементы которой определяют величину производственных фондов k-ой группы, непосредственно используемых при производстве единицы продукции j-й отрасли:

$$f_{kj} = \frac{\Phi_{kj}}{X_j}$$

Для каждой j-й отрасли могут быть вычислены коэффициенты полной фондоемкости F_{kj}, отражающие полную потребность в фондах k-й группы для выпуска единицы конечной продукции этой отрасли:

$$F_{kj} = \sum_{i=1}^{n} a_{ij} F_{ki} + f_{kj}; \; k = 1,\dots,m; \; j = 1,\dots,n$$

Решение систем данных уравнений позволяет представить коэффициенты полной фондоемкости по каждой из m групп фондов как функцию коэффициентов прямой фондоемкости:

$$F_{kj} = \sum_{i=1}^{n} b_{ij} f_{ki}; \; k = 1,\dots,m; \; j = 1,\dots,n$$

В этих формулах величины a_{ij} и b_{ij} — уже известные коэффициенты прямых и полных материальных затрат.

Коэффициенты фондоемкости в межотраслевом балансе позволяют увязать планируемый выпуск продукции с имеющимися производственными мощностями. Так, потребность в функционирующих фондах k-й группы для достижения заданного объема материального производства X_j по всем отраслям задается формулой:

$$\Phi_k = \sum_{j=1}^{n} f_{kj} X_j; \; k = 1,\dots,m.$$

4.6. Динамическая модель межотраслевого баланса

Рассмотренные выше межотраслевые балансовые модели являются *статическими,* т. е. такими, в которых все зависимости отнесены к одному моменту времени. Эти модели могут разрабатываться лишь для отдельно взятых периодов, причем в рамках данных моделей не устанавливается связь с предшествующими или последующими периодами. Капиталовложения вынесены из сферы производства в сферу конечного использования вместе с предметами потребления и непроизводственными затратами, т.е. включены в конечный продукт.

В отличие от статических *динамические* модели призваны отразить не состояние, а процесс развития экономики, установить непосредственную взаимосвязь между предыдущими и последующими этапами развития, и тем самым приблизить анализ на основе экономико-математической модели к реальным условиям развития экономической системы.

В рассматриваемой здесь динамической модели, являющейся развитием статической межотраслевой модели, производственные капитальные вложения выделяются из состава конечной продукции, исследуются их структура и влияние на рост объема производства. В основе построения модели в виде динамической системы уравнений лежит математическая зависимость между величиной капитальных вложений и приростом продукции. Решение системы, как и в случае статической модели, приводит к определению уровней производства, но в динамическом варианте в отличие от статического эти искомые уровни зависят от объемов производства в предшествующих периодах.

Принципиальная схема первых двух квадрантов динамического межотраслевого баланса приведена в табл. 4.6.

Таблица 4.6. Принципиальная схема динамического баланса

Производящие отрасли	*Потребляющие отрасли*									
	Межотраслевые потоки текущих затрат				*Межотраслевые потоки капитальных вложений*				*Конечный продукт*	*Валовой продукт*
	1	2	….	n	1	2	…	n		
1	x_{11}	x_{12}	…	x_{1n}	$\Delta\Phi_{11}$	$\Delta\Phi_{12}$	…	$\Delta\Phi_{1n}$	Y'_1	X_1
2	x_{21}	x_{22}	…	x_{2n}	$\Delta\Phi_{21}$	$\Delta\Phi_{22}$	…	$\Delta\Phi_{2n}$	Y'_2	X_2
…	…	…	…	…	…	…	…	…	…	…
n	x_{n1}	x_{n2}	…	x_{nn}	$\Delta\Phi_{n1}$	$\Delta\Phi_{n2}$	…	$\Delta\Phi_{nn}$	Y'_n	X_n

Модель содержит две матрицы межотраслевых потоков. Матрица текущих производственных затрат с элементами x_{ij} совпадает с соответствующей матрицей статического баланса. Элементы второй матрицы $\Delta\Phi_{ij}$ показывают, какое количество продукции i - й отрасли направлено в текущем периоде в j-ю отрасль в качестве производственных капитальных вложений в ее основные фонды. Материально это выражается в приросте в потребляющих отраслях производственного оборудования, сооружений, производственных площадей, транспортных средств и др.

В статическом балансе потоки капиталовложений не дифференцируются по отраслям-потребителям и отражаются общей величиной в составе конечной продукции Y_i каждой i-й отрасли. В динамической схеме конечный продукт Y'_i включает продукцию i-й отрасли, идущую в личное и общественное потребление, накопление непроизводственной сферы, прирост оборотных фондов, незавершенного строительства, на экспорт. Таким образом, сумма потоков капиталовложений и конечного продукта динамической модели равна конечной продукции статического баланса:

$$\sum_{j=1}^{n} \Delta\Phi_{ij} + Y'_i = Y_i,$$

поэтому уравнение распределения продукции вида (4.2) ($X_i = \sum_{j=1}^{n} x_{ij} + Y_i;\ i = 1,\dots,n$) в динамическом балансе преобразуется в следующее:

$$X_i = \sum_{j=1}^{n} x_{ij} + \sum_{j=1}^{n} \Delta\Phi_{ij} + Y_i'; \; i = 1,...,n \tag{4.28}$$

Межотраслевые потоки текущих затрат можно выразить, как в статической модели, через валовую продукцию отраслей с помощью коэффициентов прямых материальных затрат:

$$x_{ij} = a_{ij} X_j$$

Полагая, что прирост продукции пропорционален приросту производственных фондов, можно записать:

$$\Delta\Phi_{ij} = \varphi_{ij} \Delta X_j; \; i, j = 1,...,n \tag{4.29}$$

Рассмотрим в равенстве (4.29) коэффициенты пропорциональности φ_{ij}. Поскольку

$$\varphi_{ij} = \frac{\Delta\Phi_{ij}}{\Delta X_j},$$

то экономический смысл этих коэффициентов заключается в том, что они показывают, на сколько единиц надо увеличить мощности i-й отрасли для увеличения выпуска продукции j-й отрасли на единицу. Предполагается, что производственные мощности используются полностью и прирост продукции равен приросту мощности. Коэффициенты φ_{ij} называются *коэффициентами вложений, или коэффициентами приростной фондоемкости.*

С помощью коэффициентов прямых материальных затрат и коэффициентов вложений φ_{ij} систему уравнений (1) можно представить в следующем виде:

$$X_i = \sum_{j=1}^{n} a_{ij} X_j + \sum_{j=1}^{n} \varphi_{ij} \Delta X_j + Y_i'; \; i = 1,...,n \tag{4.30}$$

Система (4.30) представляет собой систему линейных разностных уравнений первого порядка. Ее можно привести к обычной системе линейных уравнений, если учесть, что все объемы валовой и конечной продукции относятся к некоторому периоду t, а прирост валовой продукции определен в сравнении с $(t-1)$-м периодом:

$$X_i^{(t)} = \sum_{j=1}^{n} a_{ij} X_j^{(t)} + \sum_{j=1}^{n} \varphi_{ij} (X_j^{(t)} - X_j^{(t-1)}) + Y_i'^{(t)}; \; i = 1,...,n \tag{4.31}$$

Пусть нам известны уровни валовой продукции всех отраслей в предыдущем периоде (величины $X_j^{(t-1)}$) и конечный продукт отраслей в t-м периоде. Тогда очевидно, что соотношения (4.31) представляют собой систему n линейных уравнений с n неизвестными уровнями производства t-го периода. Таким образом, решение динамической системы линейных уравнений позволяет определить выпуск продукции в последующем периоде в зависимости от уровня, достигнутого в предыдущем периоде. Связь между периодами устанавливается через коэффициенты вложений φ_{ij}, характеризующие фондоемкость единицы прироста продукции.

Переходя от дискретного анализа к непрерывному, вместо (4.28) будем иметь:

$$X_i = \sum_{j=1}^{n} x_{ij} + \sum_{j=1}^{n} \frac{d\Phi_{ij}}{dt} + Y_i'; \; i = 1,...,n$$

Выражение (29) в пределе дает:

$$\frac{d\Phi_{ij}}{dt} = \varphi_{ij}\frac{dX_j}{dt} \quad .$$

Окончательно для случая непрерывных изменений получим следующую систему соотношений:

$$X_i = \sum_{j=1}^{n} a_{ij}X_j + \sum_{j=1}^{n} \varphi_{ij}\frac{dX_j}{dt} + Y_i'; \; i = 1,...,n \qquad (4.32)$$

Соотношения (4.32) представляют собой систему n линейных дифференциальных уравнений первого порядка с постоянными коэффициентами. Для ее решения помимо матриц коэффициентов прямых материальных текущих затрат и коэффициентов капитальных затрат (вложений) необходимо знать уровни валового выпуска в начальный момент времени $t = 0$ и закон изменения величины конечного продукта, т.е. вид функций $Y_i'(t)$. На основе этих данных путем решения получившейся задачи Коши для системы дифференциальных уравнений (4.32) можно найти уровни валового выпуска теоретически для любого момента времени. Практически же более или менее достоверное описание валовых и конечных выпусков как функций времени может быть получено лишь для относительно небольших промежутков времени.

В динамической модели особую роль играют коэффициенты приростной фондоемкости φ_{ij}. Они образуют квадратную матрицу n-го порядка

$$(\varphi_{ij}) = \begin{pmatrix} \varphi_{11} & \varphi_{12} & \cdots & \varphi_{1n} \\ \varphi_{21} & \varphi_{22} & \cdots & \varphi_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \varphi_{n1} & \varphi_{n2} & \cdots & \varphi_{nn} \end{pmatrix}$$

каждый j-ый столбец которой характеризует для соответствующей j-и отрасли величину и структуру фондов, необходимых для увеличения выпуска продукции на единицу. Матрица коэффициентов приростной фондоемкости дает значительный материал для экономического анализа и планирования капитальных вложений.

Коэффициенты приростной фондоемкости φ_{ij} определенным образом связаны с валовыми коэффициентами прямой фондоемкости продукции f_{kj}, рассмотренными в предыдущем параграфе. Коэффициенты f_{kj} показывают, сколько всего фондов данного вида приходится на единицу валового выпуска продукции, а коэффициенты φ_{ij} отражают прирост фондов на единицу прироста продукции. Если бы технический прогресс в отраслях производства отсутствовал, то на единицу прироста продукции потребовалось бы столько же новых фондов, сколько их уже занято на единицу выпускаемой продукции, т.е. коэффициенты приростной фондоемкости и валовой прямой фондоемкости были бы равны между собой. Так как новые капитальные вложения производятся на новом более высоком техническом уровне по сравнению с объемом и структурой действующих фондов, то на практике коэффициенты приростной фондоемкости и коэффициенты прямой фондоемкости различаются по величине. Однако между этими двумя группами коэффициентов существует вполне определенная связь, и это используется при разработке динамических моделей, особенно в связи с тем, что достоверные данные о фондоемкости продукции получить легче, чем непосредственно рассчитать коэффициенты вложений.

Вопросы для самопроверки

3. Суть балансового метода исследования социально-экономических систем
4. Принципиальная схема межотраслевого статического баланса. Раскройте экономическое содержание ее разделов элементов. Материальный и стоимостной состав национального дохода
5. Коэффициенты прямых и полных материальных затрат, способы их вычисления.
6. Понятие продуктивности матрицы коэффициентов прямых материальных затрат.
7. Перечислите условия, которым должна удовлетворять продуктивная матрица.
8. Опишите способ агрегирования показателей межотраслевого баланса
9. Основное балансовое равенство в межотраслевом балансе труда
10. Экономический смысл коэффициентов прямой и полной трудоемкости. Описание экономико-математической модели межотраслевого баланса затрат труда.
11. Экономическое содержание коэффициентов прямой и полной фондоемкости. Порядок их расчета на основе экономико-математической модели МОБ.
12. Основное балансовое равенство в межотраслевом балансе фондов.
13. Содержательный смысл принципиальной схемы динамического межотраслевого баланса.
14. Запишите уравнение для валовой продукции в динамической межотраслевой балансовой модели.
15. Основное отличие статической модели межотраслевого баланса от динамической.
16. Экономический смысл коэффициентов приростной фондоемкости

Тема 5. Моделирование финансовых операций

- *Основными элементами финансовых моделей являются время и деньги. Финансовые модели в той или иной мере отражают количественные соотношения между денежными суммами, относящимися к различным моментам времени.*
- *Одна и та же сумма денег в различные моменты времени имеет различную ценность. С другой стороны, по отношению к определенным условиям разные суммы денег в различные моменты времени могут быть равноценными в финансово-экономическом смысле.*
- *Необходимость учёта фактора времени выражается в виде принципа неравноценности денег, относящихся к различным моментам времени.*

5.1. Наращение и дисконтирование

5.1.1 Проценты и процентные ставки

Под процентными деньгами или процентами $I=(S_T-P_0)$ понимают абсолютную величину дохода от предоставления денег в долг на срок T в любой его форме, а именно: выдача денежных ссуд, продажа в кредит, помещение денег на сберегательный счёт, покупка облигаций и т. д. Здесь P_0 – сумма денег, предоставленная в долг, S_T – сумма денег, подлежащая возврату через время T.

Под процентной ставкой i понимают отношение суммы процентных денег, выплачиваемых за единичный промежуток времени (обычно, это год), к величине ссуды:

$$i=\frac{S_1-P_0}{P_0} \tag{5.1}$$

Под учетной ставкой d понимают отношение суммы процентных денег, выплачиваемых за фиксированный промежуток времени, к величине погашаемого долга:

$$d=\frac{S_1-P_0}{S_1} \tag{5.2}$$

Формулы (5.1), (5.2) означают существование двух принципов расчета процентов. Рассмотрим инвестирование суммы P_0 в момент $t=0$ на один период. Как следует из (5.1), в момент $t=1$, т.е. в конце периода, инвестору будет возвращена сумма $S_1=P_0+iP_0$. При этом сумма iP_0, выплачиваемая в момент $t=1$, это проценты $I=S_1-P_0=iP_0$ за время [0, 1] на заем величиной P_0 в момент $t=0$. Таким образом, проценты по ставке i начисляются на сумму первоначального долга P_0 в момент $t=1$.

Согласно (5.2), в обмен на возврат суммы S_1 в момент $t=1$ инвестор даст взаймы сумму $P_0=S_1-dS_1$. В этом случае проценты по ставке d начисляются в начальный момент времени $t=0$ на сумму погашаемого долга S_1. Сумма P_0 может рассматриваться как заем суммы S_1, возвращаемой через единицу времени, при котором проценты величиной dS_1 выплачиваются заранее, в момент $t=0$, и составляют доход кредитора $D=S_1-P_0=dS_1$ за время [0, 1].

Таким образом, проценты по ставке i начисляются в конце периода начисления процентов, а проценты по учетной ставке d – в начале периода начисления процентов.

Существуют различные способы начисления процентов. Соответственно, применяют различные виды процентных ставок. Основное отличие связано с выбором исходной суммы (базы) для начисления. В связи с этим различают **простые и сложные** проценты. При начислении простых процентов базой для начисления служит начальная сумма на протяжении всего срока ссуды. При начислении сложных процентов базой служит сумма с начисленными в предыдущем периоде процентами.

В финансовом анализе процентная ставка применяется не только как инструмент наращения суммы долга, но и в более широком смысле, а именно, как универсальный показатель степени доходности любой финансовой операции.

5.1.2 Наращение по простым процентам

По определению, наращенная сумма – это первоначальная сумма с начисленными на эту сумму процентами. Введем следующие обозначения.

Тогда имеем

$$S = P + I.$$

где P – начальная сумма, I – проценты, начисленные за весь срок.

При начислении простых процентов за базу принимается первоначальная сумма. Проценты начисляются n раз, поэтому $I = P \cdot n \cdot i$ и формула простых процентов запишется в виде

$$S = P \cdot (1 + n \cdot i). \tag{5.3}$$

Величина $(1 + n \cdot i)$ называется **множителем наращения** по простой процентной ставке, т.е. множитель наращения показывает накопленную к моменту n будущую стоимость 1 д.е., вложенной в момент $t = 0$ на срок n.

Процесс увеличения суммы денег в связи с присоединением процентов к первоначальной сумме называется **наращением** или **ростом первоначальной суммы.**

Первоначальная сумма с начисленными на нее процентами называется **наращенной суммой.**

Простые проценты чаще всего используются, когда срок ссуды меньше одного года. Тогда $n = \dfrac{t}{K}$, где t – количество дней ссуды, а K – количество дней в году.

На практике используют обыкновенный или коммерческий процент, когда K=360 дней или точный процент – K *=365* (366) дней.

Пример 5.1. Пусть P=1000 руб., годовая ставка i =10%. Получим наращенные по простым процентам суммы.

1 год: 1000 + 0,1·1000 = 1100 руб; 2 год: 1100 + 100 = 1200 руб; 3 год: 1200 + 100 = =1300 руб. ►

Переменные ставки

Предположим, что весь срок ссуды n разбит на s промежутков длительностью n_t, каждый, $n = \sum_{t=1}^{s} n_t$. В каждом промежутке действует ставка i_t. Тогда формула начисления простых процентов при переменной ставке будет иметь вид

$$S = P \cdot (1 + n_1 \cdot i_1 + n_2 \cdot i_2 + \ldots + n_s \cdot i_s)$$

или

$$S = P \cdot (1 + \sum_{t=1}^{s} n_t \cdot i_t).$$

Пример 5.2. Пусть P=1000 руб., ставка в первый год равна i_1=10%, во второй год – i_2=12%, в третий год – i_3=15%. Получим наращенную за три года по простым процентам сумму.
1000 + 100 + 120 + 150 = 1370 руб. ►

5.1.3. Сложные проценты

Начисление сложных годовых процентов (формула наращения)

В долгосрочных финансовых операциях для наращения первоначальной суммы применяют сложные проценты. При начислении сложных процентов за базу принимают не первоначальную сумму, а сумму, получившуюся после начисления процентов и присоединения их к сумме долга в предыдущих периодах.

Присоединение начисленных процентов к сумме, которая служила базой для их начисления, называют **капитализацией процентов.** Процесс капитализации происходит по следующей схеме:

$$P + P \cdot i = P(1+i) \to P(1+i) + P(1+i)i = P(1+i)^2 \to P(1+i)^2 +$$
$$+P(1+i)^2 \cdot i = P(1+i)^3 \to \ldots$$

В общем виде формула наращения по сложным процентам запишется так:

$$S = P(1+i)^n. \tag{5.4}$$

Множитель $(1+i)^n$ называется множителем **наращения** по формуле сложных процентов.

Пример 5.3. Пусть P=1000, i = 10%, т.е. как доля i = 0,l. Следовательно, наращенные по сложным процентам суммы таковы:

1000,1000+0,1·1000=1000+100=1100, 1100+0,1·1100=1210, 1210+0,1·1210 = 1331,1 и т.д. ►

Пример 5.4. Годовая ставка сложных процентов равна 8%. Через сколько лет начальная сумма удвоится?

Решение. Надо решить неравенство: $(1+0.08)^n$ > 2. Логарифмируем по основанию натуральных логарифмов и получаем $n > \ln(2)/\ln(1.08)$.

Ответ: через 9 лет. ►

Соотношение множителей наращения по простым и сложным процентам

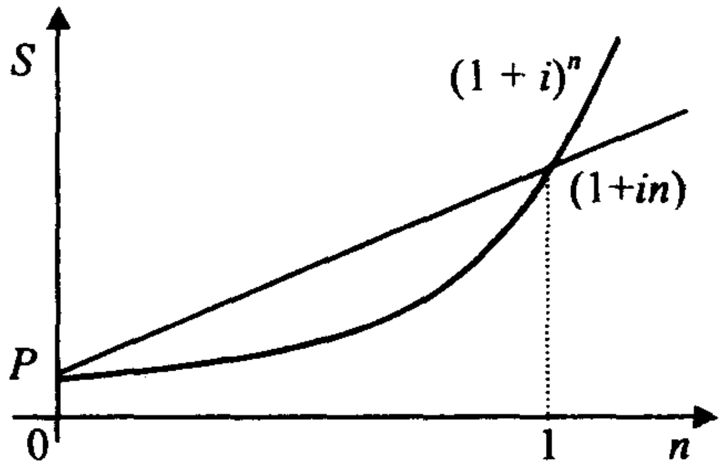


Рис 5.1

Если срок $n<1$, то множитель наращения по простым процентам $(1+n\cdot i)>(1+i)^n$, если $n>1$, то $(1+n\cdot i)<(1+i)^n$. Кривые, изображенные на рис. 5.1, иллюстрируют процессы наращения по простым и сложным процентам в зависимости от срока.

Переменные ставки

Пусть весь срок ссуды n разбит на интервалы длительностью n_t, $t=1,2,...,s,\ n=\sum_{t=1}^{s} n_t$, и в каждом интервале начисление процентов производится по ставкам $i_1, i_2,...,i_s$ соответственно. Тогда формула наращения по сложным процентам будет иметь вид

$$S=P(1+i_1)^{n_1}(1+i_2)^{n_2}\cdots(1+i_s)^{n_s}.$$

5.1.4. Номинальная и эффективная ставки процентов

Номинальная ставка

На практике часто при объявлении условий финансовой операции оговаривается годовая ставка процентов и указывается количество выплат процентов в год (например, может быть ежеквартальное начисление - четыре раза в год). В этом случае используют понятие **номинальной ставки,** обозначим ее символом j. Пусть количество выплат процентов в год равно m. Тогда начисление процентов осуществляется по ставке

j/m, а общее количество интервалов выплат за n лет будет равно $m \cdot n$. Наращенная сумма определяется по формуле

$$S = P(1+\frac{j}{m})^{m \cdot n}. \quad (5.5)$$

Таким образом, номинальная ставка - это годовая ставка процентов при начислении процентов m раз в год. Чем больше число m, тем быстрее идет процесс наращения первоначальной суммы.

Эффективная ставка

Для сравнения различных условий начисления процентов (при различных номинальных ставках и различном количестве начислений) используют понятие **эффективной ставки.** Эффективная ставка - это годовая ставка процентов, начисляемых один раз в год, которая дает тот же финансовый результат, что и m – разовое начисление в год с использованием номинальной ставки j. Таким образом, по определению, должно выполнятся равенство множителей наращения

$$(1+i) = (1+\frac{j}{m})^{m}, \quad (5.6)$$

где i – эффективная ставка. Отсюда получаем

$$i = (1+\frac{j}{m})^{m} - 1. \quad (5.7)$$

Замена в договоре номинальной ставки j при начислении процентов m раз в год на эффективную ставку по формуле (5.7) не меняет финансовых обязательств сторон.

При сравнении различных предлагаемых вариантов начисления процентов достаточно вычислить и сравнить эффективные ставки для каждого из них.

5.1.5. Понятие дисконтирования

Дисконтирование по простым ставкам

Рассмотрим следующую задачу. По заданной сумме S, которую следует уплатить через время n, требуется определить первоначальную сумму P. В этом случае говорят, что сумма S *дисконтируется***.**

Термин *дисконтирование* в широком смысле означает определение стоимости денежной суммы в данный момент времени при условии, что в будущем она составит величину S. Такой расчёт называется *приведением стоимостного показателя к заданному моменту времени***.**

Величину P, найденную дисконтированием суммы S, называют *современной или приведённой величиной* S.

Это одно из важнейших понятий при моделировании и анализе финансовых операций, так как именно с помощью дисконтирования учитывается фактор времени. Решая задачу, получим

$$P = S\frac{1}{1+n\cdot i}. \qquad (5.8)$$

Данная формула называется формулой *математического дисконтирования* (в отличие от банковского дисконтирования или учета, которое здесь не рассматривается).

Величина $\frac{1}{1+n\cdot i}$ называется *дисконтным множителем*, разность $(S-P)$ — *дисконтом* суммы S.

Дисконтирование по сложной ставке процентов

Найдем первоначальную сумму P по известной конечной сумме S с использованием сложной ставки.

Очевидно, что

$$P = \frac{S}{(1+i)^n} = S\cdot v^n, \qquad (5.9)$$

где $v^n = \frac{1}{(1+i)^n}$ — дисконтный множитель (множитель дисконтирования).

Если проценты начисляются m раз в год с использованием номинальной ставки j, то

$$P = \frac{S}{(1+\frac{j}{m})^{m\cdot n}} = S\cdot w^{m\cdot n}, \qquad (5.10)$$

где дисконтный множитель $w^{m\cdot n} = \frac{1}{(1+\frac{j}{m}i)^{m\cdot n}}$.

В финансовой и экономической литературе для современной величины P часто используют сокращенное обозначение - PV, от английского Present Value

Разность $S-P$ называется *дисконтом* суммы S, обозначим эту величину символом D. Из формулы (5.13) и (5.14) следует, что $D=S(1-v^n)$, если проценты начисляются один раз в год, и $D=S(1-w^{m\cdot n})$, если проценты начисляются m раз в год.

5.1.6. Учет инфляции при наращении процентов

Существует множество различных способов учета инфляции при наращении сложных процентов. Рассмотрим один из них, основанный на применении формулы Фишера. Пусть h - ожидаемый годовой темп инфляции в виде ставки сложных процентов (мы не касаемся здесь методики определения этого показателя), i – ставка процентов без учета инфляции, r – реальная ставка с учетом инфляции. Тогда реальная ставка определяется из уравнения, которое называется уравнением Фишера:

$$1+r=\frac{1+i}{1+h}. \qquad (5.11)$$

Решая это уравнение относительно r, получим

$$r=\frac{i-h}{1+h}. \qquad (5.12)$$

Ставка без учета инфляции (которую называют также номинальной ставкой) $i=r+h+r\cdot h$. При малых значениях h используют приближенную формулу $i=r+h$, а для реальной ставки: $r=i-h$.

5.1.7. Эквивалентность простых и сложных процентных ставок

Ставки являются эквивалентными, если в конкретной финансовой операции они приводят к одному и тому же финансовому результату .

Задача определения эквивалентных ставок возникает, например, при сравнении ставок, применяемых в различных сделках и соглашениях, определении эффективности кредитно-финансовых операций, безубыточной замене одного вида ставок другим или одного метода их начисления другим.

Для определения эквивалентных простых и сложных ставок необходимо приравнять соответствующие множители наращения

$$(1+n\cdot i_n)=(1+i)^n$$

где i_n – простая ставка, i – сложная. Из этого соотношения легко получаем

$$i_n=\frac{(1+i)^n-1}{n}, \qquad (5.12)$$

$$i=(1+n\cdot i_n)^{1/n}-1. \qquad (5.13)$$

Если сложная ставка начисляется m раз в год, то также приравнивая множители наращения, получим

$$i_n=\frac{(1+\frac{j}{m})^{m\cdot n}-1}{n}, \qquad (5.14)$$

$$j=m((1+n\cdot i_n)^{1/m\cdot n}-1). \qquad (5.15)$$

Таким образом, замена в любой финансовой операции сложных ставок процентов эквивалентными им простыми по формулам (5.12), (5.14) или простой ставки эквивалентными ей сложными по формулам (5.13), (5.15) не изменяют конечного результата финансовой операции.

Пример 5.5. Какая сумма предпочтительнее при ставке 6%: \$1000 сегодня или \$2000 через 8 лет?

Решение. Найдем современную величину \$2000 через 8 лет при ставке 6%:

$A=2000\cdot(1+0.06)^{-8}=2000\cdot 0.627=1254$.

Итак, $A=1254>1000$. Следовательно, надо предпочесть сумму \$2000 через 8 лет. ►

5.1.8. Дисконтирование и наращение по учетной ставке

Коммерческий (банковский) учет. Сформулируем задачу банковского дисконтирования. По заданной сумме S, которая будет выплачена через n периодов, требуется определить сумму займа P в настоящий момент, при котором проценты за пользование ссудой выплачиваются заранее, в момент предоставления денег в долг $t=0$. Для начисления и удержания процентов применяется учетная ставка d.

1) Простая ставка дисконтирования d.

Дисконтирование проводится по формуле

$$P_0=S_n(1-nd). \qquad (5.16)$$

(5.16) – формула современной величины суммы S_n при банковском ее учете простыми дисконтами по ставке d в течение n периодов. Суммы $D_n,\ D_{n-1},...,D_1$ – дисконты за каждый период (единицу времени). Выражение (5.16) означает, что в обмен на выплату суммы S_n через время n кредитор даст взаймы сумму $S_n(1-nd)$ в начале этого срока. Заметим, что формула (5.26) справедлива, если срок долга n и учетная ставка d удовлетворяют условию $nd < 1$. Дисконтирование по простой учетной ставке применяют, как правило, в случае краткосрочных сделок, когда $0 < n \le 1$.

Пример 5.6. Вексель, погашаемый 1 января 2002 года, учтен за 10 месяцев до его погашения на сумму 180 д.е. Какова величина годовой учетной ставки, если ежемесячный дисконт составляет 2 д.е.?

Так как проценты удерживаются за каждый месяц, то за единицу измерения времени можно принять 1 месяц. Тогда в начале каждого месяца проценты начисляются по ежемесячной учетной ставке $d/12$ где d - годовая учетная ставка. Срок погашения векселя $n = 10$ единиц времени. Сумма P_0 = 180 – приведенная (к моменту учета векселя t = 0) величина суммы S_n, погашаемой по векселю. Дисконты за каждый период (единицу времени) составляют $D_{10} = D_9 = ... = D_1 =$ 2 $= \widehat{D}$. Следовательно, вексель учтен по простой учетной ставке. Размер дисконта за весь срок $D(n) = n\widehat{D}$. Так как $P_0 = S_n - n\widehat{D}$, то сумма, погашаемая по векселю, S_n = 200 д.е. Поскольку $\widehat{D} = S_n \dfrac{d}{12}$, то годовая учетная ставка d = 0,12. ►

2) Сложная ставка дисконтирования d.

Приведем без вывода формулу дисконтирования по сложной учетной ставке

$$P_0 = S_n(1-d)^n. \tag{5.17}$$

(5.17) – формула современной величины суммы S_n при банковском ее учете сложными процентами по учетной ставке d в течение n периодов.

Пример 5.7. Государственная облигация учтена за пять лет до погашения. Какова сумма, погашаемая по облигации, если дисконты за последний и предпоследний годы до погашения составили соответственно 2000 и 1600 д.е. ?

Используем полученные соотношения для сложных дисконтов. Если единицей измерения времени является 1 год, то срок долга n = 5 лет, D_4 = 1600 д.е., D_5 = 2000 д.е., $D_4 = D_5(1-d)$, где d - годовая учетная ставка. Отсюда d = 0,2. Так как $D_5 = dS_5$, то погашаемая сумма S_5 = 10000 д.е. ►

3) Дисконтирование по номинальной учетной ставке.

Если дисконтирование по сложной учетной ставке производится не один, а m раз в году, то годовую учетную ставку называют номинальной и обозначают через g.

Определение. Годовая учетная ставка g называется номинальной, если для дисконтирования в течение $1/m$ части года применяется сложная учетная ставка g/m.

Таким образом, если дисконтирование по сложной учетной ставке производится через равные промежутки времени m раз в году, то в начале каждого периода длиной $1/m$ начисляются и удерживаются проценты по ставке g/m. Если срок долга n лет, то nm – число периодов применения ставки g/m в сроке долга. Из формулы (5.27) получаем

$$P_0 = S_n(1-\frac{g}{m})^{nm}, \tag{5.18}$$

где $m \geq 1$. Если $m = 1$, то $g = d$, т.е. номинальная учетная ставка совпадает с годовой учетной ставкой сложных процентов, применяемой раз в году. (5.28) – формула учета суммы S_n при m-разовом дисконтировании в году по номинальной учетной ставке g в течение n лет.

5.1.9. Наращение по учетной ставке

Если решается задача, обратная банковскому дисконтированию, то для нахождения суммы погашаемого долга пользуются учетной ставкой. Например, в этом возникает необходимость при определении суммы, которую надо проставить в векселе, если задана текущая сумма долга. Из формул (5.26), (5.27), (5.28), находим

$$S_n = \frac{P_0}{1-nd}, \tag{5.19}$$

$$S_n = \frac{P_0}{(1-d)^n}, \tag{5.20}$$

$$S_n = \frac{P_0}{(1-\frac{g}{m})^{nm}}. \tag{5.21}$$

При непрерывном наращении по сложной учетной ставке справедлива формула (5.29) (при $m \to \infty$ номинальные процентные ставки j и g перестают различаться).

Если учетная ставка переменная, то вместо формул (5.19), (5.20) получим

$$S_n = \frac{P_0}{1-\sum_{j=1}^{k} n_j d_j},$$

$$S_n = \frac{P_0}{(1-d_k)^{n_k}(1-d_{k-1})^{n_{k-1}}....(1-d_1)^{n_1}}.$$

5.1.10. Сравнение методов наращения

Все рассмотренные методы наращения приведены в таблице.

Метод наращения	Формула	Множитель наращения
По простой процентной ставке i	$S_n = P_0(1+in)$	$1+in$
По сложной процентной ставке i	$S_n = P_0(1+i)^n$	$(1+i)^n$
По номинальной процентной ставке j	$S_n = P_0(1+\frac{j}{m})^{nm}$	$(1+\frac{j}{m})^{nm}$
По постоянной силе роста δ	$S_n = P_0e^{\delta n}$	$e^{\delta n}$
По номинальной учетной ставке g	$S_n = \dfrac{P_0}{(1-\frac{g}{m})^{nm}}$	$\dfrac{1}{(1-\frac{g}{m})^{nm}}$
По сложной учетной ставке d	$S_n = \dfrac{P_0}{(1-d)^n}$	$\dfrac{1}{(1-d)^n}$
По простой учетной ставке d	$S_n = \dfrac{P_0}{1-nd}$	$\dfrac{1}{1-nd}$

Определение. Число, показывающее во сколько раз наращенная сумма долга больше первоначальной, называется множителем наращения (или множителем накопления).

Экономический смысл множителя наращения заключается в следующем. Если срок долга n единиц времени, то множитель наращения показывает накопленную к моменту n будущую стоимость 1 д.е., вложенной в момент $t = 0$ на срок n. Очевидно, что множитель наращения больше 1. Интенсивность процесса наращения определяется множителем наращения. Сравнивая эти множители для каждого значения срока n, считая равными процентные ставки, можно сравнить темпы наращения по различным ставкам.

5.1.11. Сравнение методов дисконтирования

Все полученные методы дисконтирования показаны в таблице.

Метод дисконтирования	Формула	Дисконтный множитель
По простой учетной ставке d	$P_0 = S_n(1-nd)$	$1-nd$
По сложной учетной ставке d	$P_0 = S_n(1-d)^n$	$(1-d)^n$
По номинальной учетной ставке g	$P_0 = S_n(1-\frac{g}{m})^{nm}$	$(1-\frac{g}{m})^{nm}$
По постоянной силе роста δ	$P_0 = S_ne^{-\delta n}$	$e^{-\delta n}$

По номинальной процентной ставке j	$P_0 = \frac{S_n}{(1+\frac{j}{m})^{nm}}$	$\frac{1}{(1+\frac{j}{m})^{nm}}$
По сложной процентной ставке i	$P_0 = \frac{S_n}{(1+i)^n}$	$\frac{1}{(1+i)^n}$
По простой процентной ставке i	$P_0 = \frac{S_n}{1+in}$	$\frac{1}{1+in}$

Определение. Число, показывающее какую долю от суммы погашаемого долга составляет его современная величина, называется дисконтным множителем.

Экономический смысл дисконтного множителя заключается в следующем. Если срок долга *n* единиц времени, то дисконтный множитель - это современная стоимость 1 д.е., подлежащей выплате через время n. Очевидно, что дисконтный множитель меньше 1. Интенсивность процесса дисконтирования определяется дисконтным множителем. Сравнивая эти множители для каждого значения срока *n*, считая равными процентные ставки за 1 времени, можно сравнить темпы дисконтирования по различным процентным ставкам.

5.2. Потоки платежей, ренты

5.2.1. Основные определения

На практике финансовые операции, как правило, предусматривают распределённые во времени выплаты и поступления денежных сумм.

Потоком платежей будем называть последовательность (ряд) выплат и поступлений, приуроченных к разным моментам времени.

Поток платежей, все элементы которого - положительные величины, а временные интервалы между двумя последовательными платежами постоянны, называют **финансовой рентой** или **аннуитетом** вне зависимости от цели, назначения и происхождения этих платежей.

Финансовая рента описывается следующими параметрами:

1. Член ренты - величина каждого отдельного платежа.
2. Период ренты - временной интервал между платежами.
3. Срок ренты - время от начала ренты до конца её последнего периода.
4. Процентная ставка - это ставка, которая используется при наращении или дисконтировании платежей, из которых состоит рента.

Кроме того, могут быть дополнительные параметры: количество платежей в год, количество начислений процентов в год, моменты выплаты платежей (в начале или в конце периода ренты) и т. д.

Виды финансовых рент

Рента называется **годовой,** если ее период равен одному году.

Рента называется p – срочной, если ее период меньше года и количество платежей в год равно p.

Эти ренты относятся к **дискретным,** поскольку выплаты приурочены к дискретным моментам времени. Бывают ренты **непрерывные,** когда поток платежей описывается непрерывной функцией.

Ренты бывают **постоянные и переменные.**
Рента называется **постоянной,** если все ее платежи одинаковы и не меняются во времени. Если размеры платежей зависят от времени, то это **переменная рента.**

По вероятности уплаты ренты делят на **верные и условные.**

Рента называется **условной,** если выплата очередного платежа ставится в зависимость от наступления некоторого случайного события. Количество выплат в такой ренте заранее не известно. Примерами таких рент могут служить выплаты пенсий, платежи по личному страхованию.

Если условия выплаты платежей по ренте (их размеры, количество и сроки) определены заранее, то такая рента называется **верной.**

Рента называется **ограниченной,** если количество платежей конечно, в противном случае рента называется **бесконечной** или **вечной.** Например, долгосрочное обязательство, когда срок финансовой операции продолжителен и заранее не оговаривается, представляет собой вечную ренту.

Ренты бывают **немедленные и отложенные (отсроченные).** Срок немедленных рент начинается с момента заключения контракта. Если рента отложенная, то срок начала выплат отодвигается на какое-то время.

Если платежи осуществляются в конце периода, то такая рента называется **обычной** или **постнумерандо.** Если выплаты осуществляются в начале периода, то такая рента называется **пренумерандо.**

Обобщающие характеристики потоков платежей

Для анализа потоков платежей необходимо уметь рассчитывать их основные обобщающие характеристики. Таких характеристик две: наращенная сумма и современная величина ренты.

Наращенной суммой потока платежей называют сумму всех последовательных платежей с начисленными на них процентами к концу срока ренты.

Современной величиной потока платежей называют сумму всех платежей, дисконтированных на некоторый момент времени, совпадающий с началом потока платежей или упреждающий его.

Наращенную сумму определяют, например, чтобы знать общую сумму задолженности на какой-либо момент времени, итоговый объём инвестиций, накопленный на момент оценки денежный резерв и т. д.

Современная величина является важнейшим показателем при оценке финансовой эффективности коммерческих сделок, реальных и финансовых инвестиций и т. д.

Рассмотрим задачи определения наращенных сумм для различных видов рент постнумерандо.

5.3. Наращенная сумма потока платежей

5.3.1. Наращенная сумма годовой ренты

При выводе формул для наращенных сумм рент важно правильно установить момент поступления очередного платежа и определить срок (в годах), в течение которого на этот платеж будут начисляться проценты. После этого, чтобы определить размер платежа с начисленными на него процентами за этот срок, достаточно применить формулу начисления сложных процентов. Наиболее просто данная задача решается для годовой ренты с начислением процентов один раз в год. Такая рента называется обычной.

Пусть

S – наращенная сумма ренты;

R – размер отдельного платежа;

i – ставка процентов в виде десятичной дроби;

n –срок ренты в годах.

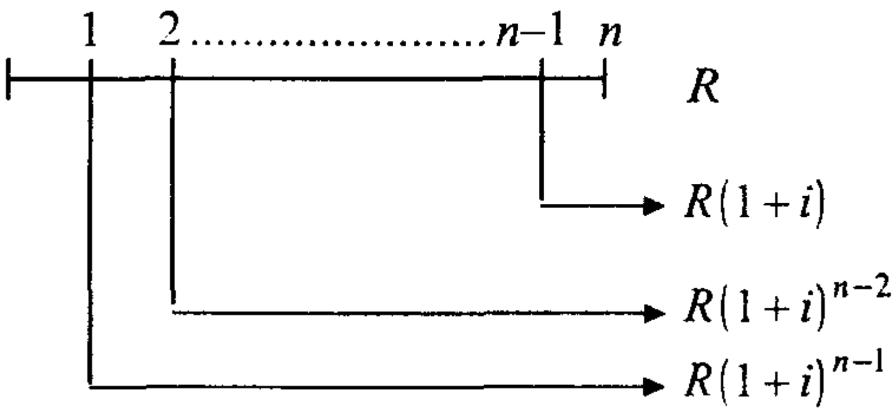


Рис 5.2

На рис. 5.2 изображена диаграмма, иллюстрирующая процесс формирования потока платежей с начисленными на них процентами. На горизонтальной оси указаны моменты поступления платежей, показатели степени в соответствующих формулах – периоды времени, в течение которых на данные платежи начисляются проценты.

Для определения наращенной суммы S необходимо просуммировать получившийся ряд платежей. Заметим, что этот ряд образует геометрическую прогрессию с первым членом R и знаменателем, равным $1+i$.

Формула для суммы n членов геометрической прогрессии имеет вид

$$S_n = a_1(q^n - 1)/(q-1), \tag{5.22}$$

где a_1 – первый член прогрессии, q – знаменатель прогрессии.

Используя эту формулу, получим соотношение для вычисления наращенной суммы годовой ренты:

$$S = R\frac{(1+i)^n - 1}{(1+i)-1} = R\frac{(1+i)^n - 1}{i}.$$

Величина

$$s_{n,i} = \frac{(1+i)^n - 1}{i} \tag{5.23a}$$

называется **коэффициентом наращения годовой ренты**.

С использованием введенного обозначения выражение для наращенной суммы годовой ренты можно записать так:

$$S = Rs_{n,i}.$$

Если платежи поступают в начале периода, то коэффициент наращения ренты равен

$$\hat{s}_{n,i} = (1+i)\cdot\frac{(1+i)^n - 1}{i} = (1+i)\cdot s_{n,i}. \qquad (5.23б)$$

В случае простых процентов наращенная коэффициент наращения годовой ренты равен

$$s_{n,i} = \left[n + i\frac{n\cdot(n-1)}{2}\right]. \qquad \textbf{\textit{(5.24a)}}$$

Для ренты пренумерандо с простой процентной ставкой для коэффициента наращения можно получить

$$\hat{s}(n,i) = \left[n + i\frac{n\cdot(n+1)}{2}\right]. \qquad (5.24б)$$

5.3.2.Наращенная сумма годовой ренты с начислением процентов *m* раз в год

Пусть платежи поступают один раз в год в конце года (то есть интервал ренты равен одному году), а начисление процентов происходит m раз в год. Процесс формирования платежей вместе с начисленными на них процентами изображен на диаграмме (рис.2.2).

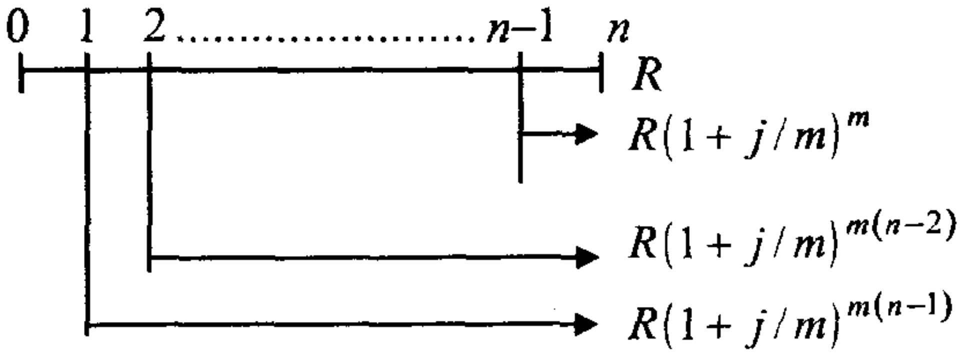


Рис 5.3

При начислении процентов здесь используется формула начисления сложных процентов по ставке j/m.

На основе формулы (5.22) для коэффициента наращения получим

$$.\,s_{n,j/m} = \frac{(1+\frac{j}{m})^{m\cdot n} - 1}{(1+\frac{j}{m})^m - 1}. \qquad (5.25a)$$

Для ренты пренумерандо коэффициент наращения равен

$$\hat{s}_{n,j/m} = (1+\frac{j}{m})^m\frac{(1+\frac{j}{m})^{m\cdot n} - 1}{(1+\frac{j}{m})^m - 1} = (1+\frac{j}{m})^m\cdot s_{n,j/m}. \qquad (5.25б)$$

5.3.3. Наращенная сумма p – срочной ренты

Пусть платежи поступают p раз в год равными суммами, проценты начисляются один раз в год в конце года $(m=1)$.

Если R – годовая сумма, то отдельный платеж равен R/p. Поскольку в год поступает p платежей, то интервал между платежами будет равен $1/p$ лет. Первый платеж поступит в момент времени $1/p$. Процесс формирования платежей с процентами изображен на диаграмме (рис.2.3).

Используя формулу (5.22) для наращенной суммы получим

$$S = \frac{R}{p}\frac{((1+i)^{1/p})^{np}-1}{(1+i)^{1/p}-1} = R\cdot s_{n,i}^{(p)},$$

где $s_{n,i}^{(p)}$ – коэффициент наращения p *-срочной* ренты с начислением процентов один раз в год, равный

$$s_{n,i}^{(p)} = \frac{(1+i)^n-1}{p\cdot\left[(1+i)^{1/p}-1\right]}. \tag{5.26а}$$

Для p – *срочной* ренты пренумерандо имеем

$$\hat{s}_{n,i}^{(p)} = (1+i)^{1/p}\frac{(1+i)^n-1}{p\cdot\left[(1+i)^{1/p}-1\right]} = (1+i)^{1/p}\cdot s_{n,i}^{(p)}. \tag{5.26б}$$

5.3.4. Наращенная сумма p – срочной ренты при начислении процентов m раз в год

Параметры такой ренты: p платежей в год и m раз в год начисление процентов по номинальной ставке j. Рента с такими условиями называется **общей.**

Принцип получения формулы для наращенной суммы аналогичен вышеприведенным случаям.

В результате получим для коэффициента наращения

$$s_{n,j/m}^{(p)} = \frac{(1+\frac{j}{m})^{mn}-1}{p\left[(1+\frac{j}{m})^{\frac{m}{p}}-1\right]} \tag{5.27а}$$

Тогда формула для наращенной суммы примет вид

$$S = Rs_{n,j/m}^{(p)}.$$

Для p – срочной ренты пренумерандо с начислением процентов m раз в год коэффициент наращения равен

$$\hat{s}_{n,j/m}^{(p)} = (1+\frac{j}{m})^{\frac{m}{p}}\frac{(1+\frac{j}{m})^{mn}-1}{p\left[(1+\frac{j}{m})^{\frac{m}{p}}-1\right]} = (1+\frac{j}{m})^{\frac{m}{p}}\cdot s_{n,j/m}^{(p)}. \tag{5.27б}$$

5.4. Современная величина потока платежей

5.4.1. Современная величина годовой ренты

Современная величина ренты является важнейшей характеристикой потока платежей, которая определяет стоимость будущего денежного потока на настоящий момент времени. Эта характеристика служит основой для многих методов финансового анализа. По определению, современная величина - это сумма всех дисконтированных членов потока платежей на начальный или предшествующий ему момент времени. Иногда вместо термина **современная величина** используют термины **приведенная** или **капитализированная** сумма платежей. При определении современной величины потока платежей важно правильно установить период времени от начала потока (момента времени, на который производится оценка) до момента поступления платежа (в годах). После этого можно применять формулы дисконтирования. Обозначим v – множитель дисконтирования, $v = \frac{1}{1+i}$, где i – годовая ставка.

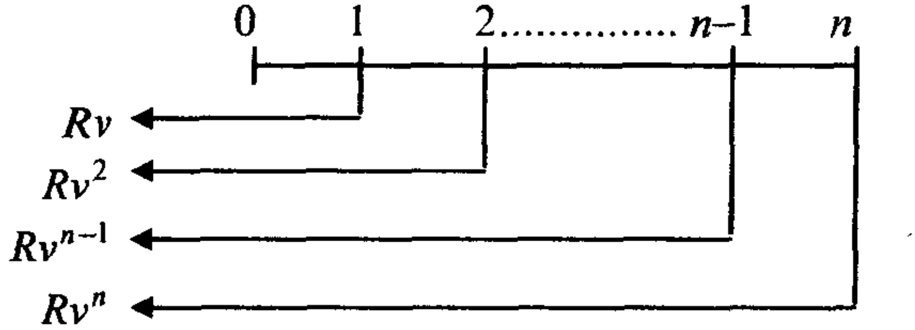


Рис 5.4

Диаграмма на рис.5.4 иллюстрирует процесс формирования потока дисконтированных платежей на начало срока ренты. Показатели степеней – периоды времени в годах от начала ренты до момента поступления платежа.

Чтобы получить современную величину потока, просуммируем все члены получившегося ряда. Используя формулу (5.22) для суммы ряда геометрической прогрессии, получим, что современная величина ренты (обозначим ее символом A) равна

$$A = \frac{Rv(v^n - 1)}{v - 1} = R\frac{1 - v^n}{i} = R\frac{1-(1+i)^{-n}}{i}. \tag{5.28}$$

Величина

$$a_{n,i} = \frac{1-(1+i)^{-n}}{i} \tag{5.29а}$$

называется **коэффициентом приведения годовой ренты**.

С учетом этого выражение для современной величины примет вид

$$A = Ra_{n,i}.$$

Если платежи поступают в начале периода, то коэффициент приведения ренты равен

$$\hat{a}_{n,i} = (1+i)\frac{1-(1+i)^{-n}}{i} = (1+i)\cdot a_{n,i}. \tag{5.29б}$$

В случае простых процентов для ренты постнумерандо мы должны просуммировать поток

$$\frac{R}{1+i} + \frac{R}{1+2\cdot i} + \ldots + \frac{R}{1+n\cdot i} = R\cdot\sum_{k=1}^{n}\frac{1}{1+i\cdot k}$$

или

$$A = R\cdot a_{n,i},$$

где

$$a_{n,i} = \sum_{k=1}^{n} \frac{1}{1+i \cdot k} \tag{5.30а}$$

— коэффициентом приведения годовой ренты по ставке простых процентов.
Для ренты пренумерандо коэффициент приведения равен

$$\hat{a}_{n,i} = \sum_{k=0}^{n-1} \frac{1}{1+i \cdot k}. \tag{5.30б}$$

5.4.2. Современная величина годовой ренты с начислением процентов *m* раз в год

В полученную формулу для современной величины годовой ренты вместо множителя дисконтирования $(1+i)^{-1}$ подставим множитель $(1+\frac{j}{m})^{-m}$. Получим

$$A = R\frac{1-(1+j/m)^{-mn}}{(1+j/m)^{m}-1}.$$

Обозначим

$$a_{n,j/m} = \frac{1-(1+j/m)^{-mn}}{(1+j/m)^{m}-1}, \tag{5.31а}$$

тогда

$$A = Ra_{n,j/m}.$$

Выражение коэффициента приведения ренты пренумерандо имеет вид

$$\hat{a}_{n,j/m} = (1+j/m)^{m}\frac{1-(1+j/m)^{-mn}}{(1+j/m)^{m}-1} = (1+j/m)^{m} \cdot a_{n,j/m}. \tag{5.31б}$$

5.4.3. Современная величина *p* – срочной ренты (*m* = 1)

Интервал между платежами у такой ренты равен $1/p$, размер платежа R/p. Порядок формирования дисконтированных платежей показан на рис. 5.5.

Имеем геометрическую прогрессию с количеством членов, равным np, первый член прогрессии равен $\frac{R}{p}v^{1/p}$, знаменатель прогрессии – $v^{1/p}$. Используя формулу (5.22) для суммы членов геометрической прогрессии, получим

$$A = \frac{R}{p}\frac{v^{\frac{1}{p}}(v^{\frac{1}{p}np}-1)}{v^{\frac{1}{p}}-1} = \frac{R}{p}\frac{(v^{n}-1)}{(1-\frac{1}{v^{1/p}})} = R\frac{1-(1+i)^{-n}}{p\left[(1+i)^{1/p}-1\right]}.$$

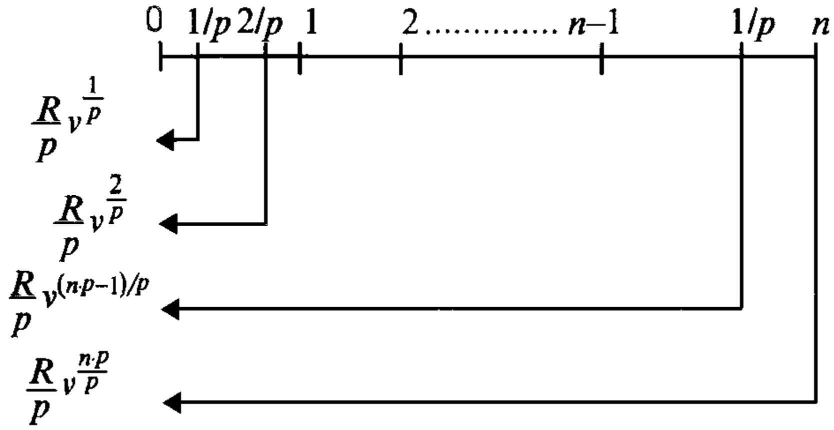


Рис 5.5

Величина

$$a_{n,i}^{(p)} = \frac{1-(1+i)^{-n}}{p\left[(1+i)^{1/p}-1\right]} \tag{5.32а}$$

называется коэффициентом приведения p – срочной ренты.

Тогда формула для расчета современной величины p – срочной ренты будет иметь вид

$$A = Ra_{n,i}^{(p)}.$$

Для ренты пренумерандо процесс дисконтирования платежей показан на рис. 2.7. Коэффициент приведения определяется формулой:

$$\hat{a}_{n,i}^{(p)} = (1+i)^{1/p}\frac{1-(1+i)^{-n}}{p\left[(1+i)^{1/p}-1\right]} = (1+i)^{1/p}\cdot a_{n,i}^{(p)} \tag{5.32б}$$

5.4.4. Современная величина p – срочной ренты при начислении процентов m раз в год

В данном случае коэффициент дисконтирования $v = \dfrac{1}{1+j/m}$. Далее первый дисконтированный платеж равен $\dfrac{R}{p}v^{m/p}$, второй дисконтированный платеж равен $\dfrac{R}{p}v^{2m/p}$, предпоследний дисконтированный платёж равен $\dfrac{R}{p}v^{\frac{(np-1}{p}m}$ и, наконец, последний платёж равен $\dfrac{R}{p}v^{mn}$. Напомним, что показатель степени у множителя дисконтирования – это интервал времени (измеряемый в годах) от начала ренты до момента поступления платежа с учетом m – разового в год начисления процентов. Имеем геометрическую прогрессию с количеством членов np, первым членом $\dfrac{R}{p}v^{m/p}$ и знаменателем прогрессии $v^{m/p}$. Сумма членов этой прогрессии равна

$$A = \frac{R}{p}\frac{v^{m/p}(v^{mn}-1)}{v^{m/p}-1} = \frac{R}{p}\cdot\frac{1-(1+j/m)^{-mn}}{(1+j/m)^{m/p}-1}.$$

Здесь

$$a^{(p)}_{mn,j/m} = \frac{1-(1+j/m)^{-mn}}{p\left[(1+j/m)^{m/p}-1\right]} \tag{5.33а}$$

– коэффициент приведения общей ренты.
Тогда окончательно формула для современной величины данной ренты примет вид

$$A = Ra^{(p)}_{n,j/m}.$$

Если платежи поступают в начале периода (рента пренумерандо), то коэффициент приведения вычисляется по формуле

$$\hat{a}^{(p)}_{mn,j/m} = (1+j/m)^{m/p}\,\frac{1-(1+j/m)^{-mn}}{p\left[(1+j/m)^{m/p}-1\right]} = (1+j/m)^{m/p}\cdot a^{(p)}_{mn,j/m}. \tag{5.33б}$$

Соотношение между наращенной и современной величинами ренты

Пусть A – современная величина годовой ренты на начало срока с начислением процентов один раз в год, S – наращенная сумма этой ренты. Тогда можно показать, что $A(1+i)^n = S$, то есть начисление процентов на сумму A за n периодов даёт наращенную сумму ренты. Действительно

$$A(1+i)^n = R\frac{1-(1+i)^{-n}}{i}(1+i)^n = R\frac{(1+i)^n-1}{i} = Rs_{n,i} = S.$$

Кроме этого, очевидно, имеем еще ряд соотношений:

$$Sv^n = A, \text{ где } v = \frac{1}{1+i},\ a_{n,i}(1+i)^n = s_{n,i},\ s_{n,i}v^n = a_{n,i}.$$

5.5 Доходность финансовой операции

Финансовой называется операция, начало и конец которой имеют денежную оценку – $P(0)$ и $S(T)$ соответственно, а цель проведения которой заключается в максимизации разности $S(T)-P(0)$ или другого подобного показателя. Важнейшей характеристикой операции является ее доходность.

В определении под $P(0)$ понимают реально вложенные средства в момент $t = 0$, под $S(T)$ – реально вырученные денежные средства в результате операции, срок которой T единиц времени. Эффект от вложения естественно измерять в виде процентной ставки наращения, которую в этом случае называют доходностью.

5.5.1. Различные виды доходности операций

Под денежной оценкой начала операции обычно понимают размер вложенных инвестиций, затраты или просто наличный капитал, под денежной оценкой конца операции – наращенный капитал, полученный доход и т.п.

Абсолютная доходность d (доходность за весь срок) операции определяется из уравнения $P(0)(1+d) = S(T)$ или $d = (S(T)-P(0)/P(0) = S(T)/P(0)-1$.

Величина $S(T)/P(0)$ *называется коэффициентом* или *множителем наращения.* Ясно, что $S(T)/P(0)=1+d$.

Определенную выше доходность будем называть еще и *номинальной* или *расчетной,* чтобы отличить от других видов доходности.

Средняя доходность r финансовой операции (доходность за единицу времени) – это ставка простых или сложных процентов, с помощью которой измеряют эффективность финансовой операции.

Согласно определению, доходность финансовой операции за единицу времени - это положительное число r, удовлетворяющее равенству:

$$P(0)(1+r\cdot T)=S(T) \tag{6.1}$$

или

$$P(0)(1+r)^T=S(T). \tag{6.2}$$

Из (6.1) и (6.2) найдем связь между d и r: $1+r\cdot T=1+d$ (для простой ставки) и $(1+r)^T=1+d$ (для сложной ставки).

Если время измеряется в годах, то r - среднегодовая доходность операции.

Таким образом, финансовой операции ставится в соответствие эквивалентная операция наращения суммы $P(0)$ по ставке r в течение времени T. Такой подход позволяет сравнить полученное значение доходности с доходностями по альтернативным вложениям средств.

5.5.2. Учет налогов и инфляции

Налоги и инфляция заметно влияют на эффективность финансовой операции. Рассмотрим учет налогов. Налог начисляется, как правило, на проценты, получаемые при размещении денежной суммы в рост. Предположим, на сумму P_0 в течение времени n начислялись проценты по ставке i, g - ставка налога на проценты. Тогда величина процентов

$$I(n)=S_n-P_0,$$

а сумма налога $G_n=g\cdot I(n)$. Наращенная сумма после выплаты налога составляет

$$S(n)=S_n-G_n.$$

Так как $S(n)<S_n$, то учет налогов фактически сокращает ставку наращения. Итак,

$$S(n)=S_n-G_n=S_n-g\cdot I(n)=S_n-g\cdot(S_n-P_0)=S_n(1-g)+gP_0$$

Если i - простая процентная ставка, то $S_n=P_0(1+i\cdot n)$. Тогда

$$S(n)=P_0(1+i\cdot(1-g)n).$$

Видим, что фактически наращение производится по ставке $i(1-g)<i$.

Если i – сложная процентная ставка, то $S_n=P_0(1+i)^n$. Тогда

$$S(n) = P_0\left((1+i)^n(1-g)+g\right).$$

Пример 6.1. При выдаче кредита на 2 года под годовую сложную процентную ставку 0,08 кредитор удерживает комиссионные в размере 0,5% от суммы кредита. Ставка налога на проценты 10%. Какова доходность операции для кредитора?

Если P_0 - сумма кредита, а S_n - сумма погашаемого долга, то $S_n = P_0(1+i)^n$, где i = 0,08 , n = 2. Сумма комиссионных cP_0, где c = 0,005. Тогда сумма, фактически выданная в долг, составит $P(0) = P_0(1-c)$. После выплаты налога у кредитора останется $S(n) = P_0\left((1+i)^n(1-g)+g\right)$, где g = 0,1 - ставка налога. Уравнение доходности имеет вид $S(n) = P(0)(1+r)^n$. Разрешая это уравнение относительно r, получим

$$r = \left(\frac{S(n)}{P(0)}\right)^{\frac{1}{n}} - 1 = \left(\frac{(1+i)^n(1-g)+g}{1-c}\right)^{\frac{1}{n}} - 1 = 0,07496.$$

Заметим, что без учета налога (g = 0) доходность операции составила бы 0,08271.

►

Инфляция – обесценение денег, проявляющееся в росте цен на товары и услуги, что влечет за собой снижение покупательной способности денег.

Инфляцию характеризуют два количественных показателя – индекс цен и темп инфляции. Предположим, выбрана единица времени. Рассмотрим отрезок времени $[0,t]$, длина которого t единиц времени от начального момента t = 0.

Индекс цен за время $[0,t]$- число

$$J(t) = \frac{K(t)}{K(0)},$$

показывающее во сколько раз выросла стоимость потребительской корзины за период времени $[0,t]$.

Темп инфляции за время $[0,t]$ - число

$$H(t) = \frac{K(t)-K(0)}{K(0)},$$

показывающее на сколько процентов выросла стоимость потребительской корзины за период времени $[0,t]$. Так как $H(t) = \dfrac{K(t)}{K(0)} - 1$, то соотношения между темпом инфляции и индексом цен имеют вид:

$$H(t) = J(t) - 1 \tag{6.3}$$

и

$$J(t) = 1 + H(t) \tag{6.4}$$

для любого периода времени $[0,t]$.

Пусть $[0,t]=\bigcup_{k=1}^{n}[t_{k-1},t_k]$, где $[0,t_1]$, … , $[t_{n-1},t_n]$ - отрезки времени в сроке $[0,t]$ ($t_0=0,\ t_n=t$), длины которых $t_1,(t_2-t_1),...,(t_n-t_{n-1})$ единиц времени; $j(0,t_1),...,j(t_{n-1},t_n)$ и $h(0,t_1),...,h(t_{n-1},t_n)$ – индексы цен и темпы инфляции за периоды $j(0,t_1),...,j(t_{n-1},t_n)$ соответственно.

Пусть j_k и h_k - индекс цен и темп инфляции за 1 единицу времени на временном отрезке $[t_{k-1},t_k]$. Тогда

$$j_k=1+h_k,\quad k=1,2,...,n,$$

а индекс цен за $[t_{k-1},t_k]$ равен

$$j(t_{k-1,}t_k)=j_k^{(t_k-t_{k-1})},\quad k=1,2,...,n.$$

В соответствии с определением индекса цен получим

$$J(t)=j_1^{t_1}\cdot j_2^{(t_2-t_1)}\cdots j_n^{(t_n-t_{n-1})}.$$

Тогда

$$J(t)=(1+h_1)^{t_1}\cdot(1+h_2)^{(t_2-t_1)}\cdots(1+h_n)^{(t_n-t_{n-1})},\tag{6.5}$$

$$1+H(t)=(1+h_1)^{t_1}\cdot(1+h_2)^{(t_2-t_1)}\cdots(1+h_n)^{(t_n-t_{n-1})}.\tag{6.6}$$

Если $h_1=h_2=...=h_n=h$, то

$$J(t)=(1+h)^t\tag{6.7}$$

$$1+H(t)=(1+h)^t.\tag{6.8}$$

Здесь h - темп инфляции за 1 единицу времени на временном отрезке $[0,t]$, $J(t)$ и $H(t)$ - индекс цен и темп инфляции за период времени $[0,t]$.

Предположим, за n единиц времени получена наращенная сумма вклада S_n. Индекс цен за период $[0,n]$ вырос до значения $J(n)$. Тогда реальная сумма вклада вследствие снижения покупательной способности денег составит

$$S(n)=\frac{S_n}{J(n)}.$$

Индекс цен $J(n)$ рассчитывается по одной из приведенных выше формул в зависимости от исходных данных. Так как $J(n)>1$, то $S(n)<S_n$, что означает фактическое снижение ставки наращения.

Пример 6.2. Ожидаемый годовой темп инфляции первых двух лет вклада составляет 3%, а следующих трех - 4%. Какую минимальную годовую ставку сложных процентов должен предложить банк клиенту, чтобы реальная годовая доходность вклада была не меньше 8% ?

Здесь $t = 0$ - момент размещения вклада, 1 год - единица измерения времени, срок вклада n = 5 лет. h_1 = 0,03 и h_2 = 0,04 – среднегодовые темпы инфляции на временных отрезках [0,2], [2,5]. Для доходности по вкладу r должно быть выполнено условие: $r \geq 0.08$. Пусть i - годовая сложная процентная ставка, под которую размещена сумма P_0. Тогда наращенная сумма вклада через n лет $S_n = P_0(1+i)^n$. С учетом инфляции реальная сумма вклада составит $S(n) = \dfrac{S_n}{J(n)}$, где индекс цен согласно (6.8) равен $J(t) = (1+h_1)^2 \cdot (1+h_2)^3$. Уравнение доходности имеет вид: $S(n) = P(0)(1+r)^n$. Разрешая это уравнение относительно r и учитывая требуемое условие для доходности, получим:

$$r = \frac{1+i}{(1+h_1)^{\frac{2}{5}}(1+h_2)^{\frac{3}{5}}} - 1 \geq 0{,}08.$$

Отсюда $i \geq 0{,}11887$. Значит, минимальная процентная ставка размещения вклада составляет 0,11887 против 0,08 без учета инфляции. ►

5.5.3. Поток платежей и его доходность

Пусть $\{R_k, t_k\}$ – поток платежей, в нем t_k – моменты времени, R_k – платежи. Будем говорить, что рассматриваемый поток имеет современную величину A при уровне доходности j, если $\sum\limits_k R_k/(1+j)^{t_k} = A$. Если поток есть годовая рента с годовым платежом R и длительностью n, то рента имеет современную величину A при уровне доходности j, если $R \cdot a_{n,j} = A$. Фиксируем A, тогда при увеличении R доходность ренты увеличивается. Можно сказать и по-другому: для увеличения доходности ренты надо увеличить годовой платеж.

Все эти соображения особенно хорошо видны на примере вечной ренты, поскольку для нее $A = R/j$, или, по-другому: доходность вечной ренты есть $j = R/A$. Важно отметить, что определенная таким образом доходность потока платежей не зависит от ставки процента, а зависит только от величины и моментов самих платежей, в силу чего ее называют *часто внутренней доходностью потока платежей.*

Более точно внутренняя доходность потока платежей есть такая его доходность в только что определенном смысле, при которой современная величина этого потока равна нулю (такая характеристика имеется не у всякого потока платежей). Равенство $A = 0$ возможно только тогда, когда в потоке платежей имеются отрицательные величины.

Пример 6.3. Вексель учтен по ставке i =10% за 160 дней до его оплаты (временная годовая база равна 360 дням). При выполнении операции учета с владельца векселя удержаны комиссионные в размере 0,5% от номинала векселя. Найти доходность операции.

Решение. При расчете доходности векселя его номинал часто не играет роли. Абсолютная доходность операции без учета комиссионных:

$$d = \frac{S}{P} - 1 = \frac{N}{N(1 - i \cdot m)} - 1,$$

где S, P — конечная и начальная стоимость векселя; N – номинал векселя; $m = 160/360$. Подставим исходные данные, получим: d = 0.046, т.е. d = 4.6%.
С учетом комиссионных абсолютная доходность равна:

$$d = \frac{S}{P} - 1 = \frac{N}{N(1 - i \cdot m - 0.005)} - 1 = 5.2\%$$

Эффективность операции (относительная доходность), т.е. доходность в процентах годовых,

$(1{,}046)^{360/160} - 1 = 0{,}106$, т.е. 10,6% —, без учета комиссионных,

$(1{,}052)^{360/160} - 1 = 0{,}1208$, т.е. 12,08% — с учетом комиссионных. ►

5.5.4. Мгновенная доходность

Пусть в момент t капитал равен $K(t)$, а через небольшое время Δt капитал равен $K(t + \Delta t)$, тогда средняя доходность d на отрезке $[t, t + \Delta t]$ в процентах годовых (в долях) равна

$$K(t + \Delta t) / K(t) = (1 + d)^{\Delta t},$$

при малом Δt величина $(1 + d)^{\Delta t}$ с точностью до бесконечно малых 2-го порядка равна $1 + d \cdot \Delta t$. Устремляя Δt к нулю, получаем

$$d = \lim_{\Delta t \to 0} [K(t + \Delta t) - K(t)] / [K(t) \cdot \Delta t] = K'(t) / K(t) = [\ln K(t)]'.$$

Итак, мгновенная доходность есть производная по времени натурального логарифма капитала или, как говорят, логарифмическая производная.

В частности, при постоянной мгновенной доходности d капитал растет во времени по экспоненте: $K(t) = K(0) \cdot e^{d \cdot t}$.

Пример 6.4. Капитал растет во времени с постоянной скоростью v, т.е. $K(t) = K_0(1 + vt)$. Найти мгновенную доходность в произвольный момент времени.

Решение. Обозначим искомую мгновенную доходность $d(t)$, тогда $d(t) = K'(t) / K(t) = K_0 v / K_0(1 + vt) = v/(1 + vt)$. ►

Итак, доходность со временем уменьшается. Это и понятно – приращение капитала за единицу времени постоянно и равно $K_0 v$, а сам капитал растет.

5.6. Кредитные расчеты

Заем, кредит, ссуда — древнейшие финансовые операции. По-латыни «creditum» означает «ссуда»; в слове «кредит» ударение на втором слоге («кредит» с ударением на первом слоге — это правая часть бухгалтерских проводок).

Все три слова — «заем», «кредит», «ссуда» — означают одно и то же - - предоставление денег или товаров в долг на условиях возвратности и, как правило, с уплатой процентов. Тот, кто выдает деньги или товары в кредит, называется кредитор, кто берет -- заемщик (или дебитор). Условия выдачи и погашения кредитов (займов, ссуд) весьма разнообразны. Здесь рассмотрены лишь самые простые и наиболее распространенные способы погашения займов.

5.6.1. Показатель полной доходности финансово-кредитной операции

Доходы от финансово-кредитных операций и различных коммерческих сделок могут иметь разную форму, а именно, это могут быть:

- проценты от выданных ссуд,
- комиссионные,
- доходы от облигаций и других ценных бумаг и т. д.

Как правило, в одной и той же операции предусматривается несколько источников дохода: ссуда приносит кредитору проценты и комиссионные, владелец облигации кроме процентов по облигации получает разницу между выкупной ценой облигации и ценой её приобретения.

В связи с этим возникает проблема измерения эффективности (доходности) операции с учётом всех источников дохода. Обобщённая характеристика доходности должна быть универсальной и применимой к любым видам финансовых операций.

Степень финансовой эффективности подобных операций обычно измеряется в виде годовой ставки сложных процентов. Данную ставку как показатель эффективности (полной доходности, т.е. доходности с учетом всех предусмотренных в операции источников дохода) получают, исходя из общего принципа, а именно: все дисконтированные по искомой ставке доходы (капитализированная величина доходов), предусмотренные в данной операции, приравниваются к приведенным по той же ставке и на тот же момент времени расходам. Из полученного уравнения определяют искомую ставку.

Эта ставка носит различные наименования: в ссудных операциях применяют термин **эффективная процентная ставка,** в анализе доходности облигаций - **доходность на момент погашения** (точнее, **обещанная доходность на момент погашения**), в анализе производственных инвестиций аналогичный показатель называется **внутренней нормой доходности** (внутренняя норма процента).

Основой для расчета полной доходности финансовой операции является соотношение, которое называется **балансом финансово-кредитной операции.**

5.6.2. Баланс финансово-кредитной операции

Необходимым условием любой финансово-кредитной операции является сбалансированность вложений и отдачи. Рассмотрим понятие баланса финансово-кредитной операции.

Пусть выдан кредит в размере K_0 на. срок T, ставка по кредиту равна i. Пусть на протяжении этого срока в счёт погашения задолженности производятся два платежа, размеры которых R_1 и R_2, а в конце срока выплачивается окончательная сумма R_3. Весь срок разбит на три периода длительностью t_1, t_2, t_3. За период времени t_1 задолженность возрастет до величины D_1, (поскольку на сумму кредита начисляются проценты). По истечении этого времени в счет погашения кредита выплачивается сумма R_1 и размер задолженности становится равным K_1. Далее, за время t_2 сумма K_1 возрастет до величины D_2. По истечении промежутка времени t_2 вносится очередная сумма R_2, размер задолженности уменьшается и становится равным сумме K_2. Наконец, за время t_3 размер задолженности возрастает до величины D_3 и по окончании этого промежутка времени, в момент времени *T*, выплачивается сумма R_3. Для сбалансированной операции размер платежа R_3 должен быть таким, чтобы задолженность была погашена. На рис. 7. 1 описанный процесс изображен в виде графика, который называют **контуром финансово-кредитной операции**.

Сбалансированная операция должна иметь замкнутый контур.

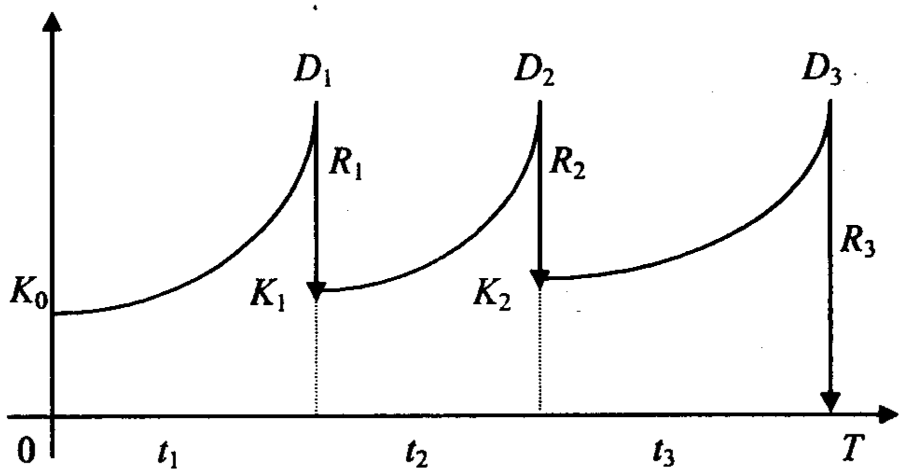


Рис 7.1

Математически процесс погашения задолженности можно описать уравнениями:

$$K_1 = K_0 q^{t_1} - R_1,$$
$$K_2 = K_1 q^{t_2} - R_2,$$
$$K_2 q^{t_3} - R_3 = 0,$$

где $q = (1+i)$ – множитель наращения по сложной ставке.

Последнее уравнение называется балансовым и описывает условие сбалансированности операции.

Определим K_2 через K_0 и подставим результат в балансовое уравнение. Получим

$$K_2 = (K_0 q^{t_1} - R_1) q^{t_2} - R_2,$$
$$[(K_0 q^{t_1} - R_1) q^{t_2} - R_2] q^{t_3} - R_3 = 0,$$
$$K_0 q^T - (R_1 q^{t_2+t_3} + R_2 q^{t_3} + R_3) = 0$$
$$T = \sum_i t_i .$$

Из последнего уравнения видно, что финансово-кредитная операция может быть условно разделена на два встречных процесса:

1) наращение первоначальной задолженности за весь период времени;
2) наращение погашающих платежей за срок от момента платежа до конца операции.

Это уравнение можно преобразовать, умножив его на дисконтный множитель

$$v^T = \left(\frac{1}{1+i}\right)^T .$$

В результате получим

$$K_0 - (R_1 v^{t_1} + R_2 v^{t_1+t_2} + R_3 v^T) = 0 .$$

Таким образом, сумма современных величин погашающих платежей на момент выдачи кредита при полной сбалансированности равна сумме кредита.

Для общего случая n погашающих платежей балансовое уравнение получается аналогично и имеет вид

$$K_0 q^T - \sum_{j=1}^{n} R_j q^{T_j} = 0,$$

где $T_j = \sum_{r=j+1}^{n} t_r$ – время от момента платежа R_j до конца срока ($T_n = 0$).

На основе балансовых уравнений можно измерить доходность финансово-кредитной операции. Для этого нужно составить балансовое уравнение, в котором наращение или дисконтирование производится по неизвестной ставке, характеризующей полную доходность, а затем решить получившееся уравнение относительно искомой ставки.

5.6.3. Определение полной доходности ссудных операций с удержанием комиссионных

Показателем доходности ссудной операции без учёта комиссионных является годовая ставка сложных процентов, эквивалентная процентной ставке, используемой в данной операции. За открытие кредита и другие услуги кредитор часто взимает комиссионные и это заметно повышает доходность операции для него, так как сумма, фактически выданная, сокращается.

Пусть ссуда в сумме D выдана на срок n. При её выдаче удерживаются комиссионные в размере G. То есть, фактически выданная ссуда равна $(D-G)$. Сделка предусматривает начисление простых процентов по ставке i. Ставку полной доходности обозначим i_e. При определении доходности данной операции в виде годовой ставки сложных процентов мы исходим из того, что наращение величины $(D-G)$ по этой ставке должно дать тот же результат, что и наращение величины D по ставке простых процентов i.

Балансовое уравнение для этой операции имеет вид

$$(D-G)(1+i_e)^n = D(1+n \cdot i). \qquad (7.1)$$

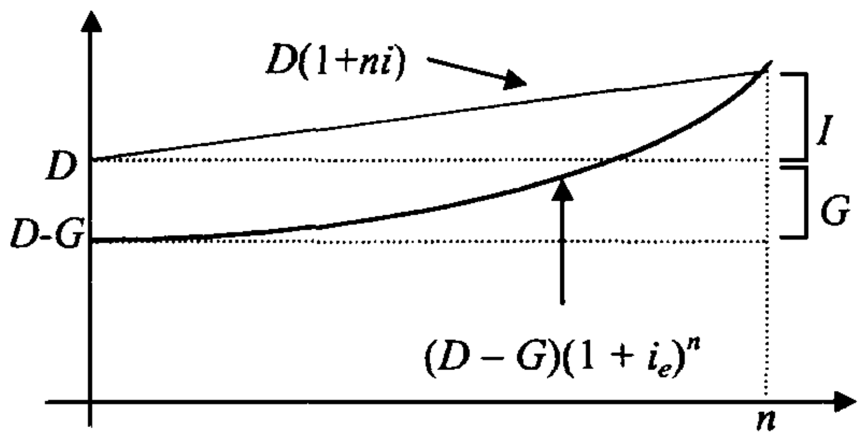


Рис 7.2

Чем больше размер комиссионных, тем больше эффективная ставка (ставка полной доходности). Рис. 7.2 графически иллюстрирует данное уравнение.

Пусть $G = D \cdot g$, где g – процент комиссионных. Тогда, решая уравнение (7.1) относительно i_e получим

$$i_e = \left(\frac{1+n \cdot i}{1-g}\right)^{1/n} - 1.$$

Если сделка предусматривает начисление сложных процентов по ставке i, то вместо уравнения (7.1) имеем $(D-G)(1+i_e)^n = D(1+i)^n$. В этом случае для ставки полной доходности i_e получим

$$i_e = \frac{(1+i)}{(1-g)^{1/n}} - 1.$$

Эффективная ставка i_e непосредственно не фигурирует в условиях операции, она полностью определяется ставкой процента по кредиту и величиной комиссионных.

Ссуды с периодической выплатой процентов

Если комиссионные не выплачиваются, то доходность такой операции равна ставке сложных процентов, эквивалентной любой применяемой в данной сделке ставке.

Предположим, что предусмотрены комиссионные. Пусть ссуда D погашается через n лет, а проценты по ставке i выплачиваются регулярно один раз в конце года. Тогда размер выплачиваемых процентов равен $D \cdot i$. С учётом комиссионных сумма ссуды равна $D(1-g)$. Запишем балансовое уравнение:

$$D(1-g) - (D \cdot i \cdot a_{n,i_e} + Dv_e^n) = 0,$$

где $v_e = \dfrac{1}{1+i_e}$; $a_{n,i_e} = \dfrac{1-(1+i_e)^{-n}}{i_e}$ – коэффициент приведения годовой ренты.

Это уравнение эквивалентно следующему:

$$f(i_e) = v_e^n + i \cdot a_{n,i_e} - (1-g) = 0.$$

Имеем нелинейное уравнение, которое необходимо решить относительно переменной i_e – ставки полной доходности. Точное решение данного уравнения в виде расчетной формулы для i_e получить невозможно. Искомую ставку можно определить приближенно, используя какой-либо численный метод.

Если проценты по кредиту выплачиваются p раз в год, то уравнение для определения доходности операции будет иметь вид

$$f(i_e) = v_e^n + i \cdot a_{n,i_e}^{(p)} - (1-g) = 0,$$

где $a_{n,i_e}^{(p)} = \dfrac{1-(1+i_e)^{-n}}{p[(1+i_e)^{1/p} - 1]}$ – коэффициент приведения p – срочной ренты.

Ссуды с периодическими расходами

Пусть по ссуде периодически выплачиваются проценты и погашается основной долг, причем сумма расходов постоянна. Предполагаем, что платежи производятся один раз в конце года. Пусть R - ежегодная сумма по обслуживанию долга.
Если долг равен D, то из условия сбалансированности операции получим

$$D = Ra_{n,i},$$

где i – сложная ставка процента по кредиту. Отсюда

$$R = \frac{D}{a_{n,i}}.$$

Если комиссионные не выплачиваются, то эффективность такой операции в виде сложной годовой ставки процентов совпадает со ставкой по кредиту i. Если комиссионные выплачиваются, то заемщик на самом деле получает сумму, равную $D(1-g)$. Тогда балансовое уравнение будет иметь вид

$$D(1-g) - Ra_{n,i_e} = 0,$$

или

$$f(i_e) = a_{n,i_e} - a_{n,i}(1-g) = 0.$$

Искомую ставку i_e можно определить, используя численные методы.

Если погасительные платежи осуществляются p раз в год, то в данном уравнении коэффициент приведения годовой ренты $a_{n,i}$, необходимо заменить на коэффициент приведения p-срочной ренты $a_{n,i}^{(p)}$.

Погашение ссуды нерегулярным потоком платежей

До сих пор мы предполагали, что задолженность погашается равными платежами. Рассмотрим случай, когда задолженность погашается путем выплаты нерегулярного потока платежей: $R_1, R_2, ..., R_n$.

Балансовое уравнение в этом случае записывается так:

$$f(i_e) = D(1-g) - \sum_{j=1}^{n} R_j v_e^j = 0,$$

где $v_e = \dfrac{1}{1+i_e}$ - множитель дисконтирования по ставке полной доходности операции.

Величина последнего взноса R_n в данном случае зависит от размеров предыдущих платежей и должна определяться из условия сбалансированности финансовой операции (последний платеж должен погашать задолженность), то есть

$$R_n = Dq^n - \sum_{j=1}^{n-1} R_j q^{T_j},$$

где T_j – срок от выплаты j-го платежа до конца сделки, $q = 1 + i$ – множитель наращения с использованием ставки по кредиту i.

5.6.4. Методы сравнения и анализа коммерческих контрактов

Анализ контрактов на основе метода капитализации платежей

Мы рассмотрим классический подход, основанный на сравнении современных величин всех платежей, предусматриваемых контрактами. Платежи приводятся к одному моменту времени, как правило, к началу срока действия соглашения. Вариант с наименьшей современной величиной с финансовой точки зрения считается предпочтительным для покупателя, при приемлемости всех прочих условий.

При расчете современных величин для сравнения контрактов центральным моментов является выбор уровня ставки процентов, по которой производится дисконтирование. Эту ставку называют **ставкой сравнения**. Чем выше эта ставка, тем более отдаленные платежи оказывают меньшее влияние на современную величину затрат. Таким образом, для инвесторов увеличение ставки сравнения делает предпочтительными контракты с более длительными сроками погашения задолженностей. Рассчитанные по принятой ставке сравнения показатели являются условными, однако установленный таким образом рейтинг контрактов оказывается устойчивым.

Рассмотрим сначала задачу, где конкурентными являются условия погашения задолженности. Пусть цена товара остается постоянной во всех вариантах. В общем случае ставки по кредиту могут быть разными. Чем больше срок кредита, тем больше ставка, поскольку необходимо компенсировать уменьшение стоимости более отдаленных платежей. Каждый вариант контракта оговаривает следующие основные условия погашения задолженности:

1) авансовые платежи, их размер и моменты выплаты;
2) продолжительность и условия выплаты процентов в льготном периоде, если он предусмотрен (в льготном периоде основной долг не погашается, а выплачиваются только проценты по основному долгу);
3) срок погашения задолженности;
4) метод погашения задолженности.

Задача состоит в том, чтобы выписать уравнения для расчета современных величин потоков платежей, предусматриваемых в конкурирующих контрактах. Поскольку в реальной практике может встретиться множество различных вариантов контрактов, то невозможно записать одну единственную формулу, пригодную для всех случаев. Поэтому рассмотрим несколько наиболее характерных вариантов.

Пример 7.1. Судостроительная фирма предлагает уплатить за стоимость заказа 8 млн. долларов. Предлагается два варианта оплаты:

1) 5% при заключении контракта, 5% при спуске судна на воду через 6 месяцев и далее в течение 5 лет равные расходы по обслуживанию долга, то есть остаток долга погашается в течение 5 лет равными суммами.
   Льготный период в данном варианте не предусмотрен.
2) 5% при заключении контракта, 10% при спуске судна на воду через 6 месяцев, предусмотрен льготный период, который начинается после выплаты второго авансового платежа и равен шести месяцам. В льготном периоде выплата процентов, начисленных по сложной ставке, предусматривается в конце этого периода. После окончания льготного периода погашение задолженности происходит в течение 8 лет равными платежами.

Процент за кредит в обоих вариантах i = 10%.

Примем ставку сравнения q = 15%. Это ставка, по которой производится дисконтирование всех платежей, предусмотренных контрактом.

Запишем уравнение для современной величины платежей по первому контракту. В нем предусмотрено два авансовых платежа Q_1 и Q_2, причем второй платеж выплачивается в момент времени t = 0,5 года от начала действия контракта. Следовательно, его необходимо дисконтировать, умножив на множитель v^t, где $v = 1/(1+q)$ - множитель дисконтирования по ставке q. Процесс погашения остатка задолженности $D_1 = P - (Q_1 + Q_2)$ (здесь P – стоимость заказа) начинается с момента выплаты второго авансового платежа. Учитывая, что платежи поступают один раз в год в конце года, поток погашающих долг платежей можно рассматривать как постоянную годовую отложенную на t лет ренту с параметрами R_1, n_1, современная величина которой определяется с использованием ставки, равной ставке сравнения q. Здесь n_1 – период времени, в течение которого погашается основной долг для первого варианта (n_1=5) Необходимо определить приведенную величину этой ренты на начальный момент времени. Современная величина ренты на момент выплаты второго авансового платежа равна $R_1 a_{n_1,q}$, а на начальный момент времени $R_1 a_{n_1,q} v^t$. Размер отдельного погашающего платежа рассчитывается по формуле $R_1 = D_1 / a_{n_1,i}$. Обратим внимание, что при расчете R_1, используется ставка по кредиту i.

Таким образом, математически процесс погашения задолженности (современная величина расходов) для первого варианта описывается уравнением

$$A_1 = Q_1 + Q_2 v^t + R_1 a_{n_1,q} v^t.$$

Рассмотрим второй вариант. В этом варианте предусмотрен льготный период. Следовательно, в уравнении для современной величины платежей необходимо учесть проценты, выплачиваемые в конце этого периода. Пусть L – продолжительность льготного периода, $D_2 = P - (Q_3 + Q_4)$, здесь Q_3, Q_4 – авансовые платежи. Тогда сумма процентов в льготном периоде равна $D_2[(1+i)^L - 1]$. Чтобы привести этот платеж к нулевому моменту, его необходимо умножить на множитель дисконтирования v^{t+L}. Погашение основного долга начинается после окончания льготного периода. Поток погашающих долг платежей - отложенная на $t+L$ лет рента. Современная величина дисконтированных на начальный момент погашающих платежей равна $R_2 a_{n_2,q} v^{t+L}$, где $R_2 = D_2 / a_{n_2,i}$. Уравнение для определения современной величины платежей по второму варианту формируется, как в первом варианте и имеет вид

$$A_2 = Q_3 + Q_4 v^t + D_2[(1+i)^L - 1]v^{t+L} + R_2 a_{n_2 q} v^{t+L},$$

здесь n_2 – период времени, в течение которого погашается основной долг для второго варианта (n_2=8).

Проведя вычисления, получим: A_1 = \$ 6 710 000, A_2 = \$ 6 382 000.

Таким образом, согласно описанному методу, второй вариант оказался выгоднее, поскольку $A_2 < A_1$. ►

Получившиеся суммы условны и используются только для установления рейтинга контрактов. Можно показать, что если современная величина платежей по одному из сравниваемых контрактов больше, чем по другому, то такое же соотношение сохраняется и для других значений ставки сравнения, в случае, если они превышают наибольшую из ставок сравниваемых контрактов или если ставки сравнения меньше наименьшей из этих ставок. Другими словами, если для каких-либо двух контрактов $A_2 < A_1$, причем в первом контракте ставка по кредиту i_1, а во втором i_2, $i_1 > i_2$ при некоторой ставке сравнения q, то это же соотношение между A_2 и A_1 сохранится для всех других значений q, таких, что $q > i_1$ или для всех значений $q < i_2$.

Рассмотрим еще один типичный вариант контракта. Пусть условия контракта следующие: аванс в сумме Q выплачивается один раз в начале сделки, предусматривается разовая поставка товара спустя период времени t от момента заключения контракта, предусмотрен льготный период продолжительностью L, который начинается с момента поставки товара, проценты в льготном периоде выплачиваются периодически один раз в конце года, погашение долга осуществляется равными выплатами, причем процесс погашения начинается с момента окончания льготного периода, общий срок контракта $t+L+n$ лет.

Учитывая все эти условия, запишем уравнение, определяющее современную величину платежей по контракту:

$$A = Q + (P-Q)\left(\frac{a_{n,q}}{a_{n,i}} v^{t+L} + i \cdot a_{L,q} v^{t}\right)$$

В данном уравнении сумма ежегодных процентных платежей в льготном периоде равна $(P-Q)i$. Поскольку проценты выплачиваются периодически начиная с момента поставки товара, поток этих платежей можно рассматривать как отложенную на t лет годовую ренту с параметрами $(P-Q)i, L, q$. Размер погашающих платежей равен $(P-Q)/a_{n,i}$. Поток этих платежей - отложенная на $t+L$ лет годовая рента.

Метод сравнения на основе определения предельных значений параметров контрактов

Задача сравнения контрактов может быть решена путём определения предельного (критического) значения цены или процентной ставки одного из двух сравниваемых контрактов.

Предельным значением параметра называется его величина, при которой сравниваемый контракт оказывается в финансовом отношении эквивалентным другому, который выбран в качестве базового при сохранении других условий.

Такой подход к анализу контракта покупатель может применить при определении допустимых значений цены или процентной ставки, когда есть возможность вести переговоры об изменении условий одного из сравниваемых контрактов. Предположим, что в контрактах предусматриваются разовые расчёты за товар в конце срока сделки без авансовых платежей, Пусть параметры первого контракта: P_1 – цена, i_1 – процентная ставка, n_1 – срок оплаты. Соответствующие параметры второго контракта – P_2, i_2, n_2. Рассмотрим две задачи.

1. Найти предельное значение ставки i_2 (значения P_2, n_2 фиксированы), при котором второй вариант будет эквивалентен первому.

2. Определить цену P_2 (i_2, n_2 фиксированы), при которой второй вариант эквивалентен первому,

Из условия равенства современных величин расходов по двум вариантам контракта можно записать:

$$P_1\left(\frac{1+i_1}{1+q}\right)^{n_1} = P_2\left(\frac{1+i_2}{1+q}\right)^{n_2}.$$

Решая первую задачу, из этого уравнения находим предельное значение ставки i_2^*:

$$i_2^* = (1+q)\left[\frac{P_1}{P_2}\left(\frac{1+i_1}{1+q}\right)^{n_1}\right]^{1/n_2} - 1.$$

Правило выбора с точки зрения покупателя: условия второго варианта будут хуже условий первого, если ставки $i_2 > i_2^*$; будут равноценны, если $i_2 = i_2^*$; будут лучше, если $i_2 < i_2^*$.

Решая вторую задачу, находим предельное значение цены P_2^*:

$$P_2^* = P_1 \frac{(1+i_1)^{n_1}}{(1+i_2)^{n_2}(1+q)^{n_1-n_2}}.$$

Правило выбора; второе соглашение эквивалентно первому, если $P_2 = P_2^*$, выгоднее для покупателя, если $P_2 < P_2^*$.

Очевидно, что i_2^* и P_2^* существенно зависят от сроков кредитования и принятой ставки сравнения q.

Если сроки одинаковы ($n_1 = n_2 = n$), то i_2^* и P_2^* не зависят от ставки q. Тогда

$$i_2^* = (1+i_1)\left(\frac{P_1}{P_2}\right)^{1/n} - 1,\; P_2^* = P_1\left(\frac{1+i_1}{1+i_2}\right)^n.$$

Мы рассмотрели случай, когда погашение кредита производится разовым платежом в конце срока. Предположим теперь, что кредит погашается равными периодическими выплатами. Тогда предельные процентные ставки по кредиту определяются из условия равенства современных величин потоков платежей, предусматриваемых контрактами:

$$P_1 \frac{a_{n_1,q}}{a_{n_1,i_1}} = P_2 \frac{a_{n_2,q}}{a_{n_2,i_2}}$$

Преобразуя это уравнение, получим

$$a_{n_2,i_2^*} = \frac{1-(1+i_2^*)^{-n_2}}{i_2^*} = \frac{P_2}{P_1}\frac{a_{n_2,q}}{a_{n_1,q}} a_{n_1,i_1}.$$

Ставку i_2^* определим, решая данное уравнение численно. Предельное значение цены определяется соотношением

$$P_2^* = P_1 \frac{a_{n_2,i_2}}{a_{n_1,i_1}}\frac{a_{n_1,q}}{a_{n_2,q}}.$$

5.6.5. Планирование погашения долгосрочной задолженности

Расходы по обслуживанию долга

Определение периодических расходов, связанных с займом, называют обслуживанием долга. Разовую сумму по обслуживанию долга называют **срочной уплатой.** Срочные уплаты включают:

- текущие процентные платежи;
- средства, предназначенные для погашения (амортизации) основного долга.

Таким образом, срочная уплата равна $Y = R + I$, где R - годовые расходы по погашению основной суммы долга, I - проценты по займу.

Обычно проценты выплачиваются на протяжении всего срока займа периодически. Иногда они начисляются и присоединяются к основной сумме долга. Основная сумма чаще всего погашается частями, причём размеры платежей по погашению (срочная уплата) могут быть одинаковы или могут изменяться.

В льготном периоде основная сумма долга не погашается, проценты могут выплачиваться периодическими платежами или в конце периода одним платежом, или вообще не выплачиваться, а присоединяться к основной сумме долга.

Методы расчёта существенно зависят от вида займа. Основной признак классификации займов - метод погашения займа. Различают следующие виды займов:

1. Займы без обязательного погашения («вечные облигации») - заёмщик обязуется выплачивать кредитору в определённые сроки доход в виде фиксированного процента, занятая сумма не возвращается. Заёмщик оставляет за собой право выкупить или погасить все выпущенные долговые обязательства в любое время.
2. Займы с обязательным погашением в один срок. Заёмщик возвращает занятую сумму в оговоренный срок и выплачивает проценты периодически или в конце срока.
3. Займы с обязательным погашением в несколько сроков. Заёмщик возвращает занятую сумму по частям и регулярно выплачивает доход от займа в виде процентов.

Погашение основного долга в рассрочку равными платежами

Один из способов погашения долга в рассрочку - погашение равными суммами.

Пусть заём D погашается в течение n лет равными платежами. Тогда сумма, ежегодно идущая на погашение долга, равна d, где $d = D/n$. Помимо погашения основного долга должник выплачивает проценты на *остаток* долга. Будем предполагать, что проценты выплачиваются один раз в конце года по ставке g. Тогда первая уплата процентов в конце первого года равна Dg. В конце второго года сумма процентов равна $(D - D/n)g$. В конце третьего года – $(D - 2(D/n))g$ и так далее.

Платежи по срочным уплатам формируются следующим образом:

Первая уплата: $Y_1 = D(1/n + g)$.

Вторая уплата: $Y_2 = D/n + (D - D/n)g$.

Третья уплата: $Y_3 = D/n + (D - 2(D/n))g$.

Срочная уплата на произвольный момент времени t:

$$Y_t = D_t g + d,$$

где D_t – остаток долга на начало года t $(t = 1, 2, \ldots, n)$, определяемый следующим рекуррентным соотношением:

$$D_{t+1} = D_t - D/n, \; D_1 = D.$$

Погашение долга в рассрочку равными срочными уплатами

На протяжении всего срока погашения регулярно выплачивается постоянная срочная уплата, часть её идёт в погашение долга, а часть - в погашение процентов за заём. Величина долга убывает после каждой выплаты. Однако в связи с уменьшением выплат по процентам с течением времени увеличиваются суммы, идущие на погашение основного долга. Такой способ называют **прогрессивным погашением.** Срочная уплата

$$Y = D_t g + d_t = const.$$

Периодически выплачиваемые суммы Y можно рассматривать как постоянную годовую ренту, член которой определяется по формуле:

$$Y = \frac{D_1}{a_{n,g}}.$$

Определим структуру срочной уплаты, то есть ту ее часть, которая идёт на погашение основного долга, и часть, которая идет на погашение процентов.

Размер первого погасительного платежа $d_1 = Y - D_1 g$.

Так как $D_1 = Ya_{n,g}$, то

$$d_1 = Y(1 - g \cdot a_{n,g}) = Yv^n \text{, где } v = \frac{1}{1+g}$$

Далее

$d_2 = Y - D_2 g = Y - (D_1 - d_1)g$ (здесь учтено, что $D_2 = D_1 - d_1$) или

$$d_2 = d_1 g + Y - D_1 g = d_1 g + Y(1 - a_{n,g} g) = Yv^n g + Yv^n = Yv^{n-1}$$
$$= d_1(1+g).$$

Для платежа в момент времени t имеем

$$d_t = Yv^{n-(t-1)} \text{, или } d_t = d_1(1+g)^{t-1}.$$

Таким образом, можно выписать следующую систему соотношений.

Размер погасительного платежа в году t:

$$d_t = Y - D_t g = d_{t-1}(1+g),\ t = 1,2,...,n;$$

$$d_1 = Y - D_1 g = Yv^n = D_1 v^n / a_{n,g} = D_1 / s_{n,g}.$$

Остаток долга на начало года $t+1$:

$$D_{t+1} = D_t - d_t = D_t(1+g) - Y.$$

Сумма погашенного долга на начало года $t+1$:

$$W_t = \sum_{k=0}^{t-1} d_1(1+g)^k = d_1 s_{t,g}$$

где $s_{t,g}$ - коэффициент наращения годовой ренты, срок которой равен t лет.

Если погасительные платежи и проценты выплачиваются p раз в год, то данная задача решается аналогично, только применяются соответствующие формулы для p–срочных рент.

Вопросы для самопроверки

1. Что понимают под процентами в финансовых операциях?
2. Что такое процентная ставка? Что называют нормированной процентной ставкой?
3. Что такое учетная ставка? Дайте экономическую интерпретацию.
4. Что такое наращение?
5. Запишите формулу наращения по простым процентам с постоянной ставкой.
6. Запишите формулу наращения по простым процентам с переменной ставкой.
7. Запишите формулу наращения для сложных процентов с постоянной ставкой.
8. Запишите формулу наращения для сложных процентов с переменной ставкой.
9. Что такое номинальная ставка и эффективная процентная ставка? Связь эффективной и номинальной ставки.
10. Что такое дисконтирование? Запишите формулу дисконтирования по простым процентам
11. Что такое современная величина суммы S? Экономический смысл этого понятия.
12. Запишите формулу для современной величины при начислении процентов один раз в год и при начислении процентов m раз в год
13. Как осуществляется учет инфляции при наращении процентов?

14. Что значит эквивалентность простых и сложных ставок?
15. Как производится дисконтирование по простой учетной ставке?
16. Как производится дисконтирование по переменной простой учетной ставке?
17. Как производится дисконтирование по переменной сложной учетной ставке?
18. Запишите формулу наращения по простой учетной ставке
19. Запишите формулу наращения по сложной учетной ставке
20. Запишите формулу наращения по номинальной учетной ставке
21. Запишите формулу наращения по переменной простой учетной ставке
22. Запишите формулу наращения по переменной сложной учетной ставке
23. Что такое финансовая рента? Какими параметрами она описывается?
24. Перечислите виды финансовых рент.
25. Дайте определение наращенной суммы и современной величины потока платежей.
26. Запишите формулу для наращенной суммы потока платежей годовой ренты. постнумерандо при начислении процентов по простой ставке.
27. Запишит формулу для наращенной суммы потока платежей годовой ренты. пренумерандо при начислении процентов по простой ставке
28. Запишит формулу для наращенной суммы потока платежей годовой ренты постнумерандо при начислении процентов по сложной ставке.
29. Запишит формулу для наращенной суммы потока платежей годовой ренты пренумерандо при начислении процентов по сложной ставке.
30. Запишите формулу для наращенной суммы потока платежей при начислении процентов m раз в год для ренты постнумерандо
31. Запишите формулу для наращенной суммы потока платежей при начислении процентов m раз в год для ренты пренумерандо
32. Запишите формулу для наращенной суммы p - срочной ренты постнумерандо
33. Запишите формулу для наращенной суммы p - срочной ренты пренумерандо
34. Запишите формулу для наращенной суммы p - срочной ренты постнумерандо при начислении процентов m раз в год
35. Запишите формулу для наращенной суммы p - срочной ренты пренумерандо при начислении процентов m раз в год
36. Запишите формулу для современной величины потока платежей годовой ренты постнумерандо по сложной ставке.
37. Запишите формулу для современной величины потока платежей годовой ренты пренумерандо по сложной ставке.
38. Запишите формулу для современной величины потока платежей годовой ренты постнумерандо по простой ставке.
39. Запишите формулу для современной величины потока платежей годовой ренты пренумерандо по простой ставке.
40. Запишите формулу для современной величины потока платежей при начислении процентов m раз в год для ренты постнумерандо
41. Запишите формулу для современной величины потока платежей при начислении процентов m раз в год для ренты пренумерандо
42. Запишите формулу для современной величины p - срочной ренты постнумерандо
43. Запишите формулу для современной величины p - срочной ренты пренумерандо
44. Запишите формулу для современной величины p - срочной ренты при начислении процентов m раз в год постнумерандо
45. Запишите формулу для современной величины p - срочной ренты при начислении процентов m раз в год пренумерандо
46. Определите зависимость между современной и наращенной величинами ренты

47. Что такое финансовая операция? Какие виды доходностей финансовой операции существуют и как они определяются?
48. Как определяется темп инфляции и индекс цен? Как влияют на доходность операции налоги и инфляция?
49. Как определяется доходность потока платежей?
50. Как определяется мгновенная доходность?
51. Какая характеристика используется для оценки доходности финансово - кредитной операции?
52. Что такое баланс кредитной операции? Запишите уравнения баланса.
53. Как определить полную доходность финансово кредитной операции без удержания комиссионных?
54. Как определить полную доходность ссудной операции с удержания комиссионных?
55. Как определить полную доходность ссудной операции с периодической выплатой процентов?
56. Как определить полную доходность ссудной операции с периодическими Расходами?
57. Как определить полную доходность ссудной операции с нерегулярным потоком платежей?
58. В чем состоит суть метода сравнения различных контрактов на основе капитализации платежей?
59. В чем состоит суть метода сравнения различных контрактов на основе определения предельных значений параметров контрактов?
60. Назовите основные цели анализа долгосрочной задолженности
61. Как классифицируются займы по способу их погашения?
62. Как погашается долг в рассрочку равными платежами?
63. Как погашается долг в рассрочку равными уплатами?

Тема 6. Математическое и компьютерное моделирование

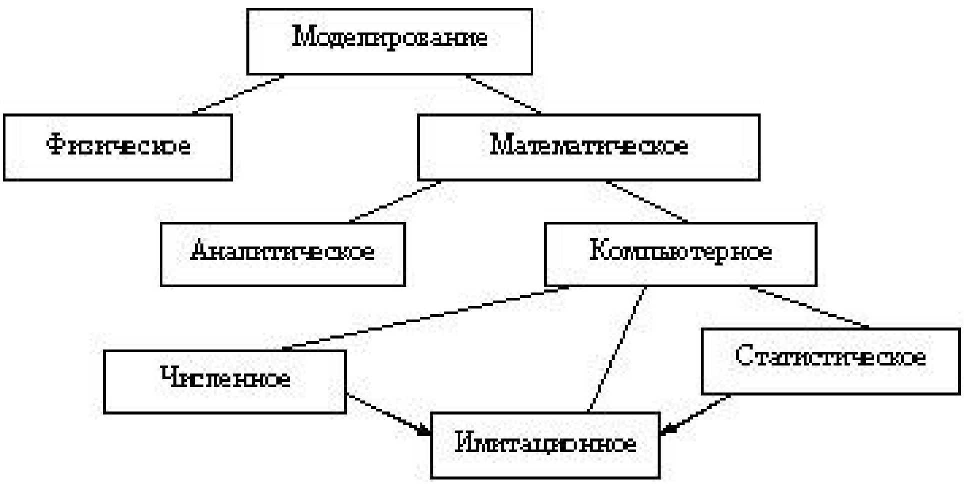
6.1. Классификация видов моделирования

В настоящее время нельзя назвать область человеческой деятельности, в которой в той или иной степени не использовались бы методы моделирования. Особенно это относится к сфере управления различными системами, где основными являются процессы принятия решений на основе получаемой информации.

Как известно, математическая модель некоторого явления или процесса может быть представлена функциональной зависимостью между совокупностью входных (независимых) переменных $x_i \, (i = 1,...,n)$ и одной или несколькими выходными (зависимыми) переменными y:

$$y = f(x_1,...,x_n)\,. \tag{6.1}$$

Различают следующие виды моделирования (см. рис. 6.1)



Физическое – используется сама система, либо подобная ей (летательный аппарат в аэродинамической трубе).

Математическое – процесс установления соответствия реальной <u>*системе*</u> <u>*S*</u> математической модели М и исследование этой модели, позволяющее получить характеристики реальной системы.

Аналитическое – процессы функционирования элементов записываются в виде явных математических соотношений (алгебраических, интегральных, дифференциальных, логических и т.д.). Аналитическая модель может быть исследована методами: а) *аналитическим* (устанавливаются явные зависимости, получаются, в основном, аналитические решения); б) *численным* (получаются приближенные решения); в) *качественным* (в явном виде можно найти некоторые свойства решения). Получить эти зависимости удается только для сравнительно простых реальных процессов и систем (РПС). В результате аналитическая модель становится слишком грубым приближением к действительности.

Компьютерное – математическое моделирование формулируется в виде алгоритма (программы для ЭВМ), что позволяет проводить над ней вычислительные эксперименты.

Численное – используются методы вычислительной математики (отличается от численного аналитического тем, что возможно задание различных параметров модели).

Статистическое – обработка данных о системе (модели) с целью получения статистических характеристик системы.

Имитационное – воспроизведение на ЭВМ (имитация) процесса функционирования исследуемой системы, соблюдая логическую и временную последовательность протекания процессов, что позволяет узнать данные о состоянии системы или отдельных ее элементов в определенные моменты времени.

Имитационное моделирование - это совокупность методов алгоритмизации функционирования объектов исследований, программной реализации алгоритмических описаний, организации, планирования и выполнения на ЭВМ вычислительных экспериментов с математическими моделями, имитирующими функционирование реальных процессов и систем в течении заданного периода.

Под алгоритмизацией функционирования реальных процессов и систем понимается пооперационное описание работы всех ее функциональных подсистем отдельных модулей с уровнем детализации, соответствующем комплексу требований к модели.

Одним из видов имитационного моделирования является статистическое имитационное моделирование, позволяющее воспроизводить на ЭВМ функционирование сложных случайных процессов.

При исследовании сложных систем, подверженных случайным возмущениям используются вероятностные аналитические модели и вероятностные имитационные модели.

В вероятностных аналитических моделях влияние случайных факторов учитывается с помощью задания вероятностных характеристик случайных процессов (законы распределения вероятностей, спектральные плотности или корреляционные функции). При этом построение вероятностных аналитических моделей представляет собой сложную вычислительную задачу. Поэтому вероятностное аналитическое моделирование используют для изучения сравнительно простых систем.

В вероятностном имитационном моделировании оперируют не с характеристиками случайных процессов, а с конкретными случайными числовыми значениями параметров ПС. При этом результаты, полученные при воспроизведении на имитационной модели рассматриваемого процесса, являются случайными реализациями. Поэтому для нахождения объективных и устойчивых характеристик процесса требуется его многократное воспроизведение, с последующей статистической обработкой полученных данных. Именно поэтому исследование сложных процессов и систем, подверженных случайным возмущениям, с помощью имитационного моделирования принято называть статистическим моделированием.

Статистическая модель случайного процесса - это алгоритм, с помощью которого имитируют работу сложной системы, подверженной случайным возмущениям; имитируют взаимодействие элементов системы, носящих вероятностный характер.

При реализации на ЭВМ статистического имитационного моделирования возникает задача получения на ЭВМ случайных числовых последовательностей с заданными вероятностными характеристиками. Численный метод, решающий задачу генерирования последовательности случайных чисел с заданными законами распределения, получил название "метод статистических испытаний" или "метод Монте-Карло".

Статистическое моделирование - это способ изучения сложных процессов и систем, подверженных случайным возмущениям, с помощью имитационных моделей.

Метод Монте-Карло - это численный метод, моделирующий на ЭВМ псевдослучайные числовые последовательности с заданными вероятностными характеристиками.

Методика статистического моделирования состоит из следующих этапов:

1. Моделирование на ЭВМ псевдослучайных последовательностей с заданной корреляцией и законом распределения вероятностей (метод Монте-Карло), имитирующих на ЭВМ случайные значения параметров при каждом испытании;
2. Преобразование полученных числовых последовательностей на имитационных математических моделях.
3. Статистическая обработка результатов моделирования.

В основе имитационного моделирования лежит метод многократного решения задач "Что будет, если?" Используя технику имитационного моделирования, мы имеем возможность непрерывно и случайным образом генерировать значения каждой входной переменной x_i $(i = 1,2,...,k)$ модели и затем рассчитывать значения выходной переменной y. Затем, используя полученные значения переменной y, мы можем оценивать закон (форму) распределения и его параметры для этой стохастической переменной. Например, изменяя случайным образом входные параметры, мы можем получить некоторое количество значений выходной переменной (выборку), затем построить частотное распределение встречающихся в выборке значений, определить пределы изменения выходной переменной, оценить среднее значение и дисперсию полученного распределения и, наконец, оценить вероятность того, что фактическое значение выходной переменной не будет больше (меньше), чем заданная величина. Все эти параметры дают возможность менеджеру точнее оценить риск, связанный с принимаемым им решением.

Имитационное (компьютерное) моделирование экономических процессов обычно применяется в двух случаях:

- для управления сложным бизнес-процессом», когда имитационная модель управляемого экономического объекта используется в качестве инструментального средства в контуре адаптивной системы управления, создаваемой на основе информационных (компьютерных) технологий;
- при проведении экспериментов с дискретно-непрерывными моделями сложных экономических объектов для получения и отслеживания их динамики в экстренных ситуациях, связанных с рисками, натурное моделирование которых нежелательно иди невозможно.

6.2. Достоинства и недостатки имитационного моделирования

Все имитационные модели представляют собой модели типа так называемого черного ящика. Это означает, что они обеспечивают выдачу выходного сигнала системы, если на ее взаимодействующие подсистемы поступает входной сигнал. Поэтому для получения необходимой информации или результатов необходимо осуществлять «прогон» имитационных моделей, а не «решать» их. Имитационные модели не способны формировать свое собственное решение в том виде, в каком это имеет место в аналитических моделях, а могут лишь служить в качестве средства для анализа поведения системы в условиях, которые определяются экспериментатором. Следовательно, имитационное моделирование — не теория, а методология решения проблем. Более того, имитационное моделирование является только одним из нескольких имеющихся в распоряжении системного аналитика важнейших методов решения проблем. Поскольку необходимо и желательно приспосабливать средство или метод к решению задачи, а не наоборот, то возникает естественный вопрос: в каких случаях имитационное моделирование полезно?

Мы определили имитационное моделирование как экспериментирование с моделью реальной системы. Необходимость решения задачи путем экспериментирования становится очевидной, когда возникает потребность получить о системе специфическую информацию, которую нельзя найти в известных источниках. Бэриш [4] указывает, что непосредственное экспериментирование на реальной системе устраняет много затруднений, если необходимо обеспечить соответствие между моделью и реальными условиями; однако недостатки такого экспериментирования иногда весьма значительны, поскольку:

1. Оно может нарушить установленный порядок работы фирмы.
2. Если составной частью системы *являются люди,* то на результаты экспериментов может повлиять так называемый хауторнский эффект, проявляющийся в том, что люди, чувствуя, что за ними наблюдают, могут изменить свое поведение.
3. Может оказаться сложным поддержание одних и тех же рабочих условий при каждом повторении эксперимента или в течение всего времени проведения серии экспериментов.
4. Для получения одной и той же величины выборки (и, следовательно, статистической значимости результатов экспериментирования) могут потребоваться чрезмерные затраты времени и средств.
5. При экспериментировании с реальными системами может оказаться невозможным исследование множества альтернативных вариантов.

По этим причинам исследователь должен рассмотреть целесообразность применения имитационного моделирования при наличии любого из следующих условий:

1. Не существует законченной математической постановки данной задачи, либо еще не разработаны аналитические методы решения сформулированной математической модели. К этой категории относятся многие модели массового обслуживания, связанные с рассмотрением очередей.
2. Аналитические методы имеются, но математические процедуры столь сложны и трудоемки, что имитационное моделирование дает более простой способ решения задачи.
3. Аналитические решения существуют, но их реализация невозможна вследствие недостаточной математической подготовки имеющегося персонала. В этом случае следует сопоставить затраты на проектирование, испытания и работу на имитационной модели с затратами, связанными с приглашением специалистов со стороны.
4. Кроме оценки определенных параметров, желательно осуществить на имитационной модели наблюдение за ходом процесса в течение определенного периода.
5. Имитационное моделирование может оказаться единственной возможностью вследствие трудностей постановки экспериментов и наблюдения явлений в реальных условиях; соответствующим примером может служить изучение поведения космических кораблей в условиях межпланетных полетов.
6. Для долговременного действия систем или процессов может понадобиться сжатие временной шкалы. Имитационное моделирование дает возможность полностью контролировать время изучаемого процесса, поскольку явление может быть замедлено или ускоренно по желанию. К этой категории относятся, например, исследования проблем упадка городов.

Дополнительным преимуществом имитационного моделирования можно считать широчайшие возможности его применения в сфере образования и профессиональной подготовки. Разработка и использование имитационной модели позволяют экспериментатору видеть и «разыгрывать» на модели реальные процессы и ситуации. Это в свою очередь должно в значительной мере помочь ему понять и прочувствовать проблему, что стимулирует процесс поиска нововведений.

6.3. Типовые задачи имитационного моделирования

Можно выделить следующие типовые задачи, решаемые средствами имитационного моделирования при управлении экономическими объектами:

- моделирование процессов логистики для определения временных и стоимостных параметров;
- управление процессом реализации инвестиционного проекта на различных этапах его жизненного цикла с учетом возможных рисков и тактики выделения денежных сумм;
- анализ клиринговых процессов в работе сети кредитных организаций (в том числе применение к процессам взаимозачетов в условиях российской банковской системы);
- прогнозирование финансовых результатов деятельности предприятия на конкретный период времени (с анализом динамики сальдо на счетах);
- бизнес-реинжиниринг несостоятельного предприятия (изменение структуры и ресурсов предприятия-банкрота, после чего с помощью имитационной модели можно сделать прогноз основных финансовых результатов и дать рекомендации о целесообразности того или иного варианта реконструкции, инвестиций или кредитования производственной деятельности);
- анализ адаптивных свойств и живучести компьютерной региональной банковской информационной системы (например, частично вышедшая из строя в результате природной катастрофы система электронных расчетов и платежей после катастрофического землетрясения 1995 г. на центральных островах Японии продемонстрировала высокую живучесть: операции возобновились через несколько дней);
- оценка параметров надежности и задержек в централизованной экономической информационной системе с коллективным доступом (на примере системы продажи авиабилетов с учетом несовершенства физической организации баз данных и отказов оборудования);
- анализ эксплуатационных параметров распределенной многоуровневой ведомственной информационной управляющей системы с учетом неоднородной структуры, пропускной способности каналов связи и несовершенства физической организации распределенной базы данных в региональных центрах;
- моделирование действий курьерской (фельдъегерской) вертолетной группы в регионе, пострадавшем в результате природной катастрофы или крупной промышленной аварии;
- анализ сетевой модели PERT (Program Evaluation and Review Technique) для проектов замены и наладки производственного оборудования с учетом возникновения неисправностей;
- анализ работы автотранспортного предприятия, занимающегося коммерческими перевозками грузов, с учетом специфики товарных и денежных потоков в регионе;
- расчет параметров надежности и задержек обработки информации в банковской информационной системе.

Приведенный перечень является неполным. Действительная область применения аппарата имитационного моделирования не имеет видимых ограничений. Например, спасение американских астронавтов при возникновении аварийной ситуации на корабле APOLLON стало возможным только благодаря «проигрыванию» различных вариантов спасения на моделях космического комплекса.

6.4. Социально-экономические процессы как объекты моделирования

Модели социально-экономических систем имеют некоторые особенности по сравнению, например, с моделями природных или технических систем. К сожалению, не всегда возможно создать математическую модель социально-экономической системы в узком значении этого слова. При изучении таких систем мы можем определить цели, указать ограничения и предусмотреть, чтобы система подчинялась техническим законам, нормативным правовым ограничениям и т.п. При этом могут быть вскрыты и представлены в той или иной математической форме существенные связи в системе. В отличие от этого решение проблем защиты от загрязнения воздушной среды, предотвращения преступлений, здравоохранения и огромное количество другим проблем связано с неясными и противоречивыми целями, а также выбором альтернатив, диктуемых политическими и социальными факторами.

Можно выделить следующие ситуации, в которых рекомендуется использовать имитационное моделирование при изучении сложных социально-экономических систем:

- экономическая система сформировалась недавно и идет процесс ее познания;
- имитационное моделирование дает более простой способ решения задачи, чем аналитический метод;
- имитационное моделирование оказывается единственным способом исследования сложной системы из-за невозможности наблюдения ее в реальных условиях или из-за того, что проведение экспериментов на реальных объектах связано с возможностями социальных и экономических потерь;
- необходимо исследовать процессы в другом масштабе времени;
- необходимо качественно подготовить специалистов для работы в какой-либо экономической системе;
-следует предсказать неочевидные нежелательные явления в системе.

Социально-экономические системы отличаются большой сложностью. Успешное их функционирование зависит от большого количества разнообразных факторов, среди которых можно выделить:

- социально-политические,
- нормативно-правовые,
- технические, технологические, зоотехнические, агрономические, климатические и т.п.,
- экологические,
- маркетинговые,
- финансовые.

Набор таких факторов может достигать десятков, сотен и более. Поэтому экспериментировать с реальными экономическими системами часто бывает невозможно, непрактично или неэкономично. Имитационный же эксперимент позволяет проводить исследование функционирования таких систем.

Система имитационного моделирования, обеспечивающая создание моделей для решения перечисленных задач, должна обладать следующими свойствами:

- возможностью применения имитационных программ совместно со специальными экономико-математическими моделями и методами, основанными на теории управления;
- инструментальными методами проведения структурного анализа сложного экономического процесса;
- способностью моделирования материальных, денежных и информационных процессов и потоков в рамках единой модели, в общем модельном времени;
- возможностью введения режима постоянного уточнения при получении выходных данных (основных финансовых показателей, временных и пространственных характеристик, параметров рисков и др.) и проведении экстремального эксперимента.

6.5. Примеры задач имитационного моделирования

Пример 6.1. Предположим, что некоторая промышленная фирма наряду с резким увеличением числа заказов на свою продукцию отметила заметное ухудшение качества обслуживания своих клиентов в части соблюдения сроков выполнения их заказов. Несоблюдение фирмой своих обязательств перед заказчиками может привести к ощутимым потерям как за счет штрафных санкций, так и за счет оттока клиентов. В рассматриваемой ситуации у фирмы может появиться желание воспользоваться компьютерным имитационным моделированием, с помощью которого можно было бы выяснить, каким образом существующие процедуры определения сроков выполнения принимаемых заказов, календарного планирования производства и оформления заявок на поставки сырья порождают наблюдаемые задержки. ►

Пример 6.2. Руководство крупной клиники разрабатывает новую систему управления запасами лекарственных препаратов. Использование имитационной модели, построенной на основе ретроспективных данных, позволит оценить, каким будет средний уровень средств, необходимых для обеспечения запасов лекарственных препаратов, и как часто будут возникать нехватки различных видов препаратов при реализации новой системы управления. ►

Пример 6.3. Рассмотрим предприятие с мелкосерийным производством, на котором производственные мощности распределены в соответствии с приоритетами, присвоенными выполняемым работам. Может быть построена имитационная модель для нахождения эффективного способа определения системы приоритетов для того, чтобы все работы могли выполняться без больших задержек и при этом коэффициент использования оборудования был бы достаточно высок.►

Пример 6.4. Биржевой игрок разработал свой порядок приобретения и продажи акций, состоящий в следующем:

1) обладая пакетом акций, необходимо продать его, как только цены на эти акции начинают падать;

2) как только цены на акции начинают возрастать, их необходимо покупать.

Игрок не желает рисковать своими ограниченными средствами в натурном эксперименте и хочет оценить прибыльность своей стратегии с помощью имитационного моделирования. ►

Пример 6.5. Рассмотрим очередь покупателей к контрольному прилавку небольшого магазина подарков (так называемая однолинейная система массового обслуживания). Предположим, что промежутки времени между последовательными появлениями покупателей распределяются равномерно в интервале от 1 до 10 мин (для простоты мы округляем время до ближайшего целого числа минут). Предположим далее, что время, необходимое для обслуживания каждого покупателя, распределяется равномерно в интервале от 1 до 6 мин. Нас интересует среднее время, которое покупатель проводит в данной системе (включая и ожидание, и обслуживание), и процент времени, в течение которого продавец, стоящий на контроле, не загружен работой. Для моделирования системы нам необходимо поставить искусственный эксперимент, отражающий основные условия ситуации. Для этого мы должны придумать способ имитации искусственной последовательности прибытий покупателей и времени, необходимого для обслуживания каждого из них. Один из способов, который мы могли бы применить, состоит в том, чтобы найти десять фишек и один кубик. Вслед за этим мы могли бы пронумеровать фишки с числами 1 по 10, положить их в шляпу и, встряхивая ее, перемешать фишки. Вытягивая фишку из шляпы и считывая выпавшее число, мы могли бы таким путем представить промежутки времени между появлением предыдущего и последующего покупателей. Бросая наш кубик и считывая с его верхней грани число очков, мы могли бы такими числами представить время обслуживания каждого покупателя. Повторяя эти операции в указанной последовательности (возвращая каждый раз фишки обратно и встряхивая шляпу перед каждым вытягиванием), мы могли бы получить временные ряды, представляющие промежутки времени между последовательными прибытиями покупателей и соответствующие им времена обслуживания. Наша задача затем сведется к простой регистрации результатов эксперимента.

Таблица 6.1 показывает, какие, например, результаты можно получить в случае анализа прибытия 20 покупателей.

Очевидно, для получения статистической значимости результатов мы должны были взять гораздо большую выборку, кроме того мы не учли некоторые важные обстоятельства, такие, например, как начальные условия (это будет обсуждаться позднее). Важным моментом является и то, что для генерирования случайных чисел мы применили два приспособления (пронумерованные покерные фишки и кубик); это было сделано с целью осуществить искусственный (имитационный) эксперимент с системой, позволяющий выявить определенные черты ее поведения.

Таблица 6.1

Имитационное моделирование работы контрольного прилавка

Покупатель	Время после прибытия предыдущего покупателя мин	Время обслуживания, мин	Текущее модельное время в моменты прибытия покупателей	Начало обслуживания	Конец обслуживания	Время пребывания покупателя у прилавка, мин	Время простоя продавца в ожидании покупателя, мин
1	---	1	0,00	0,00	0,01	1	0
2	3	4	0,03	0,03	0,07	4	2
3	7	4	0,10	0,10	0,14	4	3
4	3	2	0,13	0,14	0,16	3	0
5	9	1	0,22	0,22	0,23	1	6
6	10	5	0,32	0,32	0,37	5	9
7	6	4	0,38	0,38	0,42	4	1
8	8	6	0,46	0,46	0,52	6	4
9	8	1	0,54	0,54	0,55	1	2
10	8	3	1,02	1,02	1,05	3	7
11	7	5	1,09	1,09	1,14	5	4
12	3	5	1,12	1,14	1,19	7	0
13	8	3	1,20	1,20	1,23	3	1
14	4	6	1,24	1,24	1,30	6	1
15	4	1	1,28	1,30	1,31	3	0
16	7	1	1,35	1,35	1,36	1	4
17	1	6	1,36	1,36	1,42	6	0
18	6	1	1,42	1,42	1,43	1	0
19	7	2	1,49	1,49	1,51	2	6

20	6	2	1,55	1,55	1,57	2	5
					Всего	68	55

Среднее время пребывания покупателя у прилавка 68/20 = 3,40 мин
Процент непроизводительного времени продавца 55/117 = 47%. ►

Применение *математического моделирования* позволяет исследовать объекты, реальные эксперименты над которыми затруднены или невозможны (дорого, опасно для здоровья, однократные процессы, невозможные из-за физических или временных ограничений – находятся далеко, еще или уже не существуют и т.п.).

Экономический эффект: затраты в среднем сокращаются в 10-100 раз.

Вопросы для самопроверки

1. Какие виды моделирования вам известны? Опишите их.
2. Что понимают под алгоритмизацией функционирования реальных процессов и систем?
3. Чем отличаются вероятностные аналитические модели от вероятностных имитационных моделей?
4. Что понимают под статистическим моделированием? Перечислите этапы статистического моделирования.
5. В каких случаях применяют имитационное моделирование экономических процессов?
6. В чем состоит недостаток имитационного моделирования?
7. В чем состоит недостаток экспериментирования с реальными системами?
8. При наличии каких условий целесообразно применять имитациоенное моделирование?
9. Перечислите типовые задачи имитационного моделирования.
10. В каких ситуациях рекомендуется использовать имитационное моделирование при изучении сложных социально-экономических систем?
11. Приведите примеры задач имитационного моделирования

Тема 7. Сущность метода имитационного моделирования

7.1. Метод имитационного моделирования и его особенности

Определим **метод имитационного моделирования** в самом общем виде как *экспериментальный метод исследования реальной системы по ее имитационной модели, который сочетает особенности экспериментального подхода и специфические условия использования вычислительной техники*.

В этом определении подчеркивается, что имитационное моделирование является машинным методом моделирования, собственно без ЭВМ никогда не существовало, и только развитие информационных технологий привело к становлению этого вида компьютерного моделирования. В этом определении также акцентируется внимание на экспериментальной природе имитации, применяется имитационный метод исследования (осуществляется экспериментирование с моделью). Действительно, в имитационном моделировании важную роль играет не только проведение, но и планирование эксперимента на модели. Однако это определение не проясняет, что собой представляет сама имитационная модель. Итак, в чем же состоит сущность имитационного моделирования.

В процессе имитационного моделирования (рис. 7.1) исследователь имеет дело с четырьмя основными элементами:

- Реальная система;
- Логико математическая модель моделируемого объекта;
- Имитационная (машинная) модель;
- ЭВМ, на которой осуществляется имитация – направленный вычислительный эксперимент.

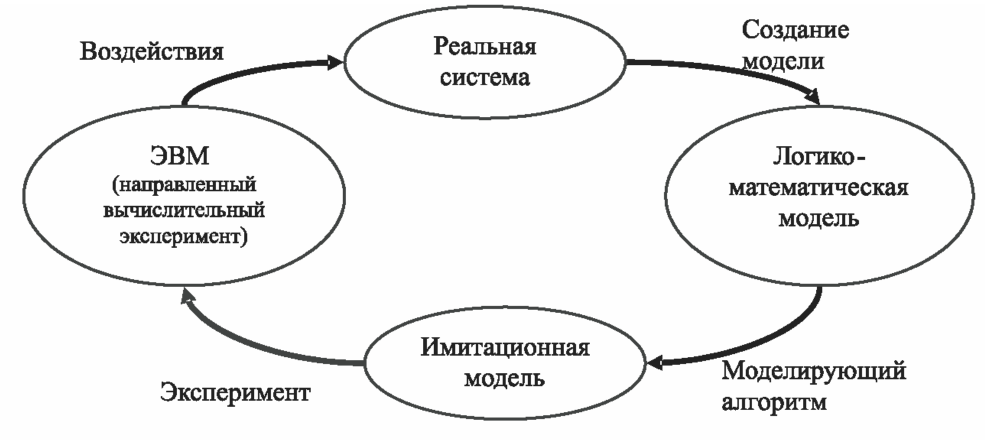


Рис. 7.1 Процесс имитационного моделирования

Исследователь изучает реальную систему, разрабатывает логико-математическую модель реальной системы. *Имитационный характер исследования предполагает наличие логико или логико-математических моделей, описываемых изучаемый процесс.*

Выше мы определяли *реальную систему как совокупность взаимодействующих элементов, функционирующих во времени.*

Составной характер сложной системы диктует представление ее модели в виде тройки:

$$< A, S, T >,$$

где

A – множество элементов (в их число включается и внешняя среда);

S – множество допустимых связей между элементами (структура модели);

T – множество рассматриваемых моментов времени.

Особенностью имитационного моделирования является то, что имитационная модель позволяет воспроизводить моделируемые объекты:

- *с сохранением их логической структуры,*
- *с сохранением поведенческих свойств* (последовательности чередования во времени событий, происходящих в системе), т.е. динамики взаимодействий.

При имитационном моделировании структура моделируемой системы адекватно отображается в модели, а процессы ее функционирования проигрываются (имитируются) на построенной модели. Поэтому построение имитационной модели заключается в описании структуры и процессов функционирования моделируемого объекта или системы.

В описании имитационной модели выделяют две составляющие:

Статическое описание системы, которое по существу является описанием ее структуры. При разработке имитационной модели необходимо выполнять структурный анализ моделируемых процессов.

Динамическое описание системы, или описание динамики взаимодействий ее элементов. При его составлении фактически требуется построение *функциональной модели* моделируемых динамических процессов.

Идея метода, с точки зрения его программной реализации, состояла в следующем. Что если элементам системы поставить в соответствие некоторые программные компоненты, а состояния этих элементов описывать с помощью переменных состояния. Элементы, по определению, взаимодействуют (или обмениваются информацией), – значит может быть реализован алгоритм функционирования отдельных элементов – моделирующий алгоритм. Кроме того, элементы существуют во времени – значит надо задать алгоритм изменение переменных состояний.

Динамика в имитационных моделях реализуется с помощью *механизма продвижения модельного времени*. Отличительной особенностью метода имитационного моделирования является возможность описания и воспроизведения взаимодействия между различными элементами системы. Таким образом, чтобы составить имитационную модель, надо:

- представить реальную систему (процесс), как совокупность взаимодействующих элементов;
- алгоритмически описать функционирование отдельных элементов;
- описать процесс взаимодействия различных элементов между собой и с внешней средой.

Ключевым моментом в имитационном моделировании является выделение и описание *состояний* системы. Система характеризуется *набором переменных состояний*, каждая комбинация которых описывает конкретное состояние. Следовательно, путем изменения значений этих переменных можно имитировать переход системы из одного состояния в другое. Таким образом, имитационное моделирование – это представление *динамического поведения* системы посредством продвижения ее от одного состояния к другому в соответствии с хорошо определенными операционными правилами. Эти изменения состояний могут происходить либо непрерывно, либо в дискретные моменты времени. Имитационное моделирование – есть динамическое отражение изменений состояния системы с течением времени.

7.2. Процесс имитации

Основные этапы процесса имитации

Роберт Шеннон [13] выделяет следующие этапы процесса имитации:

1. *Определение системы* – установление границ, ограничений и измерителей эффективности системы, подлежащей изучению.

2. *Формулирование модели* – переход от реальной системы к некоторой логической схеме (абстрагирование).

3. *Подготовка данных* – отбор данных, необходимых для построения модели, и представление их в соответствующей форме.

4. *Трансляция модели* – описание модели на языке программирования.

5. *Оценка адекватности* – повышение до приемлемого уровня степени уверенности, с которой можно судить относительно корректности выводов о реальной системе, полученных на основании обращения к модели.

6. *Стратегическое планирование* – планирование эксперимента, который должен дать необходимую информацию.

7. *Тактическое планирование* – определение способа проведения каждой серии испытаний, предусмотренных планом эксперимента.

8. *Экспериментирование* – процесс осуществления имитации с целью получения желаемых данных и анализа чувствительности.

9. *Интерпретация* – построение выводов по данным, полученным путем имитации.

10. *Реализация* – практическое использование модели и результатов моделирования.

11. *Документирование* – регистрация хода осуществления проекта и его результатов, а также документирование процесса создания и использования модели.

7.3. Формулирование модели

Модель – представление объекта, системы в некоторой форме, отличной от формы их реального существования. Модель – это прежде всего упрощенное представление реального объекта или явления, сохраняющее его основные, существенные черты. Сам процесс разработки модели, т.е. моделирование, может быть осуществлен различными способами, из которых наиболее распространено физическое и математическое моделирование.

Разработка и построение модели социально-экономической системы требует глубокого изучения свойств исходного объекта. При этом в основе построения модели лежат, как правило, определенная теоретическая концепция, включающая способ анализа ситуации, некоторые априорные представления о взаимосвязях между наблюдаемыми и изучаемыми признаками и факторами. В начале построения модели следует по возможности точно сформулировать те предпосылки (гипотезы) и отношения, на которых базируется математическая модель объекта. В основу предлагаемого описания отношений, определяющих возможный тип модели, положена классификация, приведенная в книге [27] и графически представленная на рис. 7.2.

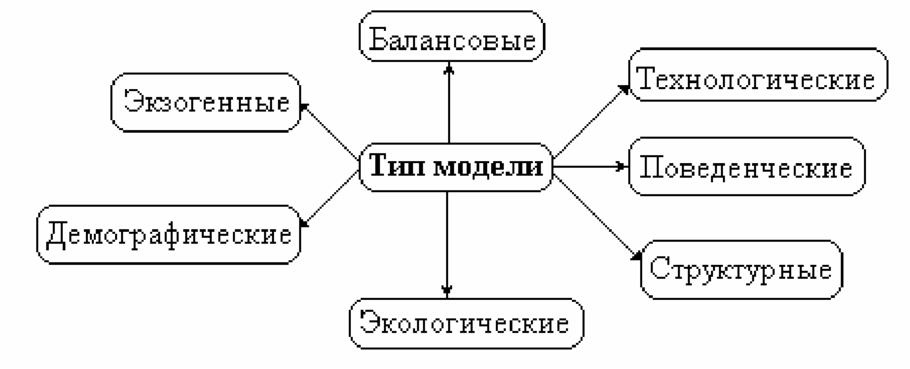


Рис. 7.2. Структура типов отношений

Балансовые отношения реализуют равенство суммы наличных объемов (товаров, ресурсов, финансовых потоков и т.п.), полученных из различных источников, сумме объемов, использованных по различным направлениям.

Технологические отношения раскрывают и детализируют технологические процессы конкретных экономических систем.

Поведенческие отношения описывают поведение элементов экономической системы, имеющих некоторую свободу выбора решений.

Структурные отношения определяют иерархию управления (взаимосвязи, соподчиненность элементов) и правила, регулирующие функционирование элементов системы в рамках этой иерархии.

Экономические отношения раскрывают взаимосвязи экономической системы с окружающей средой.

Демографические отношения учитывают влияние роста и структуры населения на производство.

Экзогенные отношения описывают взаимоотношения экономической системы с элементами внешней среды, причем влияние системы на эти элементы ограниченно (внешняя торговля, стихийные бедствия и др.).

Следует иметь в виду, что при разработке моделей достаточно сложных экономических систем необходимо учитывать несколько или даже все вышеуказанные отношения элементов системы.

7.4. Оценка адекватности модели

Какой бы сложной и полной ни была модель, она тем не менее является приближенным отображением реального объекта и отражает его при определенных принятых допущениях. Однако до тех пор пока не доказана адекватность модели реальной обстановке, нельзя с уверенностью утверждать, что с ее помощью получатся те результаты, которые действительно характеризуют функционирование исследуемого объекта. Оценка адекватности и точности математической модели любого типа, в том числе и имитационной, является важнейшей задачей моделирования, так как любые исследования на неадекватной модели теряют смысл [4].

Оценка адекватности модели заключается в повышении до приемлемого уровня степени уверенности, с которой можно судить относительно корректности выводов о реальной системе, полученных на основании обращения к модели.

Для этого могут быть использованы проверки трех видов. При проверке ***первого*** вида следует проверить: не будет ли модель давать просто абсурдные ответы. Второй метод оценки адекватности модели называется верификацией. ***Верификация имитационной модели*** – проверка соответствия ее поведения предположениям экспериментатора. Это первый этап действительной подготовки к имитационному эксперименту. Подбираются некоторые исходные данные, для которых могут быть представлены результаты просчета. Если окажется, что ЭВМ выдает данные, противоречащие тем, которые ожидались при формировании модели, значит, модель неверна. В обратном случае переходят к следующему этапу проверки работоспособности модели – ее *валидации.*

Валидация имитационной модели – проверка соответствия данных, получаемых в процессе *машинной имитации,* реальному ходу явлений, для описания которых создана *модель.* Состоит в том, что выходные данные после расчета на ЭВМ сопоставляются с имеющимися статистическими сведениями о моделируемой системе.

Таким образом, вопрос оценки адекватности модели имеет две стороны:

- приобретение уверенности в том, что модель ведет себя таким же образом, как и реальная система;

- установление того, что выводы, полученные из экспериментов с моделью, справедливы и корректны.

С ростом адекватности и точности модели возрастают как ее стоимость, так и ценность для исследования, в связи с чем приходится решать вопрос о компромиссе между стоимостью модели и последствиями ошибочных решений из-за неадекватности исследуемому процессу. Оценка адекватности и точности модели представляет собой непрерывный процесс, правильность построения модели может быть проверена только на практике за счет повторения цикла «построение модели – проверка модели». Следует отметить, что понятие адекватности модели не имеет количественного измерения: модель либо адекватна явлению, либо не адекватна. При этом, естественно, предполагается, что программа, реализующая вычисления по математической модели, не содержит ошибок, исходные данные введены в машину правильно. Таким образом, модель является достоверной, если ее концептуальная модель адекватна исследуемому процессу, математическая модель адекватна концептуальной, а точность реализации математической модели на ЭВМ соответствует заданной, т.е. погрешности расчета не превышают допустимых [4, с.118].

7.5. Экспериментирование с использованием имитационной модели

После завершения этапов разработки и планирования осуществляется процесс экспериментирования на модели с целью получения желаемой информации. На этом этапе начинают выявляться недостатки и просчеты в модели и планировании проведения экспериментов, которые должны постоянно устраняться до тех пор, пока не будут достигнуты первоначально поставленные цели.

При экспериментировании модель выступает как самостоятельный объект исследования. Одну из форм такого исследования составляет проведение модельных экспериментов, при которых целенаправленно изменяются условия функционирования модели и систематизируются данные о ее «поведении». Конечным результатом этого этапа является совокупность знаний о модели в отношении существенных сторон объекта-оригинала, которые отражены в данной модели.

На рис. 7.3…7.5. схематично показаны возможные варианты изменения входных параметров имитационной модели при проведении экспериментов на ней с целью изучения изменения выходных параметров моделируемой системы.

На рис. 7.3 показан вариант имитационного эксперимента, при котором изменяются значения только одного входного параметра x_1 имитационной модели (**ИМ**) при фиксированных значениях всех остальных входных параметров. При этом исследуется изменение выходных параметров $w_1, w_2, ..., w_n$.

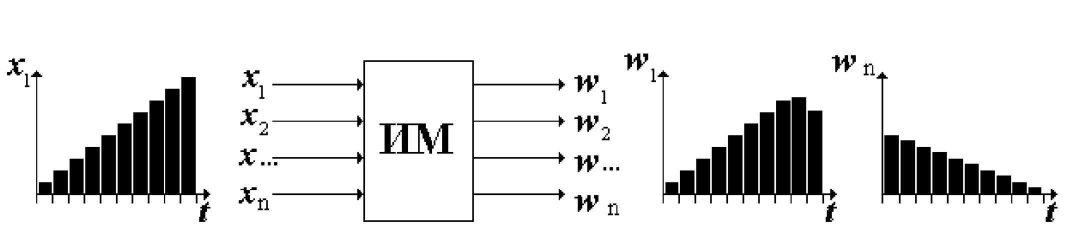


Рис. 7.3. Исследование влияния изменения одного входного параметра на выходные параметры

В варианте имитационного эксперимента, изображенном на рис. 7.4, исследуется поведение выходных параметров при изменении нескольких или всех входных параметров. Такой вид эксперимента характерен для стратегического планирования, когда выбирается та или иная стратегия изменения входных параметров.

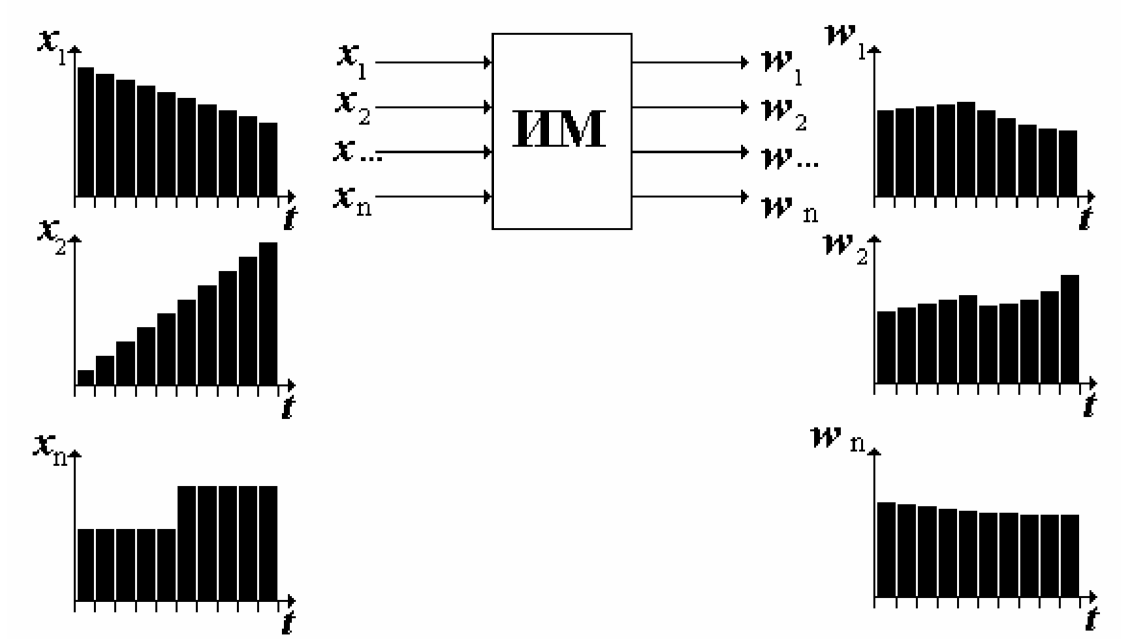


Рис. 7.4. Имитация принятой стратегии изменения входных параметров

Если одна или несколько переменных являются случайными (стохастическими) величинами, то при проведении имитационного эксперимента их значения задаются при помощи генератора случайных чисел (*ГСЧ*). Вариант эксперимента со случайными входными параметрами изображен на рис. 7.5.

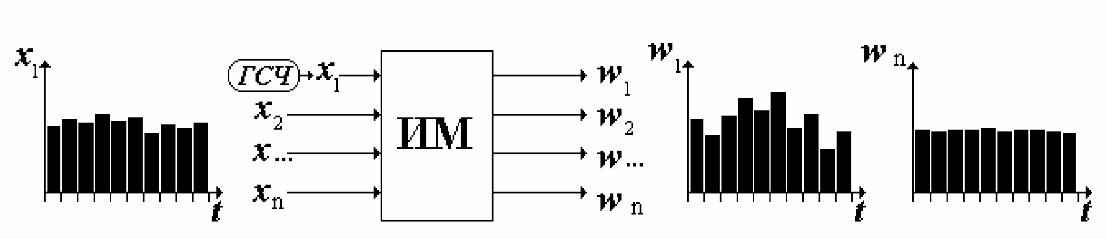


Рис. 7.5. Имитационный эксперимент со случайными входными параметрами

Если математическая модель непригодна для применения аналитических или численных методов, то для ее решения применяют метод экспериментальной оптимизации на ЭВМ. При этом методе нет необходимости в преобразовании математической модели в специальную систему уравнений. Как целевая функция, так и система ограничений могут быть заданы в виде алгоритма, позволяющего вычислять их значения в ходе моделирования. На рис. 7.6 схематично показан вариант имитационного экспериментирования с применением метода экспериментальной оптимизации.

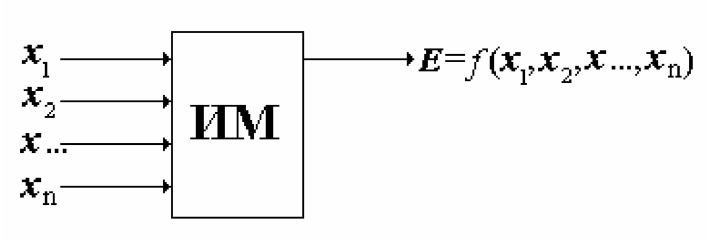


Рис. 7.6. Метод экспериментальной оптимизации с использованием имитационной модели

7.6. Понятие о модельном времени. Механизм продвижения модельного времени

Для описания динамики моделируемых процессов в имитационном моделировании реализован *механизм задания модельного времени.* Эти механизмы встроены в управляющие программы любой системы моделирования.

Если бы на ЭВМ имитировалось поведение одной компоненты системы, то выполнение действий в имитационной модели можно было бы осуществить последовательно, по пересчету временной координаты.

Чтобы обеспечить имитацию параллельных событий реальной системы вводят некоторую глобальную переменную (обеспечивающую синхронизацию всех событий в системе) *t0*, которую называют *модельным (или системным) временем.*

Существуют два основных способа изменения *t0*:

- *пошаговый* (применяются фиксированные интервалы изменения модельного времени);
- *по событийный* (применяются переменные интервалы изменения модельного времени, при этом величина шага измеряется интервалом до следующего события).

В случае *пошагового метода* продвижение времени происходит с минимально возможной постоянной длиной шага *(принцип t).* Эти алгоритмы не очень эффективны с точки зрения использования машинного времени на их реализацию.

По событийный метод (принцип “особых состояний”). В нем координаты времени меняются только когда изменяется состояние системы. В по событийных методах длина шага временного сдвига максимально возможная. Модельное время с текущего момента изменяется до ближайшего момента наступления следующего события.

Рис. 7.7. демонстрирует способы представления и управления временем при использовании обоих методов. По оси времени отложена одна и та же последовательность событий.

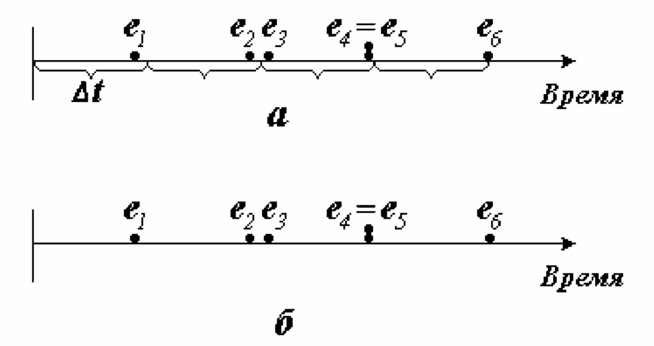


Рис. 7.7. Течение модельного времени: а – в модели с фиксированным шагом (пошаговый метод); б – по событийный метод.

У каждого из этих методов есть свои преимущества. В модели с фиксированным шагом (пошаговый метод) обработка событий происходит по группам или множествам событий. В модели, использующей по событийный метод задания шага, обработка событий идет последовательно и время имитации каждый раз смещается вперед на начало следующего события, каждое из которых обслуживается по очереди.

Использование метода фиксированных шагов более предпочтительно, если:

- события появляются регулярно и распределены во времени относительно равномерно;

- в течение цикла моделирования появляется много событий, причем математическое ожидание продолжительности событий мало;

- точная природа существенных событий не ясна, как, например, это бывает на начальном этапе имитационного исследования.

С другой стороны, по событийный метод задания шага до следующего события выгоден тем, что:

- позволяет экономить машинное время в случае статических систем, в которых существенные события могут длительное время не наступать;

- не требует определения величины приращения времени (что влияет и на продолжительность цикла моделирования, и на точность);

- может эффективно использоваться при неравномерном распределении событий во времени и (или) при большой величине математического ожидания их продолжительности.

Простейшая классификация на основные виды имитационных моделей связана с применением двух этих способов продвижения модельного времени.

Различают имитационные модели:

- *Непрерывные;*
- *Дискретные;*
- *Непрерывно дискретные.*

В *непрерывных имитационных моделях* переменные изменяются непрерывно, состояние моделируемой системы меняется как непрерывная функция времени, и, как правило, это изменение описывается системами дифференциальных уравнений. Соответственно продвижение модельного времени зависит от численных методов решения дифференциальных уравнений.

В *дискретных имитационных моделях* переменные изменяются дискретно в определенные моменты имитационного времени (наступления событий). Динамика дискретных моделей представляет собой процесс перехода от момента наступления очередного события к моменту наступления следующего события.

Поскольку в реальных системах непрерывные и дискретные процессы часто невозможно разделить, были разработаны *непрерывно дискретные модели*, в которых совмещаются механизмы продвижения времени, характерные для этих двух процессов.

7.7. Интерпретация и реализация результатов моделирования

По мнению Р. Шеннона [13], никакое задание на моделирование не может считаться успешно завершенным до тех пор, пока оно не будет принято, понято и использовано. Наибольшие неудачи, постигавшие специалистов, занимающихся проблемами управления, были связаны и восприятием и использованием их работ. Роберт Шеннон следующим образом распределяет время, необходимое на проектирование и работу с моделью: 25% на постановку задачи, 20% на сбор и анализ данных, 30% на разработку модели и 25% на реализацию. Самая лучшая в мире имитационная модель ничего не стоит, пока она не использована или не одобрена теми, для кого она была разработана. Для большинства специалистов и руководителей производства интерес представляет не изящная модель и искусное использование математических методов, а реальные проблемы и способы их решения.

Информация, получаемая с помощью имитационной модели должна быть приемлема для заказчика или пользователя, которые должны понимать, как необходимо поступить или как можно использовать полученные результаты. Если ему не ясно, как эти данные могут помочь ему или кому-либо другому принимать решения, то он их будет просто игнорировать, и вся работа по созданию модели окажется безрезультатной.

Любая имитационная модель должна позволять специалисту, работающему с ней, оценивать решения, которые удовлетворяют его собственным понятиям рациональности, а также возможные результаты сформулированных им стратегий. Пользователь модели, если он не является ее разработчиком, полагает, что именно он может наилучшим образом получать правильные решения.

Чтобы иметь максимальные шансы успешного применения результатов имитационного исследования, имитационная модель должна быть:

- понятной заказчику-пользователю;
- способной давать разумные ответы;
- способной давать информацию, которая может быть в дальнейшем использована;
- реалистичной в требованиях к данным;
- способной отвечать на вопросы типа «А что будет, если…?»;
- легко модифицируемой;
- недорогой при применении.

Вопросы для самопроверки

1. Приведите конкретные примеры задач имитационного моделирования.
2. Дайте определение метода имитационного моделирования в самом общем виде.
3. Какие элементы выделяют в процессе имитационного моделирования?

4. Какие две составляющие выделяют в описании имитационной модели? Каким образом реализуется динамика в имитационных моделях?
5. Перечислите основные этапы процесса имитации
6. Перечислите отношения, на которых базируется математическая модель объекта. Поясните эти отношения.
7. Что такое адекватность модели? Дайте понятие верификации и валидации модели.
8. Перечислите возможные схемы экспериментирования с использованием имитационной модели.
9. Назовите два основных способа задания модельного времени. Поясните суть этих способов
10. В каких случаях используют метод фиксированных шагов задания модельного времени?
11. В каких случаях используют пособытийный метод задания модельного времени?
12. Какими свойствами должна обладать имитационная модель для ее успешного использования?

Тема 8. Имитационная модель глобальной системы

Рассмотрим глобальную модель развития всей нашей цивилизации в целом, представленную в книге Дж. Форрестера «Мировая динамика» [12]. Эта книга фактически явилась первой завершенной попыткой применить точные методы для исследования мирового развития.

8.1. Основные компоненты динамической мировой модели

В динамической мировой модели Дж. Форрестера взаимоувязаны население, капиталовложения, географическое пространство, природные ресурсы, загрязнение и производство продуктов питания. Этими основными компонентами и их взаимодействиями обуславливается динамика изменений в мировой системе. Растущее население вызывает рост индустриализации, рост потребности в продуктах питания и распространение населения по все большей территории. Но рост производства продуктов питания, промышленных товаров и занимаемой территории способствует не только поддержанию, но и увеличению количества населения. Рост населения с сопровождающими его индустриализацией и загрязнением является следствием циклических процессов, в которых каждый сектор способствует росту других секторов, и обеспечивает свое развитие за их счет. Но со временем рост наталкивается на пределы, налагаемые природой. Почва и природные ресурсы истощаются, а способность биосферы Земли разлагать загрязнения не беспредельна. Задача состоит в том, чтобы выбрать наилучший из возможных вариантов перехода от динамического роста к состоянию мирового равновесия.

Модель была построена на основании ряда утверждений, наблюдений и предположений, касающихся мировой системы, в ней взаимосвязаны секторы демографии, экономики, сельского хозяйства и технологии. Модель описывает мировую систему, которая демонстрирует ряд альтернативных возможностей поведения человечества, которые оно пока еще может выбрать.

8.2. Концепция «петля обратной связи»

Решающий этап в построении машинной модели социальной системы – выбор и согласование информации о реальной системе. Обычно испытываются трудности не в нехватке информации, а в ее избытке и в необходимости избирательного подхода к ней. Разнородная информация должна быть организована.

Важной концепцией в установлении структуры системы является идея, что все изменения обуславливаются «петлями обратных связей». Петля обратной связи – это замкнутая цепочка взаимодействия, которая связывает исходное действие с его результатом, изменяющим характеристики окружающих условий и, которые в свою очередь, являются «информацией», вызывающей дальнейшие изменения. Мы часто рассматриваем причину и следствие односторонне. Мы говорим, что действие А вызывает результат В. Но такое понимание неполно. Результат В представляет новое состояние системы, изменение которой в будущем повлияют на действие А. Все процессы роста и стабилизации генерируются петлями обратных связей.

В системе с петлями обратных связей введены два типа переменных - *уровни* и *темпы*. Уровни – это накопители системы. Темпы – потоки, вызывающие изменение уровней. Уровень аккумулирует общее количество, являющееся результатом «впадающих» в него темпов, которые прибавляются или вычитаются из уровня. Системные уровни полностью описывают положение или состояние системы в любой момент времени. Уровни существуют во всех подсистемах – финансовой, физической, биологической, психологической и экономической. Население, как создающееся в результате аккумуляции «чистой» разности между темпом рождаемости и темпом смертности, рассматривается как уровень мировой системы. Изменение уровней вызывается соответствующими темпами потоков. Темп потока контролируется только одним или несколькими системными уровнями, но не другими темпами. Все системы, которые изменяются во времени, могут быть представлены как конструкции уровней и темпов, эти два типа переменных не только необходимы, но и достаточны для понимания любой системы.

В качестве основных уровней, на которых строится структура системы Дж. Форрестера, было выбрано пять следующих уровней:

- население;
- капиталовложения;
- природные ресурсы;
- часть капиталовложений, вкладываемых в сельское хозяйство;
- уровень загрязнения.

Для понимания методологии разработки модели рассмотрим примеры петлей обратных связей для различных уровней.

На рис. 8.1. показаны две основные петли, влияющие на численность населения.

(аббревиатура BRN – birth rate normal; DRN – death rate normal)

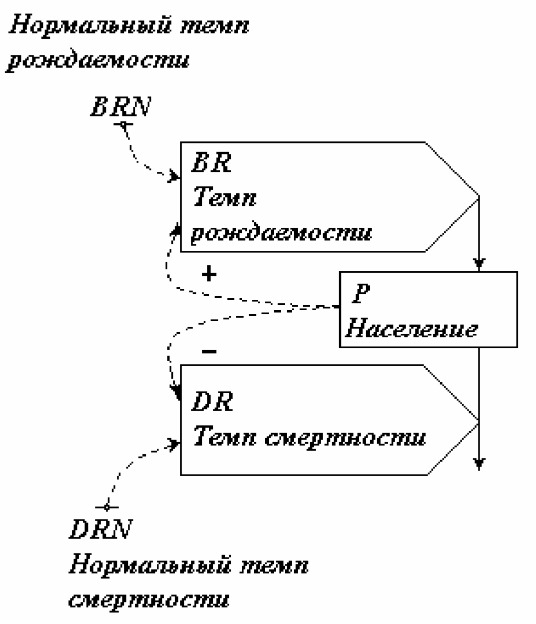


Рис.8.1. Основные петли обратных связей в секторе населения

Верхняя петля определяет темп рождаемости, который увеличивает население. Нижняя петля определяет темп смертности, уменьшающий население. Темп рождаемости и темп смертности определяются количеством людей, родившихся и умерших за год. Они определяют общий темп изменения численности населения.

Разность этих величин составляет чистый прирост населения. Эти темпы принято считать «нормальными», если они соответствуют стандартной системе мировых условий, когда величины уровня питания, материального уровня жизни, плотности и загрязнения соответствуют своим «стандартным» значениям. Однако эта же система переменных при других численных значениях может вызвать рост или падение темпов рождаемости и смертности по сравнению с их нормальными значениями. Влияние жизненных условий в мировой системе Форрестера описывается посредством «множителей», которые увеличивают или уменьшают нормальные темпы системы в зависимости от того, насколько благоприятно их воздействие в данный конкретный момент времени. Изменение этих множителей, отражающих текущее состояние мировой системы (уровень питания, материальный уровень жизни, плотность населения и уровень загрязнений) может вызывать рост населения, его стабилизацию или уменьшение.

Петля рождаемости – положительная обратная связь, вызывающая рост населения. Увеличение населения вызывает рост темпа рождаемости (число людей, родившихся в системе в течение года), который в свою очередь увеличивает население. Если бы не было сдерживающих сил, то население экспоненциально возрастало бы за счет влияния положительных обратных связей. Но петля смертности - отрицательная. Когда население растет, количество умирающих в год также возрастает. Увеличение населения увеличивает темп смертности, что ведет к уменьшению населения.

8.3. Основные петли «обратных связей» в мировой модели

Используя концепцию петель обратных связей, разработчик модели должен установить значимые зависимости между основными переменными в системе. В модели Форрестера заложены следующие логические цепочки положительных и отрицательных обратных связей мировой модели.

Капиталовложения. Генерация капиталовложений зависит от численности населения и нормальной генерации капиталовложений. Но множитель капиталовложений изменяет нормальный темп создания фондов. При очень низких значениях материального уровня жизни мотивы для потребления всей создаваемой продукции столь сильны, что накопление капитала может быть очень мало. При очень высоких значениях относительной величины фондов на душу населения падают потребности и стимулы к еще большему увеличению материального уровня жизни.

Загрязнение. Еще одна петля положительной обратной связи (вместе с отрицательной петлей) существует в секторе загрязнения. Отрицательная обратная связь представляет собой основной процесс разложения загрязнения. Чем выше уровень загрязнения в некоторый момент времени, тем большее его количество может разлагаться в единицу времени – до тех пор, пока уровень загрязнения не возрастет настолько, что будет разрушать природные процессы самоочищения. Объединение двух петель может давать либо отрицательный, либо положительный эффект в зависимости от величины загрязнения. Если общее загрязнение становится настолько большим, что подавляет природные процессы самоочищения, происходят крупномасштабные катастрофы.

Рост населения. На темпы рождаемости и смертности оказывают влияние такие факторы, как плотность населения, обеспеченность пищей, загрязнение, наличие природных ресурсов. При помощи концепции петель обратной связи можно установить взаимное влияние этих параметров на рост населения.

Ресурсы. Растущее население использует природные ресурсы до тех пор, пока падение материального уровня жизни не вызовет очередного уменьшения численности населения. Если население растет, то темп потребления природных ресурсов повышается, что приводит к истощению природных ресурсов, их запасы уменьшаются. Это, в свою очередь, понижает эффективность относительной величины фондов и материальный уровень жизни. Падение материального уровня жизни вызывает увеличение темпа рождаемости, но это увеличение компенсируется ростом темпа смертности и уменьшает численность населения. Здесь одна из петель – положительная, а другая – отрицательная.

8.4. Основные переменные в мировой модели

В этом подразделе более подробно рассмотрены некоторые переменные мировой модели Дж. Форрестера.

Население. Население в мировой модели представляет собой переменный уровень системы. Население в любой момент времени вычисляется как население в предшествующий момент времени, плюс население, которое добавляется за счет темпа рождаемости в охватываемый период, минус население, убывающее за счет смертности. В математической модели уровень «население» представляет собой вычислительную процедуру, разворачивающуюся в сторону возрастания времени, при этом его значение увеличивается или уменьшается в соответствии с темпами потоков.

Темп рождаемости. Темп рождаемости зависит от численности населения и значения нормального темпа рождаемости, иными словами, темп рождаемости есть темп увеличения населения, он измеряется количеством родившихся за год. Реальный темп рождаемости не может рассматриваться как постоянная величина на протяжении сколько-нибудь длительного периода времени. Он зависит, в частности, от фондов и природных ресурсов, поскольку эти переменные влияют на материальный уровень жизни, плотность населения, обеспеченность пищей и уровень загрязнения. Такого рода влияния со стороны других частей системы вводятся *множителями*, которые модифицируют базисный темп рождаемости. При нормальных условиях множители не должны изменять базисный темп рождаемости, т.е. они равны 1. Если условия оказываются более благоприятными, чем нормальные, то множитель должен быть больше 1, если менее – то меньше 1.

Материальный уровень жизни. Материальный уровень жизни есть безразмерная величина, которая описывает степень изменения эффективности относительной величины фондов на душу населения в сравнении с ее начальным значением на год разработки модели. Его значение определяется как отношение эффективности относительной величины фондов (на душу населения) к нормальной эффективности относительной величины фондов. Единица фондов в модельной системе определяется как количество фондов на душу населения на начальный год.

Множитель зависимости темпа рождаемости от материального уровня жизни. Этот множитель модифицирует темп рождаемости в зависимости от изменений в материальном уровне жизни. Зависимость, выбранная в модели, представлена на рис. 8.2.

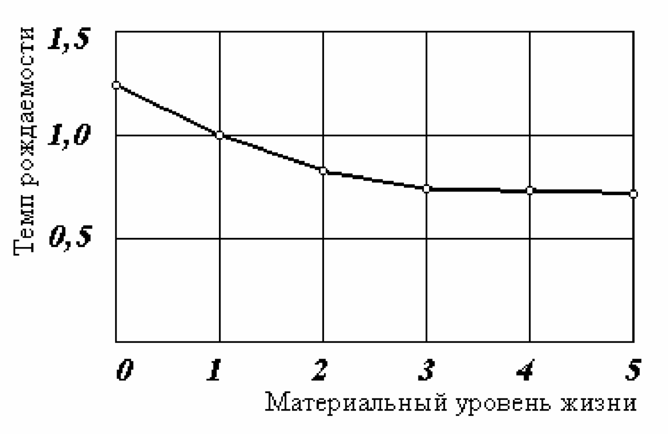


Рис.8.2. Изменение множителя темпа рождаемости в зависимости от материального уровня жизни

Выбор значения материального уровня жизни равного 1 означает, что агрегированное по всему миру количество промышленных товаров на душу населения равно среднему мировому значению на начальный год, в котором множитель темпа рождаемости принимается равным также 1. Вид оставшейся части кривой по обе стороны от этой точки зависит от предположений о том, как будет изменяться темп рождаемости при изменении материального уровня жизни.

Зависимость на рис.8..2. была построена по статистическим данным стран с разным уровнем жизни в годы разработки модели (70-е годы прошлого века). Материальный уровень жизни включает в себя и здоровье населения, и уровень медицинского обслуживания, и санитарные условия, и все другие достижения цивилизации. Однако с ростом материального уровня жизни не происходит увеличения темпов роста рождаемости, что вызвано иными причинами поведенческого свойства населения: требования к качеству питания, воспитания, образования, досуга и т.д. Эти множители тоже могут быть введены в модель мировой системы.

Природные ресурсы – системный уровень. Он связан только с одним потоком, уменьшающим его темпом потребления. В мировой модели Форрестера природные ресурсы включают в себя только невосполнимые ресурсы и не включают, например, лес и другие ресурсы, которые могут возобновляться; последние классифицируются как часть сельскохозяйственного сектора.

Темп использования природных ресурсов определяется произведением следующих сомножителей: численностью населения, нормальным потреблением природных ресурсов и множителем зависимости добычи природных ресурсов от материального уровня жизни, который растет, когда возрастает уровень жизни.

Темп смертности вводится в модель аналогично темпу рождаемости и корректируется при помощи ряда множителей: зависимости темпа смертности от материального уровня жизни, загрязнения, уровня питания, плотности населения.

Относительный уровень питания. Относительный уровень питания определяет количество пищи на душу населения и представляет собой частное от деления значения пищевого потенциала фондов на нормальный уровень питания, умноженный на три множителя: зависимости производства продуктов питания от плотности населения, от уровня загрязнения и от коэффициента питания. При нормальных условиях эти множители равны 1, более или менее благоприятные условия должны соответственно вызывать изменения в значении уровня питания на душу населения.

Относительная величина фондов в сельском хозяйстве. Количество пищи на душу населения предполагается зависящим от количества фондов на душу населения в сельском хозяйстве, ее величина определяется как относительная величина фондов (фондовооруженность), умноженная на часть фондов в сельском хозяйстве и разделенная на нормальную часть фондов в сельском хозяйстве. Фондовооруженность является относительной величиной, определяет количество единиц фондов на душу населения и измеряется в единицах количества фондов на душу населения по отношению к начальному году моделирования.

Фонды – один из системных уровней. Этот уровень образуется за счет накопления капиталовложений (фондообразования) и за счет уменьшения фондов вследствие их износа.

Помимо перечисленных в модели используются десятки переменных.

8.5. Структура модели мировой системы

Как было сказано, модель мировой системы Дж. Форрестера связывает воедино пять уровней переменных – население, природные ресурсы, капиталовложения, фонды в сельском хозяйстве и загрязнение. Подробная схема мировой модели приведена в книге ее автора [12, с.32-33], на рисунке 8.3.показана упрощенная структурная схема модели, позволяющая уяснить принципы взаимодействия уровней модели.

Каждый из уровней моделей включает в себя положительную и (или) отрицательную петлю обратной связи, определяющую темп изменения уровня. Для уровня населения положительная петля определяет темп рождаемости, отрицательная – темп смертности. Для уровня капиталовложений две петли обратных связей описывают регулирующие воздействия на капиталовложения. Аналогично вводится в модель взаимодействие прочих уровней и темпов.

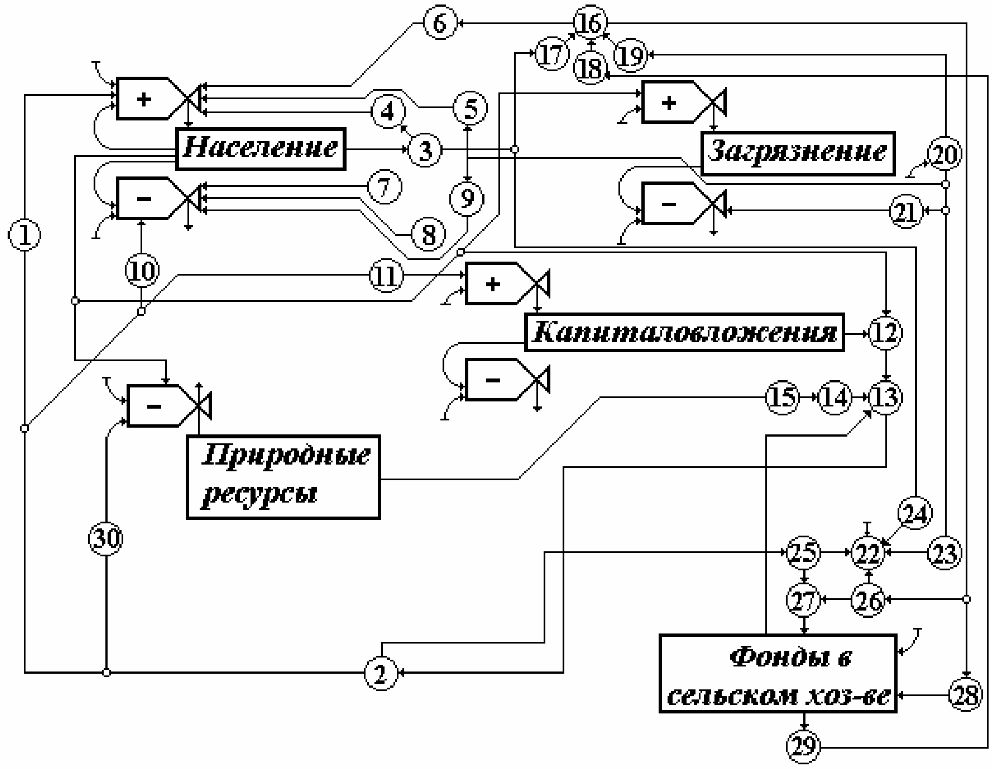


Рис.8.3. Структурная схема мировой модели. Обозначения:

1 – зависимость рождаемости от уровня жизни; 2 – уровень жизни; 3 – относительная плотность; 4 – зависимость темпа рождаемости от плотности; 5 – зависимость темпа рождаемости от загрязнения; 6 – зависимость темпа рождаемости от питания; 7 – зависимость темпа смертности от плотности; 8 – зависимость темпа смертности от питания; 9 – зависимость темпа смертности от загрязнения; 10 – зависимость темпа смертности от материального уровня жизни; 11 – зависимость капиталовложений от уровня жизни; 12 – относительная величина фондов; 13 – эффективность относительной величины фондов; 14 – зависимость добычи природных ресурсов; 15 –остающаяся часть природных ресурсов; 16 – относительный уровень питания; 17 – зависимость производства питания от плотности; 18 – пищевой потенциал фондов; 19 – зависимость производства питания от загрязнения; 20 – относительное загрязнение; 21 – время поглощения загрязнения; 22 – качество жизни; 23 – зависимость качества жизни от загрязнения; 24 – зависимость качества жизни от плотности; 25 – зависимость качества жизни от материального уровня жизни; 26 – зависимость качества жизни от питания; 27 – доля капиталовложений в зависимости от качества жизни; 28 – предписываемая относительным уровнем питания часть фондов; 29 – относительная величина фондов в сельском хозяйстве; 30 – зависимость добычи природных ресурсов от материального уровня жизни.

Все уровни и переменные так или иначе оказывают взаимное влияние друг на друга, что учитывается в модели при помощи множителей. Следует отметить, что множители в модели в свою очередь являются функциональными зависимостями, описывающими влияние одних переменных на другие. Например, изменение темпа рождаемости в зависимости от других переменных определяется множителем зависимости темпа рождаемости от плотности населения, множителем зависимости темпа рождаемости от относительного уровня питания, множителем зависимости темпа рождаемости от загрязнения, множителем зависимости темпа рождаемости от уровня жизни и т.п. Эти зависимости устанавливаются или при помощи выдвижения гипотез, или при помощи анализа предыдущих наблюдений над развитием мировой системы, или экспериментальным путем.

Обучающимся предлагается внимательно проанализировать рис. 8.3. и сделать выводы о том, взаимодействие между какими переменными учитывалось в модели мировой системы и при помощи каких множителей описывалось это взаимодействие.

8.6. Основные результаты экспериментов на модели мировой системы

Элементы модели, описанной в предыдущем разделе, определяют правила взаимодействия между ними. Эти правила предписывают, каким именно образом каждый блок системы функционирует под влиянием других блоков. Имитационная модель мировой системы представляет некоторую информацию о реальном мире. Эта информация складывается из двух частей – информации о поведении какого-то элемента системы или всей системы в целом. При проведении моделирования разработчиков модели прежде всего интересовало ее поведение при возникновении противоречий между экспоненциальным ростом уровней системы и фиксированными окружающими условиями.

Предполагается, что достаточно хорошо известны причины, вызывающие рост народонаселения и экономики. Известны более или менее точно физические пределы и естественные ресурсы планеты и предельно допустимый уровень загрязнения. Дж. Форрестер ставит вопрос: « Но что произойдет, когда рост экономики и народонаселения приблизится к фиксированным природой пределам и сменится неустойчивой формой равновесия?» Очевидно, что экспоненциальный рост населения не беспределен, вопрос состоит только в том, когда и как он прекратится, а не в том, прекратится ли оно вообще. Мировая модель Форрестера содержит четыре параметра, способных ограничить рост населения – это истощение природных ресурсов, увеличение уровня загрязнения, перенаселенность, нехватка продуктов питания. При экспериментировании было проанализировано воздействие этих параметров на рост численности населения и исследованы различные виды равновесия, которые могли бы быть созданы путем проведения соответствующих мероприятий.

Проведение экспериментов на модели позволило определить сценарии развития перечисленных ниже видов кризисов.

Истощение природных ресурсов. При достижении населением своей максимальной численности начнется процесс убывания численности населения, вызванный истощением природных ресурсов. Расчеты показывают, что при сохранении существующих темпов ресурсов может произойти их полное истощение уже к 2150 г. До опытов Форрестера анализ мировой системы часто основывался на сравнении существующего положения с ее предельными возможностями. При таком подходе сиюминутные потребности кажутся значительно меньшими в сравнении с имеющимися запасами ресурсов, но при этом упускаются из вида два фактора. Во-первых, потребности человечества возрастают в два раза каждые 20-30 лет; во-вторых, последствия начинающегося кризиса начинают проявляться значительно раньше, чем достигается сама кризисная ситуация.

Кризис загрязнения. В одном из экспериментов с моделью было сделано предположение, что скорость использования естественных ресурсов была резко снижена, например, в результате достижений научно-технического прогресса и перехода на новые технологии. В этом случае в системе возникает другая сила, подавляющая рост – развивающийся кризис загрязнения, которое резко возрастает, когда скорость загрязнения превышает скорость природных механизмов очистки.

Проблема перенаселения. Следующим этапом экспериментирования с моделью явилось предположение, что человечеству удастся справиться с проблемой загрязнения. В этом случае первоначально будет происходить экспоненциальное увеличение численности населения. Но постепенно все более и более начинают проявляться сдерживающие факторы роста. Возрастающие перенаселенность, спрос на продукты питания и необходимость использовать менее продуктивные для сельского хозяйства земли приводят к дополнительным капиталовложениям в производство продуктов питания. Кризис перенаселения по результатам моделирования не носит нарастающего характера, который наблюдался для случая загрязнения.

Уменьшение относительного уровня питания. В ряде вычислительных экспериментов не учитывалось влияние плотности населения. В этом случае фактором, ограничивающим рост численности населения, становится недостаток продуктов питания. Со снижением материального уровня жизни произойдет падение относительного уровня питания, что сдерживает рост численности населения.

Несмотря на вероятность развития мировой системы со столь негативными итогами, результаты экспериментов позволили сделать вывод, что глобальное равновесие системы в принципе возможно. Результатом имитационных экспериментов с моделью являлось не точные рекомендации по планированию развития человеческой цивилизации, а привлечение внимания к исследованиям динамики ее роста и развития.

Джей Форрестер подчеркивает, что результаты моделирования не следует воспринимать как точное предсказание пути развития сегодняшнего мира, напротив, их надо принимать как одно из возможных поведений мировой системы. Конечно, сами по себе опасности возникновения тех или иных катастрофических последствий для населения планеты прогнозировались давно. Однако, моделирование этих процессов позволяет выработать стратегию развития с целью наилучшего поведения в постоянно меняющихся условиях развития мировой системы.

Вопросы для самопроверки

1. Какие основные элементы взаимоувязаны в динамической мировой модели Дж. Форрестера?
2. Как называется замкнутая цепочка взаимодействия, которая связывает исходное действие с его результатом, изменяющим характеристики окружающих условий и которые, в свою очередь, являются «информацией», вызывающей дальнейшие изменения.
3. Какие два типа переменных включает в себя петля обратной связи?
4. Какая переменная системы называется уровнем?
5. Какая переменная системы называется темпом?
6. Какие основные уровни были выбраны Дж. Форрестером для построения структуры динамической мировой модели?
7. К каким классам моделей относится мировая модель по конкретному предназначению?
8. К каким классам моделей относится мировая модель по фактору времени?
9. К каким классам моделей относится мировая модель по целевому назначению?
10. Влияние каких параметров, способных ограничить рост населения, было исследовано при проведении имитационных экспериментов на мировой модели?

Тема 9. Метод Монте-Карло и проверка статистических гипотез

Определение: Методом Монте-Карло называется численные методы решения математических задач при помощи моделирования случайных величин.
Название "Монте-Карло" произошло от города "Монте-Карло", известного своими казино, т.к. простейшим прибором для генерирования случайных чисел служит игральная рулетка.

Статистические испытания по методу Монте-Карло представляют собой простейшее имитационное моделирование при полном отсутствии каких-либо правил поведения. Получение выборок по методу Монте-Карло - основной принцип компьютерного моделирования систем, содержащих стохастические или вероятностные элементы. Зарождение метода связано с работой фон Неймана и Улана в конце 1940-х гг., когда они ввели для него название «Монте-Карло» и применили его к решению некоторых задач экранирования ядерных излучений. Этот математический метод был известен и ранее, но свое второе рождение нашел в Лос-Аламосе в закрытых работах по ядерной технике, которые велись под кодовым обозначением «Монте-Карло». Применение метода оказалось настолько успешным, что он получил распространение и в других областях, в частности в экономике.

Поэтому многим специалистам термин «метод Монте-Карло» иногда представляется синонимом термина «имитационное моделирование», что в общем случае неверно. Имитационное моделирование - это более широкое понятие, и метод Монте-Карло является важным, но далеко не единственным методическим компонентом имитационного моделирования.

Согласно методу Монте-Карло проектировщик может моделировать работу тысячи сложных систем, управляющих тысячами разновидностей подобных процессов, и исследовать поведение всей группы, обрабатывая статистические данные. Другой способ применения этого метода заключается в том, чтобы моделировать поведение системы управления на очень большом промежутке модельного времени (несколько лет), причем астрономическое время выполнения моделирующей программы на компьютере может составить доли секунды. Рассмотрим метод Монте-Карло подробнее.

В различных задачах, встречающихся при создании сложных систем, могут использоваться величины, значения которых определяются случайным образом. Примерами таких величин являются:

- случайные моменты времени, в которые поступают заказы на фирму;
- загрузка производственных участков или служб объекта экономики;
- внешние воздействия (требования или изменения законов, платежи по штрафам и др.);
- оплата банковских кредитов;
- поступление средств от заказчиков;

В качестве соответствующих им переменных могут использоваться число, совокупность чисел, вектор или функция. Одной из разновидностей метода Монте-Карло при численном решении задач, включающих случайные переменные, является метод статистических испытаний, который заключается в моделировании случайных событий.

Метод Монте-Карло основан на статистических испытаниях и по природе своей является экстремальным, может применяться для решения полностью детерминированных задач, таких, как обращение матриц, решение дифференциальных уравнений в частных производных, отыскание экстремумов и численное интегрирование.

При вычислениях методом Монте-Карло статистические результаты получаются путем повторяющихся испытаний. Вероятность того, что эти результаты отличаются от истинных не более чем на заданную величину, есть функция количества испытаний.

В основе вычислений по методу Монте-Карло лежит случайный выбор чисел из заданного вероятностного распределения. При практических вычислениях эти числа берут из таблиц или получают путем некоторых операций, результатами которых являются псевдослучайные числа с теми же свойствами, что и числа, получаемые путем случайной выборки. Имеется большое число вычислительных алгоритмов, которые позволяют получить длинные последовательности псевдослучайных чисел.

Применение метода Монте-Карло может дать существенный эффект при моделировании развития процессов, натурное наблюдение которых нежелательно или невозможно, а другие математические методы применительно к этим процессам либо не разработаны, либо неприемлемы из-за многочисленных оговорок и допущений, которые могут привести к серьезным погрешностям или неправильным выводам. В связи с этим необходимо не только наблюдать развитие процесса в нежелательных направлениях, но и оценивать гипотезы о параметрах нежелательных ситуаций, к которым приведет такое развитие, в том числе и параметрах рисков.

Существуют различные методы проверки статистических гипотез. Наиболее широко используются на практике критерии:

- согласия χ^2 ((хи-квадрат);
- Крамера-фон Мизеса;
- Колмогорова-Смирнова.

Критерий χ^2 предпочтителен, если объемы выборок N, в отношении которых проводится анализ, велики. Это мощное средство, если N > 100 значений. Однако при анализе экономических ситуаций иногда бывает довольно трудно (или невозможно) найти 100 одинаковых процессов, развивающихся с различными исходными данными. Сложность заключается не только в том, что не бывает одинаковых объектов экономики: даже если такие объекты имеются, то к исходным данным относятся не только исходные вероятностные данные и особенности структуры объекта, но и сценарий развития процессов в этом объекте и в тех объектах внешней среды, с которыми он взаимодействует (процессы рынка, указы правительства, принятие новых законов, требования налоговых органов, платежи в бюджеты различных уровней). При относительно малых объемах выборок этот критерий вообще неприменим.

Критерий Крамера-фон Мизеса дает хорошие результаты при малых объемах выборок (при N < 10). Однако следует отметить два обстоятельства:

1) при N < 10, каким бы методом ни пользоваться, вопрос о доверительной вероятности при проверке статистической гипотезы решается плохо (эта вероятность мала при значительных размерах доверительных интервалов);

2) метод Монте-Карло используется как раз для того, чтобы недостающие данные собрать с помощью специального вычислительного статистического инструментария и компьютера.

Поэтому будем полагать, что реальные объемы выборок, которые можно получить, находятся в пределах 10 $\leq$ N < 100. Как указывают многие исследователи, для указанных пределов хорошие результаты дает критерий Колмогорова-Смирнова. Он применяется в тех случаях, когда проверяемое распределение непрерывно и известны среднее значение и дисперсия проверяемой совокупности. Рассмотрим подробнее методику использования этого критерия на конкретном примере.

Пример 9.1. Предположим, что нужно проверить данные, полученные (или наблюдаемые) при использовании метода Монте-Карло и приведенные в табл. 9.1, на их соответствие распределению Пуассона.

Эти данные имеют следующий смысл. На отрезке времени наблюдаем случайные события, число которых равно x. Если это распределение Пуассона, то вероят ность $P\{x = n\}$ того, что $x = n$, где n - заданное число, равна

$$P\{x = n\} = \frac{\lambda^n e^{-\lambda}}{n!},$$

где e = 2,71828;

λ – положительная константа, которая одновременно является и математическим ожиданием, и дисперсией.

Таблица 9.1

Данные для проверки гипотезы по критерию Колмогорова—Смирнова

Число событий	Наблюдаемая частота	Наблюдаемая вероятность	Теоретическая вероятность	Наблюдаемое распределение	Теоретическое распределение	Абсолютная разность
	1	2	3	4	5	6
0	315	0,619	0,571	0,619	0,571	0,048
1	142	0,279	0,319	0,898	0,890	0,008
2	40	0,078	0,089	0,976	0,979	0,003
3	9	0,018	0,017	0,994	0,996	0,002
4	2	0,004	0,003	0,998	0,999	0,001
5	1	0,002	0,001	1,000	1,000	0,000

Предположим, что λ = 0,55711. Сформулируем гипотезу в следующем виде: Не имеется существенных различий между наблюдаемыми данными во время эксперимента и теми данными, которые должны получаться из распределения Пуассона расчетным путем со средним значением 0,5577 и N = 509.

Рассмотрим подробнее табл. 11.1. В ней строки с номерами 0,..., 5 соответствуют числу n наблюдаемых событий, n=0, ... , 5; столбец 1 содержит наблюдаемую частоту появления n событий, а столбец 2 – наблюдаемую вероятность появления n событий. В столбце 3 представлены рассчитанные по формуле значения вероятности появления n событий.

Прежде всего нужно получить два интегральных распределения (фактически приближения функций распределения). Сначала сделаем это для наблюдаемых данных: с помощью столбца 2 получим столбец 4. Затем - для теоретических данных: с помощью столбца 3 получим столбец 5.

После этих вычислений найдем абсолютные разности для всех групп значений случайной величины и с помощью столбцов 4 и 5 получим столбец 6. В последнем столбце наибольшая абсолютная разность 0,048 получается в группе, соответствующей нулевому числу событий.

Далее необходимо найти так называемое критическое значение D_{extr} для проверки принятой гипотезы. Таблица критических чисел многократно переиздавалась. Критические числа в виде, удобном для выполняемой проверки, приведены в табл. 1.2. Абсолютную разность 0,048 необходимо сравнить с критическим значением, найденным по табл. 9.2.

При N=509 и значении индекса критического числа D_α α = 0,05 получается критическое значение

$$D_{extr} = \frac{1,36}{\sqrt{N}} = \frac{1,36}{\sqrt{509}} = 0,0603.$$

Поскольку наибольшая разность $0{,}048 < D_{extr}$, то не отказываемся от гипотезы о том, что экспериментальное распределение - пуассоновское. Проверка статистических гипотез о соответствии «событий явлению» и «явлений поведению» дает математический инструмент для оценки «рискованного поведения» исследуемого процесса.

Таблица 9.2

Критические числа Колмогорова-Смирнова

Степень свободы N	Проверка единичной выборки*			Проверка двух выборок**	
	$D_{0.10}$	$D_{0.05}$	$D_{0.01}$	$D_{0.05}$	$D_{0.01}$
1	0,950	0,975	0,995	—	—
2	0,776	0,842	0,929	—	—
3	0,642	0,708	0,828	—	—
4	0,564	0,624	0,733	1,000	1,000
5	0,510	0,565	0,669	1,000	1,000
6	0,470	0,521	0,618	0,833	1,000
7	0,438	0,486	0,577	0,857	0,857
8	0,411	0,457	0,543	0,750	0,875
9	0,388	0,432	0,514	0,668	0,778
10	0,368	0,410	0,490	0,700	0,800
11	0,352	0,391	0,468	0,636	0,727
12	0,338	0,375	0,450	0,583	0,667
13	0,325	0,361	0,433	0,538	0,692
14	0,314	0,349	0,418	0,571	0,643
15	0,304	0,338	0,404	0,533	0,600
16	0,295	0,328	0,392	0,500	0,625
17	0,286	0,318	0,381	0,471	0,588
18	0,278	0,309	0,371	0,500	0,556
19	0,272	0,301	0,363	0,474	0,526
20	0,264	0,294	0,356	0,450	0,550
25	0,240	0,270	0,320	0,400	0,480
30	0,220	0,240	0,290	0,370	0,430
35	0,210	0,230	0,270	0,340	0,390
Более 35	$\frac{1,22}{\sqrt{N}}$	$\frac{1,36}{\sqrt{N}}$	$\frac{1,63}{\sqrt{N}}$	$1,36\sqrt{\frac{N_1+N_2}{N_1 N_2}}$	$1,63\sqrt{\frac{N_1+N_2}{N_1 N_2}}$

* Применяется для оценки степени близости выборочных значений к теоретическому распределению. Здесь N – объем выборки.

** Применяется для определения принадлежности двух выборок объемами N_1 и N_2 одному и тому же распределению.

При малых размерах выборки $N = N_1 = N_2$.

Тема 10. Моделирование случайных событий

Если моделируемый процесс подвержен влиянию случайных факторов, то их действие имитируется с помощью специально организованного розыгрыша (жребия). Таким образом строится одна случайная реализация моделируемого явления, представляющая собой как бы один результат опыта. По одному опыту, конечно, нельзя судить о закономерностях изучаемого процесса. Но при большом числе реализаций средние характеристики, вырабатываемые моделью, приобретают свойство устойчивости, которое усиливается с увеличением числа реализаций.

Бросание жребия можно осуществить вручную (выбором из таблицы случайных чисел), но удобнее это делать с помощью специальных программ, входящих в состав программного обеспечения ЭВМ. Такие программы называют *датчиками* или *генераторами* случайных чисел.

В трансляторах почти всех алгоритмических языков имеются стандартные процедуры или функции, которые генерируют случайные (точнее, псевдослучайные) величины с равномерным распределением.

В трансляторе языка Visual Basic имеется стандартная функция *RND,* возвращающая случайные вещественные числа одинарной точности в интервале (0,1).
Обращение к этой функции может иметь вид: *z - RND,* где z - возможное значение случайной величины, равномерно распределенной в интервале (0,1).

10.1. Моделирование простого события

Пусть имеется событие A, вероятность наступления которого равна P_A. Требуется выработать правило, при многократном использовании которого частота появления события стремилась бы к его вероятности. Выберем с помощью датчика случайных чисел, равномерно распределенных в интервале (0,1) некоторое число z и определим вероятность того, что $z < P_A$. Для случайной величины z с равномерным распределением справедлива следующая зависимость:

$$P(z < P_A) = \int_0^{P_A} f(x)dx = P_A,$$

где $f(x)$ – плотность вероятности случайной величины с равномерным распределением.

Таким образом, вероятность попадания случайной величины в интервал $(0, P_A)$ равна величине P_A. Поэтому если при розыгрыше число z попало в этот интервал, то следует считать, что событие A произошло. Противоположное событие (не A) произойдет с вероятностью $(1 - P_A)$ в том случае, если $z \geq P_A$.

Процедура моделирования простого события в имитационной модели описывается алгоритмом, схема которого показана на рис. 10.1.

Оператор 1 обращается к датчику случайных чисел, генерирующему случайную величину z. Оператор 2 проверяет условие $z < P_A$. Если оно выполняется, считается, что произошло событие A. В противном случае считается, что произошло противоположное событие (не A).

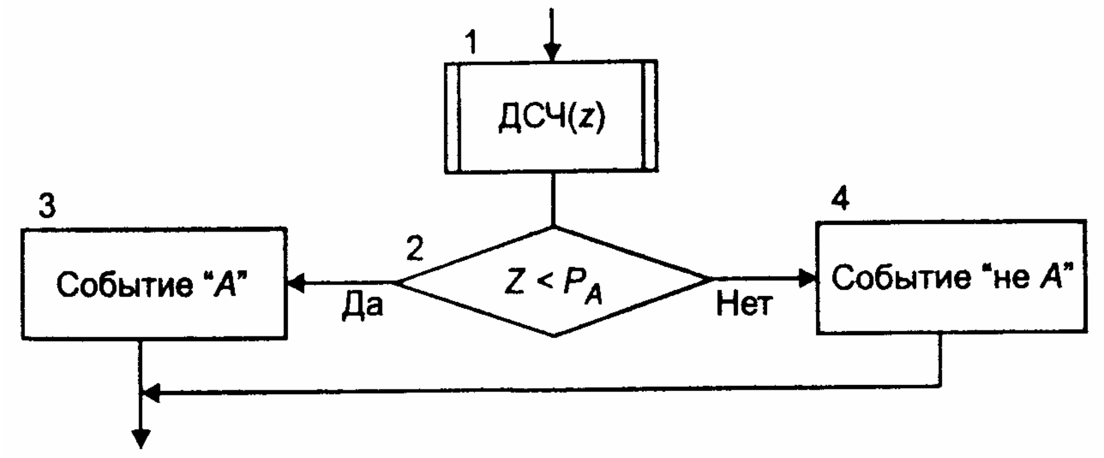


Рис. 10.1

10.2 Моделирование полной группы несовместных событий

Пусть имеется полная группа несовместных событий (ПГНС) $A_1, A_2, ..., A_k$ с вероятностями $P_1, P_2, ..., P_k$. При этом выполняется условие

$$\sum_{i=1}^{k} P_i = 1.$$

Разделим интервал (0,1) на k отрезков, длины которых составляют $P_1, P_2, ..., P_k$ (рис. 10.2).

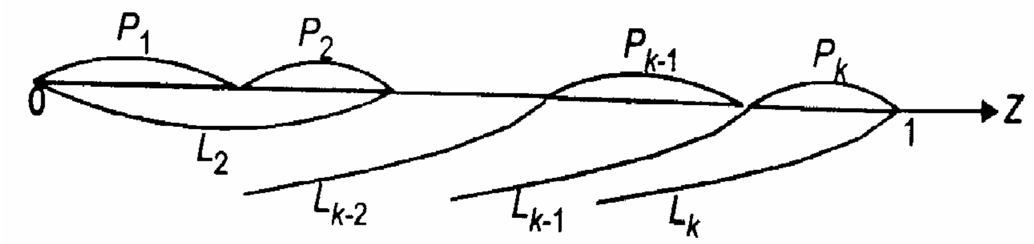


Рис. 10.2

Если случайное число z, генерированное датчиком случайных чисел с равномерным распределением в интервале (0,1), попало, например, на участок P_{k-1}, то это должно означать, что произошло событие A_{k-1} . Действительно, если обозначить

$$L_j = \sum_{i=1}^{j} P_i$$

то окажется справедливым выражение

$$P(L_{k-2} < z < L_{k-1}) = \int_{L_{k-2}}^{L_{k-1}} 1 \cdot dx = P_{k-1}.$$

Следовательно, произойдет событие, которое имеет вероятность P_{k-1}.

Процедура моделирования полной группы несовместных событий описывается алгоритмом, схема которого показана на рис. 10.3.

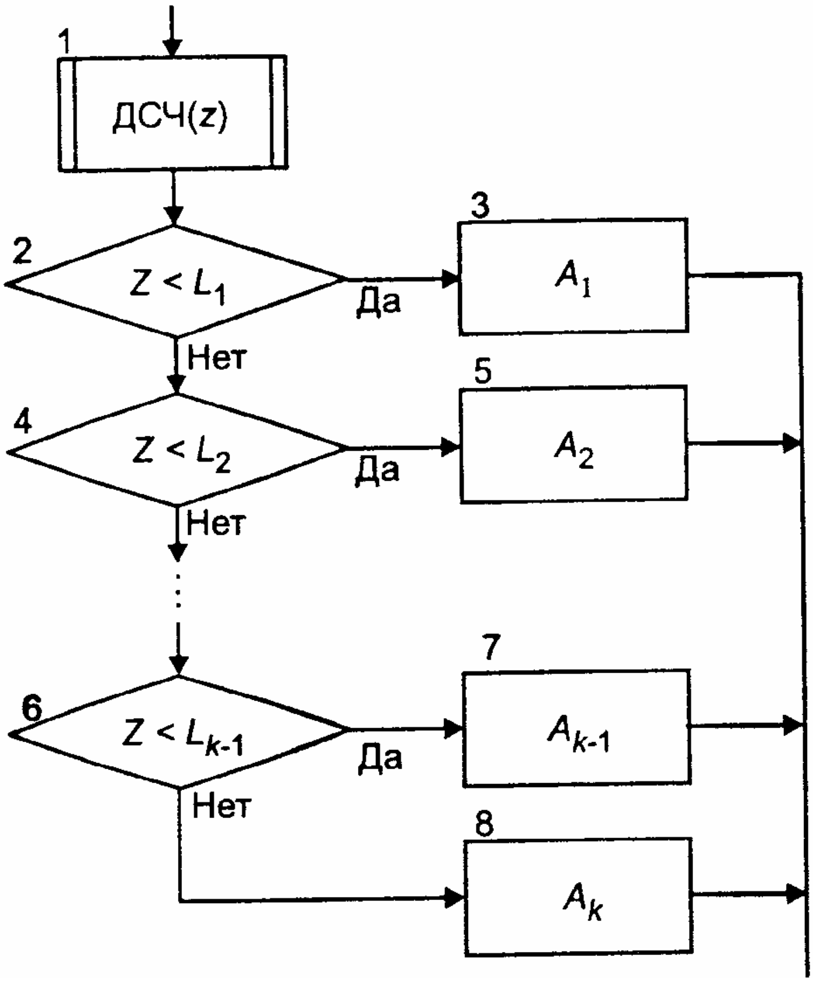


Рис. 10.3

Оператор 1 обращается к датчику случайных чисел с равномерным распределением в интервале (0,1). Условный оператор *1* проверяет условие попадания случайной величины z в интервал $(0, L_1)$. Если это условие выполняется, то считается, что произошло событие A_1. Если условие в операторе *2* не выполняется, то алгоритм осуществляет проверку условий попадания случайной величины в другие интервалы. Одно из событий $A_1, A_2, ..., A_k$ обязательно произойдет.

10.3 Моделирование дискретной случайной величины

Дискретная случайная величина может быть задана табличной зависимостью:

X	x_1	x_2	…	x_n
P	p_1	p_2	…	p_n

Здесь p_j – вероятность того, что дискретная случайная величина X примет значение x_j. При этом $p_1 + p_2 + ... + p_n = 1$. Разделим интервал (0,1) на n отрезков, длины которых пропорциональны заданным вероятностям. Если случайное число z, вырабатываемое датчиком случайных чисел, равномерно распределенных в интервале (0,1), попадет в интервал p_k, то случайная величина X примет значение x_k. Таким образом, при моделировании дискретных случайных величин фактически используется та же процедура, что и при моделировании ПГНС.

10.4 Моделирование непрерывных случайных величин

10.4.1. Метод обратной функции

Пусть имеется некоторая непрерывная случайная величина x, заданная функцией распределения $F(x)$. Можно доказать, что значения этой функции равномерно распределены в интервале (0,1). Поэтому между случайной величиной z, равномерно распределенной в том же интервале, и функцией распределения случайной величины x существует взаимно однозначное соответствие, т. е.

$$z = F(x). \tag{10.1}$$

Отсюда следует,

$$x = F^{-1}(z). \tag{10.2}$$

Следовательно, если уравнение (2.1) имеет аналитическое решение, то для моделирования случайной величины x можно использовать датчик случайных чисел, генерирующий величину z, и затем осуществить расчет по формуле (2.2).

10.4.2. Моделирование случайных величин с показательным распределением

Пусть имеется случайная величина x с показательным распределением. Функция распределения имеет вид:

$$F(x) = 1 - e^{-\lambda x},$$

где λ – параметр распределения.

Применив метод обратной функции, получим:

$$z = F(x) = 1 - e^{-\lambda x},$$

откуда

$$x = -\frac{1}{\lambda}\ln(1-z). \tag{10.3}$$

Учитывая, что случайная величина $(1-z)$ имеет также равномерное распределение в интервале (0,1), соотношение (10.3) можно заменить соотношением

$$x = -\frac{1}{\lambda}\ln(z).$$

Экспоненциальный закон распределения применяют для моделирования следующих явлений:

- времени поступления заказа на предприятие;
- посещения покупателями магазина супер-маркета;
- времени телефонных разговоров;
- срока службы деталей и узлов в компьютере.

На первый взгляд, это распределение является надуманным. Поэтому рассмотрим предельную теорему о суперпозиции потоков. Предположим, что можно наблюдать k независимых потоков событий (рис. 10.4)

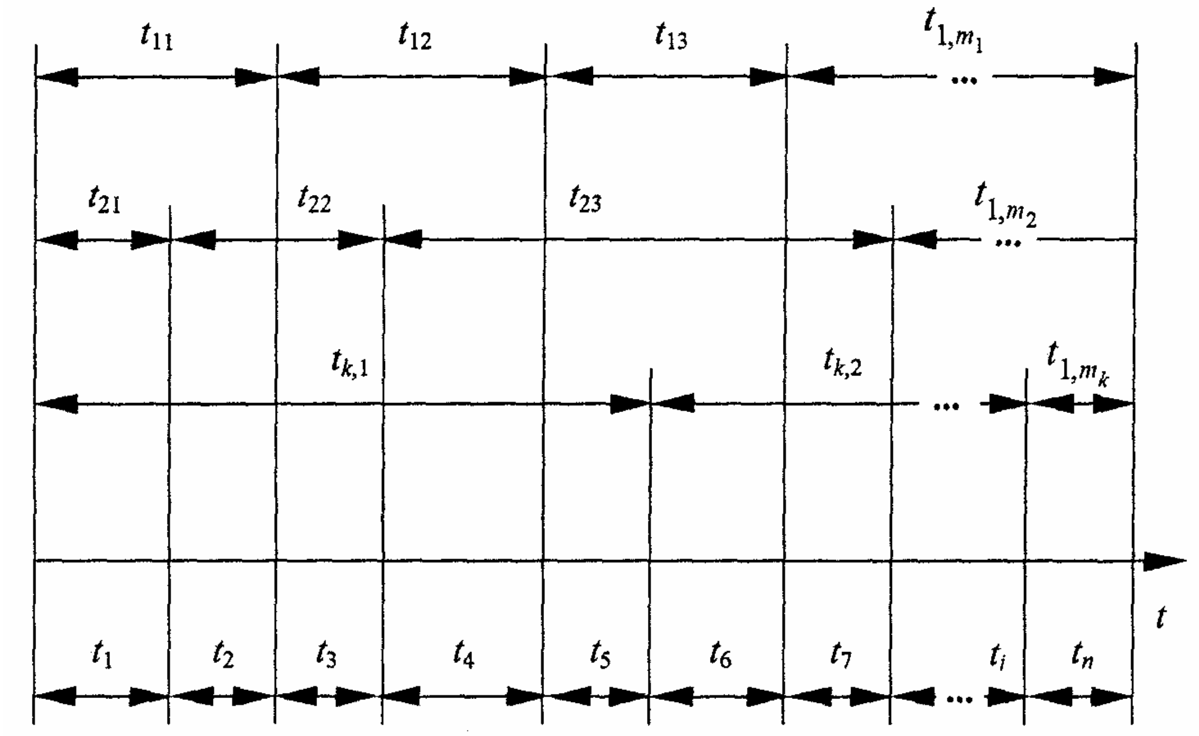


Рис. 10.4

В каждом таком потоке можно наблюдать m_j элементарных событий ($j = 1,...,k$). Интервалы между t_{kj} событиями — это независимые случайные величины, распределенные по по неизвестному закону с математическим ожиданием λ_j . Спроектируем моменты всех событий на общую ось времени и рассмотрим случайный интервал времени $t = T(k)$ между двумя событиями полученного суммарного потока, состоящего из n событий , где $n = \sum_{j=1}^{k} m_j$.

Теорема. Если сделать предельный переход и устремить $n \to \infty$, то распределение случайной величины интервала $t = T(k)$ в суммарном потоке событий, состоящем из k элементарных потоков, устремится к экспоненциальному закону с математическим ожиданием

$$M(t) = \lambda = \sum_{j=1}^{k} \lambda_j .$$

Следствие. Поток заявок, интервал поступления которых в некую систему имеет экспоненциальное распределение, является простейшим.

Пример 10.1. Допустим, что имеется некая крупная фирма. Клиенты фирмы - это физические и юридические лица. Каждый из них может иметь набор планов и расписанных дел на значительном интервале времени. Однако если рассмотреть суммарный поток обращений этих клиентов к служащим фирмы по разным вопросам, то интервал времени между двумя последовательными обращениями в соответствии с рассмотренной теоремой является случайной величиной, распределенной по экспоненциальному закону.

10.4.3. Моделирование случайных величин с равномерным распределением на произвольном интервале (*a, b*)

Датчик случайных чисел генерирует случайные величины с равномерным распределением в интервале (0,1). Если же нужно моделировать случайные величины с равномерным распределением в интервале (a,b) то можно воспользоваться методом обратной функции.

Для рассматриваемого случая выражение (2.1) примет вид:

$$z = F(x) = \frac{x-a}{b-a},$$

Откуда

$$x = a + z(b-a).$$

На практике применяется и другой способ задания равномерного распределения. Вместо границ интервала задаются среднее значение случайной величины x_{cp} и величина интервала Δx. Тогда определение возможного значения случайной величины с равномерным распределением может быть произведено по формуле

$$x = x_{cp} + \Delta x(z - 0,5).$$

Равномерное распределение можно использовать при расчетах по сетевым графикам работ, в военном деле (времени выдвижения воинской части или ее подразделения на исходный рубеж, времени марша)

10.4.4 Моделирование случайных величин с нормальным распределением

Метод обратной функции для нормального распределения неприменим, так как после подстановки соответствующей функции распределения выражение (2.2) не имеет аналитического решения. Поэтому в данном случае применяется другой метод.

Согласно центральной предельной теореме теории вероятностей при сложении достаточно большого числа одинаково распределенных независимых случайных чисел получается случайная величина, имеющая нормальное распределение.

Как показали исследования, уже при сложении более десяти случайных величин с равномерным распределением в интервале (0,1) получается случайная величина, которая с точностью, достаточной для большинства практических задач, может считаться распределенной нормально.

Процедура розыгрыша нормально распределенной случайной величины заключается в следующем.

1. Сложим 12 случайных величин с равномерным распределением в интервале (0,1), т. е. составим сумму

$$v = \sum_{i=1}^{12} z_i.$$

Использовав известные теоремы о сумме математических ожиданий и дисперсий независимых случайных величин, можно установить, что в данном случае случайная величина v имеет следующие характеристики:

математическое ожидание:

$$M(v) = \sum_{i=1}^{12} M(z_i) = 12 \cdot \frac{1}{2} = 6;$$

дисперсия:

$$D(v) = \sum_{i=1}^{12} D(z_i) = 12 \cdot \frac{1}{12} = 1.$$

среднее квадратическое отклонение: $\sigma(v) = \sqrt{D(v)} = 1$.

2. Нормируем и центрируем случайную величину v, т. е. перейдем к величине

$$\eta = \frac{v - M(v)}{\sigma(v)} = v - 6.$$

2. От нормированной и центрированной величины η перейдем к случайной величине y с заданными параметрами $M(y)$ и $\sigma(y)$ по формуле

$$y = M(y) + \sigma(y) \cdot \eta,$$

где $M(y)$ – известное математическое ожидание случайной величины y; $\sigma(y)$ – известное среднее квадратическое отклонение случайной величины y.

Любые сложные работы на объектах экономики (ввод информации из документа в компьютер, проведение переговоров, ремонт оборудования и др.) состоят из многих коротких последовательных элементарных составляющих работ. Причем количество этих составляющих настолько велико, что условие выполнимости центральной предельной теоремы не вызывает сомнений. Поэтому при оценках трудозатрат всегда справедливо предположение о том, их продолжительность — это случайная величина, которая распределена по нормальному закону.

10.4.5. Моделирование случайных величин с усеченным нормальным распределением

Усеченное нормальное распределение случайной величины x задается четырьмя параметрами: математическим ожиданием $M(x)$, средним квадратическим отклонением $\sigma(x)$, а также минимальным и максимальным значениями x_1 и x_2 (точками усечения).

Функция распределения случайной величины x определяется равенством

$$F(x) = \begin{cases} 0, & x < x_1; \\ [\Phi_0(t) - \Phi_0(t_1)] \cdot A, & x_1 < x < x_2; \\ 1, & x > x_2, \end{cases}$$

где $A = \dfrac{1}{\Phi_0(t_2) - \Phi_0(t_1)}$; $t = \dfrac{x - M(x)}{\sigma(x)}$; $t_1 = \dfrac{x_1 - M(x)}{\sigma(x)}$; $t_2 = \dfrac{x_2 - M(x)}{\sigma(x)}$.

Существуют также формулы для расчета математического ожидания, дисперсии и среднего квадратического отклонения случайной величины x. Однако с достаточной для практики точностью при моделировании случайной величины с усеченным нормальным распределением можно обойтись без расчетов по формулам.

Для определения возможных значений случайной величины с этим распределением можно использовать алгоритм, схема которого приведена на рис. 10.5.

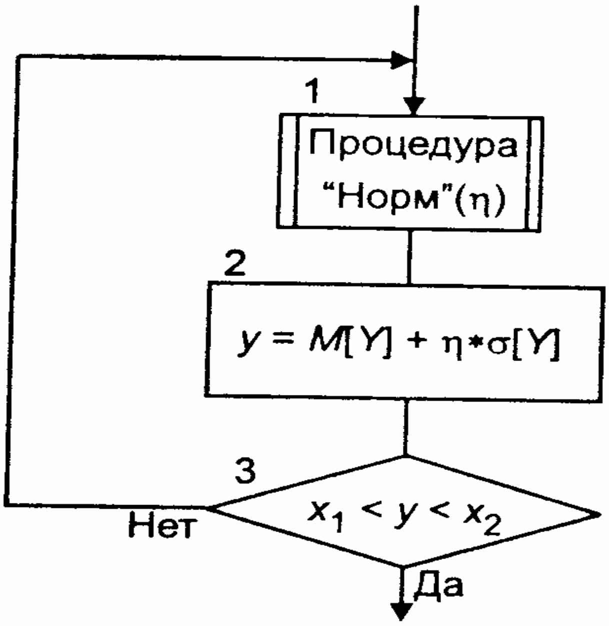


Рис. 10.5

Оператор 1 обращается к процедуре моделирования возможных значений нормированной и центрированной случайной величины η с нормальным распределением. Оператор 2 вычисляет значение случайной величины y с заданными параметрами $M(y)$ и $\sigma(y)$.

Условный оператор 3 проверяет условие попадания случайной величины y в неусеченную область. При выполнении этого условия значение случайной величины y с усеченным нормальным распределением считается найденным. В противном случае управление в алгоритме передается вновь на вход оператора 1 и генерируется другая случайная величина.

10.4.6 Моделирование случайных величин с произвольным распределением

Пусть случайная величина x задана в интервале (a_0, b_n) кусочно-постоянной функцией $f(x)$. Это значит, что интервал разбит на n частичных интервалов и плотность распределения $f_k(x)$ на каждом из них постоянна (рис. 2.6).

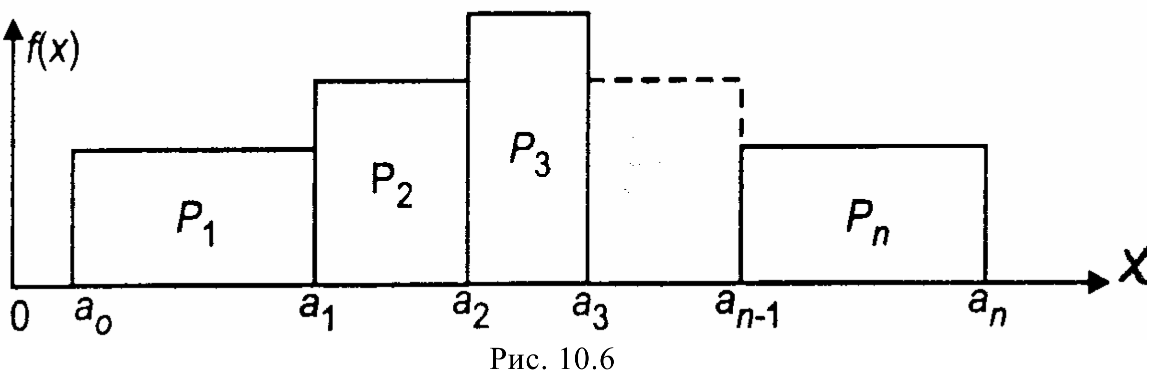


Рис. 10.6

Целесообразно выбрать величины a_k так, чтобы вероятности попадания в любой частичный интервал P_k были одинаковы, т. е.

$$\int_{a_{k-1}}^{a_k} f_k(x)dx = \frac{1}{n} \quad (k = 1, 2, ..., n)$$

Из условия постоянства функции на каждом частичном интервале следует, что случайная величина x может быть определена по формуле

$$x = a_{k-1} + z(a_k - a_{k-1}), \tag{10.4}$$

где z – возможное значение (реализация) случайной величины, равномерно распределенной в интервале (0,1); a_{k-1} – левая граница частичного интервала; a_k – правая граница частичного интервала.

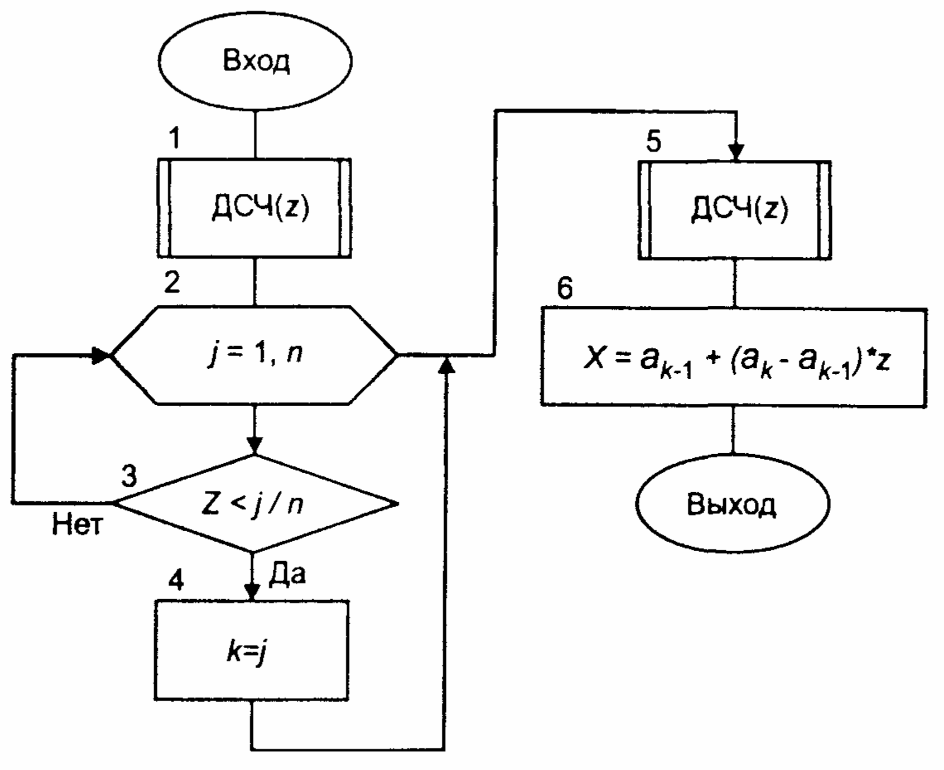


Рис. 10.7

Попадание в любой частичный интервал можно рассматривать как событие, входящее в полную группу несовместных событий. Поэтому процедура моделирования в общем случае состоит в следующем.

1. С помощью датчика случайных чисел с равномерным распределением, вырабатывающего величину z, моделируют дискретную случайную величину – номер интервала k. Вторично разыгрывают случайную величину z и определяют возможное значение случайной величины x по формуле (10.4).

Схема алгоритма показана на рис. 10.7.

Вопросы для самопроверки

1. Дайте определение метода Монте-Карло.
2. Приведите примеры характеристик систем, значения которых определяются случайным образом.
3. Перечислите критерии проверка статистических гипотез.
4. Приведите алгоритм моделирования простого события.
5. Приведите алгоритм моделирования полной группы несовместных событий.
6. Приведите алгоритм моделирования дискретной случайной величины.

7. Приведите алгоритм моделирования дискретной случайной величины.
8. В чем заключается метод обратной функции моделирования непрерывной случайной величины.
9. Приведите алгоритм моделирования случайных величин с показательным распределением.
10. Приведите алгоритм моделирования случайных величин с равномерным распределением на произвольном интервале (a,b).
11. В чем состоит суть алгоритма моделирования случайных величин с нормальным распределением.
12. Приведите алгоритм моделирования случайных величин с усеченным нормальным распределением.
13. Алгоритм моделирования случайных величин с произвольным распределением

Тема 11. Системы массового обслуживания

11.1. Основные понятия. Классификация СМО

На практике часто приходится сталкиваться с системами, предназначенными для многоразового использования при решении однотипных задач. Возникающие при этом процессы получили название процессов обслуживания, а системы — систем массового обслуживания (СМО). Примерами таких систем являются телефонные системы, ремонтные мастерские, вычислительные комплексы, билетные кассы, магазины, парикмахерские и т.п.

Каждая СМО состоит из определенного числа обслуживающих единиц (приборов, устройств, пунктов, станций), которые будем называть каналами обслуживания. Каналами могут быть линии связи, рабочие точки, вычислительные машины, продавцы и др. По числу каналов СМО подразделяют на одноканальные и многоканальные.

На вход в СМО поступает поток требований на обслуживание. В качестве таких требований могут выступать, например, клиенты или пациенты, поломки в оборудовании, готовые к упаковке изделия, телефонные вызовы, прибывающие в аэропорт самолеты.

Требования (заявки) поступают в СМО обычно не регулярно, а случайно, образуя так называемый случайный поток требований. Обслуживание заявок, вообще говоря, также продолжается какое-то случайное время. Случайный характер потока заявок и времени обслуживания приводит к тому, что СМО оказывается загруженной неравномерно: в какие-то периоды времени скапливается очень большое количество заявок (они либо становятся в очередь, либо покидают СМО необслуженными), в другие же периоды СМО работает с недогрузкой или простаивает. Предметом теории массового обслуживания является построение математических моделей, связывающих заданные условия работы СМО (число каналов, их производительность, характер потока заявок и т.п.) с показателями эффективности СМО, описывающими ее способность справляться с потоком заявок.

В качестве показателей эффективности СМО используются: среднее число заявок, обслуживаемых в единицу времени; среднее число заявок в очереди; среднее время ожидания обслуживания; вероятность отказа в обслуживании без ожидания; вероятность того, что число заявок в очереди превысит определенное значение и т.п. Здесь и в дальнейшем средние величины понимаются как математические ожидания соответствующих случайных величин.

СМО делят на два основных типа (класса): СМО с отказами и СМО с ожиданием (очередью). В СМО с отказами заявка, поступившая в момент, когда все каналы заняты, получает отказ, покидает СМО и в дальнейшем процессе обслуживания не участвует (например, заявка на телефонный разговор в момент, когда все каналы заняты, получает отказ и покидает СМО необслуженной). В СМО с ожиданием заявка, пришедшая в момент, когда все каналы заняты, не уходит, а становится в очередь на обслуживание.

СМО с ожиданием подразделяются на разные виды в зависимости от того, как организована очередь: с ограниченной или неограниченной длиной очереди, с ограниченным временем ожидания и т.п.

При анализе или прогнозировании деятельности СМО следует учитывать не только средние характеристики входящих потоков и процессов обслуживания, но и их внутреннюю структуру.

Системы массового обслуживания обладают различной структурой. Однако в них обычно можно выделить четыре основных звена: входящий поток требований, накопитель, узлы обслуживания, выходящий поток.

Характеристики входящего потока требований важны для правильной организации процесса обслуживания. Выходящий поток представляет дополнительный интерес в том случае, если он весь или какая-то его часть оказывается составной частью потока, входящего в другую СМО.

Накопитель – место, где поступившие требования ждут начала обслуживания. Накопитель может быть ограничен по объему, вмещать ограниченное число требований. Требования, для которых в накопителе не нашлось места, либо оказываются в другом, внешнем накопителе, либо вообще покидают систему. Накопителем может быть, например, зал для ожидающих клиентов в парикмахерской или бункер для ждущих обработки деталей перед станком. Иногда ограничение связано не с физическим объемом накопителя, а с очередью. Например, в очередной тур конкурса должно пройти заранее известное число кандидатов. Ограничение очереди при этом, по сути, эквивалентно физической ограниченности накопителя.

Ограниченность накопителя может проявляться не только в пространственных, но и во временных характеристиках. Требование, пробыв некоторое время в очереди, может покинуть ее, не дождавшись начала обслуживания. Оно может уйти в другую очередь или вообще погибнуть как требование на обслуживание в данной системе (например, если речь идет об обработке скоропортящихся продуктов).

Требования, находящиеся в накопителе, образуют либо одну общую очередь ко всем узлам обслуживания, либо раздельные очереди. Очереди бывают однородными или специализированными (в соответствии со специализацией узлов обслуживания); в некоторых случаях допустим переход требования из одной очереди в другую, в некоторых—запрещен.

Важное значение имеет дисциплина обслуживания, определяющая порядок выбора заявок из числа поступивших и порядок распределения их между свободными каналами. По этому признаку обслуживание заявки может быть организовано по принципу "первая пришла – первая обслужена", "последняя пришла – первая обслужена" (такой порядок может применяться, например, при извлечении для обслуживания изделий со склада, ибо последние из них оказываются часто более доступными) или обслуживание с приоритетом (когда в первую очередь обслуживаются наиболее важные заявки). Приоритет может быть как абсолютным, когда более важная заявка "вытесняет" из-под обслуживания обычную заявку (например, в случае аварийной ситуации плановые работы ремонтных бригад прерываются до ликвидации аварии), так и относительным, когда более важная заявка получает лишь "лучшее" место в очереди.

Велико разнообразие и в организации процесса обслуживания. В системе может быть как один узел обслуживания (небольшой буфет), так и несколько (отделы магазина). Число узлов может даже не быть постоянным: каждая машина такси, находящаяся в данный момент на стоянке, может рассматриваться как отдельный узел. Различают однородные (способные обслужить любое требование, поступающее в систему) и специализированные узлы. Даже будучи однородными, они могут отличаться значениями своих характеристик. Среди таких характеристик одной из наиболее существенных является интенсивность обслуживания - среднее число требований, которое способен обслужить узел в единицу времени.

Далее, узлы могут работать параллельно (кассы в универсаме, причалы в порту), последовательно (конвейер) или смешанным образом. В процессе обслуживания они могут работать независимо или взаимодействовать, помогать друг другу. Узлы могут выходить из строя и поступать на восстановление (ремонт, лечение) в другую СМО (уже в качестве требований на обслуживание).

Обычно в каждый момент времени узел обслуживает не более одного требования. Однако бывают СМО, в которых узлы обслуживают сразу группы требований: преподаватель в вузе или экскурсовод в музее.

Иногда при рассмотрении СМО нас не интересует дальнейшая судьба обслуженных требований; требования, поступающие в систему, не связаны с требованиями, уходящими из нее. В некоторых же случаях следует учитывать, что обслуженные требования после некоторой задержки (обычно со случайной, не известной заранее продолжительностью) опять поступают на вход. В первом случае СМО называются незамкнутыми, во втором – замкнутыми. Замкнутой системой является, например, поликлиника, обслуживающая данную территорию, или бригада ремонтных рабочих, закрепленная за определенной группой оборудования.

Структуры, связанные с накопителем и организацией процесса обслуживания, весьма разнообразны. Мы рассмотрим лишь некоторые «типовые» структуры СМО, которые служат основой как для изучения других, более сложных структур, так и для понимания общих задач математического моделирования работы систем обслуживания.

Почти для всех рассмотренных далее СМО предполагается, что входящий поток является пуассоновским, а длительность обслуживания распределена по экспоненциальному закону. Эти предположения, с одной стороны, выполняются для достаточно широкого класса реальных систем обслуживания и, с другой стороны, позволяют применить к изучению таких систем теорию марковских процессов.

Отметим, что в теории массового обслуживания разные авторы используют различную терминологию. Следует иметь в виду, что термины «заявка», «клиент» – синонимы термина «требование», а термины «канал», «линия», «прибор» - синонимы термина «узел обслуживания».

11.2 Понятие марковского случайного процесса

Процесс работы СМО представляет собой случайный процесс.

Под случайным (вероятностным или стохастическим) процессом понимается процесс изменения во времени состояния какой-либо системы в соответствии с вероятностными закономерностями.

Процесс называется процессом с дискретными состояниями, если его возможные состояния $S_1, S_2, S_3, \ldots$ — можно заранее перечислить, а переход системы из состояния в состояние происходит мгновенно (скачком). Процесс называется процессом с непрерывным временем, если моменты возможных переходов системы из состояния в состояние не фиксированы заранее, а случайны.

Процесс работы СМО представляет собой случайный процесс с дискретными состояниями и непрерывным временем. Это означает, что состояние СМО меняется скачком в случайные моменты появления каких-то событий (например, прихода новой заявки, окончания обслуживания и т.п.).

Математический анализ работы СМО существенно упрощается, если процесс этой работы — марковский. Случайный процесс называется марковским или случайным процессом без последствия, если для любого момента времени вероятностные характеристики процесса в будущем зависят только от его состояния в данный момент t_0 и не зависят от того, когда и как система пришла в это состояние.

Пример марковского процесса: система S – счетчик в такси. Состояние системы в момент t характеризуется числом километров (десятых долей километров), пройденных автомобилем до данного момента. Пусть в момент t_0 счетчик показывает S_0. Вероятность того, что в момент $t > t_0$ счетчик покажет то или иное число километров (точнее, соответствующее число рублей) S_1, зависит от S_0, но не зависит от того, в какие моменты времени изменялись показания счетчика до момента t_0.

Многие процессы можно приближенно считать марковскими. Например, процесс игры в шахматы: система S – группа шахматных фигур. Состояние системы характеризуется числом фигур противника, сохранившихся на доске в момент t_0. Вероятность того, что в момент $t > t_0$ материальный перевес будет на стороне одного из противников, зависит в первую очередь от того, в каком состоянии находится система в данный момент t_0, а не от того, когда и в какой последовательности исчезли фигуры с доски до момента t_0.

В ряде случаев предысторией рассматриваемых процессов можно просто пренебречь и применять для их изучения марковские модели.

При анализе случайных процессов с дискретными состояниями удобно пользоваться геометрической схемой – так называемым графом состояний. Обычно состояния системы изображаются прямоугольниками (кружками), а возможные переходы из состояния в состояние — стрелками (ориентированными дугами), соединяющими состояния.

♦**Пример 11.1**. Построить граф состояний следующего случайного процесса: устройство S состоит из двух узлов, каждый из которых в случайный момент времени может выйти из строя, после чего мгновенно начинается ремонт узла, продолжающийся заранее неизвестное случайное время.

Решение. Возможные состояния системы: S_0 – оба узла исправны; S_1 – первый узел ремонтируется, второй исправен; S_2 – второй узел ремонтируется, первый исправен; S_3 – оба узла ремонтируются. Граф системы приведен на рис. 11.1.

Стрелка, направленная, например, из S_0 в S_1, означает переход системы в момент отказа первого узла, из S_1 в S_0 — переход в момент окончания ремонта этого узла.

На графе отсутствуют стрелки из S_0 в S_3 и из S_1 в S_2. Это объясняется тем, что выходы узлов из строя предполагаются независимыми друг от друга и, например, вероятностью одновременного выхода из строя двух узлов (переход из S_0 в S_3) или одновременного окончания ремонтов двух узлов (переход из S_3 в S_0) можно пренебречь. ►

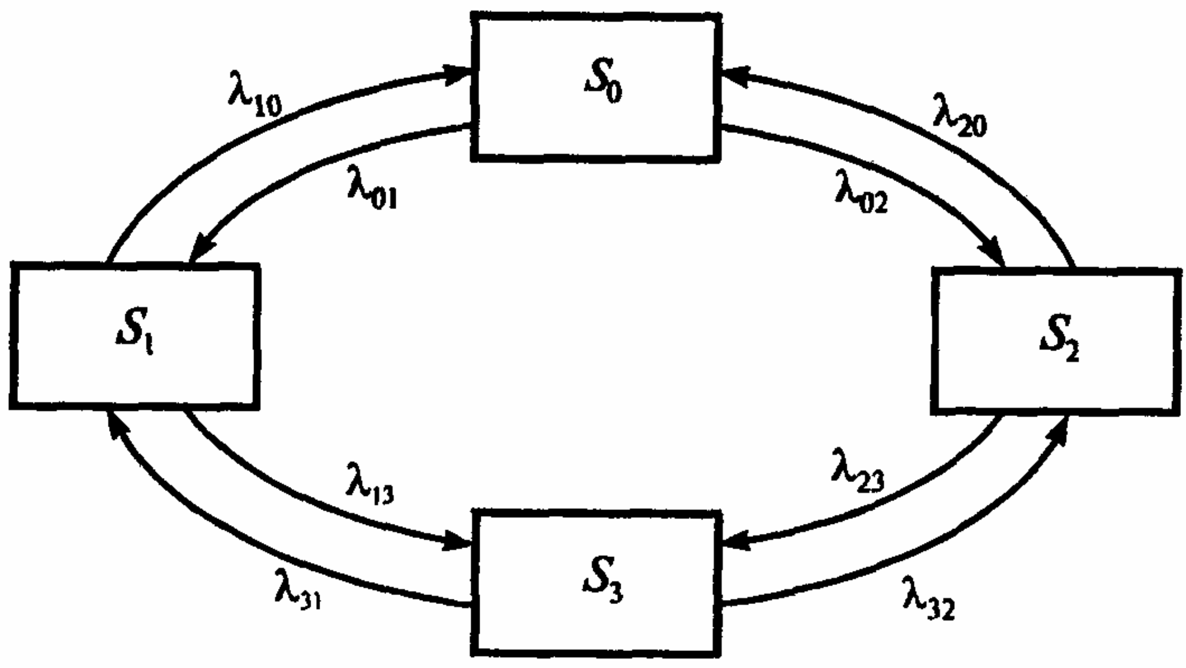


Рис. 11.1

Для математического описания марковского случайного процесса с дискретными состояниями и непрерывным временем, протекающего в СМО, познакомимся с одним из важных понятий теории вероятностей — понятием потока событий.

11.3 Потоки событий

Под *потоком событий* понимается последовательность однородных событий, следующих одно за другим в какие-то случайные моменты времени (например, поток вызовов на телефонной станции, поток отказов ЭВМ, поток покупателей и т.п.).

Поток характеризуется *интенсивностью* λ — частотой появления событий или средним числом событий, поступающих в СМО в единицу времени.

Поток событий называется *регулярным,* если события следуют одно за другим через определенные равные промежутки времени. Например, поток изделий на конвейере сборочного цеха (с постоянной скоростью движения) является регулярным.

Поток событий называется *стационарным,* если его вероятностные характеристики не зависят от времени. В частности, интенсивность стационарного потока есть величина постоянная: $\lambda(t)=\lambda$. Например, поток автомобилей на городском проспекте не является стационарным в течение суток, но этот поток можно считать стационарным в течение суток, скажем, в часы пик. Обращаем внимание на то, что в последнем случае фактическое число проходящих автомобилей в единицу времени (например, в каждую минуту) может заметно отличаться друг от друга, но среднее их число будет постоянно и не будет зависеть от времени.

Поток событий называется *потоком без последействия,* если для любых двух непересекающихся участков времени продолжительностью τ_1 и τ_2 – число событий, попадающих на один из них, не зависит от числа событий, попадающих на другие. Например, поток пассажиров, входящих в метро, практически не имеет последействия. А, скажем, поток покупателей, отходящих с покупками от прилавка, уже имеет последействие (хотя бы потому, что интервал времени между отдельными покупателями не может быть меньше, чем минимальное время обслуживания каждого из них).

Поток событий называется ординарным, если вероятность попадания на малый (элементарный) участок времени Δt двух и более событий пренебрежимо мала по сравнению с вероятностью попадания одного события. Другими словами, поток событий ординарен, если события появляются в нем поодиночке, а не группами. Например, поток поездов, подходящих к станции, ординарен, а поток вагонов не ординарен.

Поток событий называется простейшим (или стационарным пуассоновским), если он одновременно стационарен, ординарен и не имеет последействия. Название "простейший" объясняется тем, что СМО с простейшими потоками имеет наиболее простое математическое описание. Заметим, что регулярный поток не является "простейшим", так как он обладает последействием: моменты появления событий в таком потоке жестко зафиксированы.

Простейший поток в качестве предельного возникает в теории случайных процессов столь же естественно, как в теории вероятностей нормальное распределение получается в качестве предельного для суммы случайных величин: при наложении (суперпозиции) достаточно большого числа n независимых, стационарных и ординарных потоков (сравнимых между собой по интенсивностям λ_i $(i=1,2,...,n)$ получается поток, близкий к простейшему с интенсивностью λ, равной сумме интенсивностей входящих потоков, т.е.

$$\lambda=\sum_{i=1}^{n}\lambda_i .$$

Распределение интервала времени между двумя соседними событиями имеет вид

$$F(t)=P(T<t)=1-e^{-\lambda t} . \qquad (11.1)$$

Плотность вероятности случайной величины есть производная ее функции распределения (рис. 11.2), т.е.

$$\varphi(t)=F'(t)=\lambda e^{-\lambda t} . \qquad (11.2)$$

Распределение, задаваемое плотностью вероятности (11.2) или функцией распределения (11.1), наз показательным (или экспоненциальным). Таким образом, интервал времени между двумя соседними произвольными событиями имеет показательное распределение, для которого математическое ожидание равно дисперсии случайной величины

$$a = \sigma^2 = \frac{1}{\lambda} \tag{11.3}$$

и обратно пропорционально по величине интенсивности потока λ.

Важнейшее свойство показательного распределения (присущее только показательному распределению) состоит в следующем: если промежуток времени, распределенный по показательному закону, уже длился некоторое время τ, то это никак не влияет на закон распределения оставшейся части промежутка ($T-\tau$): он будет таким же, как и закон распределения всего промежутка T.

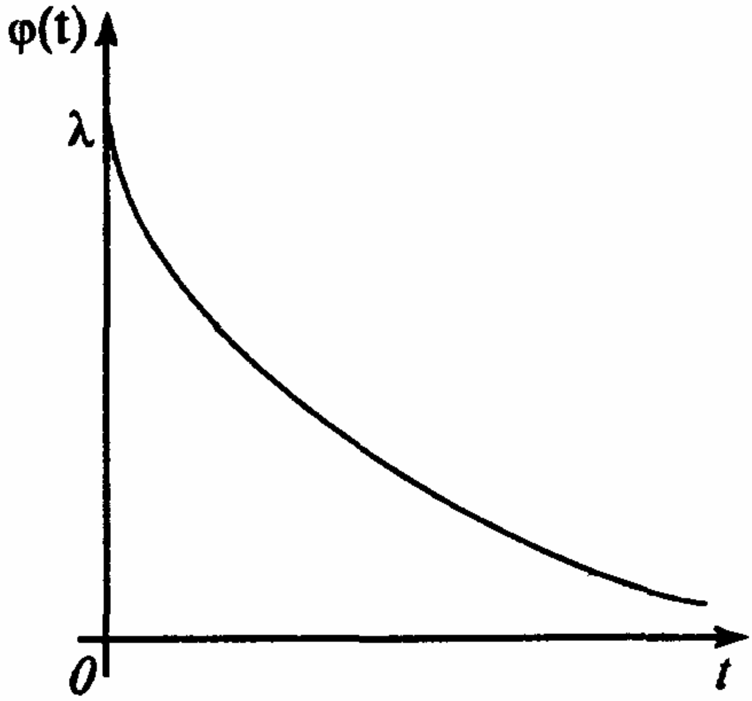


Рис. 11.2

Другими словами, для интервала времени T между двумя последовательными соседними событиями потока, имеющего показательное распределение, любые сведения о том, сколько времени протекал этот интервал, не влияют на закон распределения оставшейся части. Это свойство показательного закона представляет собой, в сущности, другую формулировку для "отсутствия последействия" – основного свойства простейшего потока.

Для простейшего потока с интенсивностью λ вероятность попадания на элементарный отрезок времени Δt хотя бы одного события потока равна согласно (11.1)

$$P_{\Delta t} = P(T < \Delta t) = 1 - e^{-\lambda \Delta t} \approx \lambda \Delta t\,. \tag{11.4}$$

(Заметим, что эта приближенная формула, получаемая заменой функции $e^{-\lambda\Delta t}$ *лишь двумя первыми членами ее разложения в ряд по степеням* Δt *, тем точнее, чем меньше* Δt).

11.4. Уравнения Колмогорова. Предельные вероятности состояний

Рассмотрим математическое описание марковского процесса с дискретными состояниями и непрерывным временем на примере случайного процесса, граф которого изображен на рис. 11.1. Будем полагать, что все переходы системы из состояния S_i в S_j происходят под воздействием простейших потоков событий с интенсивностями $\lambda_{ij}\,(i, j = 0,1,2,3)$; так, переход системы из состояния S_0 в S_1 будет происходить под воздействием потока отказов первого узла, а обратный переход из состояния S_1 в S_0 – под воздействием потока "окончаний ремонтов" первого узла и т.п.

Граф состояний системы с проставленными у стрелок интенсивностями будем называть размеченным (см. рис. 11.1). Рассматриваемая система S имеет четыре возможных состояния: S_0, S_1, S_2, S_3.

Вероятностью i -го состояния называется вероятность $p_i(t)$ того, что в момент t система будет находиться в состоянии S_i. Очевидно, что для любого момента t сумма вероятностей всех состояний равна единице:

$$\sum_{i=0}^{3} p_i(t) = 1. \tag{11.5}$$

Рассмотрим систему в момент t и, задав малый промежуток Δt, найдем вероятность $p_0(t+\Delta t)$ того, что система в момент $t+\Delta t$ будет находиться в состоянии S_0. Это достигается разными способами.

1. Система в момент t с вероятностью $p_0(t)$ находилась в состоянии S_0, а за время Δt не вышла из него.

Вывести систему из этого состояния (см. граф на рис. 11.1) можно суммарным простейшим потоком с интенсивностью ($\lambda_{01}+\lambda_{02}$), т.е. в соответствии с (11.4), с вероятностью, приближенно равной $(\lambda_{01}+\lambda_{02})\Delta t$. А вероятность того, что система не выйдет из состояния S_0, равна $[1-(\lambda_{01}+\lambda_{02})\Delta t]$. Вероятность того, что система будет находиться в состоянии S_0 по первому способу (т.е. того, что находилась в состоянии S_0 и не выйдет из него за время Δt), равна по теореме умножения вероятностей:

$$p_0(t)[1-(\lambda_{01}+\lambda_{02})\Delta t].$$

2. Система в момент t с вероятностями $p_1(t)$ (или $p_2(t)$) находилась в состоянии S_1 или S_2 и за время Δt перешла в состояние S_0.

Потоком интенсивностью λ_{10} (или λ_{20} — см. рис. 11.1) система перейдет в состояние S_0 с вероятностью, приближенно равной $\lambda_{10}\Delta t$ (или $\lambda_{20}\Delta t$). Вероятность того, что система будет находиться в состоянии S_0 по этому способу, равна $p_1(t)\lambda_{10}\Delta t$ (или $p_2(t)\lambda_{20}\Delta t$).

Применяя теорему сложения вероятностей, получим

$$p_0(t+\Delta t) = p_1(t)\lambda_{10}\Delta t + p_2(t)\lambda_{20}\Delta t + p_0(t)[1-(\lambda_{01}+\lambda_{02})\Delta t],$$

откуда

$$\frac{p_0(t+\Delta t)-p_0(t)}{\Delta t} = p_1(t)\lambda_{01} + p_2(t)\lambda_{20} - (\lambda_{01}+\lambda_{02})p_0(t).$$

Переходя к пределу при $\Delta t \to 0$ (приближенные равенства, связанные с применением формулы (2.4), перейдут в точные), получим в левой части уравнения производную $p_0'(t)$ (обозначим ее для простоты p_0'):

$$p_0' = \lambda_{10}p_1 + \lambda_{20}p_2 - (\lambda_{01}+\lambda_{02})p_0.$$

Получили дифференциальное уравнение первого порядка, т.е. уравнение, содержащее как саму неизвестную функцию, так и ее производную первого порядка.

Рассуждая аналогично для других состояний системы S, можно получить систему дифференциальных уравнений Колмогорова для вероятностей состояний:

$$\begin{cases} p_0' = \lambda_{10}p_1 + \lambda_{20}p_2 - (\lambda_{01}+\lambda_{02})p_0, \\ p_1' = \lambda_{01}p_0 + \lambda_{31}p_3 - (\lambda_{10}+\lambda_{13})p_1, \\ p_2' = \lambda_{02}p_0 + \lambda_{32}p_3 - (\lambda_{20}+\lambda_{23})p_2, \\ p_3' = \lambda_{13}p_1 + \lambda_{23}p_2 - (\lambda_{31}+\lambda_{32})p_3. \end{cases} \quad (11.6)$$

Сформулируем правило составления уравнений Колмогорова. В левой части каждого из них стоит производная вероятности i – го состояния. В правой части — сумма произведений вероятностей всех состояний (из которых идут стрелки в данное состояние) на интенсивности соответствующих потоков событий, минус суммарная интенсивность всех потоков, выводящих систему из данного состояния, умноженная на вероятность данного (i – го состояния).

В системе (11.6) независимых уравнений на единицу меньше общего числа уравнений. Поэтому для решения системы необходимо добавить уравнение (11.5).

Особенность решения дифференциальных уравнений вообще состоит в том, что требуется задать так называемые начальные условия, т.е. в данном случае вероятности состояний системы в начальный момент $t=0$. Так, например, систему уравнений (2.9) естественно решать при условии, что в начальный момент оба узла исправны и система находилась в состоянии S_0, т.е. при начальных условиях $p_0(0)=1$, $p_1(0)=p_2(0)=p_3(0)=0$.

Уравнения Колмогорова дают возможность найти все вероятности состояний как функции времени. Особый интерес представляют вероятности системы $p_i(t)$ в предельном стационарном режиме, т.е. при $t \to \infty$, которые называются предельными (или финальными) вероятностями состояний.

В теории случайных процессов доказывается, что если число состояний системы конечно и из каждого из них можно (за конечное число шагов) перейти в любое другое состояние, то предельные вероятности существуют.

Предельная вероятность состояния S_i имеет четкий смысл: она показывает среднее относительное время пребывания системы в этом состоянии. Например, если предельная вероятность состояния S_0 равна 0.5, т.е. $p_0 = 0.5$ то это означает, что в среднем половину времени система находится в состоянии S_0.

Так как предельные вероятности постоянны, то, заменяя в уравнениях Колмогорова их производные нулевыми значениями, получим систему линейных алгебраических уравнений, описывающих стационарный режим. Для системы S с графом состояний, изображенном на рис. 11.1), такая система уравнений имеет вид:

$$\begin{cases} (\lambda_{01}+\lambda_{02})p_0 = \lambda_{10}p_1 + \lambda_{20}p_2, \\ (\lambda_{10}+\lambda_{13})p_1 = \lambda_{01}p_0 + \lambda_{31}p_3, \\ (\lambda_{20}+\lambda_{23})p_2 = \lambda_{02}p_0 + \lambda_{32}p_3, \\ (\lambda_{31}+\lambda_{32})p_3 = \lambda_{13}p_1 + \lambda_{23}p_2. \end{cases} \tag{11.7}$$

Систему (11.7) можно составить непосредственно по размеченному графу состояний, если руководствоваться правилом, согласно которому слева в уравнениях стоит предельная вероятность данного состояния p_i, умноженная на суммарную интенсивность всех потоков, ведущих из данного состояния, а справа – сумма произведений интенсивностей всех потоков, входящих в i – е состояние, на вероятности тех состояний, из которых эти потоки исходят.

Пример 11.2. Найти предельные вероятности для системы S, граф состояний которой приведен на рис. 2.1, при $\lambda_{01}=1$, $\lambda_{02}=2$, $\lambda_{10}=2$, $\lambda_{13}=2$, $\lambda_{20}=3$, $\lambda_{23}=1$, $\lambda_{31}=3$, $\lambda_{32}=2$.

Решение. Система алгебраических уравнений, описывающих стационарный режим для данной системы, имеет вид (2.10) или

$$\begin{cases} 3p_0 = 2p_1 + 3p_2, \\ 4p_1 = p_0 + 3p_3, \\ 4p_2 = 2p_0 + 2p_3, \\ p_0 + p_1 + p_2 + p_3 = 1 \end{cases} \tag{11.8}$$

Здесь мы вместо одного "лишнего" уравнения системы (11.7) записали нормировочное условие (11.5)).

Решив систему (2.8), получим $p_0 = 0.40$, $p_1 = 0.20$, $p_2 = 0.27$, $p_3 = 0.13$, т.е. в предельном, стационарном режиме система S в среднем 40% времени будет находиться в состоянии S_0 (оба узла исправны), 20% — в состоянии S_1 (первый узел ремонтируется, второй работает), 27% — в состоянии S_2 (второй узел ремонтируется, первый работает) и 13% времени — в состоянии S_3 (оба узла ремонтируются). ►

Пример 11.3. Найти средний чистый доход от эксплуатации в стационарном режиме системы S в условиях задач 2.1 и 2.2, если известно, что в единицу времени исправная работа первого и второго узлов приносит доход соответственно в 10 и 6 ден.ед., а их ремонт требует затрат соответственно в 4 и 2 ден.ед. Оценить экономическую эффективность имеющейся возможности уменьшения вдвое среднего времени ремонта каждого из двух узлов, если при этом придется вдвое увеличить затраты на ремонт каждого узла (в единицу времени).

Решение. Из задачи 2.2 следует, что в среднем первый узел исправно работает долю времени, равную $p_0 + p_2 =$ 0,40+0,27=0,67, а второй узел — $p_0 + p_1 =$ 0,40+0,20=0,60. В то же время первый узел находится в ремонте в среднем долю времени, равную $p_1 + p_3 =$ 0,20+0,13=0,33, а второй узел – $p_2 + p_3 =$ 0,27+0,13=0,40. Поэтому средний чистый доход в единицу времени от эксплуатации системы, т.е. разность между доходами и затратами, равен

Д = 0,67·10 + 0,60·6 –0,33·4 –0,40·2 = 8,18 ден.ед.

Уменьшение вдвое среднего времени ремонта каждого из узлов в соответствии с (2.3) будет означать увеличение вдвое интенсивностей потока "окончаний ремонтов" каждого узла, т.е. теперь, $\lambda_{10} = 4$, $\lambda_{20} = 6$, $\lambda_{31} = 6$, $\lambda_{32} = 4$.
и система линейных алгебраических уравнений (2.10), описывающая стационарный режим системы S, вместе с нормировочным условием (2.8) примет вид:

$$\begin{cases} 3p_0 = 4p_1 + 6p_2, \\ 6p_1 = p_0 + 6p_3, \\ 7p_2 = 2p_0 + 4p_3, \\ p_0 + p_1 + p_2 + p_3 = 1 \end{cases}$$

Решив систему, получим $p_0 = 0.60$, $p_1 = 0.15$, $p_2 = 0.20$, $p_3 = 0.05$. Учитывая, что $p_0 + p_2 =$ 0,60+0,20=0,80, $p_0 + p_1 =$ 0,60+0,15=0,75, $p_1 + p_3 =$ 0,15+0,05=0,20, $p_2 + p_3 =$ 0,20+0,05=0,25, а затраты на ремонт первого и второго узла составляют теперь соответственно 8 и 4 ден. ед., вычислим средний чистый доход в единицу времени:

Д1 = 0,80·10 + 0,75·6 – 0,20·8 – 0,25·4 = 9,9 ден.ед.

Так как Д1 больше Д всего на 20%, то экономическая целесообразность ускорения ремонтов узлов очевидна. ►

11.5. Процесс гибели и размножения

В теории массового обслуживания широкое распространение имеет специальный класс случайных процессов — так называемый *процесс гибели и размножения.* Название этого процесса связано с рядом биологических задач, где он является математической моделью изменения численности биологических популяций.

Граф состояний процесса гибели и размножения имеет вид, показанный на рис. 2.4.

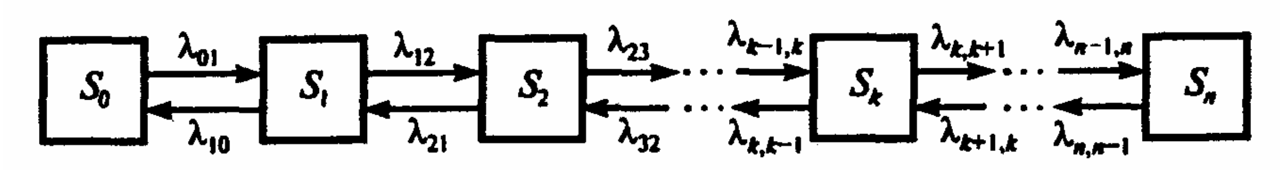


Рис. 11.4

Рассмотрим упорядоченное множество состояний системы $S_0, S_1, ..., S_k$. Переходы могут осуществляться из любого состояния только в состояния с соседними номерами, т.е. из состояния S_k возможны переходы только либо в состояние S_{k-1} либо в состояние S_{k+1}.

Замечание. При анализе численности популяций считают, что состояние S_k соответствует численности популяции, равной k, и переход системы из состояния S_k в состояние S_{k+1}, происходит при рождении одного члена популяции, а переход в состояние S_{k-1} – при гибели одного члена популяции.

Предположим, что все потоки событий, переводящие систему по стрелкам графа, простейшие с соответствующими интенсивностями $\lambda_{k,k+1}$ или $\lambda_{k+1,k}$.

По графу, представленному на рис. 2.4, составим и решим алгебраические уравнения для предельных вероятностей состояний (их существование вытекает из возможности перехода из каждого состояния в каждое другое и конечности числа состояний).

В соответствии с правилом составления таких уравнений *(см.* 2.7) получим:
для состояния S_0

$$\lambda_{01}p_0 = \lambda_{10}p_1, \qquad (11.9)$$

для состояния S_1 — $(\lambda_{12} + \lambda_{10})p_1 = \lambda_{01}p_0 + \lambda_{21}p_2$, которое с учетом (2.12) приводится к виду

$$\lambda_{12}p_1 = \lambda_{21}p_2. \qquad (11.10)$$

Аналогично, записывая уравнения для предельных вероятностей других состояний, можно получить следующую систему уравнений:

$$\begin{cases} \lambda_{01}p_0 = \lambda_{10}p_1, \\ \lambda_{12}p_1 = \lambda_{21}p_2, \\ \dots\dots\dots\dots\dots\dots\dots \\ \lambda_{k-1,k}p_{k-1} = \lambda_{k,k-1}p_k, \\ \dots\dots\dots\dots\dots\dots\dots \\ \lambda_{n-1,n}p_{n-1} = \lambda_{n,n-1}p_n. \end{cases} \qquad (11.11)$$

к которой добавляется нормировочное условие

$$p_0 + p_1 + p_2 + \ldots + p_n = 1. \qquad (11.12)$$

Решая систему (11.11), (11.12), можно получить

$$p_0 = \left(1 + \frac{\lambda_{01}}{\lambda_{10}} + \frac{\lambda_{12}\lambda_{01}}{\lambda_{21}\lambda_{10}} + \ldots + \frac{\lambda_{n-1,n}\ldots\lambda_{12}\lambda_{01}}{\lambda_{n,n-1}\ldots\lambda_{21}\lambda_{10}}\right)^{-1}, \qquad (11.13)$$

$$p_1 = \frac{\lambda_{01}}{\lambda_{10}}p_0,\ p_2 = \frac{\lambda_{12}\lambda_{01}}{\lambda_{21}\lambda_{10}}p_0,\ \ldots,\ p_n = \frac{\lambda_{n-1,n}\ldots\lambda_{12}\lambda_{01}}{\lambda_{n,n-1}\ldots\lambda_{21}\lambda_{10}}p_0. \qquad (11.14)$$

Легко заметить, что в формулах (2.14) для $p_1, p_2, \ldots, p_n$ коэффициенты при p_0 есть слагаемые, стоящие после единицы в формуле (2.13). Числители этих коэффициентов представляют произведение всех интенсивностей, стоящих у стрелок, ведущих слева направо от состояния S_0 до данного состояния S_k ($k = 1, 2, \ldots, n$), а знаменатели — произведение всех интенсивностей, стоящих у стрелок, ведущих справа налево от состояния S_k до состояния S_0.

Пример 11.4. Процесс гибели и размножения представлен графом (рис. 2.5). Найти предельные вероятности состояний.

Решение. По формуле (11.13) найдем

$$p_0 = \left(1 + \frac{1}{4} + \frac{2 \cdot 1}{3 \cdot 4}\right)^{-1} = 0.706,$$

по (2.17) – $p_1 = \frac{1}{4} 0.706 = 0.176$, $p_2 = \frac{2 \cdot 1}{3 \cdot 4} 0.706 = 0.118$, т.е. в установившемся, стационарном режиме в среднем 70,6% времени система будет находиться в состоянии S_0, 17,6% — в состоянии S_1 и 11,8% — в состоянии S_2. ►

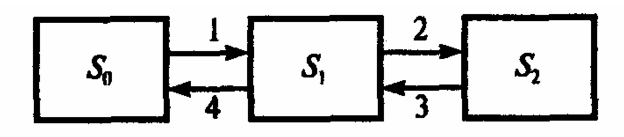


Рис. 11.5

11.6. СМО с отказами

В качестве показателей эффективности СМО с отказами будем рассматривать:

A — *абсолютную пропускную способность СМО,* т.е. среднее число заявок, обслуживаемых в единицу времени;

Q — *относительную пропускную способность,* т.е. среднюю долю пришедших заявок, обслуживаемых системой;

P_{OTK}. — *вероятность отказа,* т.е. *вероятность* того, что заявка покинет СМО необслуженной;

$\overline{k}$ — *среднее число занятых каналов (для многоканальной системы).*

Одноканальная система с отказами.

Рассмотрим задачу. Имеется один канал, на который поступает поток заявок с интенсивностью λ. Поток обслуживании имеет интенсивность μ.

Примечание. Здесь и в дальнейшем предполагается, что все потоки событий, переводящие СМО из состояния в состояние, будут простейшими. К ним относится и поток обслуживания – поток заявок, обслуживаемых одним непрерывно занятым каналом. Среднее время обслуживания $\bar{t}_{об}$ обратно по величине интенсивности μ, т.е. $\bar{t}_{об} = \frac{1}{\mu}$.

Найти предельные вероятности состояний системы и показатели ее эффективности.

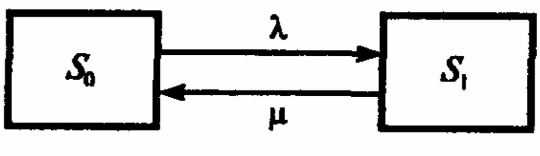


Рис. 11.6

Система S (СМО) имеет два состояния: S_0 — канал свободен, S_1 — канал занят. Размеченный *граф* состояний представлен на рис. 11.6.

В предельном, стационарном режиме система алгебраических уравнений для вероятностей состояний имеет вид *(см.* правило составления таких уравнений (11.7).

$$\begin{cases} \lambda p_0 = \mu p_1, \\ \mu p_1 = \lambda p_0. \end{cases} \qquad (11.15)$$

т.е. система вырождается в одно уравнение. Учитывая нормировочное условие до $p_0 + p_1 = 1$, найдем из (2.15) предельные вероятности состояний

$$p_0 = \frac{\mu}{\lambda + \mu}, \; p_1 = \frac{\lambda}{\lambda + \mu}, \qquad (11.16)$$

которые выражают среднее относительное время пребывания системы в состоянии S_0 (когда канал свободен) и S_1 (когда канал занят), т.е. определяют соответственно относительную пропускную способность Q системы и вероятность отказа P_{OTK}:

$$Q = \frac{\mu}{\lambda + \mu}, \qquad (11.17)$$

$$P_{OTK} = \frac{\lambda}{\lambda + \mu}. \qquad (11.18)$$

Абсолютную пропускную способность найдем, умножив относительную пропускную способность Q на интенсивность потока заявок

$$A = \frac{\lambda\mu}{\lambda + \mu}. \qquad (11.19)$$

Пример 11.5. Известно, что заявки на телефонные переговоры в телевизионном ателье поступают с интенсивностью λ, равной 90 заявок в час, а средняя продолжительность разговора по телефону $\bar{t}_{об}$~2 мин. Определить показатели эффективности работы СМО (телефонной связи) при наличии одного телефонного номера.

Решение. Имеем λ=90 (1/ч), $\bar{t}_{об}$ = 2 мин. Интенсивность потока обслуживания $\mu = \frac{1}{\bar{t}_{об}} = \frac{1}{2} = 0{,}5$ (1/мин)=30 (1/ч). По (11.17) относительная пропускная способность СМО $Q = 30/(90+30) = 0.25$, т.е. в среднем только 25% поступающих заявок осуществят переговоры по телефону. Соответственно вероятность отказа в обслуживании составит $P_{OTK} = 0.75$ *(см.* (11.18)). Абсолютная пропускная способность СМО по (11.19) $A = 90{\cdot}0{,}25$=22.5, т.е. в среднем в час будут обслужены 22.5 заявки на переговоры. Очевидно, что при наличии только одного телефонного номера СМО будет плохо справляться с потоком заявок. ►

Многоканальная система с отказами. Рассмотрим классическую *задачу Эрланга.*

Имеется n каналов, на которые поступает поток заявок с интенсивностью λ. Поток обслуживаний имеет интенсивность μ. Найти предельные вероятности состояний системы и показатели ее эффективности.

Система S (СМО) имеет следующие состояния (нумеруем их по числу заявок, находящихся в системе): $S_0, S_1, S_2, ..., S_n$, где S_k — состояние системы, когда в ней находится k заявок, т.е. занято k каналов.

Граф состояний СМО соответствует процессу гибели и размножения и показан на рис. 13.7.

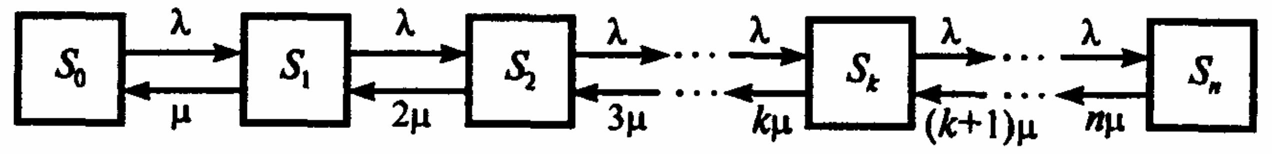


Рис. 13.7

Поток заявок последовательно переводит систему из любого левого состояния в соседнее правое с одной и той же интенсивностью λ. Интенсивность же потока обслуживаний, переводящих систему из любого правого состояния в соседнее левое состояние, постоянно меняется в зависимости от состояния. Действительно, если СМО находится в состоянии S_2 (два канала заняты), то она может перейти в состояние S_1 (один канал занят), когда закончит обслуживание либо первый, либо второй канал, т.е. суммарная интенсивность их потоков обслуживаний будет 2μ. Аналогично суммарный поток обслуживаний, переводящий СМО из состояния S_3 (три канала заняты) в S_2, будет иметь интенсивность 3μ, т.е. может освободиться любой из трех каналов и т.д.

В формуле (13.13) для схемы гибели и размножения получим для предельной вероятности состояния

$$p_0 = \left(1 + \frac{\lambda}{\mu} + \frac{\lambda^2}{2!\mu^2} + ... + \frac{\lambda^k}{k!\mu^k} + ... + \frac{\lambda^n}{n!\mu^n}\right)^{-1}, \tag{13.20}$$

где члены разложения $\frac{\lambda}{\mu}, \frac{\lambda^2}{2!\mu^2}, ..., \frac{\lambda^n}{n!\mu^n}$ будут представлять собой коэффициенты при p_0 в выражениях для предельных вероятностей $p_1, p_2, ..., p_k, ..., p_n$. Величина

$$\rho = \frac{\lambda}{\mu} \tag{13.21}$$

называется *приведенной интенсивностью потока заявок* или *интенсивностью нагрузки канала.* Она выражает среднее число заявок, приходящее за среднее время обслуживания одной заявки. Теперь

$$p_0 = \left(1 + \rho + \frac{\rho^2}{2!} + ... + \frac{\rho^k}{k!} + ... + \frac{\rho^n}{n!}\right)^{-1}, \tag{13.22}$$

$$p_1 = \rho p_0,\ p_2 = \frac{\rho^2}{2!} p_0, ..., p_k = \frac{\rho^k}{k!} p_0, ..., p_n = \frac{\rho^n}{n!} p_0 \tag{13.23}$$

Формулы (13.22) и (13.23) для предельных вероятностей получили названия *формул Эрланга (датский инженер, математик)* в честь основателя теории массового обслуживания.

Вероятность отказа СМО есть предельная вероятность того, что все n каналов системы будут заняты, т.е.

$$P_{OTK} = \frac{\rho^n}{n!} p_0 . \qquad (13.24)$$

Относительная пропускная способность — вероятность того, что заявка будет обслужена:

$$Q = 1 - P_{OTK} = 1 - \frac{\rho^n}{n!} p_0 . \qquad (13.25)$$

Абсолютная пропускная способность:

$$A = \lambda Q = \lambda \left(1 - \frac{\rho^n}{n!} p_0 \right). \qquad (13.26)$$

Среднее число занятых каналов $\bar{k}$ есть математическое ожидание числа занятых каналов:

$$\bar{k} = \sum_{k=1}^{n} k p_k ,$$

где p_k — предельные вероятности состояний, определяемых по формулам (13.22), (13.23).

Однако среднее число занятых каналов можно найти проще, если учесть, что абсолютная пропускная способность системы A есть не что иное, как среднее число обслуженных системой заявок (в единицу времени). Так как каждый занятый канал обслуживает в среднем μ заявок (в единицу времени), то среднее число занятых каналов

$$\bar{k} = \frac{A}{\mu} \qquad (13.27)$$

или, учитывая (13.26), (13.21):

$$\bar{k} = \rho \left(1 - \frac{\rho^n}{n!} p_0 \right) . \qquad (13.28)$$

◄*Справка*. Покажем другой вывод формулы (13.28).

$$\bar{k} = \sum_{k=1}^{n} k p_k = p_0 \sum_{k=1}^{n} k \frac{\rho^k}{k!} = p_0 \rho \sum_{k=1}^{n} k \frac{\rho^{k-1}}{(k-1)!} =$$
$$= p_0 \rho \sum_{j=0}^{n-1} k \frac{\rho^j}{j!} = p_0 \rho \sum_{k=0}^{n-1} k \frac{\rho^k}{k!} = p_0 \rho \left(\sum_{k=0}^{n} k \frac{\rho^k}{k!} - \frac{\rho^n}{n!} \right).$$

Далее, из (13.22) следует $p_0 = \left(\sum_{k=0}^{n} k \frac{\rho^k}{k!} \right)^{-1}$. Поэтому $\sum_{k=0}^{n} k \frac{\rho^k}{k!} = p_0^{-1}$. С учетом этого получим

$$\bar{k} = p_0\rho\left(p_0^{-1} - \frac{\rho^n}{n!}\right) = \rho\left(1 - p_0\frac{\rho^n}{n!}\right)$$ ►

♦ **Пример 13.6**. В условиях задачи 13.5 определить оптимальное число телефонных номеров в телевизионном ателье, если условием оптимальности считать удовлетворение в среднем из каждых 100 заявок не менее 90 заявок на переговоры.

Решение. Интенсивность нагрузки канала по формуле (13.21) ρ=90/30=3, т.е. за время среднего (по продолжительности) телефонного разговора $\bar{t}_{об}=2$ мин. поступает в среднем 3 заявки на переговоры.

Будем постепенно увеличивать число каналов (телефонных номеров) n=2, 3, 4, ... и определим по формулам (13.22), (13.25), (13.26) для получаемой n-канальной СМО характеристики обслуживания. Например, при n = 2

$p_0 = \left(1+3+\frac{3^2}{2!}\right)^{-1} = 0.118 \approx 0.12$; $Q = 1-\left(\frac{3^2}{2!}\right)\cdot 0.118 = 0.471 \approx 0.47$; $A = 90\cdot 0{,}471$ = 42,4 и т.д. Значение характеристик СМО сведем в табл. 91.

Таблица 13.1

Характеристика обслуживания	Число каналов (телефонных номеров)					
	1	2	3	4	5	6
Относительная пропускная способность Q	0,25	0,47	0,65	0,79	0,90	0,95
Абсолютная пропускная способность A	22,5	42,4	58,8	71,5	80,1	85,3

По условию оптимальности $Q \geq 0.9$, следовательно, в телевизионном ателье необходимо установить 5 телефонных номеров (в этом случае Q = 0,90 — *см.* табл. 13.1). При этом в час будут обслуживаться в среднем 80 заявок (A = 80,1), а среднее число занятых телефонных номеров (каналов) по формуле (2.27) $\bar{k} = \frac{80.1}{30} = 2.67$. ♦

♦ **Пример 13.7**. В вычислительный центр коллективного пользования с тремя ЭВМ поступают заказы от предприятий на вычислительные работы. Если работают все три ЭВМ, то вновь поступающий заказ не принимается, и предприятие вынуждено обратиться в другой вычислительный центр. Среднее время работы с одним заказом составляет 3 ч. Интенсивность потока заявок 0,25 (1/ч). Найти предельные вероятности состояний и показатели эффективности работы вычислительного центра.

Решение. По условию n=3, λ=0,25 (1/ч), $\bar{t}_{об}$ = 3 (ч). Интенсивность потока обслуживаний $\mu = 1/\bar{t}_{об}$ =1/3=0,33. Интенсивность нагрузки ЭВМ по формуле (13.21) ρ =0,25/0,33=0,75 =0,25/0,33=0,75. Найдем предельные вероятности состояний:

по формуле (2.22) $p_0 = \left(1 + 0.75 + \frac{0.75^2}{2!} + \frac{0.75^3}{3!}\right)^{-1} = 0.476$;

по формуле (2.23) p_1=0,75·0,476=0,357; $p_2 = \left(\frac{0.75^2}{2!}\right) \cdot 0.476 = 0.134$;

$p_3 = \left(\frac{0.75^3}{3!}\right) \cdot 0.476 = 0.033$, т.е. в стационарном режиме работы вычислительного центра в среднем 47,6% времени нет ни одной заявки, 35,7% — имеется одна заявка (занята одна ЭВМ), 13,4% — две заявки (две ЭВМ), 3,3%-времени — три заявки (заняты три ЭВМ).

Вероятность отказа (когда заняты все три ЭВМ), таким образом, $P_{OTK} = p_3 = 0,033$.

По формуле (13.25) относительная пропускная способность центра Q = 1 – 0,033 = 0,967, т.е. в среднем из каждых 100 заявок вычислительный центр обслуживает *96,7* заявок.

По формуле (13.26) абсолютная пропускная способность центра A = 0,25·0,967 = 0,242, т.е. в один час в среднем обслуживается 0,242 заявки.

По формуле (13.27) среднее число занятых ЭВМ $\bar{k}$ = 0,242/0,33 = 0,725, т.е. каждая из трех ЭВМ будет занята обслуживанием заявок в среднем лишь на 72,5/3 = 24,2%.

При оценке эффективности работы вычислительного центра необходимо сопоставить доходы от выполнения заявок с потерями от простоя дорогостоящих ЭВМ (с одной стороны, у нас высокая пропускная способность СМО, а с другой стороны — значительный простой каналов обслуживания) и выбрать компромиссное решение. ♦

11.7. СМО с ожиданием (очередью)

В качестве показателей эффективности СМО с ожиданием, кроме уже известных показателей — абсолютной A и относительной Q пропускной способности, вероятности отказа P_{OTK}, среднего числа занятых каналов $\bar{k}$ (для многоканальной системы) будем рассматривать также следующие: $L_{сист}$ *– среднее число заявок в системе,* $T_{сист}$ *— среднее время пребывания заявки в системе;* $L_{оч}$ *— среднее число заявок в очереди (длина очереди);* $T_{оч}$*— среднее время пребывания заявки в очереди;* $P_{зан}$ *— вероятность того, что канал занят (степень загрузки канала).*

Одноканальная система с неограниченной очередью. На практике часто встречаются одноканальные СМО с неограниченной очередью (например, телефон-автомат с одной будкой). Рассмотрим задачу.

Имеется одноканальная СМО с очередью, на которую не наложены никакие ограничения (ни по длине очереди, ни по времени ожидания). Поток заявок, поступающих в СМО, имеет интенсивность λ., а поток обслуживаний — интенсивность μ. Необходимо найти предельные вероятности состояний и показатели эффективности СМО.

Система может находиться в одном из состояний $S_0, S_1, S_2, ..., S_k$, по числу заявок, находящихся в СМО: S_0 — канал свободен; S_1 — канал занят (обслуживает заявку), очереди нет; S_2 — канал занят, одна заявка стоит в очереди; ... S_k — канал занят, $(k-1)$ заявок стоят в очереди и т.д.

Граф состояний СМО представлен на рис. 13.8.

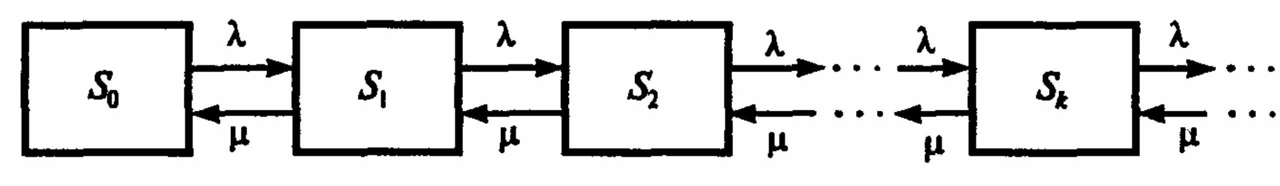


Рис. 13.8

Это процесс гибели и размножения, но с бесконечным числом состояний, в котором интенсивность потока заявок равна λ, а интенсивность потока обслуживаний μ.

Прежде чем записать формулы предельных вероятностей, необходимо быть уверенным в их существовании, ведь в случае, когда время $t \to \infty$, очередь может неограниченно возрастать. Доказано, что *если* $\rho < 1$, *т.е. среднее число приходящих заявок меньше среднего числа обслуженных заявок (в единицу времени), то предельные вероятности существуют. Если* $\rho \geq 1$, *очередь растет до бесконечности.*

Для определения предельных вероятностей состояний воспользуемся формулами (13.13), (13.14) для процесса гибели и размножения (здесь мы допускаем известную нестрогость, так как ранее эти формулы были получены для случая конечного числа состояний системы). Получим:

$$p_0 = \left[1 + \frac{\lambda}{\mu} + \left(\frac{\lambda}{\mu}\right)^2 + \ldots + \left(\frac{\lambda}{\mu}\right)^k + \ldots\right]^{-1} = (1 + \rho + \rho^2 + \ldots + \rho^k + \ldots)^{-1}. \qquad (13.29)$$

Так как предельные вероятности существуют лишь при $\rho < 1$, то геометрический ряд со знаменателем $\rho < 1$, записанный в скобках в формуле (2.32), сходится к сумме, равной, $\frac{1}{1-\rho}$. Поэтому

$$p_0 = 1 - \rho, \qquad (13.30)$$

И с учетом соотношений (2.14)

$$p_1 = \rho p_0,\ \ p_2 = \rho^2 p_0, \ldots, p_k = \rho^k p_0, \ldots$$

или

$$p_1 = \rho(1-\rho),\ \ p_2 = \rho^2(1-\rho), \ldots, p_k = \rho^k(1-\rho), \ldots \qquad (13.31)$$

Предельные вероятности $p_0, p_1, p_2, \ldots, p_k$ образуют убывающую геометрическую прогрессию со знаменателем $\rho < 1$, следовательно, вероятность p_0 — наибольшая. Это означает, что если СМО справляется с потоком заявок (при $\rho < 1$), то наиболее вероятным будет отсутствие заявок в системе.

Среднее число заявок в системе $L_{сист}$ определим по формуле математического ожидания, которая с учетом (13.31) примет вид

$$L_{сист} = \sum_{k=1}^{\infty} k p_k = (1-\rho)\sum_{k=1}^{\infty} k\rho^k \qquad (13.32)$$

(суммирование от 1 до ∞, так как нулевой член $0 \cdot p_0 = 0$).

Так как $\sum_{k=1}^{\infty} k\rho^k = \frac{\rho}{1-\rho}$, то формула (13.32) преобразуется (при $\rho < 1$) к виду

$$L_{сист} = \frac{\rho}{1-\rho}. \qquad (13.33)$$

Найдем среднее число заявок в очереди $L_{оч}$. Очевидно, что

$$L_{оч} = L_{сист} - L_{об}, \tag{13.34}$$

где $L_{об}$ — среднее число заявок, находящихся под обслуживанием.

Среднее число заявок под обслуживанием определим по формуле математического ожидания числа заявок под обслуживанием, принимающего значения 0 (если канал свободен) либо 1 (если канал занят):

$$L_{об} = 0 \cdot p_0 + 1 \cdot (1 - p_0),$$

т.е. среднее число заявок под обслуживанием равно вероятности того, что канал занят:

$$L_{об} = P_{зан} = 1 - p_0. \tag{13.35}$$

С учетом (2.30) получим

$$L_{об} = P_{зан} = \rho. \tag{13.36}$$

Теперь по формуле (2.34) с учетом (2.33) и (2.36)

$$L_{оч} = \frac{\rho^2}{1-\rho}. \tag{13.37}$$

Доказано, что *при любом характере потока заявок, при любом распределении времени обслуживания, при любой дисциплине обслуживания среднее время пребывания заявки в системе (очереди) равно среднему числу заявок в системе (в очереди), деленному на интенсивность потока заявок,* т.е.

$$T_{сист} = \frac{1}{\lambda} L_{сист}, \tag{13.38}$$

$$T_{оч} = \frac{1}{\lambda} L_{оч}. \tag{13.39}$$

Формулы (13.38) и (13.39) называются *формулами Литтла.* Они вытекают из того, что *в предельном, стационарном режиме среднее число заявок, прибывающих в систему, равно среднему числу заявок, покидающих ее:* оба потока заявок имеют одну и ту же интенсивность λ.

На основании формул (13.38) и (13.39) с учетом (13.33) и (13.37) среднее время пребывания заявки в системе определится по формуле:

$$T_{сист} = \frac{\rho}{\lambda(1-\rho)}, \tag{13.40}$$

а среднее время пребывания заявки в очереди

$$T_{оч} = \frac{\rho^2}{\lambda(1-\rho)}. \tag{13.41}$$

♦ **Пример 13.8.** В порту имеется один причал для разгрузки судов. Интенсивность потока судов равна 0,4 (судов в сутки). Среднее время разгрузки одного судна составляет 2 суток. Предполагается, что очередь может быть неограниченной длины. Найти показатели эффективности работы причала, а также вероятность того, что ожидают разгрузки не более чем 2 судна.

Решение. Имеем $\rho = \lambda / \mu = \lambda \bar{t}_{об} = 0.4 \cdot 2$ = 0,8. Так как $\rho =$ 0,8 < 1, то очередь на разгрузку не может бесконечно возрастать и предельные вероятности существуют. Найдем их.

Вероятность того, что причал свободен, по (13.30) $p_0 =$ 1– 0,8 = 0,2, а вероятность того, что он занят, $P_{зан} =$ 1 – 0,2 = 0,8. По формуле (13.31) вероятности того, что у причала находятся 1, 2, 3 судна (т.е. ожидают разгрузки 0, 1, 2 судна), равны $p_1 = 0{,}8(1\text{-}0{,}8)$ = 0,16; $p_2 = 0{,}8^2 \cdot (1\text{-}0{,}8) = 0{,}128$; $p_3 = 0{,}8^3 \cdot (1\text{-}0{,}8) = 0{,}1024$.

Вероятность того, что ожидают разгрузку не более чем 2 судна, равна

$$P = p_1 + p_2 + p_3 = 0{,}16+0{,}128+0{,}1024 = 0{,}3904.$$

По формуле (13.37) среднее число судов, ожидающих разгрузки,

$$L_{оч} = 0{,}8^2/(1\text{-}0{,}8) = 3.2 \text{ судов},$$

а среднее время ожидания разгрузки по формуле (13.39)

$$T_{оч} = 3{,}2/0{,}4 = 8 \text{ (суток)}.$$

По формуле (13.33) среднее число судов, находящихся у причала, $L_{сист} = 0{,}8/(1\text{-}0{,}8) = 4$ (судов) (или проще по (13.34) $L_{сист} = 3{,}2+0{,}8 = 4$ (судов), а среднее время пребывания судна у причала по формуле (13.38) $T_{сист} = 4/0{,}4 = 10$ (суток).

Очевидно, что эффективность разгрузки судов невысокая. Для ее повышения необходимо уменьшение среднего времени разгрузки судна $\bar{t}_{об}$ либо увеличение числа причалов. ♦

Многоканальная СМО с неограниченной очередью. Рассмотрим задачу. Имеется n - канальная СМО с неограниченной очередью. Поток заявок, поступающих в СМО, имеет интенсивность λ, а поток обслуживании — интенсивность μ. Необходимо найти предельные вероятности состояний СМО и показатели ее эффективности.

Система может находиться в одном из состояний $S_0, S_1, S_2, \ldots, S_k, \ldots, S_n, \ldots$, нумеруемых по числу заявок, находящихся в СМО: S_0 — в системе нет заявок (все каналы свободны); S_1 — занят один канал, остальные свободны;. S_2 — заняты два канала, остальные свободны; ..., S_k — занято k каналов, остальные свободны; …, S_n — заняты все n каналов (очереди нет); S_{n+1} — заняты все n каналов, в очереди одна заявка; ..., S_{n+r} — заняты все n каналов, r заявок стоит в очереди.

Граф состояний системы показан на рис. 13.9. Обратим внимание на то, что в отличие от предыдущей СМО, интенсивность потока обслуживаний (переводящего систему из одного состояния в другое справа налево) не остается постоянной, а по мере увеличения числа заявок в СМО от 0 до n увеличивается от величины μ до $n\mu$, так как соответственно увеличивается число каналов обслуживания. При числе заявок в СМО большем, чем n, интенсивность потока обслуживании сохраняется равной $n\mu$.

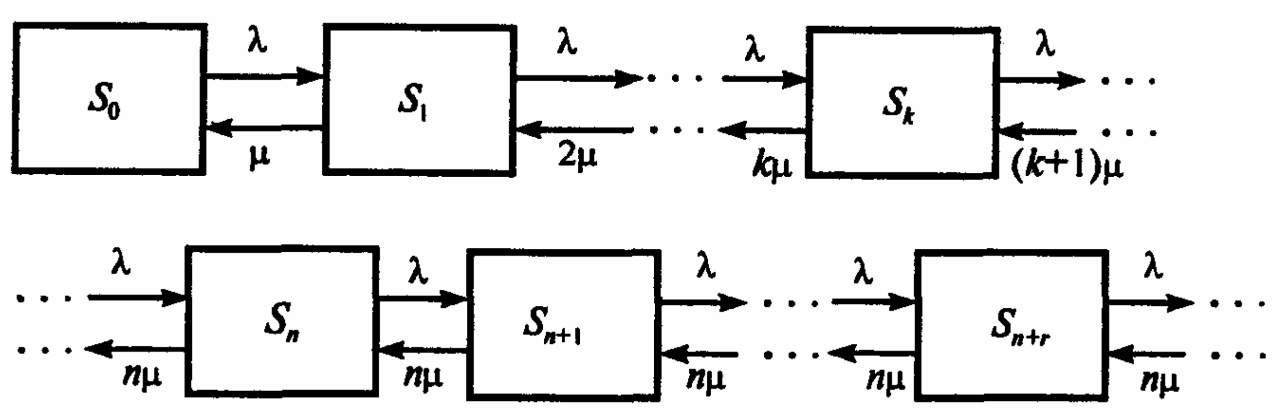


Рис. 13.9

Можно показать, что при $\rho/n < 1$ предельные вероятности существуют. Если $\rho/n \geq 1$, очередь растет до бесконечности. Используя формулы (13.13) и (13.14) для процесса гибели и размножения, можно получить следующие формулы для предельных вероятностей состояний n - канальной СМО с неограниченной очередью

$$p_0 = \left(1 + \rho + \frac{\rho^2}{2!} + ... + \frac{\rho^k}{k!} + ... + \frac{\rho^n}{n!} + \frac{\rho^{n+1}}{n!(n-\rho)}\right)^{-1}, \tag{13.42}$$

$$p_1 = \frac{\rho}{1!} p_0, ..., p_k = \frac{\rho^k}{k!} p_0, ..., p_n = \frac{\rho^n}{n!} p_0, \tag{13.43}$$

$$p_{n+1} = \frac{\rho^{n+1}}{n \cdot n!} p_0, ..., p_{n+r} = \frac{\rho^{n+r}}{n^r \cdot n!} p_0, ... \tag{13.44}$$

Вероятность того, что заявка окажется в очереди,

$$P_{оч} = \frac{\rho^{n+1}}{n!(n-\rho)} p_0. \tag{13.45}$$

Вывод формулы (13.45)

$$p_0 = \left(1 + \rho + \frac{\rho^2}{2!} + ... + \frac{\rho^k}{k!} + ... + \frac{\rho^n}{n!} + \frac{1}{n!} \cdot (\frac{\rho^{n+1}}{n} + \frac{\rho^{n+2}}{n^2} + ... + \frac{\rho^{n+r}}{n^r} + ...)\right)^{-1} =$$

$$= \left(1 + \rho + \frac{\rho^2}{2!} + ... + \frac{\rho^k}{k!} + ... + \frac{\rho^n}{n!} + \frac{1}{n!} \cdot \frac{\rho^{n+1}}{n} (1 + \frac{\rho}{n} + \frac{\rho^2}{n^2} ... + \frac{\rho^{r-1}}{n^{r-1}} + ...)\right)^{-1}$$

(A)

Сумма $1 + \frac{\rho}{n} + \frac{\rho^2}{n^2} ... + \frac{\rho^{r-1}}{n^{r-1}} + ...$ равна $\frac{1}{1-\rho/n} = \frac{n}{n-\rho}$. Подставим это в формулу (A), получим искомый результат.

Для n-канальной СМО с неограниченной очередью, используя прежние приемы, можно найти:

среднее число занятых каналов

$$\bar{k} = \frac{\lambda}{\mu} = \rho, \tag{11.46}$$

среднее число заявок в очереди

$$L_{оч} = \frac{\rho^{n+1}}{n \cdot n! \left(1 - \frac{\rho}{n}\right)^2} p_0, \tag{11.47}$$

среднее число заявок в системе

$$L_{сист} = L_{оч} + \rho. \tag{11.48}$$

Среднее время пребывания заявки в очереди и среднее время пребывания заявки в системе, как и ранее, находятся по формулам Литтла (11.42) и (11.41).

◀*Справка.* Вывод формулы (11.47).

$L_{оч} = 1 \cdot p_{n+1} + 2 \cdot p_{n+2} + ... + r \cdot p_{n+r} + ... =$

$= \frac{1}{n!} p_0 \sum_{r=1}^{\infty} \frac{r \cdot \rho^{n+r}}{n^r} = \frac{1}{n!} p_0 \frac{\rho^{n+1}}{n} \sum_{r=1}^{\infty} r \cdot (\rho/n)^{r-1} \cdot$

Рассмотрим $\sum_{r=1}^{\infty} r\cdot(\rho/n)^{r-1}$. Пусть $q=\rho/n<1$. Тогда

$$\sum_{r=1}^{\infty} r\cdot(\rho/n)^{r-1} = \sum_{r=1}^{\infty} r\cdot q^{r-1} = \frac{d}{dq}\left(\sum_{r=1}^{\infty}(\rho/n)^{r}\right) =$$

$$= \frac{d}{dq}\left(\frac{q}{1-q}\right) = \frac{1-q-q(-1)}{(1-q)^2} = \frac{1}{(1-q)^2} = \frac{1}{(1-\rho/n)^2}.$$

Таким образом, $L_{оч} = \dfrac{\rho^{n+1}}{n\cdot n!\left(1-\dfrac{\rho}{n}\right)^2} p_0$.►

Замечание. Для СМО с неограниченной очередью при $\rho<1$ любая заявка, пришедшая в систему, будет обслужена, т.е. вероятность отказа $P_{отк}$ =0, относительная пропускная способность Q = 1, а абсолютная пропускная способность равна интенсивности входящего потока заявок, т.е. $A=\lambda$.

♦ ***Пример 11.9.*** В универсаме к узлу расчета поступает поток покупателей с интенсивностью λ = 81 чел. в час. Средняя продолжительность обслуживания контролером-кассиром одного покупателя $\bar{t}_{об}$ = 2 мин. Определить:

а) Минимальное количество контролеров-кассиров $n_{\min}$, при котором очередь не будет расти до бесконечности, и соответствующие характеристики обслуживания при $n=n_{\min}$.

б) Оптимальное количество $n_{опт}$ контролеров-кассиров, при котором относительная величина затрат $C_{отн}$, связанная, с издержками на содержание каналов обслуживания и с пребыванием в очереди покупателей, задаваемая, например, как $C_{отн}=\frac{1}{\lambda}n+3T_{оч}$, будет минимальна, и сравнить характеристики обслуживания при $n=n_{\min}$ и $n=n_{опт}$.

в) Вероятность того, что в очереди будет не более трех покупателей.

Решение. а) По условию λ = 81(1/ч) = 81/60 = 1,35 (1/мин.). По формуле (11.24) $\rho=\lambda/\mu=\lambda\bar{t}_{об}$ =1.35·2 = 2,7. Очередь не будет возрастать до бесконечности при условии ρ/n < 1, т.е. при $n>\rho$ = 2,7. Таким образом, минимальное количество контролеров-кассиров $n_{\min}$ = 3.

Найдем характеристики обслуживания СМО при n = 3.

Вероятность того, что в узле расчета отсутствуют покупатели, по формуле (11.42) p_0= $(1+2{,}7+2{,}7^2/2!+2{,}7^3/3!+2{,}7^4/3!(3-2{,}7))^{-1}$ = 0,025, т.е. в среднем 2,5% времени контролеры-кассиры будут простаивать.

Вероятность того, что в узле расчета будет очередь, по (11.45)

$$P_{оч} = (2{,}7^4/3!(3-2{,}7))\cdot 0{,}025 = 0{,}735.$$

Среднее число покупателей, находящихся в очереди, по (11.47)

$$L_{оч} = (2{,}7^4/3\cdot 3!(1-2{,}7/3)^2)\cdot 0{,}025 = 7{,}35.$$

Среднее время ожидания в очереди по (11.41)

$$T_{оч} = 7{,}35/1{,}35 = 5{,}44 \text{ (мин)}.$$

Среднее число покупателей в узле расчета по (11.48)

$$L_{сист} = 7{,}35+2{,}7 = 10{,}05.$$

Среднее время нахождения покупателей в узле расчета по (11.38)

$$T_{сист} = 10{,}05/1{,}35 = 7{,}44 \text{ (мин)}.$$

Среднее число контролеров-кассиров, занятых обслуживанием покупателей, по (11.46) $\bar{k}$ = 2,7.

Коэффициент (доля) занятых обслуживанием контролеров-кассиров

$$k_{зан} = \rho / n = 2{,}7/3 = 0{,}13.$$

Абсолютная пропускная способность узла расчета A = 1,35 (1/мин), или 81 (1/ч), т.е. 81 покупатель в час.

Анализ характеристик обслуживания свидетельствует о значительной перегрузке узла расчета при наличии трех контролеров-кассиров.

б) Относительная величина затрат при n = 3

$$C_{отн} = \frac{1}{\lambda} n + 3T_{оч} = 3/1{,}35+3\cdot 5{,}44 = 18{,}54.$$

Рассчитаем относительную величину затрат при других значениях n (табл. 11.2).

Таблица 11.2

Характеристика обслуживания	Число контролеров-кассиров				
	3	4	5	6	7
Вероятность простоя контролеров -кассиров p_0	0,025	0,057	0,065	0,067	0,067
Среднее время ожидания в очереди $T_{оч}$	5,44	0,60	0,15	0,03	0,01
Относительная величина затрат $C_{отн}$	18,54	4,77	4,14	4,53	5,22

Как видно из табл. 11.2, минимальные затраты получены при $n = n_{опт}$ = 5 контролерах-кассирах.

Определим характеристики обслуживания узла расчета при $n = n_{опт}$ = 5. Получим $P_{оч}$ = 0,091; $L_{оч}$ = 0,198; $T_{оч}$ = 0,146 (мин); $L_{сист}$ = 2,90; $T_{сист}$ = 2,15 (мин); $\bar{k}$ = 2,7; $k_{зан}$ = 0,54.

Как видим, при n = 5 по сравнению с n = 3 существенно уменьшились вероятность возникновения очереди $P_{оч}$, длина очереди $L_{оч}$ и среднее время пребывания в очереди $T_{оч}$ и соответственно среднее число покупателей $L_{сист}$ и среднее время нахождения в узле расчета $T_{сист}$, а также доля занятых обслуживанием контролеров $k_{зан}$. Но среднее число занятых обслуживанием контролеров-кассиров $\bar{k}$ и абсолютная пропускная способность узла расчета A естественно не изменились.

в) Вероятность того, что в очереди будет не более 3 покупателей, определится как

$$P(r \le 3) = \underbrace{p_1 + p_2 + p_3 + p_4 + p_5}_{\substack{(когда\ заняты\ от\ 1\ до\ 5\\ контролеров-кассиров)}} + \underbrace{p_{5+1} + p_{5+2} + p_{5+3}}_{\substack{(когда\ в\ очереди\ стоят\\ от\ 1\ до\ 3\ покупателей)}} =$$

$=1-P_{оч}+p_{5+1}+p_{5+2}+p_{5+3}$, где каждое слагаемое найдем по формулам (11.42) – (11.45). Получим при n=5:

$$P(r\le 3)=1-\frac{2.7^6}{(5-2.7)\cdot 5!}0.065+\frac{2.7^6}{5\cdot 5!}0.065+\frac{2.7^7}{5^2\cdot 5!}0.065+\frac{2.7^8}{5^3\cdot 5!}0.065=0.986.$$

(Заметим, что в случае n=3 контролеров-кассиров та же вероятность существенно меньше: $P(r\le 3)=0.464$. ►

♦ ***Пример 11.10.*** Железнодорожная касса с двумя окошками продает билеты в два пункта *А* и *В*. Интенсивность потока пассажиров, желающих купить билеты, для обоих пунктов одинакова: $\lambda_A=\lambda_B=$ 0,45 (пассажиров в минуту). На обслуживание пассажиров кассир тратит в среднем 2 мин. Рассматриваются два варианта продажи билетов: первый — билеты продаются в одной кассе с двумя окошками одновременно в оба пункта *А* и *В;* второй — билеты продаются в двух специализированных кассах (по одному окошку в каждой), одна только в пункт *А,* другая — только в пункт *В.* Необходимо:

а) Сравнить два варианта продажи билетов по основным характеристикам обслуживания.

б) Определить, как надо изменить среднее время обслуживания одного пассажира, чтобы по второму варианту продажи пассажиры затрачивали на приобретение билетов в среднем меньше времени, чем по первому варианту.

Решение, а) По первому варианту имеем двухканальную СМО, на которую поступает поток заявок интенсивностью λ = 0,45+0,45 = 0,9; интенсивность потока обслуживании μ = 1/2 = 0,5; $\rho=\lambda/\mu$ = 1,8. Так как ρ/n = 1,8/2 = 0,9 < 1, то предельные вероятности существуют.

Вероятность простоя двух кассиров по (11.42)

$$p_0=\left(1+\frac{1.8}{1!}+\frac{1.8^2}{2!}+\frac{1.8^3}{2!(2-1.8)}\right)^{-1}=0.0526.$$

Среднее число пассажиров в очереди по (11.47)

$$L_{оч}=\frac{1.8^3}{2\cdot 2!(1-1.8/2)^2}0.0526=7.67.$$

Среднее число пассажиров у кассы по (11.48)

$$L_{сист} = 7{,}67+1{,}8 = 9{,}47.$$

Среднее время на ожидание в очереди и покупку билетов равно соответственно (по формулам (11.39) и (11.38)): $T_{оч}$ = 7,67/0,9 = = 8,52 (мин) и $T_{сист}$ = 9,47/0,9 = 10,5 (мин).

По второму варианту имеем две одноканальные СМО (два специализированных окошка); на каждую поступает поток заявок с интенсивностью λ = 0,45. По-прежнему μ = 0,5; $\rho=\lambda/\mu=$ = 0,9 < 1, предельные вероятности существуют. По формулам (11.37), (11.33), (11.39), (11.38)

$L_{оч}$ = $0{,}9^2$/(1-0,9) = 8,1; $L_{сист}$ = 0,9/(1-0,9) = 9,0;

$T_{оч}$ = 8,1/0,45 = 18,0 (мин), $T_{сист}$ = 9,0/0,45 = 20,0 (мин).

Итак, по второму варианту увеличились и длина очереди, и среднее время ожидания в ней и в целом на покупку билетов. Такое различие объясняется тем, что в первом варианте (двухканальная СМО) меньше средняя доля времени, которую простаивает каждый из двух кассиров: если кассир не занят обслуживанием пассажира, покупающего билет в пункт А, он может заняться обслуживанием пассажира, покупающего билет в пункт В, и наоборот. Во втором варианте такой взаимозаменяемости нет. Можно заметить, что среднее время на покупку билетов по второму варианту увеличилось более чем в 2 раза. Такое значительное увеличение связано с тем, что СМО работает на пределе своих возможностей (ρ = 0,9): достаточно незначительно увеличить среднее время обслуживания $\bar{t}_{об}$, т.е. уменьшить μ, и ρ превзойдет 1, т.е. очередь начнет неограниченно возрастать.

б) Выше было получено, что по первому варианту продажи билетов при среднем времени обслуживания одного пассажира $\bar{t}_{об}$ = 2 (мин) среднее время на покупку билетов составит $T_{сист_1}$ =10.5 (мин). По условию для второго варианта продажи $T_{сист_2} < T_{сист_1}$ или с учетом (11.33) и (11.38): $\dfrac{1}{\lambda}\dfrac{\rho}{1-\rho} < T_{сист_1}$, откуда найдем $\bar{t}_{об} < \dfrac{T_{сист_1}}{1+\lambda T_{сист_1}}$ или

$\bar{t}_{об}$<10.5/(1+0.45·10.5)=1.83 (мин).

Итак, средние затраты времени на покупку билетов по второму варианту продажи уменьшатся, если среднее время обслуживания одного пассажира уменьшится более чем на 0,17 мин, или более чем на 8,5%. ►

СМО с ограниченной очередью. СМО с ограниченной очередью отличаются от рассмотренных выше задач лишь тем, что число заявок в очереди ограничено (не может превосходить некоторого заданного m). Если новая заявка поступает в момент, когда все места в очереди заняты, она покидает СМО необслуженной, т.е. получает отказ.

Очевидно: для вычисления предельных вероятностей состояний и показателей эффективности таких СМО может быть использован тот же подход, что и выше, с той разницей, что суммировать надо не бесконечную прогрессию (как, например, мы делали при выводе формулы (11.29)), а конечную. Соответствующие формулы сведем в табл. 11.3.

Среднее время пребывания заявки в очереди и в системе, как и ранее, определяем по формулам Литтла (11.38) и (11.39).

♦ ***Пример 11.11.*** По условию задачи 11.8 найти показатели эффективности работы причала. Известно, что приходящее судно покидает причал (без разгрузки), если в очереди на разгрузку стоит более 3 судов.

Решение. По условию m = 3. Используем формулы, приведенные во второй графе табл. 11.3.

Вероятность того, что причал свободен: $p_0 = \dfrac{1-0.8}{1-0.8^{3+2}} = 0.297$.

Вероятность того, что приходящее судно покинет причал без разгрузки:

$$P_{отк} = 0.8^{3+1}\cdot 0.297 = 0.122.$$

Относительная пропускная способность причала: Q =1-0,122 = 0,878.

Абсолютная пропускная способность причала A = 0,4· 0,878 = 0,351, т.е. в среднем в сутки разгружается 0,35 судна. Среднее число судов, ожидающих разгрузку

$$L_{оч} = \frac{0.8^2[1-0.8^3(3+1-3\cdot 0.8)]}{(1-0.8^{3+2})(1-0.8)} = 0.861,$$

а среднее время ожидания разгрузки по (11.38) $T_{оч}$ = 0.861/0.4 = 2,15 (сутки).

Среднее число судов, находящихся у причала $L_{сист}$ = 0,861+(1-0,297) = 1,564, а среднее время пребывания судна у причала по (11.38): $T_{сист}$ = 1,564/0.4 = 3,91 (сутки). ►

Таблица 11.3

Показатели	Одноканальная СМО с ограниченной очередью	Многоканальная СМО с ограниченной очередью

Предельные вероятности	$p_0=\dfrac{1-\rho}{1-\rho^{m+2}}$ $p_1=\rho p_0,\ p_2=\rho^2 p_0,...,$ $p_k=\rho^k p_0$	$p_0=\left(1+\dfrac{\rho}{1!}+...+\dfrac{\rho^n}{n!}+...+\right.$ $\left.+\dfrac{\rho^{n+1}(1-(\rho/n)^m)}{n\cdot n!(1-\rho/n)}\right)^{-1}$ $p_1=\dfrac{\rho}{1!}p_0,...,p_k=\dfrac{\rho^k}{k!}p_0,...,p_n=\dfrac{\rho^n}{n!}p_0,$ $p_{n+1}=\dfrac{\rho^{n+1}}{n\cdot n!}p_0,...,p_{n+r}=\dfrac{\rho^{n+r}}{n^r\cdot n!}p_0$ $(r=1,...,m)$
Вероятность отказа	$P_{отк}=P_{m+1}=\rho^{m+1}p_0$	$P_{отк}=P_{n+m}=\dfrac{\rho^{m+n}}{n^m\cdot n!}p_0$
Абсолютная пропускная способность	$A=\lambda Q=\lambda(1-\rho^{m+1}p_{0)}$	$A=\lambda Q=\lambda(1-\dfrac{\rho^{m+n}}{n^m n!}p_0)$
Относительная пропускная способность	$Q=1-P_{отк}=1-\rho^{m+1}p_0$	$Q=1-P_{отк}=1-\dfrac{\rho^{m+n}}{n^m n!}p_0$
Среднее число заявок в очереди	$L_{оч}=\rho^2\dfrac{[1-\rho^m(m+1-m\rho)]}{(1-\rho^{m+2})(1-\rho)}$	$L_{оч}=\dfrac{\rho^{n+1}p_0\left[1-\left(m+1-m\dfrac{\rho}{n}\right)\left(\dfrac{\rho}{n}\right)^m\right]}{n\cdot n!\left(1-\dfrac{\rho}{n}\right)}$
Среднее число заявок под обслуживанием (среднее число занятых каналов)	$L_{об}=1-p_0$	$\bar{k}=\rho\left(1-\dfrac{\rho^{n+m}}{n^m\cdot n!}p_0\right)$
Среднее число заявок в системе	$L_{сист}=L_{оч}+L_{об}$	$L_{сист}=L_{оч}+\bar{k}$

СМО с ограниченным временем ожидания. На практике часто встречаются СМО с так называемыми "нетерпеливыми» заявками. Такие заявки могут уйти из очереди, если время ожидания превышает некоторую величину. В частности, такого рода заявки возникают в различных технологических системах, в которых задержка с началом обслуживания может привести к потере качества продукции, в системах оперативного управления, когда срочные сообщения теряют ценность (или даже смысл), если они не поступают на обслуживание в течение определенного времени.

В простейших математических моделях таких систем предполагается, что заявка может находиться в очереди случайное время, распределенное по показательному закону с некоторым параметром ν, т.е. можно условно считать, что каждая заявка, стоящая в очереди на обслуживание, может покинуть систему с интенсивностью ν.

Соответствующие показатели эффективности СМО с ограниченным временем получаются на базе результатов, полученных для процесса гибели и размножения.

В заключение отметим, что на практике часто встречаются замкнутые системы обслуживания, у которых входящий поток заявок существенным образом зависит от состояния самой СМО. В качестве примера можно привести ситуацию, когда на ремонтную базу поступают с мест эксплуатации некоторые машины: понятно, что чем больше машин находится в состоянии ремонта, тем меньше их продолжает эксплуатироваться и тем меньше интенсивность потока вновь поступающих на ремонт машин. Для замкнутых СМО характерным является ограниченное число источников заявок, причем каждый источник "блокируется" на время обслуживания его заявки (т.е. он не выдает новых заявок). В подобных системах при конечном числе состояний СМО предельные вероятности будут существовать при любых значениях интенсивностей потоков заявок и обслуживании. Они могут быть вычислены, если вновь обратиться к процессу гибели и размножения.

11.8. Понятие о статистическом моделировании СМО (методе Монте-Карло)

Основное допущение, при котором анализировались рассмотренные выше СМО, состоит в том, что все потоки событий, переводящие их из состояния в состояние, были простейшими. При нарушении этого требования общих аналитических методов для таких систем не существует. Имеются лишь отдельные результаты, позволяющие выразить в аналитическом виде характеристики СМО через параметры задачи.

В случаях, когда для анализа работы СМО аналитические методы не применимы (или же требуется проверить их точность), используют универсальный метод статистического моделирования, или, как его называют, метод Монте-Карло.

Идея метода Монте-Карло состоит в том, что вместо аналитического описания СМО производится "розыгрыш" случайного процесса, проходящего в СМО, с помощью специально организованной процедуры. В результате такого "розыгрыша" получается каждый раз новая, отличная от других реализация случайного процесса. Это множество реализаций можно использовать как некий искусственно полученный статистический материал, который обрабатывается обычными методами математической статистики. После такой обработки могут быть получены приближенно любые характеристики обслуживания.

Например, необходимо проанализировать очереди, возникающие в магазине, для решения вопроса о расширении магазина. Время подхода покупателей и время их обслуживания носят случайный характер, и их распределения могут быть установлены по имеющейся информации. В результате взаимодействия этих случайных процессов создается очередь.

Согласно методу Монте-Карло перебирают (с помощью ЭВМ) все возможные состояния системы с различным числом покупателей в час, временем их обслуживания и т.п., сохраняя те же характеристики распределения. В результате многократного искусственного воссоздания работы магазина рассчитывают характеристики обслуживания, как если бы они были получены при наблюдении над реальным потоком покупателей.

При моделировании случайных явлений методом Монте-Карло мы пользуемся самой случайностью как аппаратом исследования. Заметим, что для сложных систем обслуживания с немарковским случайным процессом метод статистического моделирования, как правило, оказывается проще аналитического.

Вопросы для самопроверки

1. Что понимают под системой массового обслуживания? Из каких обслуживающих единиц она состоит? Что такое поток требований?
2. Какие показатели эффективности СМО используют на практике?
3. На какие два основных типа делят СМО?
4. Назовите структуру СМО и поясните содержание ее элементов.

5. Что такое случайный процесс? Что такое процесс с дискретными состояниями и процесс с непрерывным временем?
6. Какой процесс называют Марковским?
7. Что такое граф состояний?
8. Дайте понятие потока событий? Какие потоки событий вам известны?
9. Простейший поток.
10. Какие величины описывают уравнения Колмогорова?
11. Сформулируйте правило составления уравнений Колмогорова.
12. Что такое предельные вероятности состояний, каков их смысл?
13. Сформулируйте правило составления уравнений Колмогорова для предельных вероятностей.
14. Процесс гибели и размножений, граф состояний.
15. Запишите уравнения Колмогорова для процесса гибели и размножения.
16. Решение системы уравнений для процесса гибели и размножения.
17. Перечислите показатели эффективности СМО с отказами.
18. Запишите формулы для показателей эффективности одноканальной СМО с отказами.
19. В чем состоит отличие СМО с ограниченной очередью от СМО с неограниченной очередью? Как изменятся формулы для показателей эффективности СМО с ограниченной очередью?

Тема 12. Модели управления запасами

12.1. Основные понятия

Задачи управления запасами составляют один из наиболее многочисленных классов экономических задач исследования операций, решение которых .имеет важное народнохозяйственное значение. Правильное и своевременное определение оптимальной стратегии управления запасами, а также нормативного уровня запасов позволяет высвободить значительные оборотные средства, замороженные в виде запасов, что в конечном счете повышает эффективность используемых ресурсов.

Рассмотрим основные характеристики моделей управления запасами.

Спрос. *Спрос* на запасаемый продукт может быть *детерминированным* (в простейшем случае — постоянным во времени) или *случайным.* Случайность спроса описывается либо случайным моментом спроса, либо случайным объемом спроса в детерминированные или случайные моменты времени.

Пополнение склада. Пополнение склада может осуществляться либо периодически через определенные интервалы времени, либо по мере исчерпания запасов, т. е. снижения их до некоторого уровня.

Объем заказа. При периодическом пополнении и случайном исчерпании запасов объем заказа может зависеть от того состояния, которое наблюдается в момент подачи заказа. Заказ обычно подается на одну и ту же величину при достижении запасом заданного уровня — так называемой *точки заказа.*

Время доставки. В идеализированных моделях управления запасами предполагается, что заказанное пополнение доставляется на склад мгновенно. В других моделях рассматривается задержка поставок на фиксированный или случайный интервал времени.

Стоимость поставки. Как правило, предполагается, что стоимость каждой поставки слагается из двух компонент — разовых затрат, не зависящих от объема заказываемой партии, и затрат, зависящих (чаще всего — линейно) от объема партии.

Издержки хранения. В большинстве моделей управления запасами считают объем склада практически неограниченным, а в качестве контролирующей величины служит объем хранимых запасов. При этом полагают, что за хранение каждой единицы запаса в единицу времени взимается определенная плата.

Штраф за дефицит. Любой склад создается для того, чтобы предотвратить дефицит определенного типа изделий в обслуживаемой системе. Отсутствие запаса в нужный момент приводит к убыткам, связанным с простоем оборудования, неритмичностью производства и т. п. Эти убытки в дальнейшем будем называть *штрафом за дефицит.*

Номенклатура запаса. В простейших случаях предполагается, что на складе хранится запас однотипных изделий или однородного продукта. В более сложных случаях рассматривается *многоменклатурный запас.*

Структура складской системы. Наиболее полно разработаны математические модели одиночного склада. Однако на практике встречаются и более сложные структуры: иерархические системы складов с различными периодами пополнения и временем доставки заказов, с возможностью обмена запасами между складами одного уровня иерархии и т. п.

В качестве критерия эффективности принятой стратегии управления запасами выступает *функция затрат (издержек),* представляющая суммарные затраты на хранение и поставку запасаемого продукта (в том числе потери от порчи продукта при хранении и его морального старения, потери прибыли от омертвления капитала и т. п.) и затраты на штрафы.

Управление запасами состоит в отыскании такой стратегии пополнения и расхода запасами, при котором функция затрат принимает минимальное значение.

Ниже рассматриваются простейшие модели управления запасами.

Пусть функции $A(t)$, $B(t)$ и $R(t)$ выражают соответственно пополнение запасов, их расход и спрос на запасаемый продукт за промежуток времени $[0,t]$. В моделях управления запасами обычно используются производные этих функций по времени $a(t)$, $b(t)$, $r(t)$, называемые соответственно *интенсивностями пополнения, расхода и спроса.*

Если функции $a(t)$, $b(t)$, $r(t)$ — не случайные величины, то модель управления запасами считается *детерминированной,* если хотя бы одна из них носит случайный характер — *стохастической.* Если все параметры модели не меняются во времени, она называется *статической,* в противном случае — *динамической.* Статические модели используются, когда принимается разовое решение об уровне запасов на определенный период, а динамические — в случае принятия последовательных решений об уровнях запаса или корректировке ранее принятых решении с учетом происходящих изменений.

Уровень запаса в момент t определяется основным уравнением запасов

$$J(t) = J_0 + A(t) - B(t), \tag{12.1}$$

где J_0 — начальный запас в момент $t = 0$.

Уравнение (12.1) чаще используется в интегральной форме:

$$J(t) = J_0 + \int_0^t a(t)dt - \int_0^t b(t)dt. \tag{12.2}$$

♦ ***Пример 12.1.*** Интенсивность поступления деталей на склад готовой продукции цеха составляет в начале смены 5 дет./мин, в течение первого часа линейно возрастает, достигая к концу его 10 дет./мин, и затем остается постоянной. Полагая, что поступление деталей на склад происходит непрерывно в течение всех семи часов смены, а вывоз деталей со склада производится только в конце работы, записать выражение для уровня запаса в произвольный момент времени и, используя его, найти количество деталей на складе: а) через 30 мин после начала работы; б) в конце смены.

Решение. По условию в течение смены не происходит выдачи деталей со склада, т. е. $b(t)$ = 0. Интенсивность пополнения запаса в течение первого часа линейно возрастает, т. е. $a(t) = vt + c$. Учитывая, что $a(0)$ =5, получаем c = *5.* Так как в конце часа, т. е. при t = 60 $a(60)$ = 10, то 10 = $v \cdot 60$ + 5, откуда v = 1/12. Таким образом, для первого часа смены $a(t) = (1/12)t + 5$, а затем $a(t) = 10$.

Учитывая продолжительность смены (7 ч = 420 мин) и соотношение (12.2), получаем:

$$J(t) = \int_0^t (t/12 + 5)dt = t^2/24 + 5t,$$

Если $0 \le t \le 60$, и

$$J(t) = \int_0^{60} (t/12 + 5)dt + \int_{60}^t 10dt = \left(t^2/24 + 5t\right)\Big|_0^{60} + 10t\Big|_{60}^t =$$

$$= 450 + 10t - 600 = 10t - 150$$

Количество деталей на складе через 30 мин после начала работы: $J(30)$ = 900/24 + 5·30 = 187,5, а в конце смены: $J(420)$ = 10·420 - 150 = 4050. ►

12.2. Статическая детерминированная модель без дефицита

Предположение о том, что дефицит не допускается, означает полное удовлетворение спроса на запасаемый продукт, т.е. совпадение функций $r(t)$ и $b(t)$. Пусть общее потребление запасаемого продукта за рассматриваемый интервал времени θ равно N.

Рассмотрим простейшую модель, в которой предполагается, что расходование запаса происходит непрерывно с постоянной интенсивностью, т.е. $b(t)=b$. Эту интенсивность можно найти, разделив общее потребление продукта на время, в течение которого он расходуется:

$$b=\frac{N}{\theta}. \tag{12.3}$$

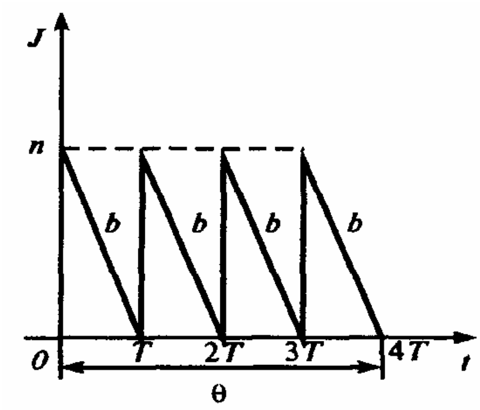


Рис. 12.1

Пополнение заказа происходит партиями одинакового объема, т.е. функция $a(t)$ не является непрерывной: $a(t)=0$ при всех t, кроме моментов поставки продукта, когда $a(t)=n$, где n — объем партии. Так как интенсивность расхода равна b, то вся партия будет использована за время

$$T=\frac{n}{b}. \tag{12.4}$$

Если отсчет времени начать с момента поступления первой партии, то уровень запаса в начальный момент равен объему этой партии n, т.е. $J(0)=n$. Графически уровень запаса в зависимости от времени представлен на рис 12.1.

На временном интервале $[0,T]$ уровень запаса уменьшается по прямой $J(t)=n-bt$ от значения n до нуля. Так как дефицит не допускается, то в момент T уровень запаса мгновенно пополняется до прежнего значения n за счет поступления партии заказа. И так процесс изменения $J(t)$ повторяется на каждом временном интервале продолжительностью T *(см.* рис. 12.1).

Задача управления запасами состоит в определении такого объема партии n, при котором суммарные затраты на создание и хранение запаса были бы минимальными.

Обозначим суммарные затраты через C, затраты на создание запаса — через C_1, затраты на хранение запаса — через C_2 и найдем эти величины за весь промежуток времени T.

Пусть затраты на доставку одной партии продукта, не зависимые от объема партии, равны c_1, а затраты на хранение одной единицы продукта в единицу времени — c_2. Так как за время θ необходимо запастись N единицами продукта, который доставляется партиями объема n, то число таких партий k равно:

$$k=\frac{N}{n}=\frac{\theta}{T}. \tag{12.5}$$

Отсюда получаем

$$C_1 = c_1 k = c_1 \frac{N}{n}. \qquad (12.6)$$

Мгновенные затраты хранения запаса в момент времени t равны $c_2 J(t)$. Значит, за промежуток времени $[0,T]$ они составят

$$c_2 \int_0^T J(t)dt = c_2 \int_0^T (n - bt)dt$$

Или, учитывая (12.4):

$$c_2 \int_0^T J(t)dt = c_2 \int_0^T \left(n - \frac{n}{T} t \right) dt = c_2 \left(nt - \frac{n}{2T} t^2 \right) \Bigg|_0^T = \frac{c_2 nT}{2}.$$

Средний запас за промежуток $[0,T]$ равен $nT/2$, т.е. *затраты на хранение всего запаса при линейном (по времени) его расходе равны затратам на хранение среднего запаса.*

Учитывая периодичность функции $J(t)$ (всего за промежуток времени θ будет $k = N/n$ «зубцов», аналогичных рассмотренному на отрезке $[0,T]$), и формулу (12.5), получаем, что затраты хранения запаса за промежуток времени θ равны:

$$C_2 = \frac{c_2 nT}{2} k = \frac{c_2 nT}{2} \frac{N}{n} = \frac{c_2 TN}{2} = \frac{c_2 \theta n}{2}. \qquad (12.7)$$

Нетрудно заметить, что затраты C_1 обратно пропорциональны, а затраты C_2 прямо пропорциональны объему партии n. Графики функций $C_1(n)$ и $C_2(n)$, а также функции суммарных затрат

$$C = \frac{c_1 N}{n} + \frac{c_2 \theta}{2} n \qquad (12.8)$$

приведены на рис. 12.2.

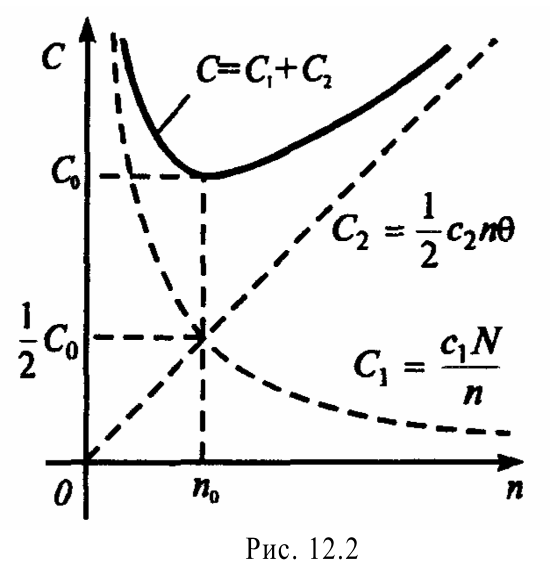


Рис. 12.2

В точке минимума функции $C(n)$ ее производная $C'(n) = -\frac{c_1 N}{n^2} + \frac{c_2 \theta}{2} = 0$, откуда

$$n = n_0 = \sqrt{\frac{2c_1 N}{c_2 \theta}} \qquad (12.9)$$

Или учитывая (12.3):

$$n_0 = \sqrt{\frac{2c_1 b}{c_2}}. \tag{12.10}$$

Формула (12.10), называемая *формулой Уилсона* или *формулой наиболее экономичного объема партии,* широко используется в экономике. Эта формула может быть получена и другим способом, если учесть, что произведение $C_1 C_2 = 0.5 c_1 c_2 N\theta$ есть величина постоянная, не зависящая от n. В этом случае, как известно, сумма двух величин принимает наименьшее значение, когда они равны, т. е. $C_1 = C_2$ или

$$\frac{c_1 N}{n} = \frac{c_2 \theta}{2} n, \tag{12.11}$$

откуда получаем (12.9).

Из (12.11) следует, что *минимум общих затрат задачи управления запасами достигается тогда, когда затраты на создание запаса равны затратам на хранение запаса.* При этом минимальные суммарные затраты

$$C_0 = C(n_0) = \frac{2c_1 N}{n}, \tag{12.12}$$

Откуда, учитывая (12.9) и (12.3), получим $C_0 = \sqrt{2c_1 c_2 \theta N}$ или

$$C_0 = \sqrt{2c_1 c_2 b}. \tag{12.13}$$

Число оптимальных партий за время θ с учетом (12.5), (12.9) и (12.3) равно:

$$k_0 = \frac{N}{n_0} = \sqrt{\frac{c_2 N\theta}{2c_1}} = \theta\sqrt{\frac{c_2 b}{2c_1}}.$$

Время расхода оптимальной партии на основании (12.4) с учетом (12.9) и (12.3) равно

$$T_0 = \frac{n_0}{b} = n_0 \frac{\theta}{N} \tag{12.14}$$

или

$$T_0 = \sqrt{\frac{2c_1 \theta}{c_2 N}} = \sqrt{\frac{2c_1}{c_2 b}}. \tag{12.15}$$

♦ ***Пример 12.2.*** Потребность сборочного предприятия в деталях некоторого типа составляет 120 000 деталей в год, причем эти детали расходуются в процессе производства равномерно и непрерывно. Детали заказываются раз в год и поставляются партиями одинакового объема, указанного в заказе. Хранение детали на складе стоит 0,35 ден. ед. в сутки, а поставка партии — 10 000 ден. ед. Задержка производства из-за отсутствия деталей недопустима. Определить наиболее экономичный объем партии и интервал между поставками, которые нужно указать в заказе (предполагается, что поставщик не допускает задержки поставок).

Решение. По условию затраты на поставку одной партии составляют c_1 = 10 000 ден. ед., затраты хранения единицы запаса в сутки c_2 = 0,35 ден. ед. Общий промежуток времени $\theta = 1$ год = 365 дней, а общий объем запаса за этот период N = 120 000 деталей. По формуле (12.9) $n_0 = \sqrt{\dfrac{2 \cdot 10000 \cdot 120000}{0.35 \cdot 365}} \approx 4335$ дет, а по (12.14) $T_0 = n_0 \dfrac{\theta}{N} = 13.2 \approx 13$ дней.

Итак, наиболее экономичный объем партии равен 4335 деталей, а интервал между поставками 13 дней. ►

На практике, естественно, объем партии может отличаться от оптимального n_0, вычисленного по (12.9). Так, в предыдущей задаче может оказаться удобным заказывать партии по 4 500 или даже по 5 000 деталей и возникает вопрос, как при этом изменятся суммарные затраты.

Для ответа на этот вопрос разложим функцию $C(n)$ в ряд Тейлора в окрестности точки n_0, ограничившись первыми тремя членами ряда при достаточно малых изменениях объема партии Δn:

$$C(n) = C(n_0) + C'(n_0)\Delta n + \frac{C''(n_0)}{2!}\Delta n^2 + \ldots .$$

Учитывая, что при $n = n_0$ $C'(n_0) = 0$, $C''(n_0) = \dfrac{2c_1 N}{n_0^3}$, а $C_0 = C(n_0)$ определяется по формуле (12.12), найдем

$$\frac{\Delta C}{C_0} = \frac{C(n) - C(n_0)}{C(n_0)} \approx \frac{C''(n_0)\Delta n^2}{2C(n_0)} = \frac{2c_1 N \Delta n^2}{n_0^3 (2c_1 N / n_0)}$$

или

$$\frac{\Delta C}{C_0} \approx \frac{1}{2}\left(\frac{\Delta n}{n_0}\right)^2 . \qquad (12.16)$$

Формула (12.16) свидетельствует об определенной устойчивости суммарных затрат по отношению к наиболее экономичному объему партии, ибо при малых Δn относительное изменение затрат примерно на порядок меньше относительного изменения объема партии по сравнению с оптимальным.

♦ ***Пример 12.3.*** По условию задачи **12.2** определить, на сколько процентов увеличатся затраты на создание и хранение запаса по сравнению с минимальными затратами при объеме заказываемых партий 5 000 деталей.

Решение. Относительное изменение объема партии по сравнению с оптимальным n_0 = 4335 составляет $\Delta n / n_0$ = (5000 – 4335)/4335=0,1510. В соответствии с (12.16) относительное изменение суммарных затрат составит $\Delta C / C_0$ =0,153^2/2 = 0,012, или лишь 1,2%. ►

♦ ***Пример 12.4.*** В условиях задачи **12.3** предположим, что заказываются не все партии сразу, а каждая отдельно, причем срок выполнения заказа равен 16 дней. Определить точки заказа, т. е. при каком уровне запаса следует заказывать следующую партию.

Решение. Так как по результатам решения задачи **12.2** длина интервала между поставками равна 13,2 дней, то заказ в условиях налаженного производства следует возобновить, когда уровень запаса достаточен для удовлетворения потребности на 16 - – 13,2 = 2,8 дня. Так как ежедневная потребность (интенсивность расхода запаса) равна по формуле (12.3) b = 120 000/365 = 329 деталей, то заказы должны делаться регулярно при достижении уровня запаса 329·2,8 = 922 деталей. ►

12.3. Статическая детерминированная модель с дефицитом

В рассматриваемой модели будем полагать наличие *дефицита.* Это означает, что при отсутствии запасаемого продукта, т.е. при $J(t)=0$ спрос сохраняется с той же интенсивностью $r(t)=b$, но потребление запаса отсутствует — $b(t)=0$, вследствие чего накап*ливается дефицит со скоростью* b. *График изменения уровня запаса в этом случае представлен на рис. 12.3. Убывание графика ниже оси абсцисс в область отрицательных значений в отличие от графика на рис. 12.2 характеризует накопление дефицита.*

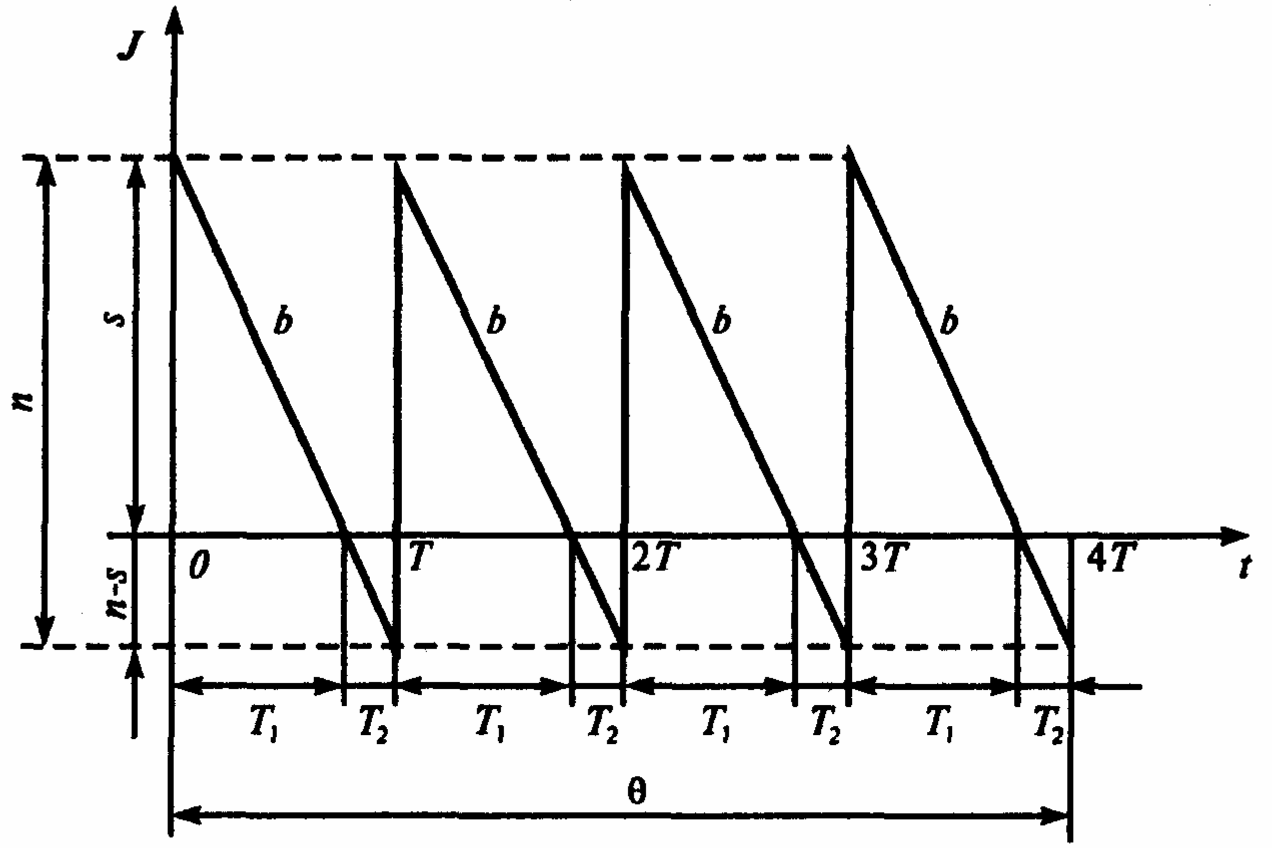


Рис. 12.3

Из рис. 12.3 видно, что каждый период "пилы" $T=\dfrac{n}{b}$ разбивается на два временных интервала, т. е. $T=T_1+T_2$, где T_1 — время, в течение которого производится потребление запаса, T_2 — время, когда запас отсутствует и накапливается дефицит, который будет перекрыт в момент поступления следующей партии.

Необходимость покрытия дефицита приводит к тому, что максимальный уровень запаса s в момент поступления каждой партии теперь не равен ее объему n, а меньше его на величину дефицита $n-s$, накопившегося за время T_2 (см. рис. 12.3).

Из геометрических соображений легко установить, что

$$T_1=\frac{s}{n}T\,,\; T_2=\frac{n-s}{n}T\,. \qquad (12.17)$$

В данной модели в функцию суммарных затрат C наряду с затратами C_1 (на пополнение запаса) и C_2 (на хранение запаса) необходимо ввести затраты C_3 — на штраф из-за дефицита, т.е.

$$C=C_1+C_2+C_3\,.$$

Затраты C_1 как и ранее, находим по формуле (12.6). В разд. 12.2 было показано, что затраты C_2 при линейном расходе запаса равны затратам на хранение среднего запаса, который за время потребления T_1 равен $sT_1/2$; поэтому с учетом (12.17) и (12.5) эти затраты составят

$$C_2 = \frac{c_2 s T_1}{2} k = \frac{c_2 s T_1 \cdot sT}{2} \cdot \frac{\theta}{T} = \frac{c_2 s^2 \theta}{2n}. \qquad (12.18)$$

При расчете затрат C_3 будем считать, что штраф за дефицит составляет в единицу времени c_3 на каждую единицу продукта. Так как средний уровень дефицита за период T_2 равен $(n-s)T_2/2$, то штраф за этот период T_2 составит $\frac{1}{2}c_3(n-s)T_2$, , а за весь период θ с учетом (12.7) и (12.17) —

$$C_3 = \frac{1}{2}c_3(n-s)T_2 k = \frac{1}{2}c_3(n-s)\frac{n-s}{n}T\frac{\theta}{T} = \frac{c_3\theta(n-s)^2}{2n}. \qquad (12.19)$$

Теперь, учитывая, (12.18) и (12.19), суммарные затраты равны

$$C = c_1\frac{N}{n} + \frac{c_2\theta s^2}{2n} + \frac{c_3\theta(n-s)^2}{2n}. \qquad (12.20)$$

Нетрудно заметить, что при $n = s$ формула (2.20) совпадает с ранее полученной (12.8) в модели без дефицита.

Рассматриваемая задача управления запасами сводится к отысканию такого объема партии n и максимального уровня запаса s, при которых функция C (12.20) принимает минимальное значение. Другими словами, необходимо исследовать функцию двух переменных $C(n,s)$ на экстремум. Приравнивая частные производные $\partial C/\partial n$ и $\partial C/\partial s$ к нулю, получим после преобразований систему уравнений:

$$\begin{cases} n^2 c_3 - (c_2 + c_3)s^2 = 2c_1 N/\theta, \\ s = n\dfrac{c_3}{c_2 + c_3}. \end{cases} \qquad (12.21)$$

Решая систему, получаем формулы наиболее экономичного объема партии $\tilde{n}_0$ и максимального уровня запаса $\tilde{s}_0$ для модели с дефицитом:

$$\tilde{n}_0 = \sqrt{\frac{2c_1 N}{c_2\theta}}\sqrt{\frac{c_2 + c_3}{c_3}} = \sqrt{\frac{2c_1 b}{c_2}}\sqrt{\frac{c_2 + c_3}{c_3}}, \qquad (12.22)$$

$$\tilde{s}_0 = \sqrt{\frac{2c_1 N}{c_2\theta}}\sqrt{\frac{c_3}{c_2 + c_3}} = \tilde{n}_0\frac{c_3}{c_2 + c_3}. \qquad (12.23)$$

Величина

$$\rho = \frac{c_3}{c_2 + c_3} \qquad (12.24)$$

называется *плотностью убытков из-за неудовлетворенного спроса* и играет важную роль в управлении запасами. Заметим, что $0 \le \rho \le 1$. Если значение c_3 мало по сравнению с c_2 то величина ρ близка к нулю; когда c_3 значительно превосходит c_2, то ρ близка к 1. Недопустимость дефицита равносильна предположению, что $c_3 = \infty$ или ρ = 1.

Используя (12.24), основные формулы (12.22) и (12.23) можно записать компактнее:

$$\tilde{n}_0 = \sqrt{\frac{2c_1 b}{c_2 \rho}}\,, \tag{12.25}$$

$$\tilde{s}_0 = \tilde{n}_0 \rho\,. \tag{12.26}$$

Следует учесть, что в силу (12.17) и (12.26) $T_1 / T = \tilde{s}_0 / \tilde{n}_0 = \rho$ и $T_2 / T = (\tilde{n}_0 - \tilde{s}_0) / \tilde{n}_0 = 1 - \rho$. Поэтому утверждение о том, что плотность убытков из-за неудовлетворенного спроса равна ρ, означает, что в течение (1-ρ)100% времени от полного периода T запас продукта будет отсутствовать.

Из сравнения формул (12.25) и (12.10) следует, что оптимальные объемы партий для задач с дефицитом и без дефицита при одинаковых параметрах связаны соотношением

$$\tilde{n}_0 = \frac{n_0}{\sqrt{\rho}}\,, \tag{12.27}$$

откуда вытекает, что *оптимальный объем партии в задаче с дефицитом всегда больше (в* $1/\sqrt{\rho}$ *раз), чем в задаче без дефицита.*

♦ ***Пример 12.5.*** Найти наиболее экономичный объем партии и интервал между поставками, сохраняя условия задачи **12.2**, кроме недопустимости дефицита, если известно, что отсутствие на сборке каждой детали приносит в сутки убытки в размере 3,5 ден. ед.

Решение. По условию c_3 = 3,5. Ранее было получено по формуле (12.9) n_0 =4335 и по (12.15) T_0 = 13,2. Найдем плотность убытков из-за неудовлетворенного спроса по формуле (12.24): ρ = 3,5/(0,35 + 3,5) = 0,909, т.е. 100·(1-0,909) = 9,1% времени между поставками детали на сборке будут отсутствовать.

Теперь оптимальный размер партии по формуле (12.27) $\tilde{n}_0$ = 4335/$\sqrt{0,909}$ = 4547. В силу (12.15) пропорционально увеличению $\tilde{n}_0$ должен увеличиться интервал между поставками, т.е. $\tilde{T}_0 = T_0 / \sqrt{\rho} = 13,2 / \sqrt{0,909} = 13,8 \approx 14$ дней. ►

12.4. Стохастические модели управления запасами

Рассмотрим *стохастические модели управления запасами,* у которых спрос является *случайным.* Этот факт существенным образом сказывается на характере соответствующих моделей и значительно усложняет их анализ, в связи с чем ограничимся рассмотрением наиболее простых моделей.

Предположим, что спрос r за интервал времени T является случайным и задан его закон (ряд) распределения $p(r)$ или плотность вероятностей $\varphi(r)$ (обычно функции $p(r)$ и $\varphi(r)$ оцениваются на основании опытных или статистических данных). Если спрос r ниже уровня запаса s, то приобретение (хранение, продажа) излишка продукта требует дополнительных затрат c_2 на единицу продукта; наоборот, если спрос r выше уровня запаса s, *то* это приводит к штрафу за дефицит c_3 на единицу продукции.

В качестве .функции суммарных затрат, являющейся в стохастических моделях случайной величиной, рассматривают ее среднее значение или математическое ожидание.

В рассматриваемой модели при дискретном случайном спросе r, имеющем закон распределения $p(r)$, математическое ожидание суммарных затрат имеет вид (учитываем только расходы на неиспользованные единицы продукта):

$$C(s) = c_2 \sum_{r=0}^{s} (s-r) p(r) + c_3 \sum_{r=s+1}^{\infty} (r-s) p(r). \qquad (12.28)$$

В выражении (12.28) первое слагаемое учитывает затраты на приобретение (хранение) излишка $(s-r)$ единиц продукта (при $r \le s$), а второе слагаемое — штраф за дефицит на $(r-s)$ единиц продукта (при $r > s$).

В случае непрерывного случайного спроса, задаваемого плотностью вероятностей $\varphi(r)$, выражение $C(s)$ принимает вид:

$$C(s) = c_2 \int_{0}^{s} (s-r)\varphi(r)dr + c_3 \int_{s}^{\infty} (r-s)\varphi(r)dr. \qquad (12.29)$$

Задача управления запасами состоит в отыскании такого запаса s, при котором математическое ожидание суммарных затрат (12.28) или (12.29) принимает минимальное значение.

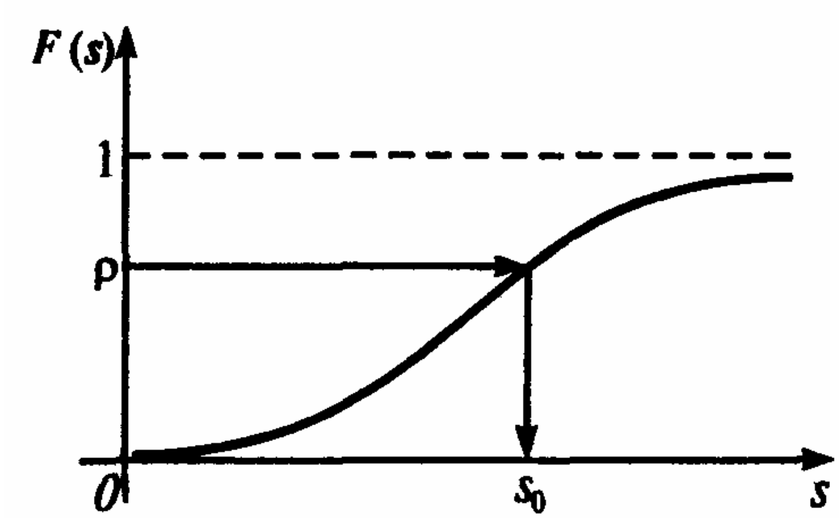


Рис. 12.4

Доказано, что при дискретном случайном спросе r выражение (12.28) минимально при запасе s_0, удовлетворяющем неравенствам

$$F(s_0) \le \rho \le F(s_0 + 1) \qquad (12.30)$$

а при непрерывном случайном спросе r выражение (12.29) минимально при значении s_0, определяемом из уравнения

$$F(s_0) = \rho, \qquad (12.31)$$

где

$$F(s) = p(r < s) \qquad (12.32)$$

есть функция распределения спроса r, $F(s_0)$ и $F(s_0+1)$ — ее значения; ρ — плотность убытков из-за неудовлетворенного спроса, определяемая по (12.24).

Оптимальный запас s_0 при непрерывном спросе по данному значению ρ может быть найден и графически (рис. 12.4).

♦ ***Пример 12.6.*** Предприятие закупает агрегат с запасными блоками к нему. Стоимость одного блока равна 5 ден. ед. В случае выхода агрегата из строя из-за поломки блока, отсутствующего в запасе, простой агрегата и срочный заказ нового блока к нему обойдется в 100 ден. ед. Опытное распределение агрегатов по числу блоков, потребовавших замену, представлено в табл. 12.1.

Таблица 12.1

Число замененных блоков r	0	1	2	3	4	5	6
Статистическая вероятность (доля) агрегатов $p(r)$, которым потребовалась замена r блоков	0,90	0,05	0,02	0,01	0,01	0,01	0,00

Необходимо определить оптимальное число запасных блоков, которое следует приобрести вместе с агрегатом.

Решение. По условию $c_2 = 5$, $c_3 = 100$. Вычислим плотность убытков из-за нехватки запасных блоков по формуле (12.24) $\rho = 100/(5+100) = 0{,}952$.

Учитывая (12.32), найдем значения функции распределения спроса (табл. 12.2).

Таблица 12.2

r	0	1	2	3	4	5	6
$F(r)$	0,00	0,90	0,95	0,97	0,98	0,99	1,00

Очевидно (см. табл. 12.2), что оптимальный запас составит $s_0 = 2$, ибо он удовлетворяет неравенству (12.30): $F(2) < 0{,}952 < F(3)$. ►

Пример 12.7. Решить задачу **12.6** при условии непрерывного случайного спроса r, распределенного по показательному закону с функцией распределения $F(r) = 1 - e^{-\lambda r}$ при $\lambda = 0{,}98$.

Решение. Оптимальное число запасных блоков s_0 найдем из уравнения (12.31):

$$1 - e^{-\lambda s_0} = \rho, \quad \text{откуда} \quad e^{-\lambda s_0} = 1 - \rho \quad \text{и} \quad s_0 = \frac{1}{\lambda}\ln(1-\rho).$$

При $\lambda = 0{,}98$ $\quad s_0 = -(1/0{,}98)\ln 0{,}048 \approx 3$ (блока). ►

В условиях рассматриваемой модели предположим, что расходование запаса происходит *непрерывно с* одинаковой интенсивностью. Такую ситуацию можно представить графически (рис. 12.5).

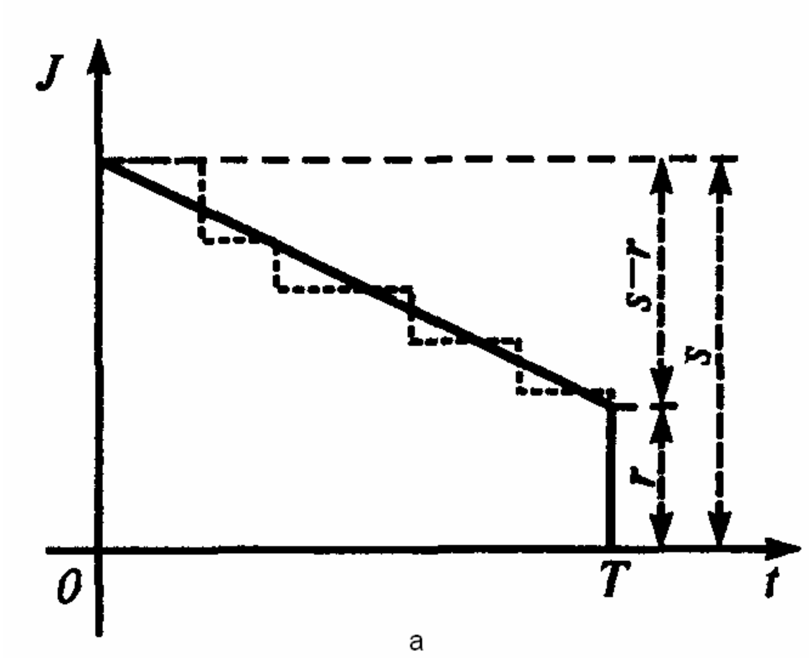


Рис. 12.5а

Рис. 12.5, *а* соответствует случаю $r \le s$, когда спрос не превосходит запаса, а рис. 12.5*б* — случаю, когда спрос превышает запас, т.е. $r > s$. Следует отметить, что на самом деле график $J(t)$ представляет ступенчатую ломаную, показанную на рис. 12.5 пунктиром, но для исследования модели нам проще рассматривать $J(t)$ *в* виде прямой, сглаживающей эту ломаную.

Средний запас, соответствующий рис. 12.5*а,* равен

$$\bar{s}_1 = \frac{1}{2}(s + (s - r)) = s - \frac{1}{2}r. \tag{12.33}$$

Средний запас, соответствующий рис. 12.5*б* с учетом формулы (12.17), в которой полагаем $n = r$, составляет

$$\bar{s}_2 = \frac{1}{2}s\frac{T_1}{T} = \frac{1}{2}\frac{s^2}{r}. \tag{12.34}$$

Формула (12.34) получается следующим образом. Из (12.17) имеем: $sT_1 = \dfrac{s^2 T}{n}$. Так как $n = r$, то $sT_1 = \dfrac{s^2 T}{r}$. Отсюда следует (12.34).

Средний дефицит продукта за период T_2 для случая, соответствующего рис. 12.5*б с* учетом (12.17), где $n = r$, равен

$$\bar{s}_3 = \frac{1}{2}(r - s)\frac{T_2}{T} = \frac{1}{2}\frac{(r - s)^2}{r}. \tag{12.35}$$

Математическое ожидание суммарных затрат составит:

$$C(s) = c_2\sum_{r=0}^{s}(s - \frac{r}{2})p(r) + c_2\sum_{r=s+1}^{\infty}\frac{1}{2}\frac{s^2}{r}p(r) + c_3\sum_{r=s+1}^{\infty}\frac{1}{2}\frac{(r - s)^2}{r}p(r). \tag{12.36}$$

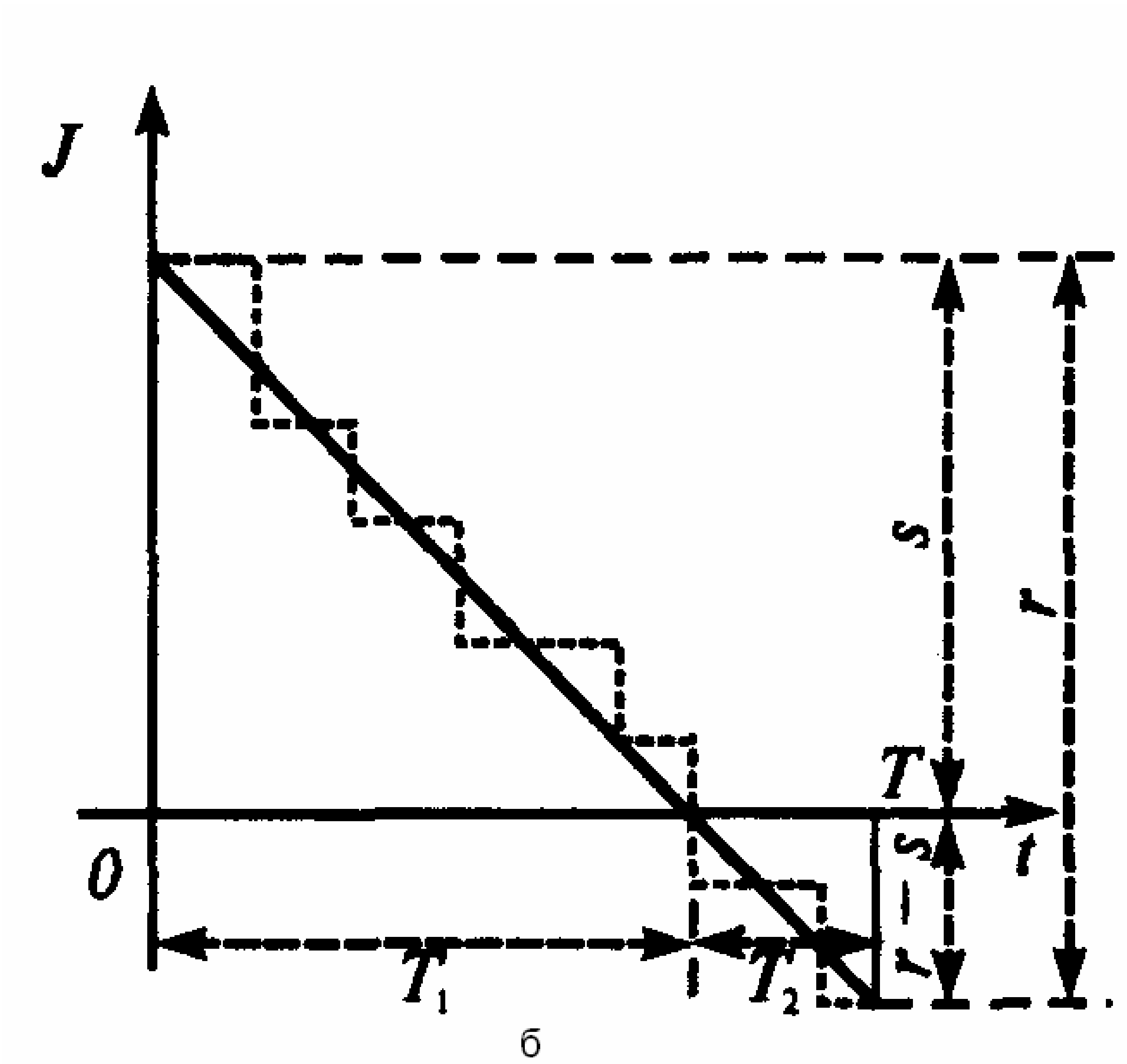


Рис. 12.5б

Доказано, что в этом случае математическое ожидание (12.36) минимально при запасе s_0, удовлетворяющем неравенству

$$L(s_0) < \rho < L(s_0+1), \tag{12.37}$$

где ρ по-прежнему определяется по формуле (12.24);

$$L(s) = F(s) + \left(s - \frac{1}{2}\right)\sum_{r=s}^{\infty}\frac{p(r)}{r}, \tag{12.38}$$

$L(s_0)$ и $L(s_0+1)$ — значения функции (12.38), а $F(s)$ находится в соответствии с определением (12.32).

♦ **Пример 12.8**. Имеющиеся на складе изделия равномерно расходуются в течение месяца. Затраты на хранение одного изделия составляют 5 ден. ед., а .штраф за дефицит одного изделия обходится в 100 ден. ед. Изучение спроса дало распределение числа потребляемых за месяц изделий, представленное в табл. 12.10.

Таблица 12.3

Спрос r	0	1	2	3	4	5	>6
Статистичская вероятность $p(r)$	0.1	0.2	0,2	0,3	0,1	0,1	0,0

Необходимо, определить оптимальный месячный запас склада.

Решение. Так же как в задаче **12.6**, c_2 = 5, c_3 = 100, ρ = 0,952.

Значения функции $L(r)$ определим с помощью табл. 12.4.

Таблица 12.4

r	$p(r)$	$\frac{p(r)}{r}$	$\sum_{r=s}^{\infty}\frac{p(r)}{r}$	$\left(s-\frac{1}{2}\right)\sum_{r=s}^{\infty}\frac{p(r)}{r}$	$F(r)$	$L(r)$
0	0,1	—	—	—	0,0	—
1	0,2	0,200	0,445	0,2225	0,1	0,3225
2	0,2	0,100	0,245	0,3675	0,3	0,6675
3	0,3	0,100	0,145	0,3625	0,5	0,8625
4	0,1	0,025	0,045	0,1575	0,8	0,9575
5	0,1	0,020	0,020	0,0900	0,9	0,9900
≥6	0,0	0,000	0,000	0,0000	1,0	1,0000

Очевидно, что оптимальный запас изделий $s_0 = 3$, так как он удовлетворяет условию (12.37): $L(3) < 0,952 < L(4)$. ♦

12.5. Стохастические модели управления запасами с фиксированным временем задержки поставок

В рассмотренных выше идеализированных моделях управления запасами предполагалось, что пополнение запаса происходит практически мгновенно. Однако в ряде задач *время задержки* поставок может оказаться настолько значительным, что его необходимо учитывать в модели.

Пусть за время задержек поставок θ уже заказаны n партий по одной в каждый из n периодов продолжительностью $T = \theta / n$.

Обозначим:

$s_{нз}$ — первоначальный уровень запаса (к началу первого периода);

s_i — запас за i-й период;

r_i — спрос за i-й период;

q_i — пополнение запаса за i-й период.

Тогда к концу n-го периода на склад поступит $\sum_{i=1}^{n} q_i$ единиц продукта, а будет израсходовано $\sum_{i=1}^{n} r_i$ единиц, т.е.

$$s_n = s_{нз} + \sum_{i=1}^{n} q_i - \sum_{i=1}^{n} r_i \tag{12.39}$$

Или

$$s_n = s - r, \tag{12.40}$$

где

$$s = s_{нз} + \sum_{i=1}^{n} q_i, \tag{12.41}$$

$$r = \sum_{i=1}^{n} r_i. \tag{12.42}$$

Требуется найти оптимальный объем партии заказа, который необходимо сделать за последний n-й период, предшествующий поступлению сделанного ранее заказа.

Математическое ожидание суммарных затрат в этом случае определяется по формуле (12.28), а оптимальный запас s находится по формуле (12.30), т.е.

$$F(s_0) < \rho < F(s_0 + 1). \tag{12.43}$$

Найдя оптимальный запас s_0 и зная $q_1, q_2, ..., q_{n-1}$, можно вычислить q_n по формуле (12.41), т.е.

$$q_n = s_0 - \left(s_{нз} + \sum_{i=1}^{n-1} q_i \right). \tag{12.44}$$

♦ ***Пример 12.9.*** Ежедневно заказываемый скоропортящийся товар поступает в магазин спустя 7 дней после заказа. В момент очередного заказа запас товара составил в стоимостном выражении 10 ден. ед. Проданный товар в день изготовления, приносит прибыль 0,95 ден, ед., а не проданный в этот день товар может быть затем реализован с убытком 0,10 ден. ед.

На основании опытных данных получено следующее распределение спроса на данный товар (табл. 12.5). Кроме того, дано пополнение запаса за 6 дней

Таблица 12.5

r	0	10	20	30	40	50	60	70	80	90	100

$p(r)$	0,00	0,00	0,01	0,02	0,05	0,08	0,11	0,12	0,14	0,13	0,10
r	110	120	130	140	150	160	170	180	190	200	>200
$p(r)$	0,08	0,05	0,03	0,02	0,02	0,01	0,01	0,01	0,01	0,00	0,00

Таблица 12.6

i	1	2	3	4	5	6
q_i	10	20	10	10	20	10

Необходимо определить оптимальный объем заказанного товара q_7 на седьмой день после заказа.

Решение. Плотность убытков из-за дефицита товара по формуле (12.24) равна ρ = 0,95/(0,10 + 0,95) = 0,905. Учитывая условия (12.32), найдем значения функции распределения спроса (табл. 12.7).

Таблица 12.7.

s	r	$F(s)$	s	r	$F(s)$	s	r	$F(s)$	s	r	$F(s)$
0	0	0,00	50	50	0,16	100	100	0,86	150	150	0,96
10	10	0,00	60	60	0,27	110	100	0,84	160	160	0,97
20	20	0,01	70	70	0,39	120	120	0,89	170	170	0,98
30	30	0,03	80	30	0,53	130	130	0,92	180	180	0,99
40	40	0,08	90	90	0,76	140	140	0,94	≥ 190	≥ 190	1,00

Условию (12.43) удовлетворяет s_0 = 120, ибо F *(120)* < 0,905 < F (130).

Таким образом, оптимальный запас товара за 7 дней должен быть на сумму 120 ден. ед., откуда оптимальный объем заказанного товара на седьмой день по (12.44) составит: q_7= 120 - (10 + (10 + 20 +10 +10 + 20 +10)) = 30 ден. ед. ►

Следует отметим, что найти аналитически оптимальные значения точки запаса s_0 и объема партии n удается только в относительно простых случаях. Если же система хранения запасов имеет сложную структуру (много видов хранимой продукции, иерархическая система складов), используемые стохастические модели сложны, а их параметры меняются во времени, то единственным средством анализа такой системы становится *имитационное моделирование,* позволяющее имитировать ("проигрывать") на ЭВМ функционирование системы, исследуя ее поведение при различных условиях, значениях параметров, отражая их случайный характер, изменение во времени и т.п.

Вопросы для самопроверки

1. Перечислите основные характеристики моделей управления запасами.
2. Дайте определения основных характеристик моделей управления запасами.
3. Запишите уравнение запаса.
4. Запишите формулу наиболее экономичного объема партии в статической детерминированной модели без дефицита.

5. Запишите формулу затрат на создание запаса в статической детерминированной модели без дефицита.
6. Запишите формулу затрат на хранение запаса в статической детерминированной модели без дефицита.
7. Запишите формулу минимальных суммарных затрат на создание и хранение запаса в статической детерминированной модели без дефицита.
8. Запишите формулу времени расхода оптимальной партии.
9. Запишите формулу наиболее экономичного объема партии в статической детерминированной модели с дефицитом.
10. Запишите формулу затрат на создание запаса в статической детерминированной модели с дефицитом.
11. Запишите формулу затрат на хранение запаса в статической детерминированной модели с дефицитом.
12. Запишите формулу затрат на штраф за отсутствие запаса в статической детерминированной модели с дефицитом.
13. Запишите формулу минимальных суммарных затрат на создание, хранение запаса и штраф из-за дефицита запаса в статической детерминированной модели с дефицитом.
14. Что такое плотность убытков из-за неудовлетворенного спроса.
15. Запишите формулу времени между поставками партии.
16. Как определить оптимальный запас в стохастической модели при дискретном случайном спросе.
17. Как определить оптимальный запас в стохастической модели при непрерывном случайном спросе.
18. Запишите формулу для среднего запаса, когда спрос не превосходит запаса.
19. Запишите формулу для среднего запаса, когда спрос превосходит запас.
20. Запишите формулу суммарных затрат в стохастической модели управления запасами.
21. Как определить оптимальный запас в стохастической модели с фиксированным временем задержки поставок.
22. Как определить пополнение запаса в последний период в стохастической модели с фиксированным временем задержки поставок.

ЛИТЕРАТУРА

1. Гультяев А.К. MATLAB 5.3. Имитационное моделирование в среде Windows : практическое пособие. - СПб. : КОРОНА принт, 2001. - 400 с.
2. Лоу А.М., Кельтон В.Д. Имитационное моделирование / Пер. с англ. ред. пер. : В. Н. Томашевский. - 3-е изд. - СПб. : Питер, Киев : BHV, 2004. - 847 с
3. Кобелев Н.Б. Основы имитационного моделирования сложных экономических систем : учебное пособие для вузов. - М. : Дело, 2003. – 335с.
4. Кундышева Е.С. Математическое моделирование в экономике: Учебное пособие/ Под науч. ред. проф. Б.А. Суслакова. – М.:Издательско-торговая корпорация «Дашков и К», 2004. –352 с.
5. Минюк С.А. Математические методы в экономике: Учеб. пособие/ Минюк С.А., Ровба Е.А., Кузьмич К.К. – Мн.: ТетраСистемс, 2002. – 432 с.
6. Домбровский В.В. Методы количественного анализа финансовых операций. – Томск: ТГУ, 1998. – 104с.
7. Малыхин В.И. Финансовая математика. – М.:ЮНИТИ-ДАНА, 2000. – 247с.
8. Экономико – математические методы и прикладные модели. Под ред. Федосеева В.В. – М.: ЮНИТИ, 2001. – 391с
9. Замков О.О., Черемных Ю.А., Толстопятенко А.В. Математические методы в экономике. –М.: «Дело и сервис», 1999. –365с.
10. Буров А.В., Миньков С.Л., Ушаков В.М. Моделирование экономических процессов и систем. Учебное пособие. Часть 1. – Томск. Изд – во ТГПУ, 2001. –158с.
11. Лукасевич И.Я. Анализ финансовых операций. Методы, модели, техника вычислений. – М.: ЮНИТИ, 1998. – 400 с.
12. Емельянов А.А., Власова Е.А., Дума Р.В. Имитационное моделирование экономических процессов. – М.: Финансы и статистика, 2002. - 368 с.
13. Борщев А.В. Применение имитационного моделирования в России – состояние на 2007 г.// Материалы III Всероссийской научно-практической конференции ИММОД-2007. - Санкт-Петербург, 17-19 октября 2007 г.
14. Paul Klemperer. Auctions: Theory and Practice. - Princeton University Press, 2004. - 256 pp.
15. Варфоломеев В.И. Алгоритмическое моделирование элементов экономических систем. – М.: Финансы и статистика, 2000. - 203 с.
16. Кельтон В., Лоу А. Имитационное моделирование. – СПб.: Питер; Киев: Издательская группа BHV, 2004. – 847 с.
17. Мицель А.А., Грибанова Е.Б. Сборник задач по имитационному моделированию экономических процессов (с грифом СибРУМЦ). – Томск: изд-во ТУСУР, 2007. – 218с. (50 экз.)
18. Мицель А.А., Грибанова Е.Б. Имитационное моделирование экономических процессов. Часть 1. Теоретические основы имитационного моделирования экономических процессов. Учебное пособие. Томск: Изд-во ТМЦ ДО, 2005. – 137с. (2 экз)
19. Мицель А.А., Грибанова Е.Б. Имитационное моделирование экономических процессов. Часть 2. Алгоритмические модели экономических систем. Учебное пособие. Томск: Изд-во ТМЦ ДО, 2005. – 99с. (3 экз)
20. Кремер Н.Ш. и др. Исследование операций в экономике. Учебное пособие для вузов/ ред. : Н. Ш. Кремер. - М. : ЮНИТИ, 2006. - 407 с (20 экз)
21. Бережная Е.В., Бережной В.И. Математические методы моделирования экономических систем. — М.:Финансы и статистика, 2002. – 368с. (7 экз)
22. Мицель А.А. Математическая экономика. Лабораторный практикум. – Томск: Изд-во НТЛ, 2006. – 184 с. (65 экз)

23. Мицель А.А. Математическая экономика: методические указания по самостоятельной и индивидуальной работе студентов всех форм обучения для специальности 080801.65 «Прикладная информатика в экономике». – Томск: ТУСУР, 2012. – 35 с. (электронный ресурс). – Режим доступа:
24. http://asu.tusur.ru/learning/spec080801/d34/s080801_d34_work.doc
25. Мицель А.А., Грибанова Е.Б. Имитационное моделирование экономических объектов. Лабораторный практикум. (с грифом СибРУМЦ)– Томск: Изд-во НТЛ, 2005. – 160с. (141 экз)
26. Мицель А.А., Грибанова Е.Б. Имитационное моделирование экономических процессов. Методические указания по выполнению лабораторных работ и курсового проекта. – Томск: Изд-во ТУСУР, 2006. – 108с. (80 экз)
27. Мицель А.А., Грибанова Е.Б. Имитационное моделирование экономических процессов. Учебное методическое пособие. Томск: Изд-во ТМЦ ДО, 2007. – 143с. (8 экз)
28. Мицель А.А. Имитационное моделирование экономических процессов: методические указания по самостоятельной и индивидуальной работе студентов всех форм обучения для специальности 080801 – Прикладная информатика в экономике. – Томск: ТУСУР, 2012. – 7с. (электронный ресурс). – Режим доступа: http://asu.tusur.ru/learning/spec080801/d35/s080801_d35_work.doc