ОГЛАВЛЕНИЕ

Тема 1. Модели финансовых ситуаций

Операции дисконтирования

Дисконтирование связано с распространенным в коммерческой сфере утверждением «время – это тоже деньги», что обусловлено неравноценностью одинаковых по абсолютной величине сумм денежных средств сегодня и в будущем. Это объясняется, например, возможностью инвестировать сегодня капитал и в будущем получить доход; кроме того, инфляционный процесс обесценивает денежную массу. Таким образом, можно утверждать, что «деньги сегодня» ценнее «будущих денег». Дисконтирование позволяет учитывать в операциях фактор времени, то есть решать вопрос, как соотносятся между собой суммы денег, полученные в различные моменты времени.

Различают математическое дисконтирование, коммерческий или банковский учет.

Математическое дисконтирование связано с определением так называемого современного, или приведенного, значения P на некоторый момент времени, которое соответствует заданному значению S в другой момент времени. Простейшая задача связана с определением суммы вклада P на основе заданной конечной величины в будущем S через временной период начислений n под заданную ставку процентов, например, начисленную без учета капитализации:

$$P = \frac{S}{1+ni} = S \cdot k_Д, \qquad (1.1)$$

где i – годовая процентная ставка;

n – количество периодов начисления процентов;

$k_Д$ – коэффициент дисконтирования (приведения).

Дисконтированное значение будущей суммы вклада с учетом капитализации равно:

$$P = \frac{S}{(1+i_c)^n} = S \cdot k_{ДС}, \qquad (1.2)$$

где i_C – годовая процентная ставка,

а по номинальной ставке процентов i_n

$$P = \frac{S}{(1+i_n/m)^{mn}}, \qquad (1.3)$$

где m – количество начислений процентов за год;

i_n – процент в пересчете на год.

Обычно понятие современной стоимости применяется к потоку платежей (во времени).

Пример решения задачи. Сколько нужно положить денег в банк под 20 % годовых (при условии ежегодной капитализации), чтобы через 2 года получить 250 тыс. рублей.

По формуле (1.2) находим: $P=\frac{250}{(1+0,2)^2}=\frac{250}{1,44}=173{,}61$, то есть в банк нужно положить 173,61 тыс. рублей.

Банковский учет заключается в покупке денежных обязательств банком, например векселя, по цене меньше номинальной указанной в нем суммы. В этом случае вексель учитывается, и клиент получит сумму:

$$P = S - D,$$

где S – номинальная стоимость данного обязательства;

P – цена покупки векселя банком;

D – дисконт, сумма процентных денег.

Вексель – письменное долговое обязательство строго установленной законом формы, которое выдается заемщиком (векселедателем) кредитору (векселедержателю) и предоставляет право векселедержателю требовать с заемщика уплаты к определенному сроку суммы займа и вознаграждения.

Процентный доход покупателя векселя определяется, например, по простой учетной ставке:

$$d\% = \frac{D}{S}100\ \%.$$

Если срок n от даты учета до даты погашения будет составлять часть года, то дисконт определяется по формуле:

$$D = n \cdot d \cdot S = \frac{t}{T} \cdot d \cdot S,$$

где d – относительная величина учетной ставки;

t – период начисления в днях;

T – количество дней в году.

Предъявителю учитываемого денежного обязательства будет выдана сумма:

$$P = S - D = S(1 - nd) = S(1 - \frac{t}{T}d). \qquad (1.4)$$

Пример решения задачи. Банк учитывает вексель под 25 % годовых, до погашения осталось 90 дней, номинальная стоимость 100 тыс. рублей. Какую сумму получит предъявитель учитываемого векселя.

По формуле (1.4): $P = 100 \cdot \left(1 - \frac{90}{365} \cdot 0{,}25\right) \approx 100 \cdot 0{,}93836 \approx 93{,}836$. Т. е. предъявитель получит 93,836 тыс. рублей.

Следует заметить, что дисконтирование может быть связано и с проведением кредитной операции. В таком случае проценты начисляются в начале интервала начисления и заемщик получает сумму P за вычетом процентных денег D из суммы кредита S, подлежащей возврату. В таком случае при проведении операции по простой учетной ставке d следует пользоваться такой формулой:

$$S = \frac{P}{1 - nd}.$$

При проведении операции по сложной учетной ставке d_c% следует пользоваться формулой:

$$S = \frac{P}{(1-d_c)^n}. \qquad (1.5)$$

При разработке условий контрактов или их анализе иногда возникает необходимость в решении обратных задач – определении срока ссуды или уровня учетной ставки.

Формулы для расчета продолжительности ссуды и величины учетной ставки получаем, решив уравнение (4.5) относительно n и d.

$$n = \frac{\ln {}^{P}\!/_{S}}{\ln(1-d_c)}; \qquad d_c = 1 - \sqrt[n]{\frac{P}{S}}.$$

Выгодность такого метода начисления процентов по учетной ставке для кредитора или заемщика зависит от величины процентной ставки и срока кредита.

<u>*Пример решения задачи.*</u> На сколько лет нужно взять кредит в 390,625 тыс. рублей под 20 % годовых с учетом капитализации, чтобы получить 200 тыс. рублей.

$$n = \frac{\ln {}^{200}\!/_{390,625}}{\ln(1-0.2)} = \frac{\ln 0.512}{\ln 0.8} = 3$$, т. е. кредит нужно взять на 3 года.

Легко заметить, что при ставке 20 % годовых без учета капитализации, этот срок составит $n = \frac{1 - {}^{P}\!/_{S}}{d} = \frac{1-0.512}{0.2} = 2{,}44$, т. е. 2,44 лет.

В операциях используется и номинальная годовая учетная ставка i_n, по которой при начислении процентов m раз в году можно определить сумму кредита:

$$S = \frac{P}{(1 - {}^{i_n}\!/_{m})^{mn}},$$

из которой находим следующие модели расчета продолжительности ссуды и величины учетной ставки:

$$n = \frac{\ln {}^{p}\!/_{S}}{m \ln(1 - {}^{i_n}\!/_{m})}; \qquad i_n = m(1 - \sqrt[mn]{\frac{P}{S}}).$$

Из приведенных моделей путем несложных преобразований можно получить формулы для расчета различных показателей финансовых операций.

Эквивалентность процентных ставок

При заключении финансовых контрактов каждый участник сделки стремится заключить контракт на наиболее выгодных для себя условиях. Условия контракта могут быть различными, и надо иметь возможность сравнивать контракты. При этом различные контракты могут предусматривать различные виды начисления процентов и для сравнения таких контрактов надо разработать способы приведении различных процентных ставок к одному виду. Напомним формулы для вычисления наращенной суммы S для всех видов процентных ставок:

$S = P(1+ni)$ – начисление простых процентов;

$S = P(1+i_c)^n$ – начисление сложных процентов;

$S = P\left(1+\frac{i_n}{m}\right)^{m\cdot n}$ – начисление процентов m раз в год;

$S = \frac{P}{1-nd}$ – простой дисконт (процентный доход, вычитаемый из ссуды в момент ее выдачи);

$S = \frac{P}{(1-d_c)^n}$ – учет по сложной учетной ставке;

$S = \frac{P}{\left(1-\frac{d_n}{m}\right)^{m\cdot n}}$ – учет по сложной учетной ставке m раз в году.

Во всех формулах есть число лет (*n*), оно может быть дробным.

Две процентные ставки называются <u>эквивалентными</u>, если применение их к одинаковым суммам в течение одинаковых промежутков времени дает одинаковые наращенные суммы.

Приравнивая правые части каких-либо двух из приведенных выше формул и выражая из этого равенства одну процентную ставку через другую, получаем условие эквивалентности соответствующих процентных ставок за *n* лет.

<u>*Пример решения задачи*</u>

Определите значение учетной ставки банка, эквивалентной ставке процентов, равной 40 % годовых.

Решение. Рассмотрим будущую стоимость этих сумм для одного года: $S = P(1+i)$ и $S = \frac{P}{1-d}$. Так как левые части равны, то равны и правые: $\frac{P}{1-d} = P(1+i)$. Выражая из этого равенства учетную ставку d, получаем: $d = 1-\frac{1}{1+i}$ или $d = \frac{i}{1+i}$.

Таким образом, для заданной i = 40 %

$$d = \frac{0{,}4}{1+0{,}4} = 0.2857 \quad (28{,}5\,\%).$$

Для расчета эффективности финансовых операций используют сравнительную доходность, которая на основе допущения о равенстве финансовых результатов различных вариантов инвестиций приводит к понятию эквивалентных ставок простых или сложных процентов. Это позволяет получить инструмент корректного сравнения финансовых операций.

Эффективная ставка процентов измеряет относительный доход, который получают в целом за год от начисления процентов несколько раз в год:

$$j = \left(1 + \frac{i_n}{m}\right)^m - 1 \qquad (1.6)$$

Иначе говоря, эффективная ставка показывает, какая годовая ставка сложных процентов дает такой же финансовый результат, что и m-разовое начисление в год по ставке i/m.

Под эффективной учетной ставкой понимают сложную годовую учетную ставку, эквивалентную номинальной при заданном значении m и вычисляют по формуле:

$$f = \left(1 - \frac{d_c}{m}\right)^m - 1$$

Пример решения задачи. Банк начисляет проценты на вклад исходя из номинальной ставки 12 % годовых. Определить эффективную (годовую) ставку при ежедневной капитализации процентов.

Решение: m = 365. По формуле (1.6) получим:

$$j = \left(1 + \frac{i_n}{m}\right)^m - 1 = \left(1 + \frac{0{,}12}{365}\right)^{365} - 1 = 0{,}12747,$$ то есть 12,747 %.

Модели денежных потоков

Денежные потоки являются составной и неотъемлемой частью практически любой сферы деятельности. В коммерции они образуют питательную среду товародвижения. В экономической, финансовой, производственной и других сферах эти потоки также порождают интерес. Примерами таких потоков являются: денежные поступления от продажи товаров и предоставления услуг; денежные поступления от аренды, гонорары, комиссионные и другие доходы; денежные выплаты поставщикам товаров и услуг; денежные выплаты работникам; денежные выплаты в качестве авансов; денежные поступления и выплаты по контрактам, заключенным для коммерческих или торговых целей; проценты по кредитам.

При этом может возникать целый ряд последовательных, например равновеликих, платежей R, которые и образуют денежный поток.

Ряд последовательных платежей, производимых через равные промежутки времени τ, называется финансовой рентой или аннуитетом. Примером аннуитета могут быть регулярные взносы в пенсионный или другие фонды, выплаты процентов по ценным бумагам, например по акциям и т. д. Финансовая рента определяется следующими основными характеристиками:

- члены ренты R_j – величина каждого отдельного платежа;
- интервал ренты τ_j – временной интервал между двумя платежами;
- срок ренты t – время от начала реализации ренты до момента последнего платежа (бывают вечные ренты);

- процентная ставка для расчета наращения или дисконтирования платежей;
- наращенная будущая сумма ренты S, включающая все члены потока платежей с процентами на дату последней выплаты;
- современная (приведенная) величина ренты A – сумма всех членов потока платежей, дисконтированная (уменьшенная) на величину процентной ставки на начальный момент времени ренты.

Ренты подразделяются на постоянные, когда члены ренты равны: $R_1=R_2=R_3=...=R_n$, и переменные, когда величины платежей различны.

Рассмотрим модели потоков ежегодных платежей, на которые начисляются проценты в конце каждого года методом сложных процентов.

Выражение для наращенной суммы ренты имеет вид:

$$S = R\frac{(1+i_c)^n - 1}{i_c},$$

из которого следует, что коэффициент наращения можно определить следующим выражением:

$$k_{на} = \frac{(1+i_c)^n - 1}{i_c}.$$

Полученные модели позволяют определить, например, величину платежа

$$R = \frac{S}{k_н} = \frac{S \cdot i_c}{(1+i_c)^n - 1},$$

Для определения срока ренты можно получить следующие формулы:

$$n = \frac{\ln\left[\left(S/R\right)i_c + 1\right]}{\ln(1+i_c)}.$$

Приведенная стоимость ренты будет вычисляться по формуле:

$$A = R\frac{1-(1+i_c)^{-n}}{i_c}.$$

Эту формулу можно использовать, когда заемщик берет кредит на условиях его погашения в будущем равными платежами ежегодно:

$$R = \frac{A \cdot i_c}{1-(1+i_c)^{-n}}. \quad (1.7)$$

Пример решения задачи. Заемщик берет кредит на сумму 100 тыс. рублей на пять лет под 25 % годовых. Собирается погашать ежегодно равными частями. Какова величина этих выплат.

По формуле (4.7): $R = \dfrac{100 \cdot 0{,}25}{1-\dfrac{1}{(1+0{,}25)^5}} = \dfrac{25}{0{,}67232} \approx 37{,}18467$.

Т. е. заемщик в конце каждого года будет выплачивать сумму 37,18467 тыс. рублей в течение пяти лет.

В зависимости от исходных данных для решения каждой задачи формируется соответствующий набор моделей для определения количественных значений показателей контракта.

Модели инфляции

Инфляция характеризуется обесцениванием национальной денежной единицы, снижением ее покупательной способности и общим повышением цен в стране. В таком случае инвестор может потерять часть дохода, а заемщик, соответственно, может выиграть за счет погашения задолженности деньгами сниженной покупательной способности. На этом основании необходимо установить количественные соотношения по определению влияния инфляции на показатели финансовой операции. Следует заметить, что если наблюдается общее снижение цен, то происходит дефляция.

Все показатели финансовой операции можно разделить на две группы: номинальные, рассчитанные в текущих ценах, и реальные, учитывающие влияние инфляции, рассчитанные в сопоставимых ценах базисного периода.

Для количественной оценки упомянутых процессов формируют определенный набор товаров и услуг, называемый потребительской корзиной, и фиксируют изменение ее стоимости в различные моменты времени. Состав потребительской корзины математически можно представить в виде совокупности различных товаров:

$$(x_1, x_2, x_3, ..., x_i, ..., x_n),$$

где x_i – количество i-го вида товара или услуги в корзине;

n – количество товаров и услуг потребительской корзины.

В базисном периоде t_0 цены товаров потребительской корзины обозначим $P_1^0, P_2^0, P_3^0, ..., P_n^0$, а в аналогичном периоде t_j, соответственно, $P_1^j, P_2^j, P_3^j, ..., P_n^j$.

Тогда стоимость потребительской корзины в базисном периоде t_0 составит:

$$S_0 = P_1^0 x_1 + P_2^0 x_2 + P_i^0 x_i + ... + P_n^0 x_n = \sum_{i=1}^{n} P_i^0 x_i,$$

а в анализируемом периоде t_i

$$S_j = P_1^j x_1 + P_2^j x_2 + P_i^j x_i + ... + P_n^j x_n = \sum_{i=1}^{n} P_i^j x_i.$$

На этом основании полагают, что изменение (рост или падение) потребительских цен определяется безразмерным показателем, называемым индексом инфляции, который показывает, во сколько раз изменились цены (темп роста цен, если выразить в процентах):

$$I_И = \frac{S_j}{S_0},$$

а относительная величина – темп прироста цен – это уровень инфляции:

$$\alpha = \frac{S_j - S_0}{S_0} = \frac{\Delta S}{S_0} = (I_И - 1),$$

откуда следует, что индекс инфляции равен: $I_И = 1 + \alpha$.

Уровень инфляции в процентах определяется так:

$$\alpha\% = \left(\frac{S_j - S_0}{S_0}\right) \cdot 100\ \%.$$

Индекс инфляции показывает, во сколько раз выросли цены, а уровень инфляции – на сколько процентов выросли цены за рассматриваемый период. При проведении исследования стоимость потребительской корзины фиксируется через, например, равные промежутки времени:

$$t_0, t_1, t_2, t_3, \ldots, t_j, \ldots, t_N,$$

что можно записать таким образом:

$$S_0, S_1, S_2, S_3, \ldots, S_j, \ldots, S_N.$$

Аналогично для темпов инфляции на этих интервалах:

$$\alpha_{0,1}, \alpha_{1,2}, \alpha_{2,3}, \ldots, \alpha_{l-1,l}, \ldots, \alpha_{N-1,N}.$$

Тогда можно записать следующие уравнения связи между членами ряда:

$$S_1 = S_0(1+\alpha_{0,1}); \; S_2 = S_1(1+\alpha_{1,2}); \; S_3 = S_2(1+\alpha_{2,3});$$

отсюда после подстановок получим:

$$S_3 = S_0(1+\alpha_{0,1})\,(1+\alpha_{1,2})\,(1+\alpha_{2,3}).$$

На этом основании нетрудно получить выражение для определения стоимости потребительской корзины в общем виде:

$$S_N = S_0 \cdot \prod_{i=1}^{N}(1+\alpha_{i-1,i}),$$

тогда индекс инфляции за весь период будет равен:

$$I_И = \frac{S_n}{S_0} = \prod_{i=1}^{N}(1+\alpha_{i-1,i}).$$

Кроме того, как было показано выше, индекс инфляции связан с уровнем инфляции выражением $I_И = 1+\alpha$, отсюда можно определить уровень инфляции за весь период:

$$\alpha = I_И - 1.$$

Следует заметить, что при равенстве значений уровней инфляции на всех интервалах

$\alpha_{0,1} = \alpha_{1,2} = \alpha_{2,3} = \alpha_{3,4} = \ldots = \alpha_{i-1,i} = \alpha$ индекс инфляции определяется по формуле:

$$I_и = (1+\alpha)^N.$$

Рассмотрим различные варианты начисления процентов с учетом инфляции.

Для процентов, начисленных без учета капитализации, обозначим i_α ставку процентов, учитывающую инфляцию, тогда наращенную сумму можно определить по формуле:

$$S_\alpha = P(1+ni_a).$$

Затем, воспользовавшись уравнением связи S_α с S с помощью индекса инфляции:

$$S_\alpha = S \cdot I_и = P\,(1+ni)\,I_и,$$

запишем равенство

$$P(1+ni_\alpha) = P(1+ni)I_и,$$

откуда и получим модель определения номинальной ставки процентов:

$$i_\alpha = \frac{(1+ni)I_и - 1}{n}.$$

Если рассматривать в пределах одного периода, то номинальная процентная ставка выражается через реальную ставку следующим соотношением:

$$i_\alpha = i + \alpha + i \cdot \alpha.$$

Реальная доходность операции при заданных i_a и $I_и$ определяется по формуле:

$$i = \frac{ni_\alpha + 1 - I_и}{nI_и}.$$

При достаточно большом темпе инфляции реальная ставка процентов может стать даже отрицательной.

Для процентов, начисленных с учетом ежегодной капитализации, аналогично запишем два выражения:

$$S_\alpha = P(1+i_{c\alpha})^n, \quad S_\alpha = P(1+i_c)^n \cdot I_и,$$

из которых определим номинальную ставку процента:

$$i_{c\alpha} = (1+i_c)\cdot\sqrt[n]{I_и} - 1.$$

и реальную доходность операции:

$$i_c = \frac{1+i_{c\alpha}}{\sqrt[n]{I_и}} - 1 = \frac{1+i_{c\alpha}}{1+\alpha} - 1.$$

По этим формулам можно сравнивать $i_{c\alpha}$ и α (больше, равно или меньше), проводить экономический анализ эффективности вложений и установить, поглощается ли доход инфляцией или происходит реальный прирост вложенного капитала, а не убыток.

При начислении процентов несколько раз в году запишем аналогичные модели:

$$S = P(1+\frac{j_\alpha}{m})^{mn}, \qquad S = P(1+\frac{j}{m})^{mn}\cdot I_и.$$

откуда получим выражение для вычисления номинальной процентной ставки, учитывающей инфляцию, при начислении процентов несколько раз в год:

$$j_\alpha = m\left[\left(1+\frac{j}{m}\right)\sqrt[mn]{I_и} - 1\right],$$

а также уравнение для определения реальной доходности:

$$j = m\left[\frac{1+\dfrac{j_\alpha}{m}}{\sqrt[mn]{I_и}} - 1\right].$$

Приведенные модели позволяют проводить взаиморасчеты с клиентами по экономическим показателям в контрактах с учетом инфляции.

Пример решения задачи. Найти реальную стоимость накоплений с учетом инфляции, если инвестор имеет 5 млн рублей и вкладывает сроком на два года с учетом ежеквартальной капитализации. Доход исчисляется из расчета 48 % годовых. Индекс инфляции за год составил 120 %.

$$S = \frac{P \cdot (1+i)^n}{(1+\alpha)^n} = \frac{5 \cdot (1+0{,}12)^8}{(1{,}2)^2} = \frac{5 \cdot 2{,}476}{1{,}44} \approx 8{,}597.$$

То есть реальная стоимость накоплений составит 8,597 млн рублей.

Изменение условий контракта

В практике нередко возникают случаи, когда необходимо заменить одно финансовое обязательство другим, объединить несколько обязательств в одно и т. п. В таких ситуациях неизбежно возникает вопрос о принципе изменения условий контрактов.

Общим принципом такого изменения является безубыточность, другими словами, финансовые отношения сторон после изменения условий должны сохраниться на прежнем уровне, т. е. новые финансовые обязательства должны быть эквивалентны старым.

Рассмотрим две постановки задачи по изменению условий контрактов: объединение (консолидирование) платежей и сбалансированное изменение сроков платежей.

1. При объединении платежей $S_1, ..., S_k$ со сроками выплат $t_1, ..., t_k$, соответственно, в один платеж S_0. При этом могут возникнуть две задачи: определить величину объединенного платежа S_0, если он должен быть сделан в заданный момент времени t_0, либо определить срок t_0 платежа S_0. Если срок больше, чем сроки объединяемых платежей t_j, то размер нового платежа равен сумме консолидируемых платежей, наращенных по принятой ставке на момент выплаты S_0. Таким образом, сумма консолидированного платежа составит:

$$S_0 = \sum_j S_j (1+i)^n, \text{ где } \quad t_0 < n < t_j, \quad n_j = t_0 - t_j. \tag{1.8}$$

В общем случае искомую величину S_0 находим как сумму наращенных или дисконтированных платежей S_j:

$$S_0 = \sum_j S_j (1+i)^n + \sum_k S_k (1+i)^{-n},$$

где $\quad t_j < t_0, \quad t_k > t_0, \quad n_k = t_k - t_0, \quad n_j = t_0 - t_j.$

<u>*Пример решения задач.*</u>

Два платежа S_1=100 тыс. руб. и S_2=50 тыс. руб. со сроками 150 и 180 дней (отсчитываемыми от одной базы) заменяются одним со сроком 200 дней. Найти сумму объединенного платежа, если стороны согласились на замену при использовании сложной ставки, равной 6 % годовых.

Решение. Согласно формуле (1.8) имеем:

$$S_{.0} = 100 \cdot (1+0{,}06)^{50/365} + 50 \cdot (1+0{,}06)^{20/365} = 150{,}82 \text{ тыс. руб.}$$

2. Более общий случай изменения условий контрактов: расчет искомой суммы S_0 осуществляется на основе уравнения эквивалентности, в котором сумма приведенных платежей по старым условиям контракта равна сумме приведенных на тот же момент времени платежей по новому (измененному) со-

глашению. Если приведение осуществляется на начальный момент времени, то уравнение эквивалентности в общем виде записывается как

$$\sum_q S_q V^{t_q} = \sum_k S_k V^{t_k} ,$$

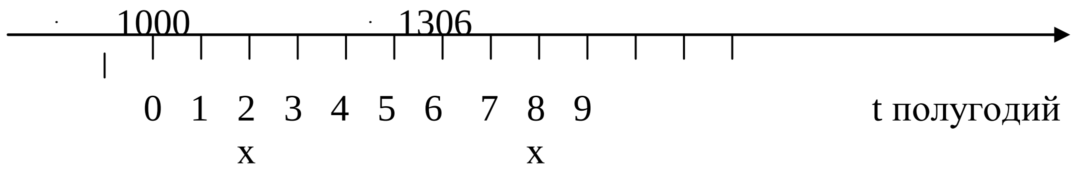
где S_k – ряд заменяемых платежей со сроками t_k,

S_q – платежи со сроками t_q, предусматриваемые новыми условиями.

Пример решения задачи.

Согласно контракту господин А обязан уплатить господину Б сумму 1000 руб. сегодня и 1306 руб. через 3 года. Господин А хочет изменить контракт, вернув долг двумя равными платежами, сделав первый через год и второй через 4 года, считая от сегодняшнего дня. Какой величины должен быть каждый из платежей, если деньги приносят кредитору 6 % годовых при начислении два раза в год?

Решение. Изобразим условие задачи на оси времени, помещая над осью платежи по первоначальному контракту, а под осью – по новому контракту. Буквой x обозначена искомая величина платежей.



Так как оба контракта должны быть равноценными для кредитора Б, то приведенные к моменту 0 (как и к любому другому моменту) ценности сумм, стоящих над осью, и сумм, стоящих под осью, должны быть равны, т. е. находим значение x из уравнения

$$1000 + 1306 \cdot \left(1 + \frac{0{,}06}{2}\right)^{-3\cdot 2} = x \cdot \left(1 + \frac{0{,}06}{2}\right)^{-1\cdot 2} + x \cdot \left(1 + \frac{0{,}06}{2}\right)^{-4\cdot 2} ;$$

$$x \cdot \left(1{,}03^{-2} + 1{,}03^{-8}\right) = 1000 + 1306 \cdot 1{,}03^{-6};$$

x = 1208,87.

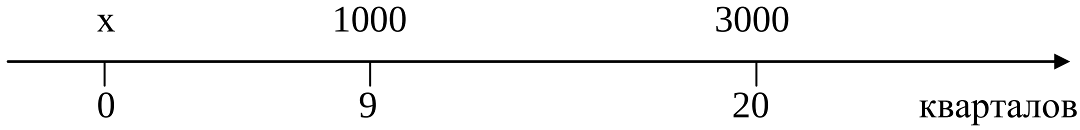
Итак, господин А должен сделать два платежа по 1208,87 руб.

Также в практике финансовых операций распространена сделка, которая называется продажей контракта. Она заключается в следующем. Некоторый субъект (или организация) имеет на руках контракт, по которому он должен получить с другого субъекта определенные суммы денег в определенные сроки. Владелец контракта желает получить деньги немедленно и для этого продает этот контракт банку или другому лицу, который получает деньги по этому контракту в будущем. Сколько следует заплатить за контракт? Очевидно, его стоимость в момент покупки.

Пример решения задачи.

Господин Иванов купил у господина Петрова некоторую вещь, заключив контракт, в соответствии с которым обязуется заплатить 1000 руб. через 27 месяцев и еще 3000 руб. – через 5 лет. Господин Петров, нуждаясь в деньгах, хочет продать этот контракт финансовой организации, которая согласна купить его при условии начисления на свои деньги процентов по ставке 8 % годовых (начисление ежеквартальное). Сколько должна заплатить компания господину Петрову за этот контракт?

Решение. Изобразим условия задачи на оси времени, на которой каждый процентный период соответствует кварталу (3 месяца). В 27 месяцах содержится 9 процентных периодов, а в 5 годах – 20 процентных периодов.



Организация должна заплатить за этот контракт его стоимость в момент 0, эта стоимость обозначена буквой x. Очевидно, что

$$x = 1000 \cdot \left(1 + \frac{0{,}08}{4}\right)^{-2{,}25\cdot 4} + 3000 \cdot \left(1 + \frac{0{,}08}{4}\right)^{-5\cdot 4},$$

$x = 1000*1{,}02^{-9} + 3000*1{,}02^{-20} = 2855{,}8$ руб.

При покупке некоторого товара продавец может заключить с продавцом контракт, включающий различные условия авансовой оплаты, получения кредита и сроков поставки товара. Чтобы выбрать наиболее выгодный для себя контракт, покупатель должен сравнить современные ценности возможных контрактов и найти контракт с наименьшей современной ценностью. Чтобы определить современную ценность тех или иных платежей, необходимо принять какую-либо ставку сравнения, то есть ставку сложных процентов i_c, по которой будет производиться дисконтирование этих платежей. В теории корпоративных финансов рассматриваются различные подходы к выбору этой ставки – это может быть и уровень ссудного процента, и уровень доходности по государственным облигациям или кредитным обязательствам и т.д.

При покупке товара покупатель делает платежи двух видов.

Во-первых, это авансовые платежи, то есть суммы, которые он выплачивает за купленный товар в обусловленные контрактом моменты t (считая от момента заключения контракта); обозначим эти платежи P_t:

$$\sum_t P_t (1+t)^{-t}.$$

Во-вторых, это платежи по погашению кредита, то есть по погашению разности между ценой товара C и авансовыми платежами:

$$C - \sum_t P_t.$$

Современная стоимость этих платежей различна при разных условиях погашения кредита. Рассмотрим два наиболее часто встречающихся случая.

а) Кредит погашения разовым платежом в конце срока; за кредит по контракту продавец получает g% годовых. Современная ценность всех платежей по контракту на момент его заключения равна:

$$A = \sum_t P_t (1+i)^{-t} + \left(C - \sum_t P_t\right)(1+g)^N (1+i)^{-(T+N)}, \qquad (1.9)$$

где T – срок поставки товара; N – срок кредита, который обычно отсчитывается от момента окончания поставки товара.

б) Кредит погашается равными срочными платежами. Современная стоимость всех платежей по контракту на момент его заключения равна:

$$A = \sum_t P_t (1+i)^{-t} + \left(C - \sum_t P_t\right) \frac{\left(1 - (1+i)^{-N}\right) \cdot g}{i \cdot \left(1 - (1+g)^{-N}\right)} (1+i)^{-T}. \qquad (1.10)$$

Пример решения задачи. Сравним два контракта.

1-й контракт: товар стоит 20 млн руб.; делается три авансовых платежа по 3 млн руб. каждый: первый – в момент заключения контракта, второй – через год, третий – еще через год. Поставка товара производится по окончанию авансовых платежей. Кредит дается на 6 лет, считая с момента поставки товара под 5 % годовых, и погашается разовым платежом в конце срока кредита.

2-й контракт: товар стоит 21 млн руб.; в момент заключения контракта делается один авансовый платеж, равный 5 млн руб. Поставка производится в момент заключения контракта. Кредит выдается на 10 лет под 5 % годовых с погашением равными ежегодными срочными уплатами.

Сравнение контрактов произвести при ставке сравнения i = 10 %.

Решение. Найдем современную ценность каждого из контрактов. Современную стоимость первого контракта вычисляем по формуле (4.9) при C = 20 млн. руб., $t_1 = 0, t_2 = 1, t_3 = 2, T = 2, N = 6, g = 5\ \%, P_1 = P_2 = P_3 = 3$ млн руб.;

$$A_1 = 3(1+i)^0 + 3(1+i)^{-1} + 3(1+i)^{-2} + (20-9)(1+0{,}05)^6(1+i)^{-(2+6)} =$$

$$= 3 + 3\cdot 1{,}1^{-1} + 3\cdot 1{,}1^{-2} + 11\cdot 1{,}05^6 \cdot 1{,}1^{-8} = 16{,}083 \text{ млн. руб.}$$

Современную стоимость второго контракта вычисляем по формуле (1.10) при $C = 21$ млн руб., $t_1 = 0, P_1 = 5, T = 0, N = 10, g = 5\ \%$:

$$A_2 = 5(1+i)^0 + (21-5)\frac{\left(1-(1+i)^{-10}\right)\cdot 0{,}05}{0{,}1\cdot\left(1-(1+0{,}05)^{-10}\right)}(1+i)^0 = 17{,}732 \text{ млн руб.}$$

Таким образом, второй контракт менее выгоден покупателю, чем первый. Однако покупатель может его предпочесть, так как поставка товара по нему производится немедленно, а по первому контракту – с отсрочкой на два года.

Модели операций с облигациями

Облигация представляет собой долговую ценную бумагу, которая является свидетельством о предоставлении займа и дает ее владельцу право на получение определенного дохода (процентного или дисконта), но не дает право голоса. По истечении срока заем подлежит возврату держателю ценной бумаги. Облигации могут быть дисконтные и процентные. По дисконтным облигациям доход представляет собой скидку с номинала. Цель операции с облигациями – использование одного из вариантов финансовых вложений для получения дохода и тем самым обеспечение защиты от обесценивания или роста капитала в условиях инфляции.

Источниками дохода от инвестирования средств в облигации являются выплачиваемые по ним проценты (купонные облигации), а также разница между ценой приобретения и их продажи или их номинальной стоимостью, по которой они выкупаются. В процессе моделирования операций с облигациями применяют различные показатели.

При расчете доходности покупки облигаций используют понятие курса:

$$P_k = \frac{P}{N}\cdot 100\ \%,$$

где P – цена облигации;

N – номинальная стоимость облигации;

$P_к$ – курс облигации.

Цена облигации при заданном курсе определяется по формуле

$$P = \frac{P_k \cdot N}{100\ \%}.$$

Если по облигациям выплачиваются проценты (ставка i годовых) , то облигации называются процентными, а доход по каждой выплате определяется от ее номинальной стоимости:

$$I = i \cdot N.$$

Если проценты по облигациям не выплачиваются, то источником дохода будет являться разность между ценой выкупа (номиналом) и ценой покупки, которая называется дисконтом, а такие облигации – дисконтными, например государственные краткосрочные обязательства (ГКО). Доход от таких облигаций составляет:

$$Д = N - P = N - \frac{P_k \cdot N}{100\ \%} = (1 - 0{,}01 \cdot P_k) \cdot N.$$

Большое значение при работе с ценными бумагами имеет такой показатель, как доходность. Доходность – это относительный показатель, который говорит о том, какой % прибыли приносит рубль инвестированных средств за год.

➢ **Важно помнить.** Доходность всегда выражается в процентах и всегда исчисляется за год.

Доходность облигации к погашению можно определить по эквивалентной ставке простых процентов:

$$i = \frac{Д}{P \cdot n} 100\ \% = \frac{N - P}{P \cdot n} 100\ \% = \frac{100\ \% - P_k}{P_k \cdot n} = \left(\frac{100\ \% - P_k}{P_k} \right) \cdot \frac{T}{t},$$

где T – число дней в году;

t – число дней до погашения облигации;

n – число периодов (если облигация на несколько лет);

Доход от покупки долгосрочных облигаций с выплатой процентов будет состоять из суммы полученных процентов и разницы между ценой их погашения (номиналом) и ценой покупки.

Если проценты по облигациям выплачиваются в конце срока, например, методом сложных процентов i_c, то сумма процентных денег при погашении облигации через n лет определяется выражением:

$$I = N(1 + i_c)^n - N = N\left[(1 + i_c)^n - 1\right].$$

Общий доход можно определить по формуле:

$$Д = I + N - P = N(1 + i_c)^n - P = N \cdot \left[(1 + i_c)^n - P_k / 100\right].$$

➢ **Важно помнить.** Не путайте понятие дохода и доходности. Доход – это прибыль, выраженная в денежных единицах, разница между полученным и инвестируемым капиталом. Доходность – это доход, выраженный в процентах и скорректированный на время.

Пример решения задачи.

Средняя аукционная цена облигации 92 %, срок жизни 1 месяц. Первый покупатель приобрел за 93 %, продержал 10 дней и продал за 97%, второй держал до погашения и погасил. Определить доходность к аукциону, к погашению, доходность, которую обеспечил себе каждый покупатель.

Доходность к аукциону: $i = \frac{100\,\% - 92\,\%}{92\,\%} \cdot \frac{365}{30} \cdot 100\,\% \approx 105{,}8\,\%$.

Доходность к погашению: $i = \frac{100\,\% - 93\,\%}{93\,\%} \cdot \frac{365}{30} \cdot 100\,\% \approx 91{,}6\,\%$.

Доходность первого покупателя:

$$i = \frac{97\,\% - 93\,\%}{93\,\%} \cdot \frac{365}{10} \cdot 100\,\% \approx 157\,\%.$$

Доходность второго покупателя:

$$i = \frac{100\% - 97\%}{97\%} \cdot \frac{365}{20} \cdot 100\% \approx 56.44\%.$$

При определении общего дохода следует учитывать возможность реинвестирования, если проценты выплачиваются периодически.

Для решения задач по ценным бумагам необходимо знать следующие понятия:

Ажио – превышение курса акции над ее номинальной стоимостью.

Дисконт – разница между ценой размещения и номиналом.

Модели операций с акциями

Акция представляет собой долевую ценную бумагу, свидетельствующую о внесении определенной части в капитал акционерного общества, которая дает право на получение дохода и формальное участие в управлении предприятием. В зависимости от порядка начисления и выплаты дивидендов акции делят на привилегированные и обыкновенные.

Привилегированные акции – это акции, держатель которых имеет право получения фиксированного дивиденда, но не имеет права голоса.

Обыкновенные акции – это акции, приносящие доход в зависимости от полученной предприятием прибыли и дающие право голоса на собраниях акционеров.

Дивиденды (доход, выплачиваемый на одну акцию) по привилегированным акциям объявляются в фиксированных процентах ее номинальной стоимости и определяются по формуле:

$$D_1 = f \cdot N,$$

где f – годовая ставка дивиденда;

N – номинальная стоимость акции;

D_1 – дивиденды по одной привилегированной акции.

Доход на одну обыкновенную акцию

$$Д_0 = \frac{ЧП - D_{пр}}{M_0},$$

где M_0 – количество обыкновенных акций;

$D_{пр}$ – дивиденд по всем привилегированным акциям: $D_{пр}=M_{пр}*D_1$;

$M_{пр}$ – количество привилегированных акций;

ЧП – чистая прибыль или ее часть по решению собрания акционеров.

Обычно на выплату дивидендов по обыкновенным акциям тратится не весь доход, а только его часть (по решению собрания акционеров), поэтому величина выплачиваемого дивиденда определяется дивидендным выходом.

Доходность по акциям определяется доходом от выплачиваемых дивидендов, а также разницей в цене покупки и продажи, что и определяет эффективность инвестиций.

$$Э = \frac{P_1 - P_a + D}{P_a} 100\ \%,$$

где P_a – цена покупки;

P_1 – цена продажи;

D – дивиденды за время владения акцией.

При анализе операций с акциями необходимо проводить расчеты по нескольким показателям, то есть различать разные виды доходности.

К таким показателям относятся: доходность текущая, без налогообложения, которая определяется по формуле:

$$r = \frac{D}{P_a} \cdot 100\ \%, \qquad (1.11)$$

где P_a – курсовая стоимость акции, которая рассчитывается в сравнении с банковской депозитной ставкой (i %):

$$P_a = \frac{D_1}{i\ \%} \cdot 100\ \%.$$

Доходность конечная, которая определяется суммой дивидендов и дополнительным доходом от перепродажи (в пересчете на один год):

$$r = \frac{D + \frac{(P_1 - P_a)}{n}}{P_a} \cdot 100\ \%. \qquad (1.12)$$

Пример решения задачи.

Инвестор приобрел акцию за 23 тыс. рублей. Номинал акции 20 тыс. рублей. По акции выплачивается фиксированный дивиденд 150 % годовых. Через 3 года инвестор продал ее за 20 тыс. рублей. Определить доходность на текущий момент для инвестора, доходность на текущий момент для покупателя и полную доходность.

Решение.

Используя приведенные выше формулы (1.11) и (1.12), получим:

доходность на текущий момент для инвестора:

$$r = \frac{20 \cdot 1{,}5}{23} 100\ \% \approx 130{,}4\ \%.$$

доходность на текущий момент для покупателя:

$$r = \frac{20 \cdot 1{,}5}{20} 100\,\% = 150\,\%.$$

Полная доходность в пересчете на один год:

$$r = \frac{20 \cdot 1{,}5 + \frac{(20 - 23)}{3}}{23} 100\,\% \approx 126{,}1\,\%.$$

Пользуясь приведенными моделями, можно сравнивать выгодность финансовых операций с акциями и, следовательно, решать задачу выбора инвестиционного проекта.

Тема 2. Оптимизация в условиях определенности

Линейное программирование

На практике постоянно встречаются такие ситуации, когда достичь какого-то результата можно не одним, а многими различными способами. В подобной ситуации может оказаться и одно предприятие и целая отрасль, если необходимо определить, как использовать имеющиеся в их распоряжении ресурсы, чтобы добиться максимального выхода продукции. Естественно, при большом количестве решений выбирается наилучшее. Математически это обычно сводится к нахождению наибольшего и наименьшего значения некоторой функции, то есть к задаче: найти *max (min) f(x),* при условии, что переменная x пробегает данное множество X.

Записывают так: $f(x) \to max\ (min),\ x \in X.$

Определенная таким образом задача называется задачей оптимизации. Множество X называется допустимым множеством, а *f(x)* – целевой функцией.

Очень многое зависит от того, в каком виде задается множество X.

Чаще всего это система неравенств: $$\begin{cases} g_1(x_1, x_2, \ldots, x_n) \ge 0 \\ g_2(x_1, x_2, \ldots, x_n) \ge 0 \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ g_n(x_1, x_2, \ldots, x_n) \ge 0 \end{cases},$$

где $(x_1, x_2, \ldots, x_n)$ – координаты точки x в R_n (n-мерном пространстве), а g_i – некоторые функции. Таким образом, надо найти экстремум функции *f(x)* при заданной системе ограничений. Но понятно, что следует найти не только само значение *max (min),* но и точку (или точки, если их несколько), в которых это значение достигается. Такие точки называются оптимальными решениями. Задачи подобного рода называются задачами математического программирования.

В большинстве случаев в число ограничений входят условия неотрицательности переменных: $x_1 \ge 0,\ x_2 \ge 0, \ldots,\ x_n \ge 0$, которые вытекают из реального экономического смысла этих чисел и называются тривиальными ограничениями.

В зависимости от характера функций $f, g_1, g_2, \ldots g_n$ различают разные виды математического программирования. Наиболее простой и часто встречающийся случай, который мы и будем изучать, – когда эти функции являются линейными, то есть каждая из них имеет вид $a_1x_1 + a_2x_2 + ... + a_nx_n + b$. Тогда говорят о задаче линейного программирования. Линейное программирование оформилось как отдельный раздел прикладной математики в 40–50 гг. XX века, когда выяснилось, что целый ряд задач из сферы планирования и управления может быть сформулирован в виде задачи линейного программирования. Подсчитано, что в настоящее время ≈ 80–85 % всех решаемых на практике задач оптимизации относится к задачам линейного программирования.

Формулировка задачи и ее геометрическое истолкование

Чтобы показать, как решить задачу графическим способом, необходимо ввести некоторые понятия. Любое $x \in X$ называется допустимым решением (планом). Допустимое решение, дающее *max (min) f(x),* называется оптимальным решением (планом). Неравенства называются ограничениями. Решение системы всех неравенств называется областью допустимых решений. Задачи на *max* и *min* идентичны, так как они сводятся друг к другу путем умножения целевой функции на (–1), то есть *max f(x)= – min (–f(x)).*

Таким образом, необходимо найти оптимальное значение функции (минимум или максимум) в области допустимых решений.

Наиболее часто встречаются две разновидности задач линейного программирования:

1. Каноническая (основная). Система ограничений, помимо тривиальных ограничений, включает в себя только уравнения.
2. Стандартная (симметричная). Система ограничений состоит только из неравенств.

В зависимости от способа решения задачи, необходимо, чтобы она была представлена в той или другой форме.

Эти две разновидности легко сводятся одна к другой.

Пример решения задачи:

1. Стандартный вид

$$f(x) = -2x_1 + x_2 + 1 \to \min$$

$$\begin{cases} x_1 - 2x_2 + 1 \ge 0 \\ x_1 - x_2 \ge 0 \\ x_1 \ge 0, \quad x_2 \ge 0 \end{cases}$$

Канонический вид

$$\begin{cases} x_1 - 2x_2 - x_3 + 1 = 0 \\ x_1 - x_2 - x_4 = 0 \\ x_1, x_2, x_3, x_4 \ge 0 \end{cases}$$

2. Канонический вид

$$f(x) = -x_1 + x_2 + x_3 - 2x_4 \to \min$$

$$\begin{cases} -2x_1 + 3x_2 + x_3 + x_4 = 1 \\ -x_1 + 3x_2 + 2x_3 - x_4 = 2 \\ x_1 \ge 0,\ x_2 \ge 0,\ x_3 \ge 0,\ x_4 \ge 0 \end{cases}$$

Решая систему линейных уравнений, получим: $\begin{cases} x_3 = x_1 - 2x_2 + 1 \\ x_4 = x_1 - x_2 \end{cases}$. Подставляя найденные значения в целевую функцию и исключая найденные переменные из системы ограничений, получим стандартный вид задачи линейного программирования:

$$F(x) = 1 - 2x_1 + 5x_2 \to \min$$

$$\begin{cases} x_1 - 2x_2 + 1 \geq 0 \\ x_1 - x_2 \geq 0 \\ x_1 \geq 0, \quad x_2 \geq 0 \end{cases}$$

Множество решений (планов) основной задачи линейного программирования образует выпуклый многогранник. Каждая вершина этого многогранника определяет опорный план. В одной из вершин (то есть для одного из опорных планов) значение целевой функции является оптимальным, то есть *max (min)*. Если значение *max (min)* функция принимает более чем в одной вершине, то это же значение она принимает в любой точке, являющейся выпуклой линейной комбинацией данных вершин. Иными словами, если оптимальное значений функции достигается в двух вершинах многоугольника, то оно будет одинаковым не только в этих вершинах, но и в любой точке отрезка, соединяющего эти вершины.

Вершину многогранника решений найти сравнительно просто, если задача, записанная в стандартной форме, содержит не более двух переменных, так как в этом случае многогранник можно изобразить на плоскости, вершина находится как точка пересечения двух прямых, которые задают ограничения в системе.

Пример решения задачи. $f(x) = c_1 x_1 + c_2 x_2 \to \max$

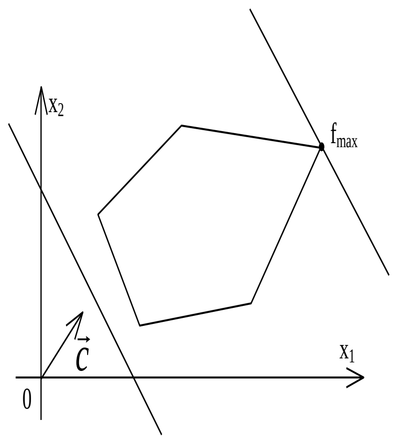


Рис. 1.1. Задача имеет единственное решение

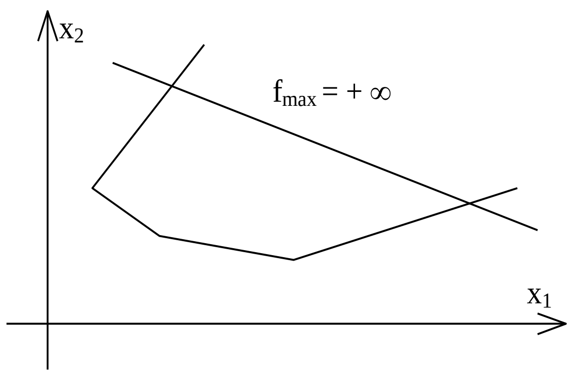


Рис. 1.2. Целевая функция не ограничена

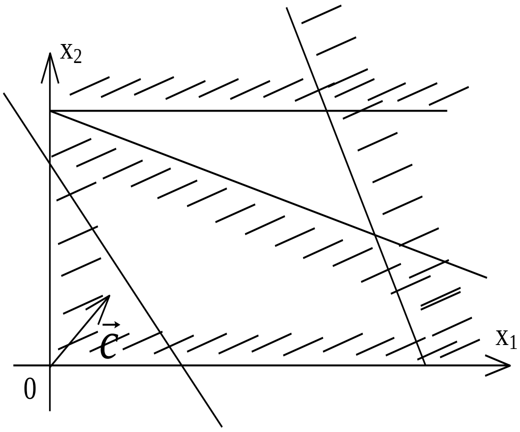


Рис. 1.3. Система ограничений несовместна

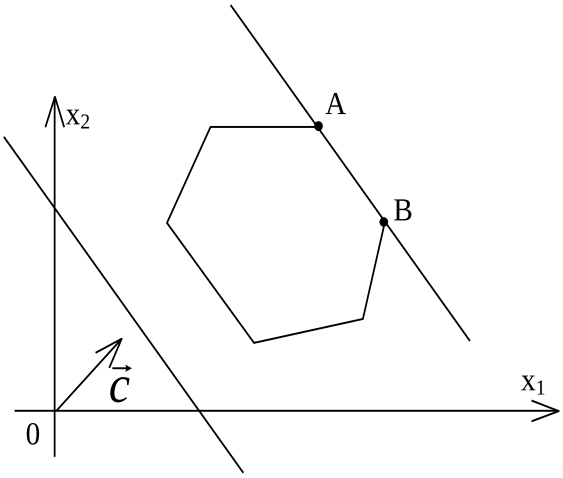


Рис. 1.4. Максимальное значение функция принимает на отрезке АВ

<u>*Пример решения задачи*</u>**.**

Найти *max* и *min* функции $f(x) = x_1 + x_2$ при заданной системе ограничений:

$$\begin{cases} 2x_1 + 4x_2 \le 16 \\ -4x_1 + 2x_2 \le 8 \\ x_1 + 3x_2 \ge 9 \\ x_1 \ge 0, \quad x_2 \ge 0 \end{cases}$$

Построим многоугольник решений:

$$\begin{cases} x_2 \le 4 - \frac{1}{2}x_1 \quad (1) \\ x_2 \le 4 + 2x_1 \quad (2) \\ x_2 \ge 3 - \frac{1}{3}x_1 \quad (3) \\ x_1 \ge 0, x_2 \ge 0 \end{cases}$$

$$f(6{,}1) = 6 + 1 = 7$$

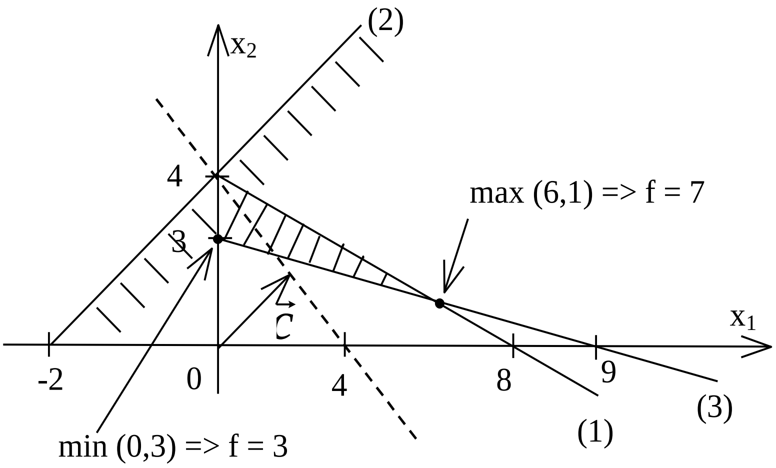


Рис. 1.5. Решение задачи графическим методом

Экономическая интерпретация задач линейного программирования

1. Задача об использовании ресурсов (планировании производства).

Необходимо составить такой план производства продукции, при котором прибыль от ее реализации будет максимальной.

Введем следующие обозначения:

x_j – число единиц продукции P_j , запланированной к производству,

b_i – запас ресурса S_i ,

a_{ij} – число единиц ресурса S_i , затрачиваемое на единицу продукции P_j (эти числа называют технологическими коэффициентами),

c_j – прибыль от реализации продукции P_j .

Тогда модель задачи будет иметь вид:

$$f(x) = c_1x_1 + c_2x_2 + \ldots\ldots + c_nx_n \to \max$$

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots\ldots\ldots + a_{1n}x_n \le b_1 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots\ldots + a_{mn}x_n \le b_m \\ x_i \ge 0, \quad i = \overline{1,n} \end{cases}$$

Рассмотрим задачу с конкретными данными.

Прибыль от реализации единицы продукции P_1 составляет 2 рубля, а от единицы продукции P_2 – 3 рубля. Данные о запасах ресурсов и количестве ресурсов, необходимых для изготовления единицы продукции, сведены в следующую таблицу.

Данные задачи об использовании ресурсов

Вид ресурса	Запас ресурса	Число единиц ресурса, затрачиваемых на изготовление единицы продукции	
		P_1	P_2
S_1	18	1	3
S_2	16	2	1
S_3	5	---	1
S_4	21	3	---

Экономико-математическая модель данной задачи будет иметь вид:

$$f(x) = 2x_1 + 3x_2 \to \max \qquad \begin{cases} x_1 + 3x_2 \le 18 \\ 2x_1 + x_2 \le 16 \\ x_2 \le 5 \\ 3x_1 \le 21 \\ x_1 \ge 0, \quad x_2 \ge 0 \end{cases}$$

2. Задача составления рациона (о диете, о смесях), или технологическая задача.

Необходимо составить дневной рацион, имеющий минимальную стоимость, в котором содержание каждого вида питательных веществ было бы не менее установленного предела.

Введем обозначения:

x_j – число единиц корма j-го вида;

b_i – необходимый минимум содержания в рационе питательного вещества S_i ;

a_{ij} – число единиц питательного вещества S_i в единице корма j-го вида;

c_j – стоимость единицы корма j-го вида.

Тогда экономико-математическая модель задачи запишется следующим образом:

$$f(x) = c_1 x_1 + c_2 x_2 + \ldots\ldots + c_n x_n \to \min$$

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots\ldots\ldots + a_{1n}x_n \ge b_1 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots\ldots + a_{mn}x_n \ge b_n \\ x_i \ge 0, \quad i = \overline{1, n} \end{cases}$$

Рассмотрим задачу с конкретными данными.

Стоимость 1 кг корма вида I – 4 рубля, а вида II – 6 рублей. Используя данные таблицы, необходимо составить такой рацион питания, чтобы стоимость его была минимальной и содержание каждого вида питательных веществ было не менее установленного предела.

Данные задачи о рационе питания

Питательное вещество	Необходимый мин. пит. веществ	Число единиц питательного вещества в 1 кг корма	
		I	II
S_1	9	3	1
S_2	8	1	2
S_3	12	1	6

Модель задачи будет иметь вид:

$$f(x) = 4x_1 + 6x_2 \to \min \qquad \begin{cases} 3x_1 + x_2 \ge 9 \\ x_1 + 2x_2 \ge 8 \\ x_1 + 6x_2 \ge 12 \\ x_1 \ge 0, \quad x_2 \ge 0 \end{cases}$$

3. Задача об использовании мощностей (задача о загрузке оборудования).

Предприятию задан план производства продукции по времени и номенклатуре: требуется за время T выпустить $n_1, n_2, \ldots\ldots n_k$ единиц продукции P_1, $P_2, \ldots\ldots P_к$. Продукция производится на станках S_1, S_2, $\ldots\ldots S_m$. Для каждого станка известны производительность a_{ij} (то есть число единиц продукции P_j , которое можно произвести на станке S_i в единицу времени) и затраты b_{ij} на изготовление продукции P_j на станке S_i в единицу времени.

Необходимо составить такой план работы станков (то есть так распределить выпуск между станками), чтобы затраты на производство всей продукции были минимальными.

Экономико-математическая модель задачи.

Обозначим x_{ij} – время, в течение которого станок S_i будет занят изготовлением продукции P_j . Так как время работы каждого станка ограничено и не превышает T, то справедливы следующие неравенства:

$$\begin{cases} x_{11} + x_{12} + ...x_{1k} \le T \\ x_{21} + x_{22} + ...x_{2k} \le T \\ .... \\ x_{m1} + x_{m12} + ...x_{mk} \le T \end{cases} \qquad \begin{cases} a_{11}x_{11} + a_{21}x_{21} + ...a_{m1}x_{m1} = n_1 \\ a_{12}x_{12} + a_{22}x_{22} + ...a_{m2}x_{m2} = n_2 \\ \\ a_{1k}x_{1k} + a_{2k}x_{2k} + ...a_{mk}x_{mk} = n_k \end{cases}$$

$$f(x) = b_{11}x_{11} + b_{12}x_{12} + ...b_{mk}x_{mk} \to \min$$

4. Технологическая задача (о раскрое материала).

Найти план раскроя, обеспечивающий максимальное число комплектов.

Для изготовления брусьев длиной 1,2 м, 3 м и 5 м в соотношении 2:1:3 на распил поступает 195 бревен длиной 6 м. Определить план распила, обеспечивающий максимальное число комплектов.

Составим модель задачи.

Определим сначала все возможные способы распила бревен, указав соответствующее число получающихся при этом брусьев и остаток.

Способы распила бревен

Способ распила	Число получающихся брусьев			Остаток
	1,2 м	3 м	5 м	
1	5	--	--	0
2	2	1	--	0,6
3	--	2	--	0
4	--	--	1	1

Через x_i обозначим число бревен распиливаемых i-м способом, $i = \overline{1,4}$, а через x – число комплектов брусьев.

Тогда экономико-математическая модель задачи будет иметь вид:

$$\begin{cases} x_1 + x_2 + x_3 + x_4 = 195 \\ 5x_1 + 2x_2 = 2x \\ x_2 + 2x_3 = x \\ x_4 = 3x \end{cases} \qquad F = x \to \max$$

Решение задачи линейного программирования методом последовательного улучшения плана (симплексным методом)

Симплексным методом можно решить любую задачу линейного программирования с произвольным числом переменных и ограничений. Для того чтобы решить задачу симплексным методом, ее надо привести к каноническому виду. Напомним, что любую задачу линейного программирования можно привести к каноническому виду:

$$f(x) = \sum_{i=1}^{n} c_i x_i \to \max$$

$$\begin{cases} \sum\limits_{j=1}^{n} a_{ij} x_j = b_i, i = \overline{1,m} \\ x_i \geq 0, i = \overline{1,n} \end{cases}$$

<u>*Пример решения задачи.*</u> Прибыль от реализации 1 т кефира составляет 2 млн руб., а от 1 т молока – 1 млн руб. Затраты рабочего времени на молоко 4 ч / т, а на кефир – 2 ч/т. Рабочий день – 8 часов. Расход сырья: 5 ед. на 1 т кефира и 1 ед. на 1 т молока. Всего имеется в запасе 5 ед. сырья. Составить такой план производства продукции, чтобы прибыль от ее реализации была максимальной.

Экономико-математическая модель задачи будет иметь вид:

$$f(x) = 2x_1 + x_2 \to \max \qquad \begin{cases} 2x_1 + 4x_2 \leq 8 \\ 5x_1 + x_2 \leq 5 \\ x_1, x_2 \geq 0 \end{cases}$$

Приведем ее к каноническому виду: $\begin{cases} 2x_1 + 4x_2 + x_3 = 8 \\ 5x_1 + x_2 + x_4 = 5 \\ x_1, x_2, x_3, x_4 \geq 0 \end{cases}$

Значения дополнительных переменных показывают разницу между запасами ресурсов каждого вида и их потреблением, то есть остатки ресурсов.

Выразим переменные:

$$x_3 = 8 - 2x_1 - 4x_2$$

$$x_4 = 5 - 5x_1 - x_2$$

Если $x_1 = 0, x_2 = 0 \quad \Rightarrow \quad x_3 = 8, x_4 = 5.$ Этот план можно взять в качестве начального решения задачи

$$X^{(0)} = \{0;0;8;5\}, \quad т.е.\ x_1^{(0)} = 0,\ x_2^{(0)} = 0,\ x_3^{(0)} = 8,\ x_4^{(0)} = 5.$$

Этот план обладает тем свойством, что число его переменных, не равных нулю, равно числу неравенств в системе ограничений задачи. Такой план называется <u>опорным</u>. Его можно получить в том случае, если часть переменных удается выразить через остальные, причем если приравнять нулю переменные, стоящие в этих выражениях справа, то переменные, стоящие слева, окажутся положительными.

Положительные переменные, стоящие слева, принято называть <u>базисными</u>, а переменные, стоящие справа и приравниваемые нулю, – <u>свободными</u>.

Для каждого опорного плана целевая функция преобразуется так, чтобы она зависела только от свободных переменных.

Для нашего опорного плана в *f(x)* базисные переменные не входят и для него *f(x)=0*.

На каждой итерации (на каждом шаге) может увеличиваться лишь одна свободная переменная. Это приводит к увеличению *f(x)*, если перед этой переменной стоит знак «+».

Так как в нашей целевой функции перед обеими свободными переменными стоит «+», то можно увеличивать любую из них и это приведет к увеличению *f(x)*. Выберем x_1, т.к. при ней больше коэффициент, следовательно скорость возрастания целевой функции выше.

При $x_2 = 0$ $\begin{cases} x_3 = 8 - 2x_1 \\ x_4 = 5 - 5x_1 \end{cases}$, следовательно, при $0 \le x_1 \le 1$ значения $x_3 \ge 0,\ x_4 \ge 0$, а при $x_1 > 1$ x_4 станет < 0, что противоречит условию. Т.е. увеличивать x_1 (и при этом уменьшать x_3 и x_4) можно не более чем на 1.

Получили новый план $X^{(1)} = \{1; 0; 6; 0\}$, т.е.

$x_1^{(1)} = 1,\ x_2^{(1)} = 0,\ x_3^{(1)} = 6,\ x_4^{(1)} = 0.$

Базисные переменные x_1 и x_3, а свободные x_2 и x_4.

Перепишем систему равенств, выразив базисные переменные и целевую функцию через свободные:

$$\begin{cases} x_1 = 1 - \frac{1}{5}x_2 - \frac{1}{5}x_4 \\ x_3 = 6 - \frac{18}{5}x_2 + \frac{2}{5}x_4 \end{cases} \qquad f(x) = 2 + \frac{3}{5}x_2 - \frac{2}{5}x_4$$

Для нового опорного плана *f(x)=2*. По *f(x)* видно, что при увеличении x_4 целевая функция будет убывать, а при увеличении x_2 – возрастать. Следовательно, x_2 вводим в базис.

При $x_4 = 0$ $\begin{cases} x_1 = 1 - \frac{1}{5}x_2 \\ x_3 = 6 - \frac{18}{5}x_2 \end{cases}$. Следовательно, при $0 \le x_2 \le \frac{5}{3}$ значения $x_1 \ge 0,\ x_3 \ge 0$, а при $x_2 > 5/3$ x_3 станет <0, что противоречит условию. Т.е. увеличивать x_2 (и при этом уменьшать x_1 и x_3) можно не более чем на 5/3.

Получили новый план $X^{(2)} = \left\{\frac{2}{3}; \frac{5}{3}; 0; 0\right\}$,

т.е. $x_1^{(2)} = \frac{2}{3}, \quad x_2^{(2)} = \frac{5}{3}, \quad x_3^{(2)} = 0, \quad x_4^{(2)} = 0.$

Базисные переменные x_1 и x_2, а свободные x_3 и x_4.

Выразим базисные переменные и линейную форму через свободные переменные

$$x_1 = \frac{2}{3} - \frac{2}{9}x_4 + \frac{1}{18}x_3 \qquad\qquad f(x) = 3 - \frac{1}{6}x_3 - \frac{1}{3}x_4$$
$$x_2 = \frac{5}{3} + \frac{1}{9}x_4 - \frac{5}{18}x_3$$

Принимая свободные переменные равными 0 ($x_3 = x_4 = 0$), получим *f(x)=3*. Все коэффициенты в целевой функции отрицательные, следовательно решение найдено (любое увеличение свободных переменных приведет к уменьшению целевой функции, уменьшать их нельзя, так как они равны 0, а отрицательными быть не могут). Базисные же переменные нельзя изменить, не меняя значения свободных. Таким образом, опорный план (2) является оптимальным и максимальное значение целевой функции равно 3.

- **Важно помнить.** Графически можно решать только задачи, которые содержат две переменные, и их необходимо привести к стандартной форме. Симплексным методом можно решать задачи с любым числом переменных, задачу необходимо привести к каноническому виду.

Практические расчеты при решении реальных задач симплексным методом выполняются в настоящее время с помощью компьютеров. В частности, в офисной программе Excel есть встроенная функция «Поиск решения» (закладка Сервис). Однако если расчеты осуществляются без компьютера, то удобно использовать так называемые симплексные таблицы, то есть преобразовывать не сами уравнения, а коэффициенты при переменных. Рассмотрим алгоритм их составления. Для определенности считаем, что решается задача на отыскание максимума.

Исходная таблица решения задачи линейного программирования симплексным методом

	β	X_1	X_2	X_3	X_4	Q
X_3	8	2	4	1	0	8/2
X_4	5	**5**	1	0	1	5/5
f(x)	0	–2	–1	0	0	–
		↑				

Значения в столбце соответствуют коэффициентам при переменных, β – столбец свободных членов.

Строки соответствуют уравнениям задачи, последняя строка – линейная форма, она записывается в виде: $f(x)=0-(-2)x_1-(-1)x_2$. Задача решается до тех пор, пока все коэффициенты в последней строке не станут больше или равны 0.

<u>1 итерация</u>.

В последней строке таблицы находим максимальное по абсолютной величине отрицательное число (= –2). Столбец, в котором стоит это число, называется <u>ведущим</u>. Соответствующая ведущему столбцу свободная переменная (x_1) станет базисной. Делим элементы столбца β на ведущий столбец – записываем значения в столбец Q. Экономический смысл элементов столбца Q: они показывают, на сколько можно увеличить переменную, вводимую в базис, чтобы остальные переменные не стали отрицательными. Из элементов столбца Q выбираем минимальное значение 8/2=4, 5/5=1 $\Rightarrow$ min=1, оно соответствует x_4

$\Rightarrow$ строка, соответствующая этой переменной называется ведущей, ведущая строка показывает, какая переменная (в данном примере переменная x_4) станет свободной. Переменная x_1 вместо x_4 становится базисной. На каждом шаге решения количество базисных переменных остается неизменным. Элемент, стоящий на пересечении ведущего столбца с ведущей строкой, называется ведущим элементом.

➢ **Важно помнить.** При заполнении столбца Q надо делить только на положительные коэффициенты ведущего столбца. Так как если бы один из коэффициентов был отрицателен или равен нулю, соответствующая базисная переменная не стала бы отрицательной ни при каком положительном значении переменной, вводимой в базис.

Заметим, что в данном примере коэффициенты ведущего столбца все положительные (стоящие в строках базисных переменных) $\Rightarrow$ обе базисные переменные убывают с увеличением x_1, Пересчитываем таблицу.

Первая итерация

базис	β	X_1	X_2	X_3	X_4	Q
X_3	6	0	**18/5**	1	–2/5	5\3
X_1	1	1	1/5	0	1/5	5
f(x)	2	0	–3/5	0	2/5	–

↑

1) Вся ведущая строка делится на ведущий элемент.

2) Преобразованная ведущая строка умножается на элемент, стоящий в ведущем столбце строки, которую хотим преобразовать, и вычитается из бывшей строки.

То же самое с линейной формой [(–2) умножить на преобразованную ведущую строку *(x_1)* и вычесть из бывшей *f(x)*].

2 итерация.

В последней строке одно число отрицательное (–3/5), следовательно это столбец ведущий. В нем оба коэффициента положительные, заполняем столбец *Q*.

Q: $\frac{6}{18/5}=\frac{5}{3}$ и $\frac{1}{1/5}=5$. $\frac{5}{3}<5$, отсюда следует, что ведущая строка первая; ведущий элемент 18/5, переменная x_2 становится базисной, а x_3 – свободной.

Пересчитываем таблицу.

Вторая итерация

базис	β	X_1	X_2	X_3	X_4
X_2	5/3	0	1	5/8	–1/9
X_1	2/3	1	0	–1/18	2/9
f(x)	3	0	0	1/6	1/3

Все элементы последней строки не отрицательные, следовательно задача решена.

$f(x)= 3 – 1/6x_3 – 1/3x_4$, $x_2=5/3$, $x_1=2/3$, $f_{max}=3$.

Алгоритм:

1) Проверка оптимальности. Если все элементы последней строки таблицы неотрицательны, то план оптимален.

Если в последней строке есть отрицательные элементы, перейти к пункту 2.

2) Выбор ведущего столбца. В последней строке таблицы найти максимальный по абсолютному значению отрицательный элемент. Столбец, в котором находится этот элемент, будет ведущим.

3) Нахождение ведущей строки. Разделить элементы столбца β на соответствующие положительные элементы ведущего столбца и найти минимальное из этих отношений. Строка, соответствующая этому минимальному отношению, будет ведущей. Элемент, расположенный на пересечении ведущего столбца и ведущей строки, – ведущий.

4) Преобразование таблицы. Разделить ведущую строку на ведущий элемент. Остальные строки таблицы преобразовываются по следующим правилам: пусть надо преобразовать i-ю строку, для этого необходимо умножить преобразованную ведущую строку на элемент i-й строки и ведущего столбца и результат вычесть из i-й строки. В преобразованной таким образом таблице изменить номера базисных переменных: ведущей строке будет соответствовать теперь номер ведущего столбца.

Положительное значение остаточной переменной указывает на неполное использование соответствующего ресурса, т. е. данный ресурс является недефицитным. Если же остаточная переменная равна нулю, это свидетельствует о полном потреблении соответствующего ресурса. В случае недефицитности любое увеличение ресурсов сверх установленного максимального значения привело бы лишь к тому, что они стали бы еще более недефицитными. Оптимальное решение задачи при этом осталось бы неизменным.

Тема 3. Транспортная задача

Важным частным случаем задачи линейного программирования является транспортная задача.

Пример решения задачи.

Построить экономико-математическую модель следующей задачи. Имеются 3 поставщика и 4 потребителя. Мощность поставщиков и спросы потребителей, а также затраты на перевозку единицы груза для каждой пары «поставщик – потребитель» сведены в таблицу поставок. В каждой клетке стоит коэффициент затрат – затраты на перевозку единицы груза от соответствующего поставщика к соответствующему потребителю.

Задача. Найти объем перевозок для каждой пары «поставщик – потребитель» так, чтобы:

1) мощность всех поставщиков были реализованы;
2) спросы всех потребителей удовлетворены;
3) суммарные затраты на перевозку были минимальными.

Транспортная задача

Поставщики	Мощность поставщиков	Потребители и их спрос			
		1	2	3	4
		20	110	40	110
1	60	1 x_{11}	2 x_{12}	5 x_{13}	3 x_{14}
2	120	1 x_{21}	6 x_{22}	5 x_{23}	2 x_{24}
3	100	6 x_{31}	3 x_{32}	7 x_{33}	4 x_{34}

Составим экономико-математическую модель задачи.

Искомый объем перевозки от i-го поставщика к j-му потребителю обозначим x_{ij} и назовем его поставкой. Тогда целевая функция, значение которой необходимо минимизировать, запишется в виде:

$$f(x) = 1x_{11}+2x_{12}+5x_{13}+\ldots+7x_{33}+4x_{34} \to min$$

Система ограничений будет иметь следующий вид:

$$\begin{cases} x_{11}+x_{12}+x_{13}+x_{14}=60 \\ x_{21}+x_{22}+x_{23}+x_{24}=120 \\ x_{31}+x_{32}+x_{33}+x_{34}=100 \end{cases} \qquad \begin{cases} x_{11}+x_{21}+x_{31}=20 \\ x_{12}+x_{22}+x_{32}=110 \\ x_{13}+x_{23}+x_{33}=40 \\ x_{14}+x_{24}+x_{34}=110 \end{cases}$$

Особенности экономико-математической модели транспортной задачи:

1) Система ограничений есть система уравнений, то есть транспортная задача задана в канонической форме.
2) Коэффициенты при переменных системы ограничений равны 1 или 0.
3) Каждая переменная входит в систему ограничений 2 раза.

Решение задачи.

Существует два метода нахождения первоначального распределения поставок (опорного плана).

1) Метод северо-западного угла.

Задаем северо-западной клетке (а именно x_{11}) максимально возможную поставку (20), после этого спрос 1-го потребителя будет полностью удовлетворен, в результате чего первый столбец поставок полностью выпадает из следующего рассмотрения. В оставшейся таблице опять выбираем северо-западную клетку.

Первый план перевозок,
построенный методом северо-западного угла

1 **20**	2 **40**	5	3	60
1	6 **70**	5 **40**	2 **10**	120
6	3	7	4 **100**	100
20	110	40	110	

Недостаток этого метода: план строится без учета стоимости (затраты на перевозку).

1) Метод минимальной стоимости (или метод наименьших затрат).

Находим клетку с наименьшим коэффициентом затрат (у нас их две, равные 1) и даем ей максимальную поставку (так как любая из них может поставить 20, выбираем любую). Таким образом, спрос 1-го потребителя удовлетворен и его (1-й столбец) вычеркиваем. В оставшейся таблице опять ищем клетку с минимальной стоимостью.

Первый план перевозок,
построенный методом минимальной стоимости

1	2 **60**	5	3	60
1 **20**	6	5	2 **100**	120
6	3 **50**	7 **40**	4 **10**	100
20	110	40	110	

➢ **Важно помнить.** Обязательно вычеркивается только один: или поставщик, или потребитель. Если на очередном шаге решения задачи совпали потребность покупателя и мощность поставщика, то одного (любого) вычеркиваем, а у второго пишем в остатке 0. На следующем шаге решения перевозим 0, тогда эта клетка участвует в плане перевозок. Если этого не сделать, то в плане будет недостаточно клеток, чтобы заполнить таблицу потенциалов.

Вычислим для обоих опорных планов суммарные затраты на перевозку.

$S_1 = 20*1 + 40*2 + 70*6 + 40*5 + 10*2 + 100*4 = 1140$

$S_2 = 60*2 + 20*1 + 100*2 + 50*3 + 40*7 + 10*4 = 810$

Во втором случае по числу шагов мы находимся ближе к оптимуму.

Решение методом потенциалов.

Выпишем отдельно полученный план перевозок X[1]

Первый план перевозок

	60 –	+	
20			100
	50 +	40 –	10

Вычисляем его стоимость: S(X[1])=60*2 + 100*1 + 20*2 + 50*3 + + 40*7 + 10*4 = 810.

Потенциалы и косвенные стоимости

α \ β	0	0	4	1
2	2' *-1*	**2**	6' *-1*	3' *0*
1	**1**	1' *5*	5' *0*	**2**
3	3' *3*	**3**	**7**	**4**

а) Вписываем в таблицу стоимости перевозок, соответствующих опорному плану.

б) Задаем произвольно один из потенциалов и вычисляем остальные, учитывая, что сумма потенциалов равна стоимости перевозки (в данной задаче задали $\beta_2 = 0$).

в) Вычисляем косвенные стоимости (суммируем соответствующие потенциалы и заполняем свободные клетки таблицы), помечаем косвенные стоимости штрихом.

г) Находим разницу между стоимостью, заданной в задаче, и косвенной стоимостью (цифры справа в клетке).

д) Выберем максимальную отрицательную разность и введем ее в опорный план, то есть увеличивая ее значение на какую-то величину, тогда значение другой переменной должно уменьшиться на эту же величину и так далее, замыкаем цикл. (Этот процесс называется <u>цикл пересчета</u>).

е) Если отрицательных значений нет, значит найденный опорный план является оптимальным.

ж) Определяем максимальную величину, на которую может быть увеличена клетка, вводимая в опорный план, так чтобы количество перевозки не стало отрицательным (в нашем случае она может быть равной 40).

Получаем новый опорный план: X[2].

Второй, улучшенный план перевозок

+	20 –	40	
20 –			100 +
	90 +		10 –

Его стоимость S(X[2])= 810 – 40*1 = 770 меньше предыдущего значения, значит полученный план ближе к оптимальному. Вычислим потенциалы для найденного опорного плана, положив $\beta_3 = 0$, косвенные стоимости и разницу между заданными и косвенными стоимостями.

Потенциалы и косвенные стоимости для второго плана перевозок

α \ β	-3	-3	0	-2
5	2’ *-1*	**2**	**5**	3’ *0*
4	**1**	1’ *5*	4’ *1*	**2**
6	3’ *3*	**3**	6’ *1*	**4**

Единственная клетка с отрицательной разностью, это клетка (1, 1). Следовательно, эту клетку можно ввести в опорный план. Максимальная величина, на которую может увеличиться клетка, чтобы опорный план не стал отрицательным 10 (эту величину выбираем из клеток, помеченных знаком «–»). Перевозка (3, 5) выйдет из опорного плана. Остальные перевозки изменятся соответственно знакам: прибавляем 10 к перевозкам, помеченным знаком «+», и вычитаем из перевозок, помеченных знаком «–».

➢ **Важно помнить.** На каждом шаге решения задачи одна перевозка входит в опорный план и одна выходит из него. Количество клеток, участвующих в плане перевозок, в ходе решения задачи не меняется. Если получается две клетки с одинаковыми минимальными значениями, помеченными знаком «–», то одна выходит из опорного плана, а во второй (в любой) остается 0, то есть клетка участвует в плане перевозок.

Получили план X[3].

Третий план перевозок

10	10	40	
10			110
	100		

Его стоимость S(X[3]) = 770 – 10*1 = 760, то есть стала еще меньше.

Чтобы выяснить, является ли полученный план оптимальным, вычислим потенциалы и косвенные стоимости для полученного плана.

Потенциалы и косвенные стоимости для третьего плана перевозок

α \ β	-4	-3	0	-2
5	**1**	**2**	**5**	3’ *0*
5	**1**	2’ 4	5’ *0*	**2**
6	2’ *4*	**3**	6’ *1*	4’ *0*

Все разности между заданными и косвенными стоимостями положительны, следовательно оптимальный план перевозок $X[3]$ найден и стоимость его равна 760 денежным единицам.

➢ **Важно помнить.** Описанный метод потенциалов позволяет решать только сбалансированные задачи, то есть задачи, в которых суммарная мощность поставщиков равна суммарному спросу потребителей.

На практике такая ситуация встречается редко, поэтому любую транспортную задачу можно привести к сбалансированной.

Если в транспортной задаче суммарная мощность поставщиков меньше суммарного спроса потребителей, то такая задача называется задачей на недостаток. Для ее решения необходимо ввести фиктивного поставщика, стоимости перевозок которого будут равны нулю, а мощность равна разности суммарного спроса потребителей и суммарной мощности действительных поставщиков, то есть размеру недостатка.

Если в транспортной задаче суммарный спрос потребителей меньше суммарной мощности поставщиков, то такая задача называется задачей на избыток. Для ее решения необходимо ввести фиктивного потребителя, стоимости перевозок которого будут равны нулю, а мощность равна разности суммарной мощности поставщиков и суммарного спроса действительных потребителей, то есть размеру избытка.

Когда задача решена, цифры в строке фиктивного поставщика показывают, какое количество продукции, кто из потребителей не получит, так как задача была на недостаток.

Когда задача решена, цифры в строке фиктивного потребителя показывают, какое количество продукции, у кого из поставщиков останется, так как задача была на избыток.

Тема 4. Игровые методы

Теория игр представляет собой теоретические основы математических моделей принятия оптимальных решений в конфликтных ситуациях рыночных отношений, носящих характер конкурентной борьбы, в которых одна противоборствующая сторона выигрывает за счет проигрыша другой. Наряду с такой ситуацией в теории принятия решений также рассматривают так называемые ситуацию риска и ситуацию неопределенности, которые имеют различные модели и требуют разных критериев выбора оптимальных решений. В ситуации риска предполагаются известными не только возможные условия, в которых нужно принимать решение, но и вероятности их появления. В ситуации неопределенности вероятности условий неизвестны и нет никакой возможности получить о них дополнительную статистическую информацию. Окружающая решение задачи среда, которая проявляется в тех или иных условиях, называется *природой*, а соответствующие математические модели называются *играми с природой* или *теорией статистических решений*.

Теория игр была систематически изложена ДЖ. Фон Нейманом и О. Монгерштерном в 1944 году. Они написали книгу, которая содержала главным образом экономические примеры, поскольку экономическому конфликту легче придать численную форму. Теория игр используется в области экономики и производства, бизнеса и финансов, сельского хозяйства и военного дела, биологии и социологии, психологии и политологии. К настоящему времени теория игр развилась в самостоятельную область математики и может рассматриваться независимо от ее приложений к реальным игровым ситуациям. По мнению лауреата Нобелевской премии по экономике за 1994 г. Нэша, теория игр вообще сыграла важную роль в интеллектуальной жизни XX века.

На практике часто приходится сталкиваться с задачами, в которых необходимо принимать решения в условиях неопределенности, то есть возникают ситуации, в которых две (или более) стороны преследуют различные цели, а результаты любого действия каждой из сторон зависят от мероприятий партнера. Такие ситуации, возникающие при игре в шахматы, шашки, домино, ситуации на рынке, когда выходит несколько производителей с одинаковым товаром (олигополия), покупатель и продавец, поставщик и потребитель, банк и клиент. То есть примером теории игр являются такие ситуации, в которых важную роль играют конфликты и совместные действия. Конфликт может возникнуть также из различия целей, которые отражают не только несовпадающие интересы различных сторон, но и многосторонние интересы одного и того же лица, например разработчик экономической политики обычно преследует разнообразные цели (рост объема производства, повышение доходов, снижение экологической нагрузки).

Таким образом, теория игр изучает и рассматривает методы определения оптимального поведения при управлении системами, в которых характерно наличие конфликтной ситуации (столкновение интересов). Для конфликтов характерно то, что ни один из его участников заранее не знает решений, принимаемых остальными участниками. Другими словами, участники конфликтов вынуждены действовать в условиях риска и неопределенности. Неопределенность исходов может проявляться не только в результате сознательных действий других участников, но и как результат действий тех или иных стихийных сил.

В свете «теории игр» можно рассмотреть экономику, общественные науки, бизнес и повседневную жизнь. К примеру, в экономике с точки зрения «теории игр» можно объяснить торговые и ценовые войны. Кроме того, некоторые обозреватели полагают, что, используя эту теорию, можно показать причины такого феномена, как «малоподвижные» цены. В соответствии с этой теорией фирмы заключают нечто вроде тайного соглашения о преобладающем значении цены (скажем, если речь идет об автомобильной или сталелитейной промышленности). После того как они пришли к соглашению, фирмы отказываются понижать или повышать цены, так как в противном случае участники рынка будут рассматривать такие изменения как сигнал объявления экономической войны. С помощью теории игр можно также объяснить, почему иностранная конкуренция может привести к более ожесточенной ценовой войне. Что случится, если японская фирма войдет на американский рынок, на котором уже существующие компании тайно договорились назначить высокую цену? Зарубежные фирмы могут «отказаться играть в эту игру». Они просто будут снижать цены в целях овладения большей долей рынка. Сговор может разрушиться.

Для грамотного решения задач с конфликтными ситуациями необходимы научно обоснованные методы. Такие методы разработаны математической теорией конфликтных ситуаций, которые носят название «теория игр».

Всякая претендующая на адекватность математическая модель социально-экономического явления должна отражать присущие ему черты конфликта, то есть описывать:

а) множество заинтересованных сторон, которые называются игроками;

б) возможные действия каждой из сторон, именуемые стратегиями или ходами;

в) интересы сторон, представленные функциями выигрыша (платежа) для каждого из игроков.

Сама модель конфликтной ситуации называется игрой.

Классификация игр

Различные виды игр можно классифицировать, основываясь на том или ином принципе: по числу игроков, по числу стратегий, по свойствам функции выигрыша, по возможности предварительных переговоров и взаимодействия между игроками в ходе игры.

1) Игра называется парной, если в ней участвуют два игрока, и множественной, если число игроков более двух.

2) Игра называется с нулевой суммой или антагонистической, если выигрыш одного из игроков равен проигрышу другого. Игры с постоянной разностью – игроки выигрывают и проигрывают одновременно.

3) Игра называется конечной, если у каждого игрока имеется конечное число стратегий, и бесконечной – в противном случае.

4) Бескоалиционными называются игры, в которых игроки не имеют права вступать в соглашения, образовывать коалиции. В коалиционной игре, наоборот, игроки могут вступать в соглашения и образовывать коалиции. В кооперативной игре коалиции определены заранее, то есть до начала игры игроки могут принимать соглашения о своих стратегиях (возможность предварительных переговоров). В некооперативной игре игроки не могут координировать свои стратегии подобным образом и принимают решения независимо друг от друга.

5) Комбинаторные игры: число исходов, стратегий, факторов конечное, не очень большое. Можно построить модель игры, выработать правила. Однако численное решение невозможно из-за большой размерности задачи. В случайных играх количество исходов не зависит от поведения игрока. В стратегических играх один участник находится в состоянии неопределенности относительно поведения других участников игры.

6) Матричная игра – это конечная игра двух игроков с нулевой суммой, в которой задаются выигрыши первого игрока в виде матрицы. Любая матричная игра имеет решение и может быть реализована методами линейного программирования. Матричные игры еще называют играми в нормальной форме. Биматричная игра – это конечная игра двух игроков с ненулевой суммой, в которой есть матрицы выигрышей (проигрышей) отдельно для каждого участника. Непрерывной считается игра, в которой функция выигрышей является непрерывной в зависимости от стратегий.

Другие термины и понятия

Выбор и осуществление одного из предусмотренных правилами действий называется ходом игрока. Ходы могут быть личными и случайными. Личный ход – это сознательный выбор игроком одного из возможных действий. Случайный ход – это случайно выбранное действие (выбор карты из перетасованной колоды).

Стратегией игрока называется совокупность правил, определяющих выбор действия при каждом ходе в зависимости от сложившейся ситуации. Чтобы решить игру или найти решение игры, следует для каждого игрока выбрать стратегию, которая удовлетворяет условию оптимальности, то есть один из игроков получает максимальный выигрыш, когда второй придерживается своей стратегии. В то же время второй игрок должен иметь минимальный проигрыш, если первый придерживается своей стратегии. Такие стратегии называются оптимальными.

Оптимальные стратегии должны удовлетворять условию устойчивости, то есть любому из игроков должно быть невыгодно отказываться от своей стратегии в этой игре. Если игра повторяется много раз, то игроков может интересовать не выигрыш и проигрыш в каждой конкретной партии, а средний выигрыш (проигрыш) во всех партиях.

Целью теории игр является определение оптимальной стратегии для каждого игрока.

Выигрыш – это мера эффекта для игрока. В теории игр выигрыш должен измеряться обязательно количественно.

В случае конечной игры двух лиц функции выигрыша каждого из игроков удобно представлять в виде матрицы выигрышей, где строки представляют собой стратегии одного игрока, а столбцы – стратегии другого. В клетках матрицы указываются выигрыши каждого из игроков в каждой из образующихся ситуаций. Данная форма представления конечных игр называется «матричные игры».

Пример решения задачи. Игра в «орлянку».

Если оба выбирают одинаковые стратегии (оба говорят «орел»), то 1-й выигрывает 1 рубль (а второй проигрывает); если выбирают разные, то 2-й выигрывает.

Матрица выигрыша первого игрока (H_1)

	Орел	Решка
Орел	1	–1
Решка	–1	1

Матрица выигрыша второго игрока (H_2)

	Орел	Решка
Орел	–1	1
Решка	1	–1

Для антагонистических игр всегда $H_1 = - H_2$.

Матрица, элементами которой являются выигрыши, соответствующие стратегиям игроков, называется платежной матрицей или матрицей игры.

Нижняя цена игры (α) (максиминный выигрыш – максимин) – это гарантированный выигрыш первого игрока при любой стратегии второго игрока (то есть из каждой строки выбираем минимальное число, а затем из всех этих минимумов берем наибольший). Стратегия, соответствующая максимину называется максиминной.

Верхняя цена игры (β) (минимаксный выигрыш – минимакс) – это гарантированный проигрыш второго игрока. (То есть из каждого столбца выбираем максимальное число, а затем из всех максимумов берем наименьший). Стратегия, соответствующая минимаксу, называется минимаксной.

Принцип, диктующий игрокам выбор наиболее «осторожных» минимаксной и максиминной стратегий, называется принципом минимакса. Этот принцип следует из разумного предположения, что каждый игрок стремится достичь цели, противоположной цели противника. Если верхняя и нижняя цены игры совпадают $\alpha=\beta=\partial$, то эта цена называется чистой ценой игры или ценой игры.

Для A=$\begin{pmatrix} -1 & +1 \\ +1 & -1 \end{pmatrix}$ $\alpha=-1, \beta=1 \qquad \alpha \neq \beta$

Для B=$\begin{pmatrix} 2 & 0 & 4 \\ 3 & 1 & 6 \end{pmatrix}$ $\alpha=1, \beta=1 \qquad \alpha=\beta=\partial=1$

Минимаксные стратегии, соответствующие цене игры, являются оптимальными стратегиями или решением игры. То есть в этом случае первый игрок получает максимальный, не зависящий от поведения второго игрока выигрыш q, а второй игрок добивается минимального гарантированного, не зависящего от поведения первого игрока проигрыша q. Такое решение обладает устойчивостью, то есть если один из игроков придерживается своей оптимальной стратегии, то для другого не может быть выгодным отклоняться от своей оптимальной стратегии.

Пара стратегий дает оптимальное решение игры тогда и только тогда, когда соответствующий элемент матрицы (размер выигрыша-проигрыша) является одновременно наибольшим в своем столбце и наименьшим в своей строке. Такая ситуация, если она существует, называется седловой точкой. То есть игра разрешима в чистых стратегиях, каждый из игроков будет на каждом шаге выбирать одну и ту же стратегию и ему не выгодно от нее отклониться: если первый игрок выберет другую стратегию, у него будет выигрыш меньше, а если второй игрок выберет другую стратегию, у него проигрыш будет больше.

Пример решения задачи.

Определить верхнюю и нижнюю цену игры, заданной платежной матрицей. Имеет ли игра седловую точку?

$$P=\begin{pmatrix} 0{,}5 & 0{,}6 & 0{,}8 \\ 0{,}9 & 0{,}7 & 0{,}8 \\ 0{,}7 & 0{,}6 & 0{,}6 \end{pmatrix}$$

Платежная матрица игры

	B_1	B_2	B_3	α
A_1	0,5	0,6	0,8	0,5
A_2	0,9	**0,7**	0,8	0,7
A_3	0,7	0,6	0,6	0,6
β	0,9	0,7	0,8	**0,7**

$\alpha = \beta = \partial = 0.7$ – цена игры. Седловая точка (A_2, B_2).

<u>*Пример решения задачи.*</u> Рассмотрим две конкурирующие финансовые компании A и B. Компания B ведёт переговоры с организаторами каждого из трёх проектов B_1, B_2, B_3 на предмет инвестирования. Задача компании B: положительный результат переговоров. Компания A ставит своей задачей свести переговоры компании B к отрицательному результату, с тем чтобы занять место компании B в инвестировании.

Компания A для достижения своей цели – срыва переговоров компании B – может применить одно из двух средств: A_1 – предложить организаторам проектов более выгодные для них условия инвестирования по сравнению с компанией B и A_2 – предоставить в распоряжение организаторов проектов материалы, компрометирующие компанию B.

Действие A_1 компании A приводит к отрицательному результату переговоров компании B с организаторами проектов B_1, B_2, B_3, соответственно, с вероятностями 0,7; 0,5; 0,3, а действие A_2 – с вероятностями 0,6; 0,9; 0,4.

Решение. Смоделируем данную ситуацию. Поскольку компании A и B преследуют противоположные цели, то рассматриваемая конфликтная ситуация является антагонистической. Игроками являются финансовые компании A и B. Игрок A имеет две чистые стратегии A_1 и A_2: $S_A^C = \{A_1, A_2\}$; множество стратегий игрока B состоит из трёх стратегий: $S_B^C = \{B_1, B_2, B_3\}$. Игрок B должен выбрать один из трех проектов, игрок A выбирает одно из двух своих действий.

В качестве выигрыша игрока A (или проигрыша игрока B) рассмотрим вероятность отрицательного результата переговоров компании B. В соответствии со своими задачами игрок A стремится максимизировать выигрыш, а игрок B – минимизировать проигрыш.

Выясним, имеет ли игра седловую точку, то есть, разрешима ли игра в чистых стратегиях.

Матрица игры с показателями эффективности стратегий A_1, A_2 и показателями неэффективности стратегий B_1, B_2, B_3 имеет следующий вид:

Платежная матрица игры

A_i \ B_j	B_1	B_2	B_3	α
A_1	0,7	0,5	0,3	0,3
A_2	0,6	0,9	**0,4**	0,4
β	0,7	0,9	0,4	**0,4** **0,4**

В данном случае максиминной стратегией игрока *A* является стратегия A_2 , а минимаксной стратегией игрока *B* – стратегия B_3.

Если игрок *A* придерживается своей максиминной стратегии A_2 , то игрок *B* должен выбрать свою минимаксную B_3, с тем чтобы выигрыш игрока *A* (или что то же – проигрыш игрока *B*) был минимальным a_{23} = 0,4 (во 2-й строке матрицы (6.1)). На это игрок *A* должен ответить выбором опять же стратегии A_2 , чтобы получить максимальный (в 3-м столбце) выигрыш: a_{23} = 0,4. Ответным ходом игрок *B* опять выбирает стратегию B_3 и т. д.

Таким образом, если игроки *A* и *B* придерживаются своих максиминной и минимаксной стратегий, то ни один из них не может увеличить свой выигрыш, отступая от своей стратегии. Ситуация (A_2 , B_3) является в данной игре *устойчивой*.

Нижняя и верхняя цены игры совпадают:

a = b = 0,4

Если игра не имеет седловой точки, то применение чистых стратегий не дает оптимального решения игры. В таком случае можно получить оптимальное решение, случайным образом чередуя чистые стратегии.

<u>Смешанной стратегией</u> S_A игрока А называется применение чистых стратегий $A_1, A_2, \ldots A_i, \ldots, A_m$ с вероятностями $p_1, p_2, \ldots, p_i, \ldots, p_m$, причем сумма вероятностей равна 1: $\sum_{i=1}^{m} p_i = 1$. Цена игры будет удовлетворять неравенству $\alpha \leq \partial \leq \beta$.

Основная теорема теории игр – <u>теорема Неймана</u>. Каждая конечная игра имеет, по крайней мере, одно оптимальное решение, возможно, среди смешанных стратегий.

В качестве примеров применения теории можно назвать решения по поводу проведения принципиальной ценовой политики, выхода на новые рынки, кооперации и создания совместных предприятий, определения лидеров и исполнителей в области инноваций, вертикальной интеграции и т.д. Положения данной теории в принципе можно использовать для всех видов решений, если на их принятие влияют другие действующие лица. Этими лицами, или игроками, необязательно должны быть рыночные конкуренты; в их роли могут выступать субпоставщики, ведущие клиенты, сотрудники организаций, а также коллеги по работе.

Геометрическая интерпретация игры 2×2

Пусть имеется два игрока А и В. У каждого из игроков по две стратегии (A_1 и A_2 у игрока А, B_1 и B_2 у игрока В). Игра с нулевой суммой.

По оси абсцисс отложим отрезок A_1A_2, то есть точка A_1 изображает стратегию A_1 (x=0), A_2 – стратегию A_2, все промежуточные точки – смешанные стратегии. На оси ординат откладываем выигрыш первого игрока, если второй применил стратегию B_1. Аналогично строим второй график, если второй график выбрал стратегию B_2.

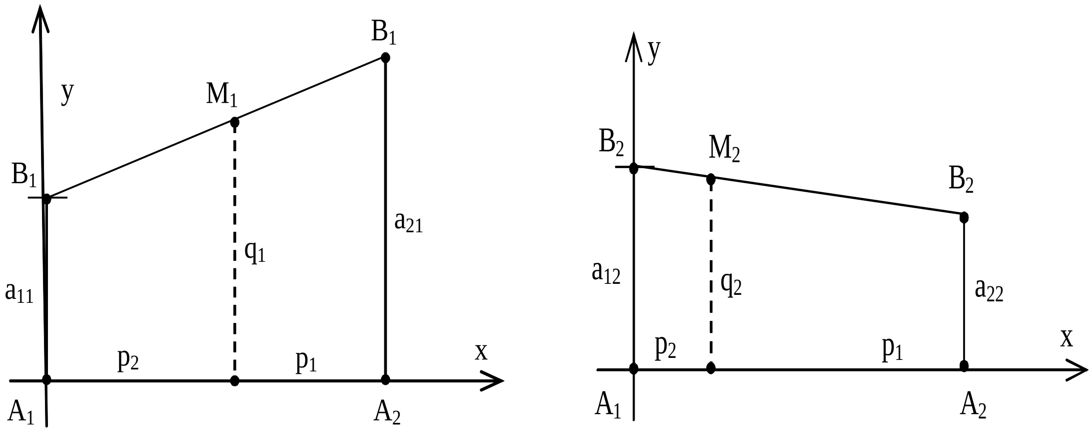


Рис. 2.1. Геометрическая интерпретация стратегий первого игрока

$q_1 = a_{11}p_1 + a_{21}p_2$

$q_2 = a_{12}p_1 + a_{22}p_2$ (ордината точки M_1 и M_2, соответственно)

$\left|\overline{A_1A_2}\right| = 1$

В соответствии с принципом минимакса оптимальная стратегия S_A^* такова, что минимальный выигрыш игрока А (при наихудшем поведении игрока В) обращается в максимум.

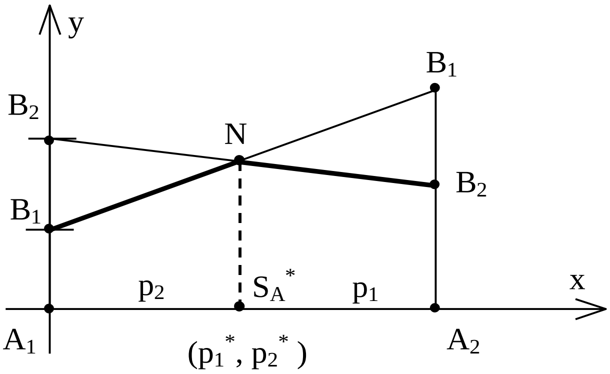


Рис. 2.2. Решение игры графическим способом

Отрезок B_1N – минимальный выигрыш игрока А при использовании любой смешанной стратегии, если игрок В выбрал стратегию B_1. Аналогично, отрезок B_2N – выигрыш игрока А, если игрок В выбрал стратегию B_2. Следовательно, оптимальную стратегию определяет точка N, то есть минимальный выигрыш достигает максимума.

<u>Пример решения задачи.</u>

Решить графически игру, заданную платежной матрицей:

$$P = \begin{pmatrix} 1{,}5 & 3 \\ 2 & 1 \end{pmatrix}$$

Определим верхнюю и нижнюю цены игры: $\alpha = 1{,}5, \quad \beta = 2$. Следовательно, седловая точка отсутствует, будем искать решение в смешанных стратегиях. Отметим на графике величину выигрыша для каждой пары стратегий.

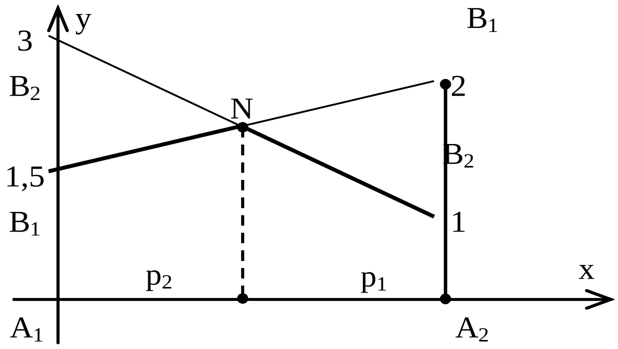


Рис. 2.3. Графическое решение игры

Запишем уравнение прямых, соответствующих величине выигрыша игрока А, если игрок В выбирает, соответственно, первую *(B_1)* $y = = 0,5\,x + 1,5$ и вторую *(B_2)* $y = -2x + 3$ стратегии. Точка их пересечения $N(0,6;\ 1,8)$. Следовательно, $p_2 = x = 0,6$ – вероятность выбора игроком А второй стратегии; $p_1 = 1 - 0,6 = 0,4$ – вероятность выбора игроком А первой стратегии. Цена игры $\partial = y = 1,8$ – максимально возможный из минимально гарантированных выигрышей игрока А (либо минимально возможный из максимально гарантированных проигрышей игрока В).

- **Важно помнить.** Графически можно решить игру, если в игре участвуют только два игрока и у одного из игроков имеется только две стратегии (у второго игрока – любое количество стратегий).

Рассмотрим ситуацию, когда игра разрешается в чистых стратегиях, т. е. есть седловая точка. Графическое решение задачи может иметь следующие виды.

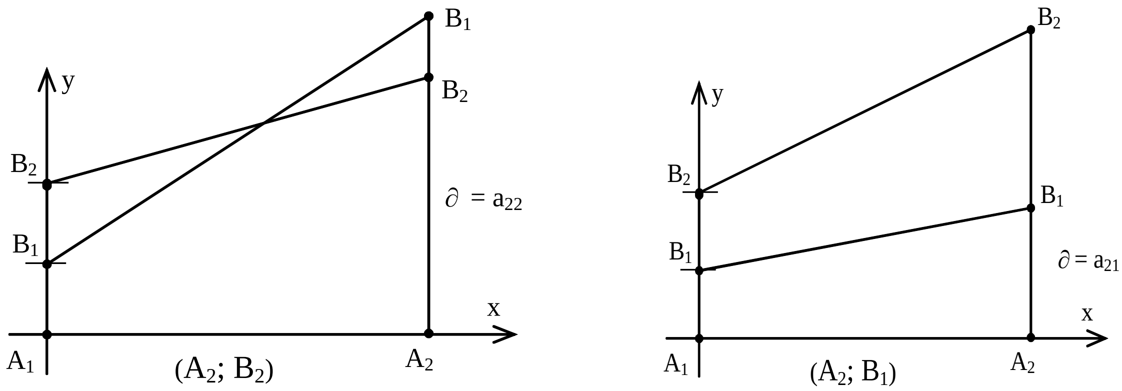


Рис. 2.4. Графики игр, разрешимых в чистых стратегиях

Приведение матричной игры к задаче линейного программирования

Игра $m \times n$ в общем случае не имеет наглядной геометрической интерпретации. Решение достаточно трудоемко при больших m и n, но может быть сведено к задаче линейного программирования.

<u>Пример решения задачи.</u> Предприятие может выпускать 3 вида продукции A_1, A_2 и A_3, получая при этом прибыль, зависящую от спроса, который может быть в одном из 4-х состояний (B_1, B_2, B_3, B_4). Элементы платежной матрицы характеризуют прибыль, которую получат при выпуске i-й продукции при j-м состоянии спроса. Определить оптимальные пропорции в выпускаемой про-

дукции, гарантирующие максимизацию средней величины прибыли при любом состоянии спроса, считая его определенным. Задача сводится к игровой модели, в которой игра предприятия А против спроса В задана платежной матрицей.

Платежная матрица игры

	B_1	B_2	B_3	B_4	α
A_1	3	3	6	8	3
A_2	9	10	4	2	2
A_3	7	7	5	4	4
β	9		6	8	4 / 6

Прежде чем решить задачу, попытаемся упростить игру, проведя анализ платежной матрицы и отбросив заведомо невыгодные и дублирующие стратегии.

Для игрока В невыгодна вторая стратегия (столбец B_2), так как все элементы этого столбца больше или равны элементам столбца B_1, то есть при любой стратегии, выбранной игроком А, игрок В проигрывает больше в случае выбора второй стратегии вместо первой.

Поскольку $\alpha \neq \beta$, следовательно седловая точка отсутствует, решение будем искать в смешанных стратегиях.

Чтобы привести игру к задаче линейного программирования, обозначим:

$$x_i = \frac{p_i}{\partial} \quad ; \quad y_i = \frac{q_i}{\partial} .$$

Получим две взаимно двойственные задачи линейного программирования:

Для первого игрока:

$$\begin{cases} 3p_1 + 9p_2 + 7p_3 \geq \partial \\ 6p_1 + 4p_2 + 5p_3 \geq \partial \\ 8p_1 + 2p_2 + 4p_3 \geq \partial \end{cases} \Rightarrow \begin{cases} 3x_1 + 9x_2 + 7x_3 \geq 1 \\ 6x_1 + 4x_2 + 5x_3 \geq 1 \\ 8x_1 + 2x_2 + 4x_3 \geq 1 \end{cases} \Rightarrow \quad z = x_1 + x_2 + x_3 \to \min$$ (та

к как $\partial \to \max$, то $\sum_{i=1}^{3} x_i \to \min$.

Поскольку $x_1 + x_2 + x_3 = \frac{p_1}{\partial} + \frac{p_2}{\partial} + \frac{p_3}{\partial} = \frac{1}{\partial}$.

И для второго игрока:

$$\begin{cases} 3y_1 + 6y_2 + 8y_3 \leq 1 \\ 9y_1 + 4y_2 + 2y_3 \leq 1 \\ 7y_1 + 5y_2 + 4y_3 \leq 1 \end{cases} \Rightarrow \quad z = y_1 + y_2 + y_3 \to \max$$

Вторую задачу на максимум решать легче. Решаем ее симплексным методом. Получаем следующий результат:

$$Y\left(\frac{1}{27}, \frac{4}{27}, 0, 0, \frac{2}{27}, 0\right), \quad z_{\max} = \frac{5}{27} .$$

Последняя запись целевой функции имеет вид:

$$z(y) = \frac{5}{27} - \frac{1}{27}y_3 - \frac{2}{27}y_4 - \frac{1}{9}y_6.$$

Следовательно, можем записать решение взаимно двойственной задачи, то есть задачи для первого игрока, используя теорию двойственности. Это решение будет иметь вид:

$$x_6 = \frac{1}{27},\ x_1 = \frac{2}{27},\ x_3 = \frac{1}{9}.$$

Цена игры $\partial = \dfrac{1}{z_{\max}} = \dfrac{27}{5}.$

Найдем вероятности выбора стратегий:

$$p_1 = x_1 \cdot \partial = \frac{2}{27} \cdot \frac{27}{5} = \frac{2}{5}, \quad p_2 = 0, \quad p_3 = x_3 \cdot \partial = \frac{1}{9} \cdot \frac{27}{5} = \frac{3}{5}.$$

То есть $S_A = \left(\frac{2}{5}, 0, \frac{3}{5}\right)$.

Данный результат означает, что предприятие должно выпустить 40 % продукции A_1 и 60 % продукции A_3, а продукцию A_2 выпускать не надо. Тогда максимально гарантированное из минимально возможных значений средней величины прибыли составит 5,4 ден. ед., независимо от спроса покупателей.

При решении произвольной конечной игры размера m×n рекомендуется придерживаться следующей **схемы:**

1) Исключить из платежной матрицы заведомо невыгодные стратегии по сравнению с другими стратегиями. Такими стратегиями для игрока А (игрока В) являются те, которым соответствуют строки (столбцы) с элементами заведомо меньшими (большими) по сравнению с элементами других строк (столбцов).
2) Определить верхнюю и нижнюю цены игры и проверить, имеет ли игра седловую точку. Если седловая точка есть, то соответствующие ей стратегии игроков будут оптимальными, а цена совпадает с верхней (нижней) ценой.
3) Если седловая точка отсутствует, то решение следует искать в смешанных стратегиях. Для игр размера m×n рекомендуется симплексный метод, для игр размера 2×2, 2×n, n×2 возможно геометрическое решение.

На практике реализация решения в смешанных стратегиях может происходить несколькими путями. Первый состоит в физическом смешении чистых стратегий в пропорциях, заданных вероятностями p_i. Другой путь – при многократном повторении игры предполагает применение в каждой партии чистых стратегий в виде случайной последовательности, причем каждая из них – с частотой, равной ее вероятности в оптимальном решении.

Однако существуют определенные границы применения аналитического инструментария теории игр. В следующих случаях он может быть использован лишь при условии получения дополнительной информации.

Во-первых, это тот случай, когда у предприятий сложились разные представления об игре, в которой они участвуют, или когда они недостаточно информированы о возможностях друг друга. Например, может иметь место неясная информация о платежах конкурента (структуре издержек). Если неполнотой характеризуется не слишком сложная информация, то можно оперировать сопоставлением подобных случаев с учетом определенных различий.

Во-вторых, теорию игр трудно применять при множестве ситуаций равновесия. Эта проблема может возникнуть даже в ходе простых игр с одновременным выбором стратегических решений.

В-третьих, если ситуация принятия стратегических решений очень сложна, то игроки часто не могут выбрать лучшие для себя варианты. Например, на рынок в разные сроки могут вступить несколько предприятий или реакция уже действующих там предприятий может оказаться более сложной, нежели быть агрессивной или дружественной.

Экспериментально доказано, что при расширении игры до десяти и более этапов игроки уже не в состоянии пользоваться соответствующими алгоритмами и продолжать игру с равновесными стратегиями.

➢ **Важно помнить.** Теория игр является очень сложной областью знания. При обращении к ней надо соблюдать известную осторожность и четко знать границы применения. Слишком простые толкования, принимаемые фирмой самостоятельно или с помощью консультантов, таят в себе скрытую опасность. Анализ и консультации на основе теории игр из-за их сложности рекомендуются лишь для особо важных проблемных областей. Опыт фирм показывает, что использование соответствующего инструментария предпочтительно при принятии однократных, принципиально важных плановых стратегических решений, в том числе при подготовке крупных кооперационных договоров.

Пример решения задачи.

Швейная фабрика выпускает брюки и шорты, сбыт которых зависит от состояния погоды. Затраты фабрики на единицу продукции составили: брюки – 15 ден. ед., шорты – 10 ден. ед. Цена реализации: брюки – 21 ден. ед., шорты – 14 ден. ед. Фабрика может реализовать при теплой погоде 120 брюк и 300 шорт, а при прохладной погоде: 370 брюк и 100 шорт.

Составим платежную матрицу игры.

Матрица выигрыша фабрики

Погода

		Теплая	Холодная
Фабрика	Теплая	a_{11}	a_{12}
	Холодная	a_{21}	a_{22}

Вычислим значения элементов матрицы:

a_{11} = 120*(21 – 15) + 300*(14 – 10) = 1920 – прибыль фабрики, если продукция выпущена по плану теплой погоды и погода оказалась теплой;

a_{22} = 370*(21 – 15) + 100*(14 – 10) = 2620 – прибыль фабрики, если продукция выпущена по плану холодной погоды и погода оказалась холодной;

a_{12} = 120*(21 – 15) + 100*(14 – 10) – 200*10 = –880 – прибыль фабрики, если продукция выпущена по плану теплой погоды, а погода оказалась холодной;

a_{21} = 120*(21 – 15) + 100*(14 – 10) – 250*15 = –2630 – прибыль фабрики, если продукция выпущена по плану холодной погоды, а погода оказалась теплой.

Таким образом, платежная матрица данной игры имеет вид:

Платежная матрица игры

Погода

		Теплая	Холодная
Фабрика	Теплая	1920	-880
	Холодная	-2630	2620

Обозначим: p_1 – вероятность выбора фабрикой первой стратегии (то есть производства продукции по плану теплой погоды), p_2 – вероятность выбора фабрикой второй стратегии (то есть производства продукции по плану холодной погоды). Тогда:

(1920* p_1 – 2630* p_2) – прибыль фабрики, если погода будет теплой;

(–880* p_1 + 2620* p_2) – прибыль фабрики, если погода будет холодной.

Чтобы прибыль фабрики не зависела от погоды, надо найти такие p_1 и p_2, что (1920* p_1 – 2630* p_2) = (–880* p_1 + 2620* p_2). Учитывая свойство вероятностей: $p_1 + p_2 = 1$, решаем уравнение.

1920* p_1 – 2630*(1 – p_1) = –880* p_1 + 2620* (1 – p_1)

Раскрывая скобки и приводя подобные, получаем:

p_1*(1920 + 2630 + 880 + 2620) = (2620 + 2630)

$p_1 = \frac{5250}{8050} = 0{,}652$. Тогда p_2 = 1–0,652 = 0,348.

Тогда план выпуска продукции для фабрики должен составить:
количество брюк = 120*0,652 + 370*0,348 = 207 шт.;
количество шорт = 300*0,652 + 100*0,348 = 230 шт.
При таком плане производства фабрика гарантирует себе прибыль в размере 1920*0,652 – 2630*0,348 = 337 ден.ед.

Тема 5. Теория массового обслуживания

Структура и классификация систем массового обслуживания

Нередко возникает необходимость в решении вероятностных задач, связанных с системами массового обслуживания (СМО), примерами которых могут быть:

- Билетные кассы;
- Ремонтные мастерские;
- Торговые, транспортные, энергетические системы;
- Системы связи;
- и т.д.

Общность таких систем выявляется в единстве математических методов и моделей, применяемых при исследовании их деятельности.

Рис. 4.1. Основные сферы применения ТМО

На вход в СМО поступает поток требований на обслуживание. Например, клиенты или пациенты, поломки в оборудовании, телефонные вызовы. Требования поступают нерегулярно, в случайные моменты времени. Случайный характер носит и продолжительность обслуживания. Это создает нерегулярность в работе СМО, служит причиной ее перегрузок и недогрузок.

Системы массового обслуживания обладают различной структурой, но обычно в них можно выделить четыре основных элемента:

1. Входящий поток требований.
2. Накопитель (очередь).
3. Приборы (каналы обслуживания).
4. Выходящий поток.

Рис. 4.2. Общая схема систем массового обслуживания

Рис. 4.3. Модель работы системы
(стрелками показаны моменты поступления требований в систему, прямоугольниками – время обслуживания)

На рис.4.3 а представлена модель работы системы с регулярным потоком требований. Поскольку известен промежуток между поступлениями требований, то время обслуживания выбрано так, чтобы полностью загрузить систему. Для системы со стохастическим потоком требований ситуация совершенно иная – требования приходят в различные моменты времени и время обслуживания тоже является случайной величиной, которое может быть описано неким законом распределения (рис.4.3 б).

В зависимости от правил образования очереди различают следующие СМО:

1) ***системы с отказами***, в которых при занятости всех каналов обслуживания заявка покидает систему необслуженной;

2) ***системы с неограниченной очередью***, в которых заявка встает в очередь, если в момент ее поступления все каналы обслуживания были заняты;

3) ***системы с ожиданием и ограниченной очередью***, в которых время ожидания ограниченно какими-либо условиями или существуют ограничения на число заявок, стоящих в очереди.

Рассмотрим характеристики входящего потока требований.

Поток требований называется ***стационарным***, если вероятность попадания того или иного числа событий на участок времени определенной длины зависит только от длины этого участка.

Поток событий называется ***потоком без последствий***, если число событий, попадающих на некоторый участок времени, не зависит от числа событий, попадающих на другие.

Поток событий называется ***ординарным***, если невозможно одновременное поступление двух или более событий.

Поток требований называется ***пуассоновским*** (или простейшим), если он обладает тремя свойствами: стационарен, ординарен и не имеет последствий. Название связано с тем, что при выполнении указанных условий число событий, попадающих на любой фиксированный интервал времени, будет распределен по закону Пуассона.

Интенсивностью потока заявок λ называется среднее число заявок, поступающих из потока за единицу времени.

Для стационарного потока интенсивность постоянна. Если τ – среднее значение интервала времени между двумя соседними заявками, то $\lambda = \frac{1}{\tau}$. В случае пуассоновского потока вероятность поступления на обслуживание m заявок за промежуток времени t определяется по закону Пуассона:

$$P_m(t) = \frac{(\lambda t)^m}{m!} \cdot e^{-\lambda t}.$$

Время между соседними заявками распределено по экспоненциальному закону с плотностью вероятности $f(t) = \lambda e^{-\lambda t}$.

Время обслуживания $t_{обсл}$ является случайной величиной и подчиняется показательному закону распределения с плотностью вероятности $f(t) = \mu e^{-\mu t}$, где μ – интенсивность потока обслуживания, т.е. среднее число заявок, обслуживаемых в единицу времени, $t_{обсл} = \frac{1}{\mu}$.

Отношение интенсивности входящего потока к интенсивности потока обслуживания называется ***загрузкой системы*** $\rho = \frac{\lambda}{\mu}$.

Загрузка – это среднее число заявок, приходящих за среднее время обслуживания одной заявки.

Марковский процесс

Система массового обслуживания представляет собой систему дискретного типа с конечным или счетным множеством состояний, а переход системы из одного состояния в другое происходит скачком, когда осуществляется какое-нибудь событие.

Процесс называется ***процессом с дискретными состояниями***, если его возможные состояния $S_1, S_2, S_3, \ldots$ можно заранее перенумеровать, и переход системы из состояния в состояние происходит практически мгновенно.

Такие процессы бывают двух типов: с дискретным или непрерывным временем.

В случае дискретного времени переходы из состояния в состояние могут происходить в строго определенные моменты времени. Процессы с непрерывным временем отличаются тем, что переход системы в новое состояние возможен в любой момент времени.

Случайным процессом называется соответствие, при котором каждому значению аргумента (в данном случае – моменту из промежутка времени проводимого опыта) ставится в соответствие случайная величина (в данном случае – состояние СМО). ***Случайной величиной*** называется величина, которая в результате опыта может принять одно, но неизвестное заранее, какое именно, числовое значение из данного числового множества.

Поэтому для решения задач теории массового обслуживания необходимо этот случайный процесс изучить, т.е. построить и проанализировать его математическую модель.

Случайный процесс называется ***марковским***, если для любого момента времени t_0 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.

Переходы системы из состояния в состояние происходит под действием каких-то потоков (поток заявок, поток отказов). Если все потоки событий, приводящие систему в новое состояние, – простейшие пуассоновские, то процесс, протекающий в системе, будет марковским, так как простейший поток не обладает последствием: в нем будущее не зависит от прошлого.

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.

Системы массового обслуживания с отказами

В системах с отказами заявка, поступившая в момент, когда все каналы обслуживания заняты, немедленно получает отказ, покидает систему и в дальнейшем в процессе обслуживания не участвует.

Имеется n каналов в обслуживании, на которые поступает поток заявок с интенсивностью λ. Поток обслуживания имеет интенсивность μ (величина, обратная среднему времени обслуживания $t_{обсл}$). Требуется найти вероятности состояний СМО и характеристики ее эффективности.

Так как оба потока – заявок и освобождений – простейшие, процесс, протекающий в системе, будет марковским. Рассмотрим ее как систему с конечным множеством состояний:

S_0 – свободны все каналы;

S_1 – занят ровно один канал;

…

S_k – заняты k каналов;

…

S_n – заняты все n каналов/

Через $P_k(t)$ обозначим вероятность того, что в момент времени t система будет находиться в состоянии S_k.

Простейшие задачи для систем массового обслуживания с отказами были впервые решены А.К. Эрлангом. Им же были выведены формулы оценки функционирования этих систем при условии поступления простейшего потока заявок и для показательного закона распределения времени обслуживания.

Для установившегося процесса обслуживания при этих условиях Эрланг получил следующие зависимости.

- <u>Вероятность того, что обслуживанием заняты k аппаратов</u> (линий, приборов и т.п.):

$$P_k = \frac{\frac{\rho^k}{k!}}{\sum_{k=0}^{n} \frac{\rho^k}{k!}}, \qquad (0 \le k \le n), \tag{4.1}$$

где $\rho = \frac{\lambda}{\mu}$, k – количество занятых аппаратов,

λ – интенсивность потока заявок,

μ – интенсивность потока обслуживания.

<u>*Частные случаи:*</u>

- <u>Вероятность простоя</u> **(**того, что все обслуживающие аппараты свободны, нет заявок):

$$P_0 = \frac{1}{\sum_{k=0}^{n} \frac{\rho^k}{k!}}, \tag{4.2}$$

- <u>Вероятность отказа</u> (вероятность того, что все обслуживающие приборы заняты):

$$P_{отк} = P_n = \frac{\frac{\rho^n}{n!}}{\sum_{k=0}^{n} \frac{\rho^k}{k!}} = \rho^n \frac{P_0}{n!}. \tag{4.3}$$

Отсюда находим ***относительную пропускную способность***, т.е. среднюю долю пришедших заявок, обслуживаемых системой, – вероятность того, что заявка будет обслужена:

$$P_{обсл} = 1 - P_{отк} = 1 - \rho^n \frac{P_0}{n!}. \tag{4.4}$$

Абсолютную пропускную способность, т.е. среднее число заявок, обслуживаемых в единицу времени, получим, умножив интенсивность потока заявок на относительную пропускную способность:

$A = \lambda P_{обсл.}$

Абсолютная пропускная способность – это интенсивность потока обслуженных системой заявок, а каждый занятый канал в единицу времени обслуживает в среднем μ заявок. Значит, ***среднее число занятых каналов*** равно

$$\bar{k} = \frac{A}{\mu} = \rho P_{обсл.} \quad (4.5)$$

Доля каналов, занятых обслуживанием (коэффициент загрузки):

$$q = \frac{\bar{k}}{n}. \quad (4.6)$$

<u>*Пример решения задачи.*</u>

На вход трехканальной СМО с отказами поступает поток заявок с интенсивностью λ = 4 заявки в минуту. Время обслуживания заявки одним каналом $t_{обсл} = \frac{1}{\mu} = 0.5$ мин.

Найти показатели эффективности работы системы.

Решение.

Находим вероятность простоя трехканальной СМО по формуле (4.2):

$\rho = \frac{\lambda}{\mu} = \frac{4}{2} = 2$ - загрузка системы (среднее число заявок, приходящих за среднее время обслуживания одной заявки).

$n = 3.$

$$P_0 = \frac{1}{1 + \frac{2}{1!} + \frac{2^2}{2!} + \frac{2^3}{3!}} = \frac{1}{1 + 2 + 2 + \frac{4}{3}} = \frac{1}{\frac{19}{3}} \approx 0{,}158.$$

Вероятность отказа определяем по формуле (4.3):

$$P_{отк} = \rho^n \frac{P_0}{n!} = 2^3 \frac{0.158}{3!} \approx 0{,}21.$$

Относительная пропускная способность системы:

$$P_{обсл} = 1 - P_{отк} = 1 - \rho^n \frac{P_0}{n!} = 1 - 0{,}21 = 0{,}79.$$

Абсолютная пропускная способность системы (среднее число заявок, обслуживаемых в единицу времени):

$A = \lambda P_{обсл.} = 4 \cdot 0{,}79 = 3{,}16.$

Среднее число занятых каналов (в ед. времени) определяем по формуле (4.5):

$$\bar{k} = \frac{A}{\mu} = \rho P_{обсл.} = 2 \cdot 0{,}79 = 1{,}58.$$

Доля каналов, занятых обслуживанием (формула (4.6)):

$$q = \frac{\bar{k}}{n} = \frac{1.58}{3} \cong 0{,}53.$$

Среднее время пребывания заявки в СМО находим как вероятность того, что заявка принимается к обслуживанию, умноженную на среднее время обслуживания:

$t_{СМО} = 0{,}79 \cdot 0{,}5 = 0{,}395$ мин.

Системы массового обслуживания с неограниченным ожиданием

Пусть имеется n-канальная СМО с очередью, на которую не наложено ограничений ни по длине очереди, ни по времени ожидания. В силу неограниченности очереди каждая заявка рано или поздно будет обслужена, поэтому

$P_{обсл} = 1, \qquad P_{отк} = 0.$

Для СМО с неограниченной очередью накладывается ограничение $\frac{\rho}{n} < 1$.

Если это условие нарушено, то очередь растет до бесконечности, наступает явление «взрыва».

- ***Вероятность простоя*** **(**того, что все обслуживающие аппараты свободны, нет заявок):

$$P_0 = \frac{1}{\sum\limits_{k=0}^{n} \frac{\rho^k}{k!} + \frac{\rho^{n+1}}{n!(n-\rho)}}, \tag{4.7}$$

- ***Вероятность занятости обслуживанием k каналов***:

$$P_k = \rho^k \frac{P_0}{n!}, \qquad 1 \le k \le n. \tag{4.8}$$

- ***Вероятность занятости обслуживанием всех каналов при отсутствии очереди***:

$$P_n = \rho^n \frac{P_0}{n!}. \tag{4.9}$$

- ***Вероятность наличия очереди*** есть вероятность того, что число требований в системе больше числа каналов:

$$P_{оч} = \rho^{n+1} \frac{P_0}{n!(n-\rho)}. \tag{4.10}$$

- ***Вероятность для заявки попасть в очередь*** есть вероятность занятости всех каналов, эта вероятность равна сумме вероятностей наличия очереди и занятости всех n каналов при отсутствии очереди:

$$P_{зан} = P_n + P_{оч} = \rho^n \frac{P_0}{(n-1)!(n-\rho)}. \tag{4.11}$$

- ***Среднее число занятых обслуживанием каналов***:

$$\bar{k} = \frac{\lambda}{\mu} = \rho. \tag{4.12}$$

- ***Доля каналов, занятых обслуживанием***:

$$q = \frac{\bar{k}}{n}. \tag{4.13}$$

- ***Среднее число заявок в очереди*** (длина очереди)

$$L = \rho^{n+1}\frac{P_0}{(n-1)!(n-\rho)^2}. \qquad (4.14)$$

- ***Среднее число заявок в системе***

$$M = L + \bar{k} = L + \rho. \qquad (4.15)$$

- ***Среднее время ожидания заявки в очереди***

$$t = \frac{L}{\lambda}. \qquad (4.16)$$

- ***Среднее время пребывания заявки в системе***

$$T = t + \frac{1}{\mu}, \qquad T = \frac{M}{\lambda}. \qquad (4.17)$$

<u>*Пример решения задачи.*</u>

На вход трехканальной СМО с неограниченной очередью поступает поток заявок с интенсивностью λ = 4 заявки в минуту. Среднее время обслуживания заявки $t_{обсл} = \frac{1}{\mu} = 0.5$ ч.

Найти показатели эффективности работы системы.

Решение

Для рассматриваемой системы

$$n = 3, \quad \lambda = 4, \quad \mu = \frac{1}{0.5} = 2, \quad \rho = \frac{\lambda}{\mu} = 2, \quad \frac{\rho}{n} = \frac{2}{3} < 1.$$

Определяем вероятность простоя по формуле (4.7):

$$P_0 = \frac{1}{\sum_{k=0}^{n}\frac{\rho^k}{k!} + \frac{\rho^{n+1}}{n!(n-\rho)}} = \left(1 + \frac{2}{1!} + \frac{2^2}{2!} + \frac{2^3}{3!} + \frac{2^4}{3!(3-2)}\right)^{-1} = \frac{1}{9}.$$

Среднее число заявок в очереди находим по формуле (4.14):

$$L = \rho^{n+1}\frac{P_0}{(n-1)!(n-\rho)^2} = 2^4\frac{\frac{1}{9}}{2!\cdot(3-2)^2} = \frac{8}{9}.$$

Среднее время ожидания заявки в очереди считаем по формуле (4.16):

$$t = \frac{L}{\lambda} = \frac{8}{9}\Big/4 = \frac{2}{9} \approx 0{,}22 \text{ ч.}$$

Среднее время пребывания заявки в системе

$$T = t + \frac{1}{\mu} = \frac{2}{9} + \frac{1}{2} = \frac{13}{18} \cong 0{,}72 \text{ ч.}$$

Системы массового обслуживания с ожиданием и ограниченной длинной очереди

Имеется n-канальная СМО с ожиданием, в которой количество заявок, стоящих в очереди, ограничено числом m, т.е. заявка, заставшая все каналы занятыми, становится в очередь, только если в ней находится менее m заявок.

Если число заявок в очереди равно m, то последняя прибывшая заявка в очередь не становится и покидает систему необслуженной.

Системы с ограниченной очередью являются обобщением двух рассмотренных ранее СМО: при $m = 0$ получаем СМО с отказами, при $m = \infty$ получаем СМО с ожиданием.

- ***Вероятность простоя*** (того, что все обслуживающие аппараты свободны, нет заявок):

$$P_0 = \frac{1}{\sum_{k=0}^{n} \frac{\rho^k}{k!} + \frac{\rho^{n+1}}{n!(n-\rho)}\left(1-\left(\frac{\rho}{n}\right)^m\right)}, \tag{4.18}$$

- ***Вероятность отказа в обслуживании*** равна вероятности P_{n+m} того, что в очереди уже стоят m заявок:

$$P_{отк} = P_{n+m} = \rho^{n+m} \frac{P_0}{n! \cdot n^m}. \tag{4.19}$$

- ***Относительная пропускная способность*** есть величина, дополняющая вероятность отказа до 1, т.е. вероятность обслуживания:

$P_{обсл} = 1 - P_{отк}$.

- ***Абсолютная пропускная способность*** определяется равенством:

$$A = \lambda(1 - P_{отк}) = \lambda P_{обсл}. \tag{4.21}$$

- ***Среднее число занятых обслуживанием каналов***:

$$\bar{k} = \frac{A}{\mu} = \rho P_{обсл}. \tag{4.22}$$

- ***Среднее число заявок в очереди*** (средняя длина очереди)

$$L = \frac{\rho^n P_0\left(\frac{\rho}{n} - (m+1)\left(\frac{\rho}{n}\right)^{m+1} + m\left(\frac{\rho}{n}\right)^{m+2}\right)}{n!\left(n - \frac{\rho}{n}\right)^2}. \tag{4.23}$$

- ***Среднее время ожидания обслуживания в очереди***

$$t = \frac{L}{\lambda}. \tag{4.24}$$

- ***Среднее число заявок в системе***

$$M = L + \bar{k}. \tag{4.25}$$

- ***Среднее время пребывания заявки в системе***

$$T = t + \frac{1}{\mu}, \qquad T = \frac{M}{\lambda}. \tag{4.26}$$

Пример решения задачи.

В парикмахерской работают 3 мастера, в зале ожидания расположено 3 стула. Поток клиентов имеет интенсивность λ = 12 клиентов в час. Среднее время обслуживания заявки $t_{обсл} = \frac{1}{\mu} = 20$ мин. Определить относительную и абсолютную пропускную способность системы, среднее число занятых кресел, среднюю длину очереди, среднее время, которое клиент проводит в парикмахерской.

Решение

Для данной задачи

$$n = 3, \quad m = 3, \quad \lambda = 12, \quad \mu = 3, \quad \rho = \frac{\lambda}{\mu} = 4, \quad \frac{\rho}{n} = \frac{4}{3}.$$

Определяем вероятность простоя по формуле (4.18):

$$P_0 = \frac{1}{\sum_{k=0}^{n} \frac{\rho^k}{k!} + \frac{\rho^{n+1}}{n!(n-\rho)}\left(1 - \left(\frac{\rho}{n}\right)^m\right)} =$$

$$= \frac{1}{1 + 4 + \frac{4^2}{2} + \frac{4^3}{3!} + \frac{4^4}{3!(3-4)}\left(1 - \left(\frac{4^3}{3^3}\right)\right)} \approx 0{,}012.$$

Вероятность отказа в обслуживании определим по формуле (4.19)

$$P_{отк} = P_{n+m} = \rho^{n+m} \frac{P_0}{n! \cdot n^m} = 4^{3+3} \frac{0.012}{3! \cdot 3^3} \approx 0{,}307.$$

Относительная пропускная способность, т.е. вероятность обслуживания:

$$P_{обсл} = 1 - P_{отк} = 1 - 0{,}307 = 0{,}693.$$

Абсолютная пропускная способность (среднее число заявок, обслуживаемых в единицу времени):

$$A = \lambda P_{обсл} = 12 \cdot 0{,}693 \approx 8{,}32.$$

Среднее число занятых обслуживанием каналов (парикмахеров):

$$\bar{k} = \frac{A}{\mu} = \frac{8.32}{3} \approx 2{,}78.$$

Среднее число заявок в очереди (средняя длина очереди)

$$L=\frac{\rho^n P_0\left(\frac{\rho}{n}-(m+1)\left(\frac{\rho}{n}\right)^{m+1}+m\left(\frac{\rho}{n}\right)^{m+2}\right)}{n!\left(n-\frac{\rho}{n}\right)^2}=$$

$$=\frac{4^3\cdot 0.012\cdot\left(\frac{4}{3}-4\cdot\left(\frac{4}{3}\right)^4+3\cdot\left(\frac{4}{3}\right)^5\right)}{3!\cdot\left(1-\frac{4}{3}\right)^2}\approx 1{,}56.$$

Среднее время ожидания обслуживания в очереди

$$t=\frac{L}{\lambda}=\frac{1.56}{12}\approx 0{,}13 \text{ ч.}$$

Среднее число заявок в системе

$$M=L+\overline{k}=1{,}56+2{,}78\approx 4{,}34.$$

Среднее время пребывания заявки в системе

$$T=\frac{M}{\lambda}=\frac{4.34}{12}\approx 0{,}36 \text{ ч.}$$

Замкнутые системы массового обслуживания

До сих пор мы рассматривали системы, в которых входящий поток никак не связан с выходящим. Такие системы называются ***разомкнутыми***. В некоторых же случаях обслуженные требования после задержки опять поступают на вход. Такие СМО называются ***замкнутыми***.

Примеры:

- Поликлиника, обслуживающая данную территорию.
- Бригада рабочих, закрепленная за группой станков.

В замкнутых СМО циркулирует одно и то же конечное число потенциальных требований. Пока потенциальное требование не реализовалось в качестве требования на обслуживание, считается, что оно находится в ***блоке задержки***.

В момент реализации оно поступает в саму систему. Например, рабочие обслуживают группу станков. Каждый станок является потенциальным требованием, превращаясь в реальное в момент своей поломки. Пока станок работает, он находится в блоке задержки, а с момента поломки до момента окончания ремонта – в самой системе. Каждый работник является каналом обслуживания.

Пусть n – число каналов обслуживания, s – число потенциальных заявок, $n < s$, λ –интенсивность потока заявок каждого потенциального требования, μ – интенсивность обслуживания, $\rho=\frac{\lambda}{\mu}$.. Поток

- ***Вероятность простоя*** (того, что все обслуживающие аппараты свободны, нет заявок):

$$P_0 = \frac{1}{s!\left(\sum_{k=0}^{n}\frac{\rho^k}{(s-k)!k!} + \sum_{k=n+1}^{s}\frac{\rho^k}{(s-k)!n^{k-n}n!}\right)}. \tag{4.27}$$

- ***Финальные вероятности состояний системы***

$$P_k = \frac{s!\rho^k P_0}{(s-k)!k!}, \qquad k < n,$$
$$P_k = \frac{s!\rho^k P_0}{(s-k)!n^{k-n}n!}, \qquad n \le k \le s. \tag{4.28}$$

Через эти вероятности выражается ***среднее число замкнутых каналов:***

$\bar{k} = P_1 + 2P_2 + \ldots + n(P_n + P_{n+1} + \ldots + P_s)$ или

$$\bar{k} = P_1 + 2P_2 + \ldots + (n-1)P_{n-1} + n(1 - P_0 - P_1 - \ldots - P_{n-1}). \tag{4.29}$$

Через $\bar{k}$ находим ***абсолютную пропускную способность системы***

$$A = \bar{k}\mu, \tag{4.30}$$

а также ***среднее число заявок в системе***

$$M = s - \frac{\mu\bar{k}}{\lambda} = s - \frac{\bar{k}}{\rho}. \tag{4.31}$$

<u>Пример решения задачи.</u>

Рабочий обслуживает 4 станка. Каждый станок отказывает с интенсивностью λ = 0,5 отказа в час. Среднее время ремонта $t_{рем} = \frac{1}{\mu} = 0{,}8$ ч. Определить пропускную способность системы.

Решение

Эта задача рассматривает замкнутую СМО,

$\mu = 1.25,\ \rho = \frac{\lambda}{\mu} = \frac{0{,}5}{1{,}25} = 0{,}4.$

Вероятность простоя рабочего определяется по формуле (4.27):

$$P_0 = \frac{1}{s!\left(\sum_{k=0}^{n}\frac{\rho^k}{(s-k)!k!} + \sum_{k=n+1}^{s}\frac{\rho^k}{(s-k)!n^{k-n}n!}\right)} =$$

$$= \frac{1}{4!\left(\frac{1}{4!} + \frac{0{,}4}{3!} + \frac{0{,}4^2}{2!} + 0{,}4^3 + 0{,}4^4\right)} \approx 0{,}15.$$

Вероятность занятости рабочего

$P_{зан} = 1 - P_0 \cong 0{,}85.$

Если рабочий занят, он налаживает μ станков в единицу времени, пропускная способность системы

$A = (1 - P_0)\mu = 0{,}85\mu \approx 1{,}06$ станков в час.

➢ **Важно помнить.** При применении экономического показателя важно правильно оценить реальные издержки, которые могут изменяться, например, от времени года, от объема запасов угля и пр.

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 значения интенсивностей потоков заявок и обслуживании. Они могут быть вычислены, если вновь обратиться к процессу гибели и размножения.

Тема 6. Прогнозирование парной, множественной и нелинейной регрессией

Парная линейная регрессия

Суть регрессионного анализа

Поведение и значение любого экономического показателя зависят практически от бесконечного количества факторов, и все их учесть нереально. Однако обычно лишь ограниченное количество факторов действительно существенно воздействует на исследуемый экономический показатель. Доля влияния остальных факторов столь незначительна, что их игнорирование не может привести к существенным отклонениям в поведении исследуемого объекта. В естественных науках большей частью имеют дело со строгими (функциональными) зависимостями, при которых каждому значению одной переменной соответствует единственное значение другой. Однако в подавляющем большинстве случаев между экономическими переменными таких зависимостей нет. Поэтому в экономике говорят не о функциональных, а о *корреляционных* либо *стохастических* зависимостях.

Если переменные обозначить X и Y, то зависимость условного математического ожидания переменной Y вида:

$$M(Y|X) = f(x) \qquad (1.1)$$

называется *функцией регрессии Y на X.* При этом X называется независимой (объясняющей) переменной (прогрессором), Y – зависимой (объясняемой) переменной. При рассмотрении двух случайных величин говорят о *парной регрессии*.

Зависимость нескольких переменных, выражаемую функцией

$$M(Y|\ x_1, x_2, \ldots, x_m) = f(x_1, x_2, \ldots, x_m), \qquad (1.2)$$

называют *множественной регрессией*.

Под *регрессией* понимается функциональная зависимость между объясняющими переменными и условным математическим ожиданием (средним значением) зависимой переменной, которая строится с целью предсказания (прогнозирования) этого среднего значения при фиксированных значениях первых.

Для отражения того факта, что реальные значения зависимой переменной не всегда совпадают с ее условными математическими ожиданиями и могут быть различными при одном и том же значении объясняющей переменной (наборе объясняющих переменных), фактическая зависимость должна быть дополнена некоторым слагаемым ε, которое, по существу, является случайной величиной и указывает на стохастическую суть зависимости. Из этого следует, что связи между зависимой и объясняющей (ими) переменными выражаются соотношениями

$$Y = M(Y|X) + \varepsilon; \tag{1.3}$$

$$Y = M(Y|\ x_1, x_2, \ldots, x_m) + \varepsilon, \tag{1.4}$$

называемыми *регрессионными моделями (уравнениями).*

Среди причин обязательного присутствия в регрессионных моделях случайного фактора (отклонения) можно выделить следующие:

1. *Невключение в модель всех объясняющих переменных.* Любая регрессионная (в частности, эконометрическая) модель является упрощением реальной ситуации. Последняя всегда представляет собой сложнейшее переплетение различных факторов, многие из которых в модели не учитываются, что порождает отклонение реальных значений зависимой переменной от ее модельных значений. Например, спрос *(Q)* на товар определяется его ценой *(P),* ценой (P_s) на товары-заменители, ценой (P_c) на дополняющие товары, доходом *(Г)* потребителей, их количеством *(N),* вкусами *(T),* ожиданиями *(W)* и т.д. Безусловно, перечислить все объясняющие переменные здесь практически невозможно. Например, мы не учли такие факторы, как традиции, национальные или религиозные особенности, географическое положение региона, погода и многие другие, влияние которых приведет к некоторым отклонениям реальных наблюдений от модельных, которые можно выразить через случайный член е: $Q = f(P, P_s, P_c, I, N, T, W, \varepsilon)$. Проблема еще и в том, что никогда заранее не известно, какие факторы при создавшихся условиях действительно являются определяющими, а какими можно пренебречь. Например, в ряде случаев учесть непосредственно какой-то фактор нельзя в силу невозможности получения по нему статистических данных.

2. *Неправильный выбор функциональной формы модели.* Из-за слабой изученности исследуемого процесса либо из-за его переменчивости может быть неверно подобрана функция, его моделирующая. Это, безусловно, скажется на отклонении модели от реальности, что отразится на величине случайного члена. Кроме того, неверным может быть подбор объясняющих переменных.

3. *Агрегирование переменных.* Во многих моделях рассматриваются зависимости между факторами, которые сами представляют сложную комбинацию других, более простых переменных. Например, при рассмотрении в качестве зависимой переменной совокупного спроса проводится анализ зависимости, в которой объясняемая переменная является сложной композицией инди-

видуальных спросов, оказывающих на нее определенное влияние помимо факторов, учитываемых в модели. Это может оказаться причиной отклонения реальных значений от модельных.

4. *Ошибки измерений.* Какой бы качественной ни была модель, ошибки измерений переменных отразятся на несоответствии модельных значений эмпирическим данным, что также отразится на величине случайного члена.

5. *Ограниченность статистических данных.* Зачастую строятся модели, выражаемые непрерывными функциями. Но для этого используется набор данных, имеющих дискретную структуру. Это несоответствие находит свое выражение в случайном отклонении.

6. *Непредсказуемость человеческого фактора.* Эта причина может «испортить» самую качественную модель. Действительно, при правильном выборе формы модели, скрупулезном подборе объясняющих переменных все равно невозможно спрогнозировать поведение каждого индивидуума.

Решение задачи построения качественного уравнения регрессии, соответствующего эмпирическим данным и целям исследования, является достаточно сложным и многоступенчатым процессом. Его можно разбить на три этапа:

1) выбор формулы уравнения регрессии;

2) определение параметров выбранного уравнения;

3) анализ качества уравнения и проверка адекватности уравнения эмпирическим данным, совершенствование уравнения.

Выбор формулы связи переменных называется *спецификацией* уравнения регрессии. В случае парной регрессии выбор формулы обычно осуществляется по графическому изображению реальных статистических данных в виде точек в декартовой системе координат, которое называется *корреляционным полем (диаграммой рассеивания)* (рис. 1.1).

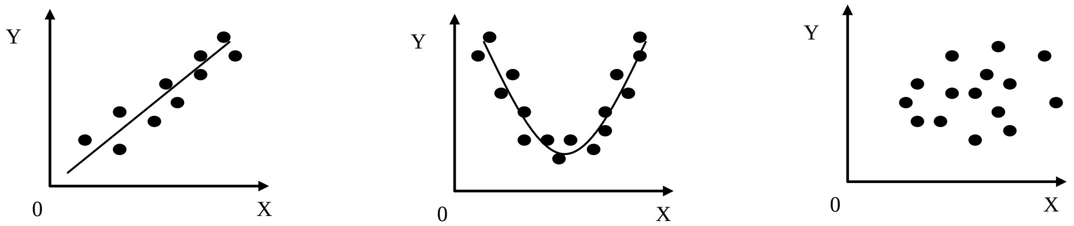


Рис. 1.1. Корреляционное поле

Если на первых двух графиках относительно четко определяется форма связи (для первого – линейная, для второго – квадратичная), то для третьего явная взаимосвязь между переменными отсутствует. В случае множественной регрессии определение подходящего вида регрессии является наиболее сложной задачей.

Парная линейная регрессия

Если функция регрессии линейна, то говорят о *линейной регрессии.* Модель линейной регрессии (линейное уравнение) является наиболее распространенным (и простым) видом зависимости между экономическими переменными. Кроме того, построенное линейное уравнение может служить начальной точкой эконометрического анализа. *Линейная регрессия (теоретическое линейное уравнение регрессии)* представляет собой линейную функцию между условным математическим ожиданием $M(Y|X=x_i)$ зависимой переменной Y и одной объясняющей переменной X (x_i – значения независимой переменной в i-ом наблюдении, $i=1,2,\ldots,n$).

$$M(Y|X=x_i)=\beta_0+\beta_1 x_i. \quad (1.5)$$

Для отражения того факта, что каждое индивидуальное значение y_i отклоняется от соответствующего условного математического ожидания, необходимо ввести в последнее соотношение случайное слагаемое ε_i.

$$y_i=M(Y|X=x_i)+\varepsilon_i=\beta_0+\beta_1 x_i+\varepsilon_i. \quad (1.6)$$

Это соотношение называется *теоретической линейной регрессионной моделью,* β_0 и β_1 *– теоретическими параметрами (теоретическими коэффициентами) регрессии,* ε_i *– случайным отклонением.*

Следовательно, индивидуальные значения y_i представляются в виде суммы двух компонент – систематической $(\beta_0+\beta_1 x_i)$ и случайной ε_i, причина появления которой достаточно подробно рассмотрена ранее. В общем виде теоретическую линейную регрессионную модель будем представлять в виде:

$$Y=\beta_0+\beta_1 X+\varepsilon. \quad (1.7)$$

Для определения значений теоретических коэффициентов регрессии необходимо знать и использовать все значения переменных X и Y генеральной совокупности, что практически невозможно.

Таким образом, задачи линейного регрессионного анализа состоят в том, чтобы по имеющимся статистическим данным (x_i, y_i), $i=1,2,\ldots,n$ для переменных X и *Y:*

а) получить наилучшие оценки неизвестных параметров β_0 и β_1;

б) проверить статистические гипотезы о параметрах модели;

в) проверить, достаточно ли хорошо модель согласуется со статистическими данными (адекватность модели данным наблюдений).

Следовательно, по выборке ограниченного объема мы сможем построить так называемое *эмпирическое уравнение регрессии*

$$\hat{y}_i=b_0+b_1 x_i, \quad (1.8)$$

где $\hat{y}_i$ – оценка условного математического ожидания $M(Y|X=x_i)$; b_0 и b_1– оценки неизвестных параметров β_0 и β_1, называемые *эмпирическими коэффициентами регрессии.* Следовательно, в конкретном случае:

$$y_i=b_0+b_1 x_i+e_i, \quad (1.9)$$

где *отклонение* e_i – оценка теоретического случайного отклонения ε_i.

В силу несовпадения статистической базы для генеральной совокупности и выборки оценки b_0 и b_1 практически всегда отличаются от истинных значений коэффициентов β_0 и β_1, что приводит к несовпадению эмпирической и теоретической линий регрессии. Различные выборки из одной и той же генеральной совокупности обычно приводят к определению отличающихся друг от друга оценок. Возможное соотношение между теоретическим и эмпирическими уравнениями регрессии схематично изображено на рисунке 1.2.

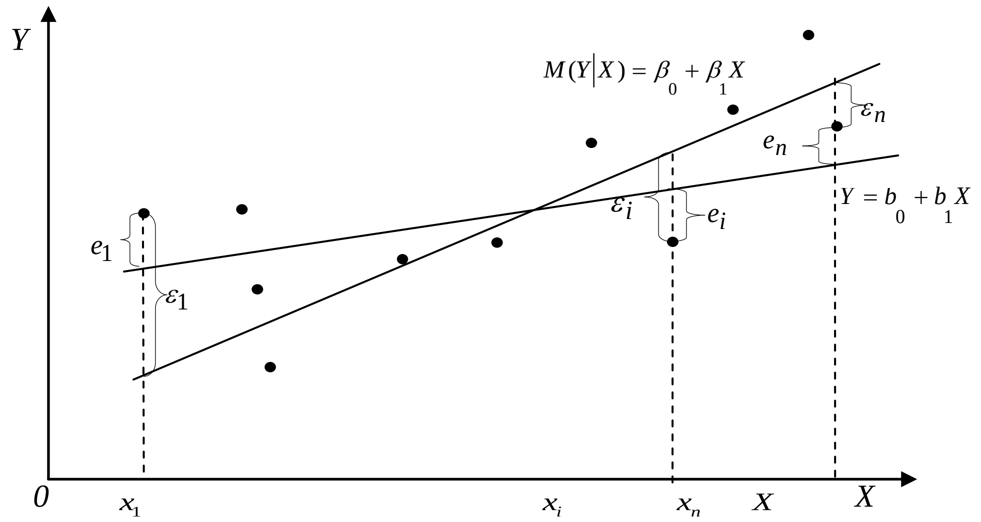


Рис. 1.2. Теоретическое и эмпирическое уравнения регрессии

Задача состоит в том, чтобы по конкретной выборке (x_i, y_i), $i = 1,2,\ldots,n$ найти оценки b_0 и b_1 неизвестных параметров β_0 и β_1 так, чтобы построенная линия регрессии являлась бы наилучшей в определенном смысле среди всех других прямых. Например, коэффициенты b_0 и b_1 эмпирического уравнения регрессии могут быть оценены исходя из условий минимизации одной из следующих сумм:

1. $\sum\limits_{i=1}^{n} e_i = \sum\limits_{i=1}^{n}(y_i - \hat{y}_i) = \sum\limits_{i=1}^{n}(y_i - b_0 - b_1 x_i)$, однако эта сумма не может быть мерой качества найденных оценок в силу того, что существует бесчисленное количество прямых, для которых $\sum\limits_{i=1}^{n} e_i = 0$.

2. $\sum\limits_{i=1}^{n}|e_i| = \sum\limits_{i=1}^{n}\lfloor y_i - \hat{y}_i \rfloor = \sum\limits_{i=1}^{n}|y_i - b_0 - b_1 x_i|$. Этот метод называется методом наименьшей суммы.

3. $\sum\limits_{i=1}^{n} e_i^2 = \sum\limits_{i=1}^{n}(y_i - \hat{y}_i)^2 = \sum\limits_{i=1}^{n}(y_i - b_0 - b_1 x_i)^2$. Это самый распростаренный и теоретически обоснованный метод, который получил название метода наименьших квадратов (МНК). Кроме того, он является наиболее простым с вычислительной точки зрения.

Найдем оценки b_0 и b_1, используя метод наименьших квадратов. При этом минимизируется следующая функция:

$$Q(b_0, b_1) = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = \sum_{i=1}^{n}(y_i - b_0 - b_1 x_i)^2 . \qquad (1.10)$$

Эта функция является квадратичной функцией двух параметров b_0 и b_1. Условием существования минимума функции двух переменных является равенство нулю ее частных производных:

$$\begin{cases} \dfrac{\partial Q}{\partial b_0} = -2\sum(y_i - b_0 - b_1 x_i) = 0; \\ \dfrac{\partial Q}{\partial b_1} = -2\sum(y_i - b_0 - b_1 x_i)x_i = 0; \end{cases} \Rightarrow \begin{cases} nb_0 + b_1 \sum x_i = \sum y_i; \\ b_0 \sum x_i + b_1 \sum x_i^2 = \sum x_i y_i. \end{cases}$$

Разделив оба уравнения системы на n, получим:

$$\begin{cases} b_0 + b_1 \bar{x} = \bar{y}; \\ b_0 \bar{x} + b_1 \overline{x^2} = \overline{xy}; \end{cases} \Rightarrow \begin{cases} b_1 = \dfrac{\overline{xy} - \bar{x} \cdot \bar{y}}{\overline{x^2} - \bar{x}^2}; \\ b_0 = \bar{y} - b_1 \bar{x}, \end{cases},$$

где
$$\begin{aligned} \bar{x} &= \frac{1}{n}\sum x_i, \quad \overline{x^2} = \frac{1}{n}\sum x_i^2; \\ \bar{y} &= \frac{1}{n}\sum y_i. \quad \overline{xy} = \frac{1}{n}\sum x_i \cdot y_i \end{aligned} \qquad (1.11)$$

Из формул статистики очевидно, что:

$$b_1 = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2} = \frac{S_{xy}}{S_x^2}.$$

Тогда
$$b_1 = \frac{S_{xy}}{S_x^2} = \frac{S_{xy}}{S_x S_y} \cdot \frac{S_y}{S_x} = r_{xy} \cdot \frac{S_y}{S_x}, \qquad (1.12)$$

где r_{xy} – выборочный коэффициент корреляции, S_x, S_y – стандартные отклонения.

Параметр b_1 называется коэффициентом регрессии, он показывает, на сколько в среднем в абсолютном выражении изменится значение результативного признака y при изменение факторного признака x на единицу. Знак при коэффициенте регрессии показывает направление связи: при b_1>0 – связь прямая, а при b_1<0 – связь обратная. Параметр b_0 формально показывает значение y при $x = 0$. Если факторный признак не имеет и не может иметь нулевого значения, то трактовка свободного члена не имеет смысла, то есть он может не иметь экономического содержания.

Проверка качества уравнения регрессии

Регрессионный анализ позволяет определить оценки коэффициентов регрессии. Но они не позволяют сделать вывод, насколько точно эмпирическое уравнение регрессии соответствует уравнению для всей генеральной совокупности, насколько близки оценки b_0 и b_1 коэффициентов к своим теоретическим прототипам β_0 и β_1, как близко оцененное значение $\hat{y}_i$ к условному математическому ожиданию $M(Y|X = x_i)$, насколько надежны найденные оценки. Для ответа на эти вопросы необходимы дополнительные исследования.

Как следует из соотношения (1.6), значения y_i зависят от значений x_i и случайных отклонений ε_i. Следовательно, переменная Y является случайной величиной, напрямую связанной с ε_i. Таким образом, пока не будет определенности в вероятностном поведении ε_i, нет уверенности в качестве оценок.

Доказано, что для получения по МНК наилучших результатов необходимо, чтобы выполнялся ряд предпосылок относительно случайного отклонения (предпосылки Гаусса-Маркова):

1. Математическое ожидание случайного отклонения ε_i равно нулю: $M(\varepsilon_i) = 0$ для всех наблюдений. Данное условие означает, что случайное отклонение в среднем не оказывает влияния на зависимую переменную. В каждом конкретном наблюдении может быть либо положительным, либо отрицательным, но он не должен иметь систематического смещения. Выполнимость $M(\varepsilon_i) = 0$ влечет выполнимость $M(Y|X = x_i) = \beta_0 + \beta_1 x_i$.

2. Дисперсия случайных отклонений ε_i постоянна: $D(\varepsilon_i) = D(\varepsilon_j) = \sigma^2$ для любых наблюдений i и j. Данное условие подразумевает, что, несмотря на то, что при каждом конкретном наблюдении случайное отклонение может быть либо большим, либо меньшим, не должно быть некой причины, вызывающей большую ошибку (отклонение).

3. Случайные отклонения ε_i и ε_j являются независимыми друг от друга для $i \neq j$. Выполнимость данной предпосылки предполагает, что отсутствует систематическая связь между любыми случайными отклонениями.

4. Случайное отклонение должно быть независимо от объясняющих переменных.

5. Модель является линейной относительно параметров.

Теорема Гаусса-Маркова. Если предпосылки 1-5 выполнены, то оценки, полученные по МНК, обладают следующими свойствами:

1. Оценки являются несмещенными, т.е. $M(b_0) = \beta_0$, $M(b_1) = \beta_1$. Это вытекает из того, что $M(\varepsilon_i) = 0$, и говорит об отсутствии систематической ошибки в определении положения линии регрессии.

2. Оценки состоятельны, так как дисперсия оценок параметров при возрастании числа наблюдений n стремится к нулю: $D(b_0) \xrightarrow[n \to \infty]{} 0$, $D(b_1) \xrightarrow[n \to \infty]{} 0$. Другими словами, при увеличении объема выборки надежность оценок увеличивается (b_0 близко к β_0, а b_1 близко к β_1).

3. Оценки эффективны, т.е. они имеют наименьшую дисперсию по сравнению с любыми другими оценками данных параметров, линейными относительно величин y_i.

Если предпосылки 2 и 3 нарушены, то свойства несмещенности и состоятельности сохраняются, но свойство эффективности – нет.

Анализ точности определения оценок коэффициентов регрессии

В силу случайного отбора элементов в выборку случайными являются также оценки b_0 и b_1 коэффициентов β_0 и β_1 теоретического уравнения регрессии. Их математические ожидания при выполнении предпосылок об отклонениях ε_i равны соответственно $M(b_0)=\beta_0$, $M(b_1)=\beta_1$. При этом оценки тем надежнее, чем меньше их разброс вокруг β_0 и β_1, т.е. чем меньше дисперсии $D(b_0)$ и $D(b_1)$ оценок, рассчитываемые по формулам:

$$D(b_0)=\frac{\sigma^2\sum x_i^2}{n\sum(x_i-\bar{x})^2},\ D(b_1)=\frac{\sigma^2}{\sum(x_i-\bar{x})^2}. \tag{1.13}$$

Из соотношений (1.13) очевидны следующие выводы:

- Дисперсии b_0 и b_1 прямо пропорциональны дисперсии случайного отклонения σ^2. Следовательно, чем больше фактор случайности, тем менее точными будут оценки.
- Чем больше число наблюдений n, тем меньше дисперсия оценок. Это очевидно, так как чем большим числом данных мы располагаем, тем вероятнее получение более точных оценок.
- Чем больше дисперсия объясняющей переменной, тем меньше дисперсия оценок коэффициентов. Другими словами, чем шире область изменений объясняющей переменной, тем точнее будут оценки.

В силу того, что случайные отклонения ε_i по выборке определены быть не могут, при анализе надежности оценок коэффициентов регрессии они заменяются отклонениями $e_i=y_i-b_0-b_1x_i$ значений y_i переменной Y от оцененной линии регрессии. Дисперсия случайных отклонений $D(\varepsilon_i)=\sigma^2$ заменяется ее несмещенной оценкой

$$S^2=\frac{1}{n-2}\sum(y_i-b_0-b_1x_i)^2=\frac{\sum e_i^2}{n-2}. \tag{1.14}$$

Тогда

$$D(b_0)\approx S_{b0}^2=\frac{S^2\sum x_i^2}{n\sum(x_i-\bar{x})^2}=\overline{x^2}S_{b1}^2. \tag{1.15}$$

$$D(b_1)\approx S_{b1}^2=\frac{S^2}{\sum(x_i-\bar{x})^2}. \tag{1.16}$$

$S^2=\dfrac{\sum e_i^2}{n-2}$ – необъясненная дисперсия (мера разброса зависимой переменной вокруг линии регрессии).

$S=\sqrt{\dfrac{\sum e_i^2}{n-2}}$ – стандартная ошибка оценки (стандартная ошибка регрессии).

$S_{b0}=\sqrt{S_{b0}^2}$ и $S_{b1}=\sqrt{S_{b1}^2}$ – стандартные отклонения случайных величин

b_0 и b_1, называемые стандартными ошибками коэффициентов регрессии.

Проверка гипотез относительно коэффициентов линейного уравнения регрессии

При проведении статистического анализа перед исследователем зачастую возникает необходимость сравнения эмпирических коэффициентов регрессии b_0 и b_1 с некоторыми теоретически ожидаемыми значениями $\tilde{\beta}_0$ и $\tilde{\beta}_1$ этих коэффициентов. Данный анализ осуществляется при помощи статистической проверки гипотез.

Для проверки гипотезы $b_1 = \tilde{\beta}_1$ используется статистика:

$$t = \frac{b_1 - \tilde{\beta}_1}{S_{b1}}, \tag{1.17}$$

которая при справедливости $b_1 = \tilde{\beta}_1$ имеет распределение Стьюдента с числом степеней свободы $\nu = n - 2$, где n – объем выборки. Следовательно, гипотеза $b_1 = \tilde{\beta}_1$ отклоняется на основании данного критерия, если:

$$|T_{набл}| = \left|\frac{b_1 - \tilde{\beta}_1}{S_{b1}}\right| \geq t_{\frac{\alpha}{2}, n-2}, \tag{1.18}$$

где α – требуемый уровень значимости. При невыполнении (1.18) считается, что нет оснований для отклонения $b_1 = \tilde{\beta}_1$.

Наиболее важной на начальном этапе статистического анализа построенной модели является задача установления значимости коэффициентов. Эта задача решается при помощи отношений, называемых t-статистикой:

$$t = \frac{b_1}{S_{b1}} \quad \text{и} \quad t = \frac{b_0}{S_{b0}}. \tag{1.19}$$

В случае, если $|t| > t_{\frac{\alpha}{2}, n-2}$, то статистическая значимость соответствующего коэффициента регрессии подтверждается. Значения $t_{\frac{\alpha}{2}, n-2}$ в зависимости от уровня значимости и числа степеней свободы ν ($\nu = n - 2$) приведены в приложении 1.

При оценке значимости коэффициента линейной регрессии на начальном этапе можно использовать следующее «грубое» правило (почти всегда работает при $n > 10$), позволяющее не прибегать к таблицам:

- Если $|t| \leq 1$, то коэффициент не может быть признан значимым, так как доверительная вероятность составит менее, чем 0,7.
- Если $1 < |t| \leq 2$, то найденная оценка может рассматриваться как относительно (слабо) значимая. Доверительная вероятность в этом случае лежит между 0,7 и 0,95.
- Если $2 < |t| \leq 3$, то доверительная вероятность составляет от 0,95 до 0,99.

- Если $|t| > 3$, то это почти гарантия значимости коэффициента.

Интервальные оценки коэффициентов
линейного уравнения регрессии

Доверительные интервалы коэффициентов b_0 и b_1, которые с надежностью $(1-\alpha)$ накрывают определяемые параметры β_0 и β_1, определяются по формулам:

$$\left(b_0 - t_{\frac{\alpha}{2},n-2} S(b_0);\ b_0 + t_{\frac{\alpha}{2},n-2} S(b_0)\right) \text{ и } \left(b_1 - t_{\frac{\alpha}{2},n-2} S(b_1);\ b_1 + t_{\frac{\alpha}{2},n-2} S(b_1)\right). \quad (1.20)$$

Таким образом с вероятностью $(1-\alpha)$ можно утверждать, что истинное значение параметров β_0 и β_1 попадают в найденный интервал.

Доверительные интервалы для зависимой переменной

Интервал, определяющий границы, за пределами которых могут оказаться не более $100 \cdot \alpha\%$ точек наблюдений при $X = x_p$, рассчитывается следующим образом:

$$\left(b_0 + b_1 x_p \pm t_{\frac{\alpha}{2},n-2} S \sqrt{1 + \frac{1}{n} + \frac{(\bar{x} - x_p)^2}{\sum (x_i - \bar{x})^2}}\right). \quad (1.21)$$

При прогнозировании на основе уравнения регрессии следует помнить, что величина прогноза зависит не только от стандартной ошибки индивидуального значения *y*, но и от точности прогноза значения *x*. Его величина может задаваться на основе анализа других моделей исходя из конкретной ситуации, а также анализа динамики данного фактора.

Проверка общего качества уравнения регрессии

Общее качество уравнения регрессии оценивается по тому, как хорошо эмпирическое уравнение регрессии согласуется со статистическими данными. Иными словами, насколько широко рассеяны точки наблюдений относительно линии регрессии. Очевидно, что если все точки лежат на построенной прямой, то регрессия Y на X «идеально» объясняет поведение зависимой переменной.

Суммарно мерой общего качества уравнения регрессии является *коэффициент детерминации* R^2. В случае парной регрессии коэффициент детерминации будет совпадать с квадратом коэффициента корреляции. В общем случае коэффициент детерминации рассчитывается по формуле:

$$R^2 = 1 - \frac{\sum e_i^2}{\sum (y_i - \bar{y})^2}. \quad (1.22)$$

Коэффициент детерминации характеризует долю дисперсии результативного признака *y*, объясняемую регрессией, в общей дисперсии результативного признака. Справедливо соотношение $0 \le R^2 \le 1$. Чем теснее линейная связь между X и Y, тем ближе коэффициент детерминации R^2 к единице. Для парной регрессии коэффициент детерминации равен квадрату коэффициента корреляции, то есть $R^2 = r^2$.

Рекомендуется после проверки общего качества уравнения регрессии провести анализ его статистической значимости. Для этого используется F-статистика:

$$F = \frac{R^2}{1-R^2} \cdot \frac{n-2}{1}. \qquad (1.23)$$

Из формулы (1.23) очевидно, что показатели F и R^2 равны или не равны нулю одновременно. Если $F = 0$, то $R^2 = 0$, следовательно, величина y линейно не зависит от x.

Расчетное значение F сравнивается с критическим $F_{кр}$. $F_{кр}$, исходя из требуемого уровня значимости α и числа степеней свободы $v_1 = 1$ и $v_2 = n-2$, определяется на основе распределения Фишера (приложение 2). Если $F > F_{кр}$, то коэффициент детерминации R^2 и уравнение регрессии статистически значимы.

Пример. Для анализа зависимости объема потребления Y (у.е.) домохозяйства от располагаемого дохода X (у.е.) отобрана выборка n=12 (помесячно в течение года). Необходимо провести регрессионный анализ. Данные и расчеты представлены в таблице

Исходные данные и расчеты

i	x_i	y_i	x_i^2	$x_i y_i$	y_i^2	$\hat{y}_i$	e_i	e_i^2
1	107	102	11449	10914	10104	103,63	-1,63	2,66
2	109	105	11881	11445	11025	105,49	-0,49	0,24
3	110	108	12100	11880	11664	106,43	1,57	2,46
4	113	110	12769	12430	12100	109,23	0,77	0,59
5	120	115	14400	13800	13225	115,77	-0,77	0,59
6	122	117	14884	14274	13689	117,63	-0,63	0,40
7	123	119	15129	14637	14161	118,57	0,43	0,18
8	128	125	16384	16000	15625	123,24	1,76	3,10
9	136	132	18496	17952	17424	130,71	1,29	1,66
10	140	130	19600	18200	16900	134,45	-4,45	19,8
11	145	141	21025	20445	19881	139,11	1,89	3,57
12	150	144	22500	21600	20736	143,78	0,22	0,05
Сумма	1503	1448	190617	183577	176834	-	0	35,3
Среднее	125,25	120,67	15884,75	15298,08	14736,17	-	-	-

Согласно $$\begin{cases} b_1 = \dfrac{\overline{xy} - \bar{x} \cdot \bar{y}}{\overline{x^2} - \bar{x}^2} = \dfrac{15298{,}08 - 125{,}25 \cdot 120{,}67}{15884{,}75 - (125{,}25)^2} = 0{,}9339; \\ b_0 = \bar{y} - b_1 \bar{x} = 120{,}67 - 0{,}9339 \cdot 125{,}25 = 3{,}699. \end{cases}$$

Таким образом, уравнение парной регрессии имеет вид:

$\hat{Y} = 3{,}699 + 0{,}9339X$. По этому уравнению рассчитаем $\hat{y}_i$, а также $e_i = y_i - \hat{y}_i$.

В данном примере коэффициент b_1 может трактоваться как предельная склонность к потреблению. Фактически он показывает, на какую величину изменится объем потребления, если располагаемый доход возрастает на одну единицу. Свободный член b_0 определяет прогнозируемое значение Y при величине располагаемого дохода X, равной нулю (т.е. автономное потребление).

Рассчитаем другие показатели.

$$S^2 = \frac{\sum e_i^2}{n-2} = \frac{35{,}3}{12-2} = 3{,}53; \quad S = \sqrt{S^2} = 1{,}88.$$

$$S_{b1}^2 = \frac{S^2}{\sum (x_i - \bar{x})^2} = \frac{3{,}53}{2366.25} = 0{,}0015; \quad S_{b1} = \sqrt{S_{b1}^2} = 0{,}039.$$

$$S_{b0}^2 = \overline{x^2} S_{b1}^2 = 15884{,}75 \cdot 0{,}0015 = 23{,}83; \quad S_{b0} = \sqrt{S_{b0}^2} = 4{,}88.$$

Проверим статистическую значимость коэффициентов b_0 и b_1:

$$t_1 = \frac{b_1}{S_{b1}} = \frac{0{,}9339}{0{,}039} = 23{,}946 \quad \text{и} \quad t_0 = \frac{b_0}{S_{b0}} = \frac{3{,}699}{4{,}88} = 0{,}76.$$

Критическое значение при уровне значимости $\alpha = 0{,}05$ равно (см. приложение 1) $t_{крит} = t_{\frac{\alpha}{2}, n-2} = t_{0,025;10} = 2{,}228$.

Так как $|t_1| = 23{,}946 > 2{,}228$, то это подтверждает статистическую значимость коэффициента регрессии b_1. Аналогично для другого коэффициента: так как $|t_0| = 0{,}76 < 2{,}228$, то гипотеза о статистической значимости коэффициента b_0 отклоняется. Это означает, что в данном случае свободным членом уравнения регрессии можно пренебречь, рассматривая регрессию как $Y = b_1 X$.

Определим доверительные интервалы коэффициентов регрессии (по формуле 1.20), которые с надежность 95% $(\alpha = 0{,}05)$ будут следующими:

для b_0: $(3{,}699 - 2{,}228 \cdot 4{,}88;\ 3{,}699 + 2{,}228 \cdot 4{,}88) = (-7{,}173;\ 14{,}572)$;

для b_1: $(0{,}9339 - 2{,}228 \cdot 0{,}039;\ 0{,}9339 + 2{,}228 \cdot 0{,}039) = (0{,}8470;\ 1{,}021)$.

Рассчитаем границы интервала, в котором будет сосредоточено 95% возможных объемов потребления при неограниченно большом числе наблюдений и уровне дохода $X = 160$ (формула 1.21):

$$3{,}699 + 0{,}9339 \cdot 160 \pm 2{,}228 \cdot 1{,}88 \cdot \sqrt{1 + \frac{1}{12} + \frac{(125{,}25 - 160)^2}{2102{,}1875}}.$$

Таким образом, этот интервал имеет вид: (147,4898; 158,7082).

Рассчитаем коэффициент детерминации (1.22):

$R^2 = 1 - \dfrac{35{,}3}{2108{,}6668} = 0{,}983$. Столь высокое значение коэффициента детерминации свидетельствует о высоком общем качестве построенного уравнения регрессии.

Множественная линейная регрессия

Определение параметров уравнения регрессии

На любой экономический показатель чаще всего оказывает влияние не один, а несколько факторов. В этом случае вместо парной регрессии $M(Y|X) = f(x)$ рассматривается *множественная регрессия*

$$M(Y|x_1, x_2, \ldots, x_m) = f(x_1, x_2, \ldots, x_m). \tag{2.1}$$

Задача оценки статистической взаимосвязи переменных Y и $X_1, X_2, \ldots, X_m$ формулируется аналогично случаю парной регрессии. Уравнение множественной регрессии может быть представлено в виде:

$$Y = f(\beta, X) + \varepsilon, \tag{2.2}$$

где $X = (X_1, X_2, \ldots, X_m)$ – вектор независимых (объясняющих) переменных; β – вектор параметров (подлежащих определению); ε – случайная ошибка (отклонение); Y – зависимая (объясняемая) переменная.

Рассмотрим самую употребляемую и наиболее простую из моделей множественной регрессии – модель множественной линейной регрессии.

Теоретическое линейное уравнение регрессии имеет вид:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_m X_m + \varepsilon \tag{2.3}$$

или для индивидуальных наблюдений $i = 1,2,\ldots,n$:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_m x_{im} + \varepsilon_i. \tag{2.4}$$

Здесь $\beta = (\beta_0, \beta_1, \ldots, \beta_m)$ – вектор размерности $(m+1)$ неизвестных параметров. β_j, $j = 1,2,\ldots,m$ называется j-тым теоретическим коэффициентом регрессии (частичным коэффициентом регрессии). Он характеризует чувствительность величины Y к изменению величины X_j, т.е. отражает влияние на условное математическое ожидание $M(Y|x_1, x_2, \ldots, x_m)$ зависимой переменной Y объясняющей переменной X_j при условии, что все другие объясняющие переменные модели остаются постоянными. β_0 – свободный член, определяющий Y в случае, когда все объясняющие переменные X_j равны нулю.

После выбора линейной функции в качестве модели зависимости необходимо оценить параметры регрессии. Пусть имеется n наблюдений вектора объясняющих переменных $X = (X_1, X_2, \ldots, X_m)$ и зависимой переменной Y:

$$(x_{i1}, x_{i2}, \ldots, x_{im}, y_i), \quad i = 1,2,\ldots,n.$$

Для того чтобы однозначно можно было решить задачу нахождения параметров $\beta_0, \beta_1, \ldots, \beta_m$ (т.е. найти некоторый наилучший вектор β), должно выполняться неравенство $n \geq m+1$. Если это неравенство не будет выполняться, то существует бесконечно много различных векторов параметров, при которых линейная формула связи между X и Y будет абсолютно точно соответствовать имеющимся наблюдениям.

Например, для однозначного определения оценок параметров уравнения регрессии $Y=\beta_0+\beta_1X_1+\beta_2X_2$ достаточно иметь выборку из трех наблюдений $(x_{i1},x_{i2},x_{i3},y_i), i=1,2,3$. В этом случае найденные значения параметров β_0,β_1,β_2 определяют такую плоскость $Y=\beta_0+\beta_1X_1+\beta_2X_2$ в трехмерном пространстве, которая пройдет именно через заданные три точки. С другой стороны, добавление в выборку к имеющимся трем наблюдениям еще одного приведет к тому, что четвертая точка $(x_{41},x_{42},x_{43},y_4)$ практически наверняка будет лежать вне построенной плоскости, что потребует определенной переоценки параметров.

Число $v=n-m-1$ называется *числом степеней свободы*. Если число степеней свободы невелико, то статистическая надежность оцениваемой формулы невысока. Например, вероятность верного вывода (получения более точных оценок) по трем наблюдениям существенно ниже, чем по тридцати. Считается, что при оценивании множественной линейной регрессии для обеспечения статистической надежности требуется, чтобы число наблюдений, по крайней мере, в три раза превосходило число оцениваемых параметров.

Самым распространенным методом оценки параметров уравнения множественной регрессии является метод наименьших квадратов (МНК).

Предпосылки МНК:

1. Математическое ожидание случайного отклонения ε_i равно нулю для всех наблюдений: $M(\varepsilon_i)=0, i=1,2,\ldots,n$.
2. Постоянство дисперсии отклонений (гомоскедастичность): $D(\varepsilon_i)=D(\varepsilon_j)=\sigma^2$ для любых наблюдений *i* и *j*.
3. Отсутствие автокорреляции: случайные отклонения ε_i и ε_j являются независимыми друг от друга для $i\neq j$.
4. Случайное отклонение должно быть независимо от объясняющих переменных.
5. Модель является линейной относительно параметров.
6. Отсутствие мультиколлинеарности: между объясняющими переменными отсутствует строгая (сильная) линейная зависимость.
7. Ошибки $\varepsilon_i, i=1,2,\ldots,n$ имеют нормальное распределение. Выполнимость данной предпосылки важна для проверки статистических гипотез и построения интервальных оценок.

Как и в случае парной регрессии, истинные значения параметров β_j по выборке получить невозможно. Находим теоретическое уравнение регрессии:

$$Y=b_0+b_1X_1+b_2X_2+\ldots+b_mX_m+e. \tag{2.5}$$

Здесь $b_0,b_1,\ldots,b_m$ – оценки $\beta_0,\beta_1,\ldots,\beta_m$ коэффициентов регрессии (эмпирические коэффициенты регрессии); e – оценка отклонения ε. Для индивидуальных наблюдений имеем:

$$y_i=b_0+b_1x_{i1}+b_2x_{i2}+\ldots+b_mx_{im}+e_i. \tag{2.6}$$

При выполнении предпосылок МНК относительно ошибок ε_i оценки $b_0, b_1, \ldots, b_m$ параметров $\beta_0, \beta_1, \ldots, \beta_m$ множественной линейной регрессии по МНК являются несмещенными, эффективными и состоятельными.

На основании (2.6):

$$e_i = y_i - b_0 - b_1 x_{i1} - \ldots - b_m x_{im}. \tag{2.7}$$

Тогда по методу наименьших квадратов для нахождения оценок $b_0, b_1, \ldots, b_m$ минимизируется следующая функция:

$$Q = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_i - (b_0 + \sum_{j=1}^{m} b_j x_{ij}))^2 . \tag{2.8}$$

Необходимым условием минимизации функции Q является равенство нулю всех ее частных производных по b_j, т.е.:

$$\begin{cases} \dfrac{\partial Q}{\partial b_0} = -2\sum_{i=1}^{n}(y_i - (b_0 + \sum_{j=1}^{m} b_j x_{ij})); \\ \dfrac{\partial Q}{\partial b_j} = -2\sum_{i=1}^{n}(y_i - (b_0 + \sum_{j=1}^{m} b_j x_{ij}))x_{ij}, \ j = 1,2,\ldots,m. \end{cases} \tag{2.9}$$

Приравнивая их к нулю, получаем систему $m+1$ линейных уравнений с $m+1$ неизвестными. Такая система обычно имеет единственное решение и называется системой нормальных уравнений. Ее решение в явном виде наиболее наглядно представимо в векторно-матричной форме.

Расчет коэффициентов множественной линейной регрессии

Данные наблюдений и соответствующие коэффициенты в матричной форме выглядят следующим образом:

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ \ldots \\ y_n \end{pmatrix}, \quad X = \begin{pmatrix} 1 & x_{11} & x_{12} & \ldots & x_{1m} \\ 1 & x_{21} & x_{22} & \ldots & x_{2m} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 & x_{n1} & x_{n2} & \ldots & x_{nm} \end{pmatrix}, \quad B = \begin{pmatrix} b_0 \\ b_1 \\ \ldots \\ b_m \end{pmatrix}, \quad e = \begin{pmatrix} e_1 \\ e_2 \\ \ldots \\ e_n \end{pmatrix}.$$

Здесь $Y - n$-мерный вектор-столбец наблюдений зависимой переменной X; X – матрица размерности $n \times (m+1)$, в которой i-тая строка $(i = 1,2,\ldots,n)$ представляет наблюдение вектора значений независимых переменных $X_1, X_2, \ldots, X_m$; единица соответствует переменной при свободном члене b_0; B – вектор-столбец размерности $(m+1)$ параметров уравнения регрессии; e – вектор-столбец размерности n отклонений выборочных (реальных) значений y_i зависимой переменной Y от значений $\hat{y}_i$, получаемых по уравнению регрессии

$$\hat{y}_i = b_0 + b_1 x_{i1} + b_2 x_{i2} + \ldots + b_m x_{im}. \qquad (2.10)$$

Функция $Q = \sum_{i=1}^{n} e_i^2$ в матричной форме представима как произведение вектор-строки $e^T = (e_1, e_2, \ldots, e_n)$ на вектор-столбец e. Вектор-столбец e может быть в свою очередь представлен в следующем виде:

$$e = Y - XB. \qquad (2.11)$$

Отсюда:

$$Q = e^T e = (Y - XB)^T (Y - XB) = Y^T Y - B^T X^T Y - Y^T XB + B^T X^T XB =$$

$$= Y^T Y - 2B^T X^T Y + B^T X^T XB.$$

Здесь e^T, B^T, X^T, Y^T – векторы и матрицы, транспонированные к e, B, X, Y соответственно. При выводе формулы использовались следующие известные соотношения линейной алгебры:

$(Y - XB)^T = Y^T - (XB)^T$; $(XB)^T = B^T X^T$; $B^T X^T Y = Y^T XB$. Необходимым условием экстремума функции Q является равенство нулю ее частных производных $\frac{\partial Q}{\partial b_j}$ по всем параметрам b_j, $j = 0, 1, \ldots, m$. Вектор-столбец $\frac{\partial Q}{\partial B}$ частных производных в матричном виде выглядит следующим образом:

$$\frac{\partial Q}{\partial B} = -2X^T Y + 2(X^T X)B. \qquad (2.12)$$

Приравняв $\frac{\partial Q}{\partial B}$ к нулю, получаем:

$$-2X^T Y + 2(X^T X)B = 0 \Rightarrow$$

$$X^T Y = (X^T X)B \Rightarrow \qquad (2.13)$$

$$B = (X^T X)^{-1} X^T Y. \qquad (2.14)$$

Здесь $(X^T X)^{-1}$ – матрица, обратная к $\left(X^T X\right)$.

Полученные общие соотношения справедливы для уравнений регрессии с произвольным количеством m объясняющих переменных. Проанализируем полученные результаты для случаев:

1. $m = 1$. Для парной регрессии $Y = b_0 + b_1 X + e$ имеем:

$$Y=\begin{pmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{pmatrix}, \quad X=\begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \dots & \dots \\ 1 & x_n \end{pmatrix}, \quad B=\begin{pmatrix} b_0 \\ b_1 \end{pmatrix}, \quad e=\begin{pmatrix} e_1 \\ e_2 \\ \dots \\ e_n \end{pmatrix}.$$

Из (2.11) следует: $e=Y-XB$, т.е.

$$\begin{pmatrix} e_1 \\ e_2 \\ \dots \\ e_n \end{pmatrix}=\begin{pmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{pmatrix}-\begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \dots & \dots \\ 1 & x_n \end{pmatrix}\cdot\begin{pmatrix} b_0 \\ b_1 \end{pmatrix}.$$

$$Z=(X^T X)^{-1}=\begin{pmatrix} \dfrac{\sum x_i^2}{n\sum x_i^2-(\sum x_i)^2} & -\dfrac{\sum x_i}{n\sum x_i^2-(\sum x_i)^2} \\ -\dfrac{\sum x_i}{n\sum x_i^2-(\sum x_i)^2} & \dfrac{n}{n\sum x_i^2-(\sum x_i)^2} \end{pmatrix}.$$

$$X^T Y=\begin{pmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \end{pmatrix}\cdot\begin{pmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{pmatrix}=\begin{pmatrix} \sum y_i \\ \sum x_i y_i \end{pmatrix}.$$

Из (2.14) следует $B=(X^T X)^{-1}X^T Y=$

$$=\begin{pmatrix} \dfrac{\sum x_i^2\sum y_i}{n\sum x_i^2-(\sum x_i)^2}-\dfrac{\sum x_i\sum x_i y_i}{n\sum x_i^2-(\sum x_i)^2} \\ -\dfrac{\sum x_i\sum y_i}{n\sum x_i^2-(\sum x_i)^2}+\dfrac{n\sum x_i y_i}{n\sum x_i^2-(\sum x_i)^2} \end{pmatrix}=$$

$$=\begin{pmatrix} \dfrac{\sum x_i^2\sum y_i-\sum x_i\sum x_i y_i}{n\sum x_i^2-(\sum x_i)^2} \\ \dfrac{n\sum x_i y_i-\sum x_i\sum y_i}{n\sum x_i^2-(\sum x_i)^2} \end{pmatrix}=\begin{pmatrix} \bar{y}-b_1\bar{x} \\ b_1 \end{pmatrix}=\begin{pmatrix} b_0 \\ b_1 \end{pmatrix}. \qquad (2.15)$$

2. $m=2$. Рассуждая аналогично, можно вывести формулы определения коэффициентов регрессии для уравнения с двумя объясняющими переменными. Соотношение (2.13) в этом случае имеет вид системы трех линейных уравнений с тремя неизвестными b_0, b_1, b_2:

$$\begin{cases} \sum y_i = nb_0 + b_1\sum x_{i1} + b_2\sum x_{i2}; \\ \sum x_{i1}y_i = b_0\sum x_{i1} + b_1\sum x_{i1}^2 + b_2\sum x_{i1}x_{i2}; \\ \sum x_{i2}y_i = b_0\sum x_{i2} + b_1\sum x_{i1}x_{i2} + b_2\sum x_{i2}^2. \end{cases} \quad (2.16)$$

Решение данной системы имеет вид:

$$b_0 = \bar{y} - b_1\bar{x}_1 - b_2\bar{x}_2;$$

$$b_1 = \frac{\sum(x_{i1}-\bar{x}_1)(y_i-\bar{y})\cdot\sum(x_{i2}-\bar{x}_2)^2 - \sum(x_{i2}-\bar{x}_2)(y_i-\bar{y})\cdot\sum(x_{i1}-\bar{x}_1)(x_{i2}-\bar{x}_2)}{\sum(x_{i1}-\bar{x}_1)^2\sum(x_{i2}-\bar{x}_2)^2 - (\sum(x_{i1}-\bar{x}_1)(x_{i2}-\bar{x}_2))^2}; \quad (2.17)$$

$$b_2 = \frac{\sum(x_{i2}-\bar{x}_2)(y_i-\bar{y})\cdot\sum(x_{i1}-\bar{x}_1)^2 - \sum(x_{i1}-\bar{x}_1)(y_i-\bar{y})\cdot\sum(x_{i1}-\bar{x}_1)(x_{i2}-\bar{x}_2)}{\sum(x_{i1}-\bar{x}_1)^2\sum(x_{i2}-\bar{x}_2)^2 - (\sum(x_{i1}-\bar{x}_1)(x_{i2}-\bar{x}_2))^2}.$$

Анализ качества эмпирического уравнения множественной линейной регрессии

Построение эмпирического уравнения регрессии является начальным этапом эконометрического анализа. Первое же построенное по выборке уравнение регрессии очень редко является удовлетворительным по тем или иным характеристикам. Поэтому следующей важнейшей оценкой является проверка качества уравнения регрессии. В эконометрике принята устоявшаяся схема такой проверки, которая проводится по следующим направлениям:

- проверка статистической значимости коэффициентов уравнения регрессии;
- проверка общего качества уравнения регрессии;
- проверка свойств данных, выполнимость которых предполагалась при оценивании уравнения (проверка выполнимости предпосылок МНК).

Прежде чем проводить анализ качества уравнения регрессии, необходимо определить дисперсии и стандартные ошибки коэффициентов, а также интервальные оценки коэффициентов.

Выборочные дисперсии эмпирических коэффициентов регрессии можно определить следующим образом:

$$S_{bj}^2 = S^2 z'_{jj} = \frac{\sum e_i^2}{n-m-1} z'_{jj}, \; j = 1,2,\ldots,m. \quad (2.18)$$

Здесь z'_{jj} – j-тый диагональный элемент матрицы $Z^{-1} = (X^T X)^{-1}$.

При этом:

$$S^2 = \frac{\sum e_i^2}{n-m-1}, \quad (2.19)$$

где m – количество объясняющих переменных модели. Иногда в формуле (2.19) знаменатель представляют в виде $n-m-1=n-k$, подразумевая под k число параметров модели (подлежащих определению коэффициентов регрессии).

В частности, для уравнения $\hat{Y} = b_0 + b_1 X_1 + b_2 X_2$ с двумя объясняющими переменными используются следующие формулы:

$$S_{b0}^2 = \left[\frac{1}{n} + \frac{\bar{x}_1^2 \sum (x_{i2} - \bar{x}_2)^2 + \bar{x}_2^2 \sum (x_{i1} - \bar{x}_1)^2 - 2\bar{x}_1 \bar{x}_2 \sum (x_{i1} - \bar{x}_1)(x_{i2} - \bar{x}_2)}{\sum (x_{i1} - \bar{x}_1)^2 \sum (x_{i2} - \bar{x}_2)^2 - (\sum (x_{i1} - \bar{x}_1)(x_{i2} - \bar{x}_2))^2}\right] \cdot S^2,$$

$$S_{b1}^2 = \frac{\sum (x_{i2} - \bar{x}_2)^2}{\sum (x_{i1} - \bar{x}_1)^2 \sum (x_{i2} - \bar{x}_2)^2 - (\sum (x_{i1} - \bar{x}_1)(x_{i2} - \bar{x}_2))^2} \cdot S^2 \Rightarrow$$

$$\Rightarrow S_{b1}^2 = \frac{S^2}{\sum (x_{i1} - \bar{x}_1)^2 \cdot (1 - r_{12}^2)},$$

$$S_{b2}^2 = \frac{\sum (x_{i1} - \bar{x}_1)^2}{\sum (x_{i1} - \bar{x}_1)^2 \sum (x_{i2} - \bar{x}_2)^2 - (\sum (x_{i1} - \bar{x}_1)(x_{i2} - \bar{x}_2))^2} \cdot S^2 \Rightarrow$$

$$\Rightarrow S_{b2}^2 = \frac{S^2}{\sum (x_{i2} - \bar{x}_2)^2 \cdot (1 - r_{12}^2)},$$

$$S_{b0} = \sqrt{S_{b0}^2}, \quad S_{b1} = \sqrt{S_{b1}^2}, \quad S_{b2} = \sqrt{S_{b2}^2}. \tag{2.20}$$

Здесь r_{12} – выборочный коэффициент корреляции между объясняющими переменными X_1 и X_2; S_{bj} – стандартная ошибка коэффициента регрессии; S – стандартная ошибка регрессии (несмещенная оценка).

По аналогии с парной регрессией после определения точечных оценок b_j коэффициентов β_j ($j = 0,1,\ldots,m$) теоретического уравнения регрессии могут быть рассчитаны интервальные оценки указанных коэффициентов. Доверительный интервал, накрывающий с надежностью $(1-\alpha)$ неизвестное значение параметра β_j, определяется как

$$\left(b_j - t_{\frac{\alpha}{2}, n-m-1} S(b_j);\ b_j + t_{\frac{\alpha}{2}, n-m-1} S(b_j)\right). \tag{2.21}$$

Проверка статистической значимости коэффициентов уравнения регрессии

Как и в случае парной регрессии, статистическая значимость коэффициентов множественной линейной регрессии с m объясняющими переменными проверяется на основе t-статистики:

$$t = \frac{b_j}{S_{bj}}, \tag{2.22}$$

имеющей в данном случае распределение Стьюдента с числом степеней свободы $\nu = n - m - 1$. При требуемом уровне значимости наблюдаемое значение t-статистики сравнивается с критической точной $t_{\frac{\alpha}{2},n-m-1}$ распределения Стьюдента.

В случае если $|t| > t_{\frac{\alpha}{2},n-m-1}$, то статистическая значимость соответствующего коэффициента регрессии подтверждается. Это означает, что фактор X_j линейно связан с зависимой переменной Y. Если же установлен факт незначимости коэффициента b_j, то рекомендуется исключить из уравнения переменную X_j. Это не приведет к существенной потере качества модели, но сделает ее более конкретной.

Проверка общего качества уравнения регрессии

Для этой цели, как и в случае парной регрессии, используется *коэффициент детерминации* R^2:

$$R^2 = 1 - \frac{\sum e_i^2}{\sum (y_i - \bar{y})^2}. \tag{2.23}$$

Справедливо соотношение $0 \le R^2 \le 1$. Чем ближе этот коэффициент к единице, тем больше уравнение регрессии объясняет поведение Y.

Для множественной регрессии коэффициент детерминации является неубывающей функцией числа объясняющих переменных. Добавление новой объясняющей переменной никогда не уменьшает значение R^2, так как каждая последующая переменная может лишь дополнить, но никак не сократить информацию, объясняющую поведение зависимой переменной.

Иногда при расчете коэффициента детерминации для получения несмещенных оценок в числителе и знаменателе вычитаемой из единицы дроби делается поправка на число степеней свободы, т.е. вводится так называемый *скорректированный (исправленный) коэффициент детерминации*:

$$\bar{R}^2 = 1 - \frac{\sum e_i^2 /(n-m-1)}{\sum (y_i - \bar{y})^2 /(n-1)}. \tag{2.24}$$

Соотношение (2.24) может быть представлено в следующем виде:

$$\bar{R}^2 = 1 - (1 - R^2)\frac{n-1}{n-m-1}. \tag{2.25}$$

Из (2.25) очевидно, что $\bar{R}^2 < R^2$ для $m > 1$. С ростом значения m скорректированный коэффициент детерминации растет медленнее, чем обычный. Очевидно, что $\bar{R}^2 = R^2$ только при $R^2 = 1$. $\bar{R}^2$ может принимать отрицательные значения.

Доказано, что $\bar{R}^2$ увеличивается при добавлении новой объясняющей переменной тогда и только тогда, когда t-статистика для этой переменной по модулю больше единицы. Поэтому добавление в модель новых объясняющих переменных осуществляется до тех пор, пока растет скорректированный коэффициент детерминации.

Рекомендуется после проверки общего качества уравнения регрессии провести анализ его статистической значимости. Для этого используется F-статистика:

$$F = \frac{R^2}{1-R^2} \cdot \frac{n-m-1}{m}. \qquad (2.26)$$

Из формулы (2.26) очевидно, что показатели F и R^2 равны или не равны нулю одновременно. Если $F = 0$, то $R^2 = 0$, следовательно, величина Y линейно не зависит от $X_1, X_2, \ldots, X_m$.

Расчетное значение F сравнивается с критическим $F_{кр}$. $F_{кр}$, исходя из требуемого уровня значимости α и чисел степеней свободы $v_1 = m$ и $v_2 = n-m-1$, определяется на основе распределения Фишера (приложение 2). Если $F > F_{кр}$, то R^2 статистически значим.

Пример. Анализируется объем S сбережений домохозяйства за 10 лет. Предполагается, что его размер s_t в текущем году t зависит от величины y_t располагаемого дохода Y и от величины z_t реальной процентной ставки Z.

Исходные статистические данные

Год	95	96	97	98	99	00	01	02	03	04	05
Y, тыс.у.е.	100	110	140	150	160	160	180	200	230	250	260
Z, %	2	2	3	2	3	4	4	3	4	5	5
S, тыс.у.е.	20	25	30	30	35	38	40	38	44	50	55

Средние значения исходных данных равны: $\bar{y} = 176{,}3636$, $\bar{z} = 3{,}3636$, $\bar{s} = 36{,}8182$.

Представим требующиеся для построения модели множественной регрессии и проведения дальнейшего анализа промежуточные вычисления

Промежуточные вычисления

Год	$(y_i-\bar{y})^2$	$(z_i-\bar{z})^2$	$(s_i-\bar{s})^2$	$(y_i-\bar{y})\times$ $\times(z_i-\bar{z})$	$(y_i-\bar{y})\times$ $\times(s_i-\bar{s})$	$(z_i-\bar{z})\times$ $\times(s_i-\bar{s})$
95	5831,4050	1,8595	282,8512	104,1322	1284,2975	22,9339
96	4404,1322	1,8595	139,6694	90,4959	784,2975	16,1157
97	1322,3140	0,1322	46,4876	13,2231	247,9339	2,4793
98	695,0413	1,8595	46,4876	35,9504	179,7521	9,2975
99	267,7686	0,1322	3,3058	5,9504	29,7521	0,6612
00	267,7686	0,4050	1,3967	-10,4132	-19,3388	0,7521
01	13,2231	0,4050	10,1240	2,3140	11,5702	2,0248
02	558,6777	0,1322	1,3967	-8,5950	27,9339	-0,4298
03	2876,8595	0,4050	51,5785	34,1322	385,2066	4,5702
Продолжение таблицы 2.2						
Год	$(y_i-\bar{y})^2$	$(z_i-\bar{z})^2$	$(s_i-\bar{s})^2$	$(y_i-\bar{y})\times$ $\times(z_i-\bar{z})$	$(y_i-\bar{y})\times$ $\times(s_i-\bar{s})$	$(z_i-\bar{z})\times$ $\times(s_i-\bar{s})$

04	5422,3140	2,6777	173,7603	120,4959	970,6612	21,5702
05	6995,0413	2,6777	330,5785	136,8595	1520,6612	29,7521
$\sum$	28654,5455	12,5455	1087,6364	524,5455	5422,7273	109,7273

Расчет коэффициентов уравнения регрессии производится по формулам (2.17):

$$b_0 = 36,8182 - 0,124189 \cdot 176,3636 - 3,553796 \cdot 3,3636 = 2,962233;$$

$$b_1 = \frac{5422,7273 \cdot 12,5455 - 109,7273 \cdot 524,5455}{28654,5455 \cdot 12,5455 - (524,5455)^2} = \frac{10473,8639}{84337,619} = 0,124189;$$

$$b_2 = \frac{109,7273 \cdot 28654,5455 - 5422,7273 \cdot 524,5455}{28654,5455 \cdot 12,5455 - (524,5455)^2} = \frac{299718,7075}{84337,619} = 3,553796.$$

Таким образом, эмпирическое уравнение регрессии имеет вид:

$$s_t = 2,962233 + 0,124189 y_t + 3,553796 z_t.$$

Подставляя соответствующие значения y_t и z_t в эмпирическое уравнение регрессии, получаем $\hat{s}_t$. Расчет отклонений e_i реальных значений от модельных

Год	S	$\hat{S}$	e_i	e_i^2	$e_i - e_{i-1}$	$(e_i - e_{i-1})^2$
95	20	22,489	-2,48873	6,19375	-	-
96	25	23,731	1,26939	1,61134	3,75811	14,1234
97	30	31,01	-1,01008	1,02026	-2,27947	5,19597
98	30	28,698	1,30183	1,69475	2,31191	5,34491
99	35	33,494	1,50614	2,26845	0,20431	0,04174
00	38	37,048	0,95234	0,90696	-0,55380	0,30669
01	40	39,531	0,46856	0,21955	-0,48378	0,23404
02	38	38,461	-0,46142	0,21291	-0,92998	0,86487
03	44	45,741	-1,74089	3,03069	-1,27947	1,63703
04	50	51,778	-1,77846	3,16293	-0,03758	0,00141
05	55	53,02	1,97965	3,91900	3,75811	14,1234
сумма	405	405	≈ 0	24,24060	4,46837	41,8734
	36,8182	36,8182				

Рассчитаем дисперсию регрессии по формуле (2.19):

$$S^2 = \frac{\sum e_i^2}{n - m - 1} = \frac{24,2406}{11 - 2 - 1} = 3,03.$$

Определим дисперсии и стандартные ошибки коэффициентов по (2.20):

$$S_{b0}^2 = \left[\frac{1}{11} + \frac{(176,3636)^2 \cdot 12,5455 + (3,3636)^2 \cdot 28654,5455 - 2 \cdot 176,3636 \cdot 3,3636 \cdot 524,5455}{28654,5455 \cdot 12,5455 - (524,5455)^2} \right] \cdot 3,03;$$

$$S_{b0} = \sqrt{S_{b0}^2} = \sqrt{3,5832} = 1,8929;$$

$$S_{b1}^2 = \frac{12,5455}{28654,5455 \cdot 12,5455 - (524,5455)^2} \cdot 3,03 = 0,00054;$$

$$S_{b1} = \sqrt{S_{b1}^2} = \sqrt{0,00054} = 0,0212;$$

$$S_{b2}^2 = \frac{28654,5455}{28654,5455 \cdot 12,5455 - (524,5455)^2} \cdot 3,03 = 1,0294; \; S_{b2} = \sqrt{S_{b2}^2} = \sqrt{1,0294} = 1,0146.$$

Рассчитаем по формуле (2.22) соответствующие t-статистики:

$t_{b0} = 1,565$, $t_{b1} = 5,858$, $t_{b2} = 3,503$.

Проверим статистическую значимость коэффициентов на основе распределения Стьюдента. По таблице, приведенной в приложении 1, определим критические значения с уровнем значимости $\alpha = 0,05$: $t_{кр} = t_{\frac{\alpha}{2}; n-m-1} = t_{0,025;\ 8} = 2,306$. Таким образом, $|t_{b0}| < t_{кр}$, $|t_{b1}| > t_{кр}$, $|t_{b2}| > t_{кр}$. То есть коэффициенты b_1 и b_2 являются статистически значимыми, а гипотеза о статистической значимости коэффициента β_0 отклоняется. Это означает, что в данном случае свободным членом уравнения регрессии можно пренебречь, но тогда придется пересчитать оценки параметров β_1 и β_2.

По формуле (2.21) определим 95%-е интервальные оценки коэффициентов:

для β_0: (2,962233–2,306·1,8929; 2,962233+2,306·1,8929),

т.е. (–1,4028; 7,3273);

для β_1: (0,124189–2,306·0,0212; 0,124189+2,306·0,0212),

т.е. (0,0753; 0,1731);

для β_2: (3,553796–2,306·1,0146; 3,553796+2,306·1,0146),

т.е. (1,2141; 5,8935).

Рассчитаем коэффициент детерминации по формуле (2.23):

$$R^2 = 1 - \frac{24,2406}{1087,6364} = 0,9777.$$

Анализ статистической значимости коэффициента детерминации осуществляется на основе F-статистики (2.26):

$$F = \frac{0,9777}{1-0,9777} \cdot \frac{11-2-1}{2} = 175,3722.$$

Определим по приложению 2 критическую точку распределения Фишера: $F_{кр} = F_{0,05;2;8} = 4,46$ с 95%-й вероятностью. Очевидно, что 175,3722>4,46, следовательно, коэффициент детерминации статистически значим, т.е. совокупное влияние переменных Y и Z на переменную S существенно.

На основе проведенных рассуждений и вычислений можно заключить, что построенное уравнение регрессии объясняет 97,77% разброса зависимой переменной S.

Нелинейная регрессия

Во многих практических случаях моделирование экономических зависимостей линейными уравнениями дает вполне удовлетворительный результат и может использоваться для анализа и прогнозирования. Однако многие экономические зависимости не являются линейными по своей сути, и поэтому их моделирование линейными уравнениями регрессии не даст положительного результата.

Построение и анализ нелинейных моделей имеют свою специфику. Рассмотрим нелинейные модели, допускающие сведение их к линейным. Такие модели называют *линейными относительно параметров моделями*. Будем рассматривать модели парной регрессии с целью простоты изложения и графической иллюстрации.

Логарифмические (лог-линейные) модели

Рассмотрим модель парной регрессии. Пусть некоторая экономическая зависимость моделируется формулой:

$$Y = AX^{\beta}, \tag{3.1}$$

где A, β – параметры модели (константы, подлежащие определению).

Это функция может отражать зависимость спроса Y на благо от его цены X (тогда $\beta < 0$); зависимость спроса Y на благо от дохода X (тогда $\beta > 0$); зависимость объема выпуска Y от использования ресурса X (тогда $0 < \beta < 1$).

Для анализа функций вида (3.1) используется логарифмирование по основанию экспонента. Прологарифмировав обе части, имеем:

$$\ln Y = \ln A + \beta \ln X. \tag{3.2}$$

Заменим $\ln A$ на β_0:

$$\ln Y = \beta_0 + \beta \ln X. \tag{3.3}$$

С целью статистической оценки коэффициентов добавим в модель случайную погрешность ε, в результате получим *двойную логарифмическую модель* (и зависимая переменная, и объясняющая переменная заданы в логарифмическом виде):

$$\ln Y = \beta_0 + \beta \ln X + \varepsilon. \tag{3.4}$$

Введем замены $Y^* = \ln Y$ и $X^* = \ln X$, получаем:

$$Y^* = \beta_0 + \beta X^* + \varepsilon. \tag{3.5}$$

Модель (3.5) является линейной моделью, подробно рассмотренной ранее. Коэффициент β определяет *эластичность* переменной Y по переменной X, т.е. процентное изменение Y для данного процентного изменения X. Действительно, продифференцировав обе части (3.4), получаем:

$$\frac{1}{Y} \cdot \frac{dY}{dX} = \beta \cdot \frac{1}{X} \Rightarrow \beta = \frac{dY/Y}{dX/X} = \frac{\text{относительное изменение } Y}{\text{относительное изменение } X} = \text{эластичность}$$

В случае парной регрессии обоснованность использования логарифмической модели проверить достаточно просто. Для этого определяются точки $(\ln x_i; \ln y_i)$, $i = 1,2,\ldots,n$, которые затем наносятся на корреляционное поле. Если их расположение соответствует прямой линии, то использование данной модели обоснованно.

При большем числе переменных:

$$\ln Y = \beta_0 + \beta_1 \ln X_1 + \ldots + \beta_m \ln X_m + \varepsilon. \tag{3.6}$$

Коэффициенты β_j, $j = 1,2,\ldots,m$ являются эластичностями переменной Y по переменной X_j.

Полулогарифмические модели

Полулогарифмическими моделями являются модели вида (в случае парной регрессии):

$$\ln Y = \beta_0 + \beta X + \varepsilon. \tag{3.7}$$

$$Y = \beta_0 + \beta \ln X + \varepsilon. \tag{3.8}$$

Лог-линейная модель

Рассмотрим известную в банковском анализе зависимость:

$$Y_t = Y_0 (1 + r)^t, \tag{3.9}$$

где Y_0 – первоначальный вклад в банке; r – процентная ставка; Y_t – вклад в банке в момент времени t.

Прологарифмировав обе части (3.9): $\ln Y_t = \ln Y_0 + t \cdot \ln(1 + r)$ и введя обозначение $\ln Y_0 = \beta_0$, $\ln(1 + r) = \beta$, а также введя случайное слагаемое ε_t, получаем модель вида (3.7):

$$\ln Y_t = \beta_0 + \beta t + \varepsilon_t.$$

Полулогарифмическая модель (3.7) легко сводится к линейной модели путем замены $Y^* = \ln Y$, т.е. $Y^* = \beta_0 + \beta t + \varepsilon_t$.

Коэффициент β в (3.7) имеет смысл темпа прироста переменной Y по переменной X. Действительно, продифференцировав обе части (3.7), получаем:

$$\frac{1}{Y} \cdot \frac{dY}{dX} = \beta \Rightarrow \beta = \frac{dY / Y}{dX} = \frac{\text{относительное изменение } Y}{\text{абсолютное изменение } X}.$$

Линейно-логарифмическая модель

Линейно-логарифмическая модель (3.8) сводится к линейной модели путем замены: $X^* = \ln X$, т.е.:

$$Y = \beta_0 + \beta X^* + \varepsilon. \tag{3.10}$$

Коэффициент β в (3.8) определяет изменение переменной Y вследствие единичного относительного прироста X. Действительно, продифференцировав обе части (3.8), получаем:

$$\frac{dY}{dX} = \beta \cdot \frac{1}{X} \Rightarrow \beta = \frac{dY}{dX / X} = \frac{\textit{абсолютное изменение } Y}{\textit{относительное изменение } X}.$$

3.3. Обратная модель

Обратной моделью называется модель вида:

$$Y = \beta_0 + \beta_1 \cdot \frac{1}{X} + \varepsilon. \qquad (3.11)$$

Она сводится к линейной путем замены: $X^* = \frac{1}{X}$. Таким образом:

$$Y = \beta_0 + \beta_1 \cdot X^* + \varepsilon. \qquad (3.12)$$

В зависимости от знаков β_0, β_1 возможны случаи, представленные на рис. 3.1.

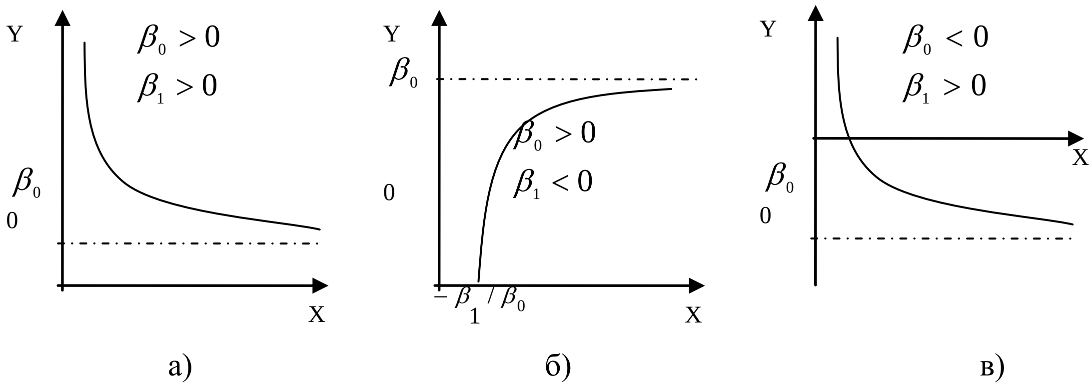


Рис. 3.1. Графики обратной линейной модели

График рис. 3.1. а) может отражать зависимость между объемом выпуска X и средними фиксированными издержками Y. График рис. 3.1. б) может отражать зависимость между доходом X и спросом на благо Y. График рис. 3.1. в) может отражать зависимость между уровнем безработицы X в процентах и процентным изменением заработной платы Y (кривая Филлипса).

Показательная модель

Показательная функция имеет вид:

$$Y = \beta_0 e^{\beta X} \qquad (3.13)$$

и сводится к лог-линейной модели путем логарифмирования:

$$\ln Y = \ln \beta_0 + \beta X. \qquad (3.14)$$

Сведение лог-линейной модели к линейной рассмотрено в главе 3.2.

Модели этого вида применяются для экономических показателей, которые характеризуются приблизительно постоянным темпом относительного прироста во времени.

Выбор формы модели

Многообразие и сложность экономических процессов предопределяют многообразие моделей, используемых для экономического анализа. Это существенно усложняет процесс нахождения максимально адекватной формулы зависимости. Для случая парной регрессии подбор модели обычно осуществляется на основе расположения наблюдаемых точек на корреляционном поле. Однако нередки ситуации, когда расположение точек приблизительно соответствует нескольким функциям и необходимо из них выявить наилучшую.

На практике неизвестно, какая модель является верной, и зачастую подбирают такую модель, которая наиболее точно соответствует реальным данным. Признаками «хорошей» модели являются:

1. Скупость (простота). Модель должна быть максимально простой. Данное свойство определяется тем фактом, что модель не отражает действительность идеально, а является ее упрощением. Поэтому из двух моделей, приблизительно одинаково отражающих реальность, предпочтение отдается модели, содержащей меньшее число объясняющих переменных.
2. Единственность. Для любого набора статистических данных определяемые коэффициенты должны вычисляться однозначно.
3. Максимально соответствие. Уравнение тем лучше, чем большую часть разброса зависимой переменной оно может объяснить. Поэтому стремятся построить уравнение с максимально возможным скорректированным коэффициентом детерминации $\overline{R}^2$.
4. Согласованность с теорией. Никакое уравнение не может быть признано качественным, если оно не соответствует известным теоретическим предпосылкам.
5. Прогнозные качества. Модель может быть признана качественной, если полученные на ее основе прогнозы подтверждаются реальностью. Другим критерием прогнозных качеств оцененной модели регрессии может служить следующее отношение:

$$V = \frac{S}{\bar{y}}. \qquad (3.15)$$

где $S = \sqrt{\dfrac{\sum e_i^2}{n-m-1}}$ – стандартная ошибка регрессии, $\bar{y}$ – среднее значение зависимой переменной. Если V мала (она определяет относительную ошибку прогноза в процентах) и отсутствует автокорреляция остатков, то прогнозные качества модели высоки.

Пример. Анализируется индекс потребительских цен Y по объему денежной массы X на основании приведенных данных. Необходимо построить логарифмическую модель.

Исходные статистические данные

Год	Y	X	Год	Y	X
1990	65	110	1993	77,5	137
1991	68	125	1994	82	160

1992	72,5	132	1995	85,5	177
1998	95	235	2001	112	250
1999	100	240	2002	115,5	275
2000	106,5	245	2003	118,5	285
1996	88,5	192	2004	120	295
1997	91	215	2005	120,5	320
			2006	121	344

Решение.

Логарифмическая модель имеет вид: $Y = AX^{\beta}$. Данная модель сводится к линейной следующим образом: $\ln Y = b_0 + b\ln X$. Для определения коэффициентов в этой модели определим логарифмы переменных Y и X, $(\ln X)^2$, $(\ln X)\cdot(\ln Y)$

Расчеты для нахождения параметров модели

Год	Y	X	$\ln Y$	$\ln X$	$(\ln X)^2$	$(\ln X)\times$ $\times(\ln Y)$
1990	65	110	4,1744	4,7005	22,0947	19,6218
1991	68	125	4,2195	4,8283	23,3125	20,3730
1992	72,5	132	4,2836	4,8828	23,8417	20,9160
1993	77,5	137	4,3503	4,9200	24,2064	21,4035
1994	82	160	4,4067	5,0752	25,7577	22,3649
1995	85,5	177	4,4485	5,1761	26,7920	23,0259
1996	88,5	192	4,4830	5,2575	27,6413	23,5694
1997	91	215	4,5109	5,3706	28,8433	24,2262
1998	95	235	4,5539	5,4596	29,8072	24,8625
1999	100	240	4,6052	5,4806	30,0370	25,2393
2000	106,5	245	4,6681	5,5013	30,2643	25,6806
2001	112	250	4,7185	5,5215	30,4870	26,0532
2002	115,5	275	4,7493	5,6168	31,5484	26,6759
2003	118,5	285	4,7749	5,6525	31,9508	26,9901
2004	120	295	4,7875	5,6870	32,3420	27,2265
2005	120,5	320	4,7916	5,7683	33,2733	27,6394
2006	121	344	4,7958	5,8406	34,1126	28,0103
Сумма	1639	3737	77,3217	90,7392	486,3122	413,8784
Среднее	96,4118	219,8235	4,5483	5,3376	28,6066	24,3458

Рассчитываем коэффициенты для этой модели следующим образом:

$$\begin{cases} b = \dfrac{\overline{(\ln X)\cdot\ln(Y)} - \overline{\ln X}\cdot\overline{\ln Y}}{\overline{(\ln X)^2} - (\overline{\ln X})^2} = \dfrac{24{,}3458 - 5{,}3376\cdot 4{,}5483}{28{,}6066 - (5{,}3376)^2} = \dfrac{0{,}0688}{0{,}1166} = 0{,}5901; \\ b_0 = \overline{\ln Y} - b\cdot\overline{\ln X} = 4{,}5483 - 0{,}5901\cdot 5{,}3376 = 1{,}3986. \end{cases}$$

Следовательно, модель имеет вид: $\ln Y = 1{,}3986 + 0{,}5901\cdot\ln X$. Если свести данную модель к виду $Y = AX^{\beta}$, то получим:

$Y = 4{,}0495 \cdot X^{0{,}5901}$ (т.к. $\ln A = b_0 = 1{,}3986$, следовательно, $A = e^{b_0} = 4{,}0495$).

Представим графически корреляционное поле для переменных $\ln Y$ и $\ln X$, а также график рассчитанной модели $\ln Y = 1{,}3986 + 0{,}5901 \cdot \ln X$ (рис. 3.2).

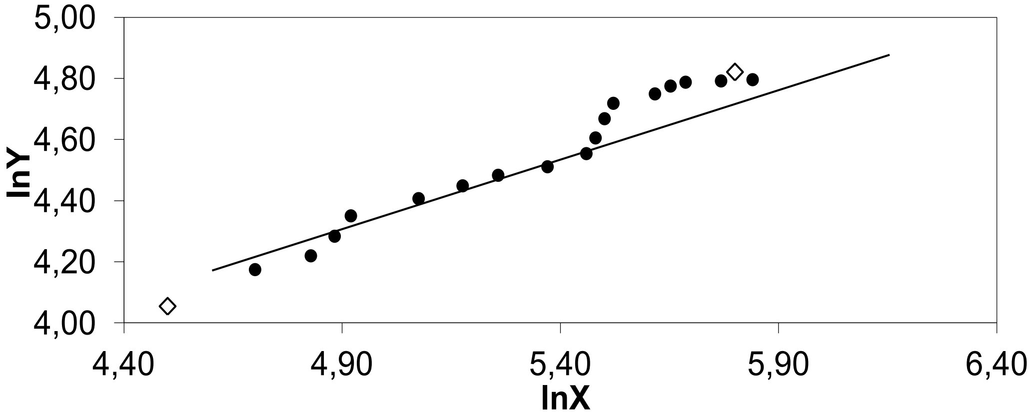


Рис. 3.2. Исходные данные и теоретическая линия регрессии

После определения коэффициентов модели необходимо проверить значимость коэффициентов уравнения регрессии, определить их интервальные оценки и рассчитать коэффициент детерминации. Расчет проводится аналогично примеру для модели вида $\ln Y = 1{,}3986 + 0{,}5901 \cdot \ln X$.

Тема 7. Фиктивные переменные в регрессионных моделях. Динамические модели

Фиктивные переменные в регрессионных моделях

Необходимость использования фиктивных переменных

В регрессионных моделях в качестве объясняющих переменных часто приходится использовать не только количественные (определяемые численно), но и качественные переменные. Например, спрос на какое-либо благо может определяться как количественными переменными (цена данного блага, цена на заменители данного блага, доход потребителя и т.д.), так и качественными (вкусы потребителей, их ожидания, национальные и религиозные особенности и т.д.). Качественные показатели в численном виде представить нельзя. Возникает проблема отражения в модели влияния таких переменных на исследуемую величину.

Обычно в моделях влияние качественного фактора выражается в виде фиктивной (искусственной) переменной, которая отражает два противоположных состояния качественного фактора. В этом случае фиктивная переменная может выражаться в двоичной форме:

$$D = \begin{cases} 0, & \text{фактор не действует;} \\ 1, & \text{фактор действует.} \end{cases}$$

Например, D=0, если потребитель не имеет высшего образования, D=1, если потребитель имеет высшее образование; D=0, если в обществе имеются инфляционные ожидания, D=1, если инфляционных ожиданий нет.

Переменная D называется *фиктивной (искусственной, двоичной) переменной (индикатором)*.

Регрессионные модели, содержащие лишь качественные объясняющие переменные, называются *моделями дисперсионного анализа* (ANOVA-моделями).

Например, пусть Y – начальная заработная плата.

$$D=\begin{cases}0, & \text{если претендент не имееет высшего образования;}\\ 1, & \text{если претендент имеет высшее образование.}\end{cases}$$

Тогда зависимость можно выразить моделью парной регрессии:

$$Y=\beta_0+\gamma D+\varepsilon; \qquad (6.1)$$

Очевидно,

$$M(Y|D=0)=\beta_0+\gamma\cdot 0=\beta_0;$$

$$M(Y|D=1)=\beta_0+\gamma\cdot 1=\beta_0+\gamma.$$

При этом коэффициент β_0 определяет среднюю начальную заработную плату при отсутствии высшего образования. Коэффициент γ указывает, на какую величину отличаются средние начальные заработные платы при наличии и при отсутствии высшего образования у претендента. Проверяя статистическую значимость коэффициента γ с помощью t-статистики либо значимость коэффициента детерминации R^2 с помощью F-статистики, можно определить, влияет или нет наличие высшего образования на начальную заработную плату.

Модели, в которых объясняющие переменные носят как количественный, так и качественный характер, называются *моделями ковариационного анализа* (ANCOVA-моделями).

Модели ковариационного анализа

Существует несколько разновидностей таких моделей.

1. Модели ковариационного анализа при наличии у фиктивной переменной двух альтернатив

Сначала рассмотрим простейшую модель с одной количественной и одной качественной переменными, имеющую два альтернативных состояния:

$$Y=\beta_0+\beta_1 X+\gamma D+\varepsilon. \qquad (6.2)$$

Пусть, например, Y заработная плата сотрудника фирмы, X – стаж сотрудника, D – пол сотрудника, т.е.

$$D=\begin{cases}0, & \text{если сотрудник - женщина;}\\ 1, & \text{если сотрудник - мужчина.}\end{cases}$$

Тогда ожидаемое значение заработной платы сотрудников при x годах трудового стажа будет:

$M(Y|x, D=0) = \beta_0 + \beta_1 x;$

$M(Y|x, D=1) = \beta_0 + \beta_1 x + \gamma = (\beta_0 + \gamma) + \beta_1 x.$

Заработная плата в данном случае является линейной функцией от стажа работы. Графически данную ситуацию можно представить следующим образом (рис. 6.1).

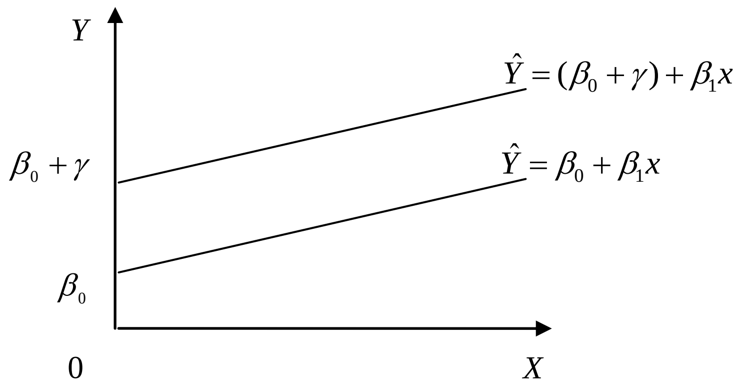


Рис. 6.1. Зависимость заработной платы от стажа работы

Очевидно, что и для мужчин, и для женщин заработная плата меняется с одним и тем же коэффициентом пропорциональности β_1. Однако свободные члены отличаются на величину γ. Проверив с помощью t-статистики статистические значимости коэффициентов β_0 и $(\beta_0 + \gamma)$, можно определить, имеет ли место в фирме дискриминация по половому признаку. Если эти коэффициенты окажутся статистически значимыми, то, очевидно, дискриминация есть. Более того, при $\gamma > 0$ она будет в пользу мужчин, при $\gamma < 0$ – в пользу женщин.

При составлении моделей с фиктивными переменными необходимо руководствоваться следующим *правилом моделирования*: если качественная переменная имеет k альтернативных значений, то при моделировании используются только $(k-1)$ фиктивных переменных. Таким образом, если переменная имеет два альтернативных значения (например, пол), то в модель можно ввести только одну фиктивную переменную.

Если не следовать данному правилу, то при моделировании исследователь попадает в ситуацию совершенной мультиколлинеарности или так называемую *ловушку фиктивной переменной*.

Значения фиктивной переменной можно изменять на противоположные. Суть модели от этого не изменится. Например, в модели (6.2) можно положить, что:

$$D = \begin{cases} 0, & \text{если сотрудник - мужчина;} \\ 1, & \text{если сотрудник - женщина.} \end{cases}$$

Однако при этом знак коэффициента γ изменится на противоположный.

Значение качественной переменной, для которого принимается $D = 0$, называется *базовым* или *сравнительным*. Выбор базового значения обычно диктуется целями исследования, но может быть и произвольным.

Коэффициент γ в модели (6.2) называется *дифференциальным коэффициентом свободного члена*, так как он показывает, на какую величину отличается свободный член модели при значении фиктивной переменной, равном единице, от свободного члена модели при базовом значении фиктивной переменной.

2. Модели ковариационного анализа при наличии у качественных переменных более двух альтернатив

Рассмотрим модель с двумя объясняющими переменными, одна из которых количественная, а другая – качественная. Причем качественная переменная имеет три альтернативы. Например, расходы на содержание ребенка могут быть связаны с доходами домохозяйства и возрастом ребенка: дошкольный, младший школьный и старший школьный. Так как качественная переменная имеет три альтернативы, то по общему правилу моделирования необходимо использовать две фиктивные переменные. Таким образом, модель может быть представлена в виде:

$$Y = \beta_0 + \beta_1 X + \gamma_1 D_1 + \gamma_2 D_2 + \varepsilon, \qquad (6.3)$$

где Y – расходы, X – доходы домохозяйств.

$$D_1 = \begin{cases} 0, & \text{если дошкольник;} \\ 1, & \text{в противоположном случае;} \end{cases}$$

$$D_2 = \begin{cases} 0, & \text{если дошкольник или младший школьник;} \\ 1, & \text{в противоположном случае.} \end{cases}$$

Образуются следующие зависимости:

1. Средний расход на дошкольника:

$$M(Y | D_1 = 0, D_2 = 0) = \beta_0 + \beta_1 X. \qquad (6.4)$$

2. Средний расход на младшего школьника:

$$M(Y | D_1 = 1, D_2 = 0) = (\beta_0 + \gamma_1) + \beta_1 X. \qquad (6.5)$$

3. Средний расход на старшего школьника:

$$M(Y | D_1 = 1, D_2 = 1) = (\beta_0 + \gamma_1 + \gamma_2) + \beta_1 X. \qquad (6.6)$$

Здесь γ_1, γ_2 – дифференциальные свободные члены. Базовым значением качественной переменной является значение «дошкольник». Таким образом, получаются три регрессионные прямые, параллельные друг другу (рис. 6.2).

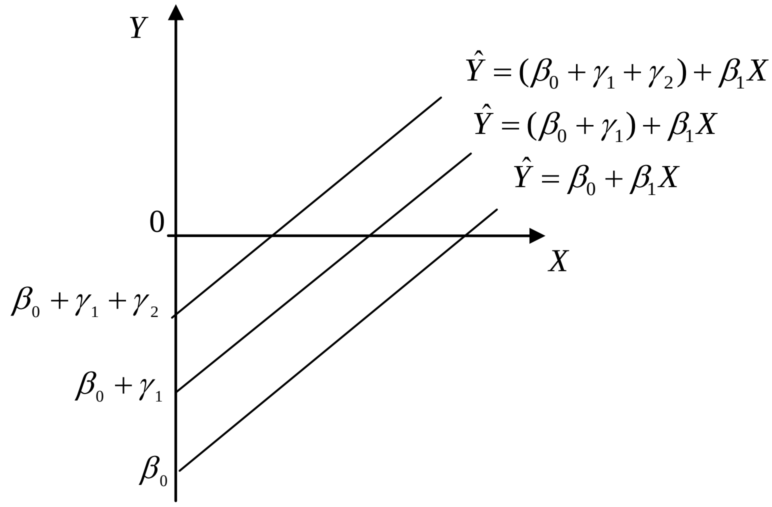


Рис. 6.2. Зависимость расходов на ребенка от его возраста

После вычисления коэффициентов уравнений регрессии (6.4) - (6.6) определяется статистическая значимость коэффициентов γ_1 и γ_2 на основе рассмотренной ранее *t*-статистики.

Если коэффициенты γ_1 и γ_2 оказываются статистически незначимыми, то можно сделать вывод, что возраст ребенка не оказывает существенного влияния на расходы по его содержанию.

3. Регрессия с одной количественной и двумя качественными переменными

Техника фиктивных переменных может быть распространена на произвольное число качественных факторов. Для простоты рассмотрим ситуацию с двумя качественными переменными.

Пусть Y – заработная плата сотрудников фирмы, X – стаж работы, D_1 – наличие высшего образования, D_2 – пол сотрудника:

$$D_1=\begin{cases}0, & \text{если сотрудник - женщина;}\\ 1, & \text{если сотрудник - мужчина;}\end{cases}$$

$$D_2=\begin{cases}0, & \text{если нет высшего образования;}\\ 1, & \text{в противоположном случае.}\end{cases}$$

Таким образом, получим следующую модель:

$$Y=\beta_0+\beta_1 X+\gamma_1 D_1+\gamma_2 D_2+\varepsilon. \qquad (6.7)$$

Из этой модели выводятся следующие регрессионные зависимости:

1. Средняя зарплата женщины без высшего образования:

$$M(Y|D_1=0, D_2=0)=\beta_0+\beta_1 X. \qquad (6.8)$$

2. Средняя зарплата женщины с высшим образованием:

$$M(Y|D_1=0, D_2=1)=(\beta_0+\gamma_2)+\beta_1 X. \qquad (6.9)$$

3. Средняя зарплата мужчины без высшего образования:

$$M(Y|D_1=1, D_2=0)=(\beta_0+\gamma_1)+\beta_1 X. \qquad (6.10)$$

4. Средняя зарплата мужчины с высшим образованием:

$$M(Y|D_1=1, D_2=1)=(\beta_0+\gamma_1+\gamma_2)+\beta_1 X. \qquad (6.11)$$

Очевидно, что все регрессии отличаются только свободными членами. Дальнейшее определение статистической значимости коэффициентов γ_1 и γ_2 позволяет убедиться, влияют ли образование и пол сотрудника на его заработную плату или нет.

Пример. Рассмотрим зависимость между весом новорожденного Y (в граммах), X – количеством сигарет, выкуриваемых в день будущей матерью во время беременности и фиктивной переменной D, которая отражает факт того, является ребенок первенцем или нет. Пусть $D=0$, если ребенок – первенец, и $D=1$, если ребенок не первенец. Рассмотрим выборку из 20 значений.

Исходные данные

Наблюдение	Y	X	D	Наблюдение	Y	X	D
1	3520	10	1	11	3210	29	1
2	3460	19	1	12	3290	15	1
3	3000	16	1	13	3190	3	0
4	3320	26	1	14	3060	12	0
5	3540	4	1	15	3270	17	0
6	3310	14	1	16	3170	14	0
7	3360	21	1	17	3230	18	0
8	3650	10	1	18	3700	11	0
9	3150	22	1	19	3300	14	0
10	3440	8	1	20	3460	9	0

Данная модель содержит одну количественную и одну качественную переменные. В общем виде запишем ее следующим образом: $Y=b_0+b_1X+b_2D$. Коэффициенты b_0, b_1, b_2 определяются из формул:

$$b_0=\bar{y}-b_1\bar{x}_1-b_2\bar{x}_2;$$

$$b_1=\frac{\sum(x_{i1}-\bar{x}_1)(y_i-\bar{y})\cdot\sum(x_{i2}-\bar{x}_2)^2-\sum(x_{i2}-\bar{x}_2)(y_i-\bar{y})\cdot\sum(x_{i1}-\bar{x}_1)(x_{i2}-\bar{x}_2)}{\sum(x_{i1}-\bar{x}_1)^2\sum(x_{i2}-\bar{x}_2)^2-(\sum(x_{i1}-\bar{x}_1)(x_{i2}-\bar{x}_2))^2};$$

$$b_2=\frac{\sum(x_{i2}-\bar{x}_2)(y_i-\bar{y})\cdot\sum(x_{i1}-\bar{x}_1)^2-\sum(x_{i1}-\bar{x}_1)(y_i-\bar{y})\cdot\sum(x_{i1}-\bar{x}_1)(x_{i2}-\bar{x}_2)}{\sum(x_{i1}-\bar{x}_1)^2\sum(x_{i2}-\bar{x}_2)^2-(\sum(x_{i1}-\bar{x}_1)(x_{i2}-\bar{x}_2))^2}.$$

Вспомогательная таблица для расчета коэффициентов имеет следующий вид

№	$(y_i-\bar{y})$	$(x_i-\bar{x})$	$(d_i-\bar{d})$	$(x_i-\bar{x})^2$	$(d_i-\bar{d})^2$	$(x_i-\bar{x})\times$ $\times(y_i-\bar{y})$	$(d_i-\bar{d})\times$ $\times(y_i-\bar{y})$	$(x_i-\bar{x})\times$ $\times(d_i-\bar{d})$
1	188,5	-4,6	188,5	21,16	0,16	-867,1	75,4	-1,84
2	128,5	4,4	128,5	19,36	0,16	565,4	51,4	1,76
3	-331,5	1,4	-331,5	1,96	0,16	-464,1	-132,6	0,56
4	-11,5	11,4	-11,5	129,96	0,16	-131,1	-4,6	4,56
5	208,5	-10,6	208,5	112,36	0,16	-2210,1	83,4	-4,24
6	-21,5	-0,6	-21,5	0,36	0,16	12,9	-8,6	-0,24

№	$(y_i-\bar{y})$	$(x_i-\bar{x})$	$(d_i-\bar{d})$	$(x_i-\bar{x})^2$	$(d_i-\bar{d})^2$	$(x_i-\bar{x})\times$ $\times(y_i-\bar{y})$	$(d_i-\bar{d})\times$ $\times(y_i-\bar{y})$	$(x_i-\bar{x})\times$ $\times(d_i-\bar{d})$
7	28,5	6,4	28,5	40,96	0,16	182,4	11,4	2,56
8	318,5	-4,6	318,5	21,16	0,16	-1465,1	127,4	-1,84
9	-181,5	7,4	-181,5	54,76	0,16	-1343,1	-72,6	2,96
10	108,5	-6,6	108,5	43,56	0,16	-716,1	43,4	-2,64
11	-121,5	14,4	-121,5	207,36	0,16	-1749,6	-48,6	5,76
12	-41,5	0,4	-41,5	0,16	0,16	-16,6	-16,6	0,16
13	-141,5	-11,6	-141,5	134,56	0,36	1641,4	84,9	6,96
14	-271,5	-2,6	-271,5	6,76	0,36	705,9	162,9	1,56
15	-61,5	2,4	-61,5	5,76	0,36	-147,6	36,9	-1,44
16	-161,5	-0,6	-161,5	0,36	0,36	96,9	96,9	0,36
17	-101,5	3,4	-101,5	11,56	0,36	-345,1	60,9	-2,04
18	368,5	-3,6	368,5	12,96	0,36	-1326,6	-221,1	2,16
19	-31,5	-0,6	-31,5	0,36	0,36	18,9	18,9	0,36
20	128,5	-5,6	128,5	31,36	0,36	-719,6	-77,1	3,36
$\sum$				856,8	4,8	-8278	272	18,8

Рассчитаем средние показатели: $\bar{y}=3331{,}5$, $\bar{x}=14{,}6$, $\bar{d}=0{,}6$.

Параметры уравнения находим по формулам (параметры уравнения множественной регрессии):

$$b_0=3331{,}5+11{,}93\cdot14{,}6-103{,}39\cdot0{,}6=3443{,}64;$$

$$b_1=\frac{(-8278)\cdot4{,}8-272\cdot18{,}8}{856{,}8\cdot4{,}8-(18{,}8)^2}=-11{,}93;$$

$$b_2=\frac{272\cdot856{,}8-(-8278)\cdot18{,}8}{856{,}8\cdot4{,}8-(18{,}8)^2}=103{,}39.$$

Таким образом, уравнение регрессии с учетом рассчитанных коэффициентов примет вид: $\hat{Y}=3443{,}64-11{,}93X+103{,}39D$.

Затем аналогично примеру, приведенному в главе 2, рассчитывается статистическая значимость коэффициентов. Рассчитанное значение t-статистики для коэффициента b_2 при фиктивной переменной D составляет $t=1{,}23$.

Из приложения 1 определим для уровня значимости $\alpha=0{,}05$ и числа степеней свободы $\nu=n-m-1=20-2-1=17$ критическое значение t-статистики: $t_{кр}=t_{\frac{\alpha}{2},n-m-1}=t_{0{,}025{,}17}=2{,}110$. Так как $t<t_{кр}$, то коэффициент b_2 при фиктивной переменной D является статистически незначимым с уровнем значимости 0,05. То есть данной фиктивной переменной в уравнении можно пренебречь.

Однако можно предположить, что это объясняется малым размером выборки (20 значений). Если рассмотреть большую выборку, то обнаружится статистическая значимость данного коэффициента.

Сравнение двух регрессий

В примерах, рассматриваемых до сих пор, предполагалось, что изменение значения качественного фактора влияет лишь на изменение свободного члена. Но это не всегда так. Изменение качественного фактора может привести также к изменению наклона прямой регрессии.

Обычно это характерно для временных рядов экономических данных при изменении институциональных условий, введении новых правовых или налоговых ограничений. Например, можно предположить, что до некоторого года в стране обменный курс валют был фиксированным, а затем плавающим. Или налог на ввозимые автомобили был одним, а затем он существенно изменился. В этом случае зависимость может быть выражена так:

$$Y_t = \beta_0 + \beta_1 X_t + \gamma_1 D_t + \gamma_2 D_t X_t + \varepsilon_t, \qquad (6.12)$$

где $D_t = \begin{cases} 0, & \text{до изменения институциональных условий;} \\ 1, & \text{после изменения институциональных условий.} \end{cases}$

В этой ситуации ожидаемое значение зависимой переменной определяется следующим образом:

$$M(Y_t | D_t = 0) = \beta_0 + \beta_1 X_t. \qquad (6.13)$$

$$M(Y_t | D_t = 1) = (\beta_0 + \gamma_1) + (\beta_1 + \gamma_2) X_t. \qquad (6.14)$$

Коэффициенты γ_1 и γ_2 в уравнении (6.12) называются *дифференциальным свободным членом* и *дифференциальным угловым коэффициентом* соответственно. Фиктивная переменная D_t в уравнении (6.12) используется как в аддитивном виде $(\gamma_1 D_t)$, так и в мультипликативном $(\gamma_2 D_t X_t)$, что позволяет фактически разбивать рассматриваемую зависимость на две части, связанные с периодами изменения некоторого рассматриваемого в модели качественного фактора. Уравнение регрессии (6.12) достаточно хорошо моделирует ситуацию, изображенную на рисунке 6.3.

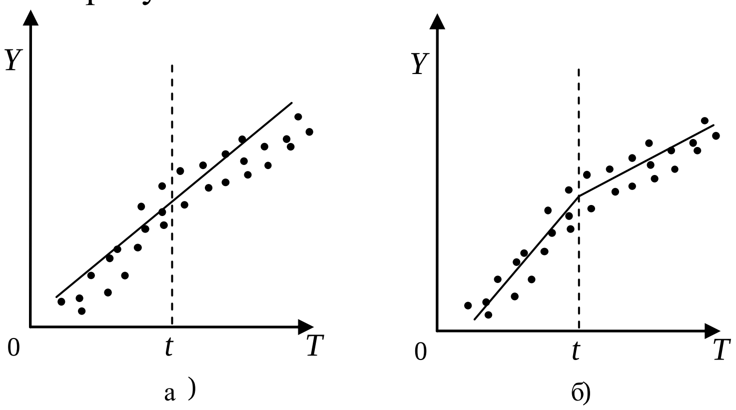


Рис. 6.3. Изменение показателя во времени

На рисунке 6.3, а) зависимость моделируется обыкновенной линейной регрессией. На рисунке 6.3, б) в модели учитываются изменения, произошедшие с некоторого времени t в характере расположения точек наблюдений. На данном примере хорошо видно, каким образом можно проанализировать, имеет ли смысл разбивать выборку на части и строить для каждой из них уравнение регрессии (т.е. фактически строить сложную регрессию с фиктивными переменными) либо можно ограничиться общей «обыкновенной» регрессией для всех точек наблюдений. Для этого можно использовать тест Чоу.

Суть теста Чоу заключается в следующем. Пусть выборка имеет объем n. Через S_0 обозначим сумму квадратов отклонений $\sum e_i^2$ значений y_i от общего уравнения регрессии. Пусть есть основание предполагать, что целесообразно общую выборку разбить на две подвыборки объемами n_1 и n_2 соответственно $(n_1 + n_2 = n)$ и построить для каждой из выборок уравнение регрессии. Обозначим через S_1 и S_2 суммы квадратов отклонений значений y_i каждой из подвыборок от соответствующих уравнений регрессий. Затем рассчитывается F-статистика, которая для теста Чоу имеет вид:

$$F = \frac{S_0 - S_1 - S_2}{S_1 + S_2} \cdot \frac{n - 2m - 2}{m + 1}, \qquad (6.15)$$

где m – число количественных объясняющих переменных в уравнении регрессии (одинаково для всех трех уравнений регрессии).

Из приложения 2 определяется $F_{кр}$ для числа степеней свободы $v_1 = m + 1$, $v_2 = n - 2m - 2$ и требуемого уровня значимости α. Если $F_{набл} < F_{кр}$ при выбранном уровне значимости, то нет смысла разбивать уравнение регрессии на части. В противном случае разбиение на подынтервалы целесообразно с точки зрения улучшения качества модели, что означает необходимость введения в уравнение регрессии соответствующей фиктивной переменной.

Использование фиктивных переменных в сезонном анализе

Многие экономические показатели напрямую связаны с сезонными колебаниями. Например, спрос на туристические путевки, охлажденную воду и мороженое существенно выше летом, чем зимой. Спрос на обогреватели, шубы выше зимой, чем летом. Некоторые показатели имеют существенные квартальные колебания и т.д.

Обычно сезонные колебания характерны для временных рядов. Устранение или нейтрализация сезонного фактора в таких моделях позволяет сконцентрироваться на других важных количественных и качественных характеристиках модели, в частности, на общем направлении развития модели, так называемом *тренде*. Такое устранение сезонного фактора называется *сезонной корректировкой*. Существует несколько методов сезонной корректировки, одним из которых является *метод фиктивных переменных.*

Пусть переменная Y определяется количественной переменной X, причем эта зависимость существенно разнится по кварталам. Тогда общую модель в этой ситуации можно представить в виде:

$$Y_t = \beta_0 + \beta_1 X_t + \gamma_1 D_{1t} + \gamma_2 D_{2t} + \gamma_3 D_{3t} + \varepsilon_t, \qquad (6.16)$$

где $D_{1t} = \begin{cases} 1, & \text{если рассматривается II квартал;} \\ 0, & \text{в противном случае;} \end{cases}$

$D_{2t} = \begin{cases} 1, & \text{если рассматривается III квартал;} \\ 0, & \text{в противном случае;} \end{cases}$

$D_{3t} = \begin{cases} 1, & \text{если рассматривается IV квартал;} \\ 0, & \text{в противном случае.} \end{cases}$

Заметим, что число кварталов равно четырем, следовательно, по правилу моделирования, число фиктивных переменных должно быть равно трем.

В нашем примере в качестве базы выбран I квартал. Если значения Y существенно различаются по кварталам (сезонам), то в уравнении (6.16) коэффициенты при фиктивных переменных окажутся статистически значимыми. Тогда ожидаемое значение Y по кварталам определяется следующими соотношениями:

$M(Y_t | D_{1t} = 0, D_{2t} = 0, D_{3t} = 0) = \beta_0 + \beta_1 X_t$ – для I квартала;

$M(Y_t | D_{1t} = 1, D_{2t} = 0, D_{3t} = 0) = (\beta_0 + \gamma_1) + \beta_1 X_t$ – для II квартала;

$M(Y_t | D_{1t} = 0, D_{2t} = 1, D_{3t} = 0) = (\beta_0 + \gamma_2) + \beta_1 X_t$ – для III квартала;

$M(Y_t | D_{1t} = 0, D_{2t} = 0, D_{3t} = 1) = (\beta_0 + \gamma_3) + \beta_1 X_t$ – для IV квартала.

В модели (6.16) рассматриваются такие ситуации, при которых квартальные различия отражаются лишь в различии свободных членов моделей. Если же различия затрагивают и изменения коэффициента пропорциональности, то этот факт может быть отражен в следующей модели:

$$Y_t = \beta_0 + \beta_1 X_t + \gamma_1 D_{1t} + \gamma_2 D_{2t} + \gamma_3 D_{3t} + \gamma_4 D_{1t} X_t + \gamma_5 D_{2t} X_t + \gamma_6 D_{3t} X_t + \varepsilon_t. \quad (6.17)$$

Выбор правильной формы модели регрессии является в данном случае достаточно серьезной проблемой, так как вполне вероятны ошибки спецификации. Наиболее рациональной практической стратегией выбора модели является следующая схема. Вначале рассматривается модель (6.17). Определяется статистическая значимость коэффициентов. Если дифференциальные угловые коэффициенты оказываются статистически незначимыми, то переходят к модели (6.16). Если в этой модели дифференциальные свободные члены оказываются статистически незначимыми, то делают вывод, что квартальные (сезонные) изменения несущественны для рассматриваемой зависимости.

Пример. Данные по расходам потребителей на газ и электричество в США в постоянных ценах 2000 года с I квартала 2000 года по IV квартал 2005 года (млрд. долл.)

	I квартал	II квартал	III квартал	IV квартал
2000год	7,33	4,70	5,10	5,46
2001год	7,65	4,92	5,15	5,55
2002год	7,96	5,01	5,05	5,59
2003год	7,74	5,10	5,67	5,92
2004год	8,04	5,27	5,51	6,04
2005год	8,26	5,51	5,41	5,83

Для модели вида

$$Y_t = \beta_0 + \beta_1 X_t + \gamma_1 D_{1t} + \gamma_2 D_{2t} + \gamma_3 D_{3t} + \varepsilon_t,$$

с соответствующими переменными D_{1t}, D_{2t}, D_{3t}

$$D_{1t} = \begin{cases} 1, & \text{если рассматривается II квартал;} \\ 0, & \text{в противном случае;} \end{cases}$$

$$D_{2t} = \begin{cases} 1, & \text{если рассматривается III квартал;} \\ 0, & \text{в противном случае;} \end{cases}$$

$$D_{3t} = \begin{cases} 1, & \text{если рассматривается IV квартал;} \\ 0, & \text{в противном случае.} \end{cases}$$

по данным, приведенным в таблице 6.3, в результате расчетов по формуле $B = (X^T X)^{-1} X^T Y$. (см. главу 2), получена модель вида:

$$\hat{Y}_t = 7{,}50 + 0{,}030t - 2{,}78D_1 - 2{,}58D_2 - 2{,}19D_3.$$

Из этого результата выводим отдельные уравнения для каждого квартала:

для I квартала: $\hat{Y}_t = 7{,}50 + 0{,}030t;$

для II квартала: $\hat{Y}_t = 4{,}72 + 0{,}030t;$

для III квартала: $\hat{Y}_t = 4{,}92 + 0{,}030t;$

для IV квартала: $\hat{Y}_t = 5{,}31 + 0{,}030t.$

Динамические модели

Временные ряды. Лаги в экономических моделях

При анализе многих экономических показателей (особенно в макроэкономике) часто используются ежегодные, ежеквартальные, ежемесячные, ежедневные данные. Для рационального анализа необходимо систематизировать моменты получения соответствующих статистических данных.

В этом случае следует упорядочить данные по времени их получения и построить так называемые *временные ряды*.

Пусть исследуется показатель Y. Его значение в текущий момент (период) времени t обозначают y_t; значения Y в последующие моменты обозначаются $y_{t+1}, y_{t+2}, \ldots, y_{t+k}, \ldots$; значения Y в предыдущие моменты времени обозначаются $y_{t-1}, y_{t-2}, \ldots, y_{t-k}, \ldots$.

Нетрудно понять, что при изучении зависимостей между такими показателями либо при анализе их развития во времени в качестве объясняющих переменных используются не только текущие значения переменных, но и некоторые предыдущие по времени значения, а также само время *T*. Модели такого типа называются *динамическими*.

В свою очередь переменные, влияние которых характеризуется определенным запаздыванием, называются *лаговыми переменными*.

Обычно динамические модели подразделяются на два класса:

1) *модели с лагами* (модели с распределенными лагами) – содержат в качестве лаговых переменных лишь независимые (объясняющие) переменные. Примером является модель:

$$y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \ldots + \beta_k x_{t-k} + \varepsilon_t; \quad (7.1)$$

2) *авторегрессионные модели* – модели, уравнения которых в качестве лаговых объясняющих переменных включают значения зависимых переменных.

$$y_t = \alpha + \beta_0 x_t + \gamma y_{t-1} + \varepsilon_t. \quad (7.2)$$

Во многих случаях воздействие одних экономических факторов на другие осуществляется не мгновенно, а с некоторым временным запаздыванием – лагом. Причин наличия лагов в экономике достаточно много, среди них можно выделить следующие:

- *психологические причины* – обычно выражаются через инерцию в поведении людей. Например, люди тратят свой доход не мгновенно, а постепенно. Привычка к определенному образу жизни приводит к тому, что люди приобретают те же блага в течение некоторого времени даже после падения реального дохода;
- *технологические причины*. Например, изобретение персональных компьютеров не привело к мгновенному вытеснению ими больших ЭВМ в силу необходимости замены соответствующего программного обеспечения, которое потребовало продолжительного времени;
- *институциональные причины*. Например, контракты между фирмами требуют определенного постоянства в течение времени контракта;
- *механизмы формирования экономических показателей*. Например, инфляция во многом является инерционным процессом.

Оценка параметров моделей с лагами в независимых переменных

Оценка модели с распределенными лагами во многом зависит от того, конечное или бесконечное число лагов она содержит.

$$y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \ldots + \beta_k x_{t-k} + \varepsilon_t \text{ (конечное число лагов).}$$

$$y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \ldots + \varepsilon_t \text{ (бесконечное число лагов).}$$

В обеих этих моделях коэффициент β_0 называется *краткосрочным мультипликатором*. Он характеризует изменение среднего значения Y под воздействием единичного изменения переменной X в тот же самый момент времени.

Сумма всех коэффициентов $\sum_j \beta_j$ называется *долгосрочным мультипликатором*. Он характеризует изменение Y под воздействием единичного изменения переменной X в каждом из рассматриваемых временных периодов.

Любую сумму коэффициентов $\sum_j^h \beta_j$, $(h < k)$ называют *промежуточным мультипликатором*.

Модель с конечным числом лагов (7.1) оценивается достаточно просто сведением ее к уравнению множественной регрессии. При этом $X_0^* = x_t$, $X_1^* = x_{t-1}$, …, $X_k^* = x_{t-k}$ и получают уравнение:

$$y_t = \alpha + \beta_0 X_0^* + \beta_1 X_1^* + \ldots + \beta_k X_k^* + \varepsilon_t. \qquad (7.4)$$

Рассмотрим пример модели с распределенным лагом:

$$y_t = -0.67 + 4.5x_t + 3x_{t-1} + 1.5x_{t-2} + 0.5x_{t-3},$$

где Y – объем продаж (млн. руб), X – расходы на рекламу (млн. руб).

Краткосрочный мультипликатор 4,5 показывает, что с увеличением расходов на рекламу на 1 млн. руб. объем продаж компании вырастет на 4,5 млн. руб.

Вычислим промежуточные мультипликаторы для разных моментов времени:

$t+1 \qquad 4.5+3=7.5$

$t+2 \qquad 7.5+1.5=9$

$t+3 \qquad 9+0.5=9.5$

То есть долгосрочный мультипликатор 9,5 показывает, что с увеличением расходов на рекламу на 1 млн. руб. объем продаж компании вырастет через 3 месяца на 9,5 млн. руб.

Вычислим относительные коэффициенты регрессии:

$$\beta_1 = \frac{b_1}{b} = \frac{4.5}{9.5} = 0.474 \qquad \beta_2 = \frac{b_2}{b} = \frac{3}{9.5} = 0.316$$

$$\beta_3 = \frac{b_3}{b} = \frac{1.5}{9.5} = 0.158 \qquad \beta_4 = \frac{b_4}{b} = \frac{0.5}{9.5} = 0.053$$

То есть 47,4% общего увеличения объема продаж, вызванного ростом затрат на рекламу, происходит в текущем моменте времени, 31,6% через 1 месяц, 15,8% через 2 месяца и только 5,3% этого увеличения происходит через 3 месяца.

Для оценки моделей с бесконечным числом лагов разработано несколько моделей. Рассмотрим две из них.

1. Метод последовательного увеличения количества лагов

По данному методу уравнение (7.3) рекомендуется оценивать с последовательно увеличивающимся количеством лагов. Признаками завершения процедуры увеличения количества лагов могут являться следующие:

- при добавлении нового лага какой-либо коэффициент регрессии β_k при переменной x_{t-k} меняет знак. Тогда в уравнении регрессии оставляют переменные $x_t, x_{t-1}, \ldots, x_{t-k+1}$, коэффициенты при которых знак не поменяли;
- при добавлении нового лага коэффициент регрессии β_k при переменной x_{t-k} становится статистически незначимым. Очевидно, что в уравнении будут использоваться только переменные $x_t, x_{t-1}, \ldots, x_{t-k+1}$, коэффициенты при которых остаются статистически значимыми.

Применение метода последовательного увеличения количества лагов весьма ограничено в силу постоянно уменьшающегося числа степеней свободы, сопровождающегося увеличением стандартных ошибок и ухудшением качества оценок, а также возможности мультиколлинеарности. Кроме того, при неправильном определении числа лагов возможны ошибки спецификации.

2. Метод геометрической прогрессии (метод Койка)

В распределении Койка предполагается, что коэффициенты («веса») β_k при лаговых значениях объясняющей переменной убывают в геометрической прогрессии:

$$\beta_k = \beta_0 \lambda^k, k = 0,1\ldots, \tag{7.5}$$

где $0 < \lambda < 1$ характеризует скорость убывания коэффициентов с увеличением лага (с удалением от момента анализа). Такое предположение достаточно логично, если считать, что влияние прошлых значений объясняющих переменных на текущее значение зависимой переменной будет тем меньше, чем дальше по времени эти показатели имели место.

В данном случае (7.3) преобразуется в:

$$y_t = \alpha + \beta_0 x_t + \beta_0 \lambda x_{t-1} + \beta_0 \lambda^2 x_{t-2} + \ldots + \varepsilon_t. \tag{7.6}$$

Параметры уравнения (7.6) можно определять различными способами.

- Одним из них является перебор значений λ из интервала (0;1) с произвольным фиксированным шагом (например, 0,01; 0,001 и др.). Для каждого λ рассчитывается:

$$z_t = x_t + \lambda x_{t-1} + \lambda^2 x_{t-2} + \lambda^3 x_{t-3} + \ldots + \lambda^p x_{t-p}. \tag{7.7}$$

Значение p определяется из условия, что при дальнейшем добавлении лаговых значений x величина изменения z_t менее любого ранее заданного числа.

Далее оценивается уравнение регрессии:

$$y_t = \alpha + \beta_0 z_t + \varepsilon_t. \tag{7.8}$$

Из всех возможных значений λ выбирается то, при котором коэффициент детерминации R^2 для уравнения (7.8) будет наибольшим. Найденные при этом параметры α, β_0, λ подставляются в (7.6).

- Более распространенной является схема вычислений на основе преобразования Койка. Для этого определим уравнение (7.6), умноженное на λ и вычисленное для предыдущего периода времени:

$$\lambda y_{t-1} = \lambda\alpha + \beta_0 \lambda x_{t-1} + \beta_0 \lambda^2 x_{t-2} + \ldots + \lambda\varepsilon_t. \tag{7.9}$$

Из уравнения (7.6) вычтем уравнение (7.9):

$$y_t - \lambda y_{t-1} = (1-\lambda)\alpha + \beta_0 x_t + (\varepsilon_t - \lambda\varepsilon_{t-1}) \Rightarrow$$

$$\Rightarrow y_t = (1-\lambda)\alpha + \beta_0 x_t + \lambda y_{t-1} + v_t, \tag{7.10}$$

где $v_t = \varepsilon_t - \lambda\varepsilon_{t-1}$ – скользящая средняя между ε_t и ε_{t-1}.

Преобразование по данному методу уравнения (7.3) в уравнение (7.10) называется *преобразованием Койка*. Таким образом, с помощью данного преобразования уравнение с бесконечным числом лагов сведено к авторегрессионному, для которого требуется определить всего три коэффициента: λ, α, β_0.

Авторегрессионные модели

Рассмотрим два вида авторегрессионных моделей.

1. Модель адаптивных ожиданий

Ожидания играют существенную роль в экономической активности, что затрудняет моделирование соответствующих экономических процессов. Особенно серьезна эта проблема на макроэкономическом уровне. Например, при прогнозировании объема инвестиций требуется учитывать не только процентную ставку, но и экономическую политику государства, на основе которой потенциальные инвесторы принимают свои решения.

Одним из направлений решения рассматриваемой задачи является модель адаптивных ожиданий. В данной модели происходит постоянная корректировка ожиданий на основе получаемой информации о реализации исследуемого показателя. Если реальное значение показателя оказалось больше ожидаемого, то ожидаемое в следующем периоде корректируется в сторону увеличения. В противном случае – наоборот. При этом величина корректировки должна быть пропорциональна разности между реальным и ожидаемым значениями.

В данной модели в уравнение регрессии в качестве объясняющей переменной вместо текущего значения x_t вводится ожидаемое значение x_t^*:

$$y_t = \alpha + \beta x_t^* + \varepsilon_t. \tag{7.11}$$

Так как ожидаемые значения не являются фактически существующими, выдвигается предположение, что эти значения связаны следующим соотношением:

$$x_t^* - x_{t-1}^* = \gamma(x_t - x_{t-1}^*). \tag{7.12}$$

Модель (7.12) называется *моделью адаптивных ожиданий* (или моделью обучения на ошибках). Коэффициент $0 \le \gamma \le 1$ называется *коэффициентом ожидания*. Иногда в модели (7.12) вместо текущего значения x_t используют предыдущее значение x_{t-1}:

$$x_t^* - x_{t-1}^* = \gamma(x_{t-1} - x_{t-1}^*). \tag{7.13}$$

Перепишем соотношение (7.12) в виде:

$$x_t^* = \gamma x_t + (1-\gamma)x_{t-1}^*. \tag{7.14}$$

Из (7.14) видно, что ожидаемое значение x_t^* является взвешенным средним между текущим значением x_t и его ожидаемым значением в предыдущий период x_{t-1}^* с весами γ и $(1-\gamma)$ соответственно. Если $\gamma = 0$, то ожидания являются неизменными: $x_t^* = \gamma x_t + (1-\gamma)x_{t-1}^*$. Если $\gamma = 1$, то $x_t^* = x_t$, что означает мгновенно реализуемые ожидания.

Подставим (7.14) в (7.11), в результате чего получим:

$$y_t = \alpha + \beta(\gamma x_t + (1-\gamma)x_{t-1}^*) + \varepsilon_t. \tag{7.15}$$

Определим по (7.15) значение в предыдущий момент времени, умноженное на $(1-\gamma)$:

$$(1-\gamma)y_{t-1} = \alpha(1-\gamma) + \beta(1-\gamma)(\gamma x_t + (1-\gamma)x_{t-1}^*) + \varepsilon_t(1-\gamma). \tag{7.16}$$

Из (7.15) отнимем (7.16): $y_t - (1-\gamma)y_{t-1}$

$$\Rightarrow y_t = \alpha\gamma + \beta\gamma x_t + (1-\gamma)y_{t-1} + v_t, \tag{7.17}$$

где $v_t = \varepsilon_t - (1-\gamma)\varepsilon_{t-1}$.

На практике при оценивании параметров авторегрессионного уравнения (7.17) сначала оценивается параметр γ, затем коэффициент при x_t $(\beta = \dfrac{\beta\gamma}{\gamma})$, затем свободный член $(\alpha = \dfrac{\alpha\gamma}{\gamma})$.

Рассмотрим случай, когда зависимая переменная y_t в текущий момент времени связана с ожидаемым в следующий период времени значением x^*_{t+1} (например, зависимость спроса на деньги от ожидаемой процентной ставки), т.е.:

$$y_t = \alpha + \beta x^*_{t+1} + \varepsilon_t. \tag{7.18}$$

Допустим, что ожидаемое в следующий период времени значение переменной определяется как взвешенное среднее ее реального и ожидаемого значения в текущий период времени (аналогично 7.14):

$$x^*_{t+1} = \gamma x_t + (1-\gamma)x^*_t. \tag{7.19}$$

Следовательно, $x^*_t = \gamma x_{t-1} + (1-\gamma)x^*_{t-1}$. Отсюда (7.19) примет вид:

$x^*_{t+1} = \gamma x_t + (1-\gamma)(\gamma x_{t-1} + (1-\gamma)x^*_{t-1}) = \gamma x_t + \gamma(1-\gamma)x_{t-1} + (1-\gamma)^2 x^*_{t-1}$. Из (7.19) следует, что $x^*_{t-1} = \gamma x_{t-2} + (1-\gamma)x^*_{t-2}$, $x^*_{t-2} = \gamma x_{t-3} + (1-\gamma)x^*_{t-3}$ и т.д. С учетом этого (7.19) примет вид:

$$x^*_{t+1} = \gamma[x_t + (1-\gamma)x_{t-1} + (1-\gamma)^2 x_{t-2} + \ldots]. \tag{7.20}$$

Подставив (7.20) в (7.18), получаем:

$$y_t = \alpha + \beta\gamma[x_t + (1-\gamma)x_{t-1} + (1-\gamma)^2 x_{t-2} + \ldots] + \varepsilon_t. \tag{7.21}$$

Обозначив $\beta\gamma$ через β_0 и $(1-\gamma)$ через λ, получаем соотношение (7.6).

Рассмотрим пример модели авторегрессии:

$$\hat{C} = 3 + 0.85Y_t + 0.1C_{t-1},$$

где C – потребление в среднем на душу населения, Y – совокупный доход в среднем на душу населения.

Краткосрочный мультипликатор 0,85 отражает предельную склонность к потреблению в краткосрочном периоде, то есть при увеличении среднедушевого совокупного дохода на 1 тыс. руб., среднедушевое потребление увеличивается на 0,85 тыс. руб.

Долгосрочную предельную склонность к потреблению отражает долгосрочный мультипликатор:

$$b = b_0 + b_0C_1 + b_0C_1^2 + b_0C_1^3 + \ldots = b_0\left(1 + C_1 + C_1^2 + C_1^3 + \ldots\right) = \frac{b_0}{1-C_1}, \text{ т. к. } C_1 < 1.$$

Подставляя значения, получим $b = \dfrac{b_0}{1-C_1} = \dfrac{0.85}{1-0.1} = 0.944$.

2. Модель частичной корректировки

В этой модели в уравнение регрессии в качестве зависимой переменной входит не фактическое значение y_t, а желаемое значение y_t^*:

$$y_t^* = \alpha + \beta x_t + \varepsilon_t. \qquad (7.22)$$

Так как гипотетическое значение y_t^* не является фактически существующим, то относительно него выдвигается предположение *частичной корректировки*:

$$y_t - y_{t-1} = \lambda(y_t^* - y_{t-1}), \qquad (7.23)$$

по которому фактическое приращение зависимой переменной пропорционально разнице между ее желаемым значением и значением в предыдущий период. Коэффициент $0 \le \lambda \le 1$ называется *коэффициентом корректировки*. Уравнение (7.23) преобразуется к виду:

$$y_t = \lambda y_t^* + (1-\lambda) y_{t-1}. \qquad (7.24)$$

Подставив (7.22) в (7.24), получаем *модель частичной корректировки*:

$$y_t = \lambda(\alpha + \beta x_t + \varepsilon_t) + (1-\lambda) y_{t-1} \Rightarrow y_t = \lambda\alpha + \lambda\beta x_t + (1-\lambda) y_{t-1} + \lambda\varepsilon_t. \qquad (7.25)$$

Очевидно, что, чем больше λ, тем быстрее идет корректировка. При $\lambda = 1$ полная корректировка происходит за один период. При $\lambda = 0$ корректировка не происходит вовсе.

Прогнозирование с помощью временных рядов

Конечной целью статистического анализа временных рядов является прогнозирование будущих значений исследуемого показателя. Различают долгосрочное и краткосрочное прогнозирование. В первом анализируется долговременная динамика исследуемого процесса, и в этом случае главным считается выделение общего направления его изменения (тренда). Для предсказания краткосрочных колебаний проводится более детальный регрессионный анализ с целью выявления большого числа показателей, определяющих поведение исследуемой величины.

Пусть оцениваются параметры модели вида $\hat{y}_t = b_0 + b_1 x_t$ в момент времени ($t + p$). Значение $\hat{y}_{t+p}$ – значение, полученное по уравнению регрессии, построенному по МНК. Тогда доверительный интервал для действительного значения y_{t+p} имеет вид:

$$\left(\hat{y}_{t+p} \pm t_{\frac{\alpha}{2}, n-2} S_e \sqrt{1 + \frac{1}{n} + \frac{(x_{t+p} - \bar{x})^2}{n S_x^2}} \right), \qquad (7.26)$$

где $t_{\frac{\alpha}{2}, n-2}$ – критическое значение, определяемое из приложения 1 для соответствующего уровня значимости α и числа степеней свободы $(n-2)$; $S_e = \sqrt{\frac{\sum e_i^2}{n-2}}$ – стандартная ошибка оценки (стандартная ошибка регрессии);

x_{t+p} – значение объясняющей переменной в момент $(t+p)$; $S_x^2 = \frac{\sum(x_i - \bar{x})^2}{n}$ – дисперсия переменной x_t.

После получения прогнозных значений необходимо проверить качество прогноза. Для этого используются следующие показатели:

- Относительная ошибка прогноза, вычисляемая по формуле:

$$\delta_{t+p} = \frac{\hat{y}_{t+p} - y_{t+p}}{y_{t+p}} \cdot 100\% \qquad (7.27)$$

или

$$\delta_{t+p} = \frac{\Delta\hat{y}_{t+p} - \Delta y_{t+p}}{\Delta y_{t+p}} \cdot 100\%, \qquad (7.28)$$

где $\Delta\hat{y}_{t+p} = \hat{y}_{t+p} - \hat{y}_t$, $\Delta y_{t+p} = y_{t+p} - y_t$.

Чем больше значение ошибки (выраженное в процентах), тем хуже качество прогноза.

- Стандартная среднеквадратическая ошибка, рассчитываемая по формуле:

$$U = \frac{\frac{1}{k}\sum_{p=1}^{k}(\Delta\hat{y}_{t+p} - \Delta y_{t+p})^2}{\frac{1}{k}\sum \Delta y_{t+p}^2}, \qquad (7.29)$$

где k – количество прогнозных периодов.

Значения показателя U лежат в интервале от нуля до единицы. При $U = 0$ прогноз абсолютно точен. Таким образом, чем ближе значение U к нулю, тем точнее прогноз.

<u>Пример.</u> В таблице приведены данные по располагаемому доходу домохозяйств (X) и затратами домохозяйств на розничные покупки (Y) за 22 года

Доходы и затраты домашних хозяйств

t	Y_t	X_t
1	5,49	9,098
2	5,54	9,137
3	5,305	9,095
4	5,505	9,28
5	5,42	9,23
6	5,32	9,348
7	5,54	9,525
8	5,69	9,755
9	5,87	10,28
10	6,157	10,665
11	6,342	11,02

t	Y_t	X_t
12	5,905	11,305
13	6,125	11,43
14	6,185	11,45
15	6,225	11,697
16	6,495	11,87
17	6,72	12,018
18	6,92	12,525
19	6,47	12,055
20	6,395	12,088
21	6,555	12,215
22	6,755	12,495

Необходимо оценить уравнение регрессии вида $\hat{y}_t = b + b_1 x_t + \gamma\, y_{t-1}$ (принять $y_0 = 5{,}4$), проверить значимость коэффициентов b_0, b_1, γ, оценить качество построенной модели при помощи коэффициента детерминации.

Для расчета коэффициентов составим вспомогательную таблицу (при этом рассчитанные средние значения равны $\bar{y}_t = 6{,}04223$, $\bar{y}_{t-1} = 5{,}98067$, $\bar{x} = 10{,}79914$).

Расчетная таблица

t	Y_t	X_t	Y_{t-1}	$(y_t - \bar{y}_t)$	$(x_t - \bar{x})$	$(y_{t-1} - \bar{y}_{t-1})$	$(y_t - \bar{y}_t)^2$	$(x_t - \bar{x})^2$	$(y_{t-1} - \bar{y}_{t-1})^2$	$(x_t - \bar{x}) \times$ $\times (y_t - \bar{y}_t)$	$(y_{t-1} - \bar{y}_{t-1}) \times$ $\times (y_i - \bar{y}_t)$	$(x_t - \bar{x}) \times$ $\times (y_{t-1} - \bar{y}_{t-1})$
1	5,49	9,098	5,4	-0,5522	-1,7011	-0,581	0,3050	2,8939	0,3371	0,9394	0,3206	0,9877
2	5,54	9,137	5,49	-0,5022	-1,6621	-0,491	0,2522	2,7627	0,2407	0,8348	0,2464	0,8155
3	5,305	9,095	5,54	-0,7372	-1,7041	-0,441	0,5435	2,9041	0,1942	1,2563	0,3248	0,7509
4	5,505	9,28	5,305	-0,5372	-1,5191	-0,676	0,2886	2,3078	0,4565	0,8161	0,3630	1,0264
5	5,42	9,23	5,505	-0,6222	-1,5691	-0,476	0,3872	2,4622	0,2262	0,9764	0,2960	0,7463
6	5,32	9,348	5,42	-0,7222	-1,4511	-0,561	0,5216	2,1058	0,3143	1,0481	0,4049	0,8136
7	5,54	9,525	5,32	-0,5022	-1,2741	-0,661	0,2522	1,6234	0,4364	0,6399	0,3318	0,8417
8	5,69	9,755	5,54	-0,3522	-1,0441	-0,441	0,1241	1,0902	0,1942	0,3678	0,1552	0,4601
9	5,87	10,28	5,69	-0,1722	-0,5191	-0,291	0,0297	0,2695	0,0845	0,0894	0,0501	0,1509
10	6,157	10,665	5,87	0,1148	-0,1341	-0,111	0,0132	0,0180	0,0122	-0,0154	-0,0127	0,0148
11	6,342	11,02	6,157	0,2998	0,2209	0,176	0,0899	0,0488	0,0311	0,0662	0,0529	0,0390
12	5,905	11,305	6,342	-0,1372	0,5059	0,361	0,0188	0,2559	0,1306	-0,0694	-0,0496	0,1828
13	6,125	11,43	5,905	0,0828	0,6309	-0,076	0,0069	0,3980	0,0057	0,0522	-0,0063	-0,0477
14	6,185	11,45	6,125	0,1428	0,6509	0,144	0,0204	0,4236	0,0208	0,0929	0,0206	0,0940
15	6,225	11,697	6,185	0,1828	0,8979	0,204	0,0334	0,8062	0,0418	0,1641	0,0374	0,1835
16	6,495	11,87	6,225	0,4528	1,0709	0,244	0,2050	1,1467	0,0597	0,4849	0,1106	0,2617
17	6,72	12,018	6,495	0,6778	1,2189	0,514	0,4594	1,4856	0,2646	0,8261	0,3486	0,6269
18	6,92	12,525	6,72	0,8778	1,7259	0,739	0,7705	2,9786	0,5467	1,5149	0,6490	1,2760
19	6,47	12,055	6,92	0,4278	1,2559	0,939	0,1830	1,5772	0,8824	0,5372	0,4018	1,1797
20	6,395	12,088	6,47	0,3528	1,2889	0,489	0,1244	1,6612	0,2395	0,4547	0,1726	0,6307
21	6,555	12,215	6,395	0,5128	1,4159	0,414	0,2629	2,0047	0,1717	0,7260	0,2125	0,5867
22	6,755	12,495	6,555	0,7128	1,6959	0,574	0,5080	2,8760	0,3299	1,2088	0,4094	0,9740
Σ							5,3998	34,1000	5,2208	13,0114	4,8397	12,5953

По формулам

$$b_0 = \bar{y} - b_1\bar{x}_1 - b_2\bar{x}_2;$$

$$b_1 = \frac{\sum(x_{i1}-\bar{x}_1)(y_i-\bar{y})\cdot\sum(x_{i2}-\bar{x}_2)^2 - \sum(x_{i2}-\bar{x}_2)(y_i-\bar{y})\cdot\sum(x_{i1}-\bar{x}_1)(x_{i2}-\bar{x}_2)}{\sum(x_{i1}-\bar{x}_1)^2\sum(x_{i2}-\bar{x}_2)^2 - (\sum(x_{i1}-\bar{x}_1)(x_{i2}-\bar{x}_2))^2};$$ (см.

$$b_2 = \frac{\sum(x_{i2}-\bar{x}_2)(y_i-\bar{y})\cdot\sum(x_{i1}-\bar{x}_1)^2 - \sum(x_{i1}-\bar{x}_1)(y_i-\bar{y})\cdot\sum(x_{i1}-\bar{x}_1)(x_{i2}-\bar{x}_2)}{\sum(x_{i1}-\bar{x}_1)^2\sum(x_{i2}-\bar{x}_2)^2 - (\sum(x_{i1}-\bar{x}_1)(x_{i2}-\bar{x}_2))^2}.$$

главу 2) находим параметры уравнения:

$$b_1 = \frac{(13{,}0114)\cdot 5{,}2208 - 4{,}8397\cdot 12{,}5953}{34{,}100\cdot 5{,}2208 - (12{,}5953)^2} = \frac{6{,}9724}{19{,}3877} = 0{,}36;$$

$$\gamma = \frac{4{,}8397\cdot 34{,}10 - 13{,}0114\cdot 12{,}5953}{34{,}100\cdot 5{,}2208 - (12{,}5953)^2} = \frac{1{,}1513}{19{,}3877} = 0{,}06;$$

$$b_0 = 6{,}04223 - 0{,}36\cdot 10{,}79914 - 0{,}06\cdot 5{,}98067 = 1{,}80.$$

Таким образом, уравнение регрессии с учетом рассчитанных коэффициентов примет вид: $\hat{y}_t = 1{,}8 + 0{,}36x_t + 0{,}06y_{t-1}$.

Для определения статистической значимости коэффициентов и оценки качества уравнения регрессии составим следующую вспомогательную таблицу:

t	Y_t	$\hat{Y}_t$	e_i	e_i^2
1	5,49	5,3960	-0,0940	0,008843
2	5,54	5,4153	-0,1247	0,015542
3	5,305	5,4032	0,0982	0,009642
4	5,505	5,4558	-0,0492	0,002422
5	5,42	5,4497	0,0297	0,00088
6	5,32	5,4871	0,1671	0,02791
7	5,54	5,5448	0,0048	2,29E-05
8	5,69	5,6406	-0,0494	0,002444
9	5,87	5,8383	-0,0317	0,001006
10	6,157	5,9874	-0,1696	0,028757
11	6,342	6,1321	-0,2099	0,044047
12	5,905	6,2456	0,3406	0,11601
13	6,125	6,2646	0,1396	0,019496
14	6,185	6,2849	0,0999	0,009975
15	6,225	6,3773	0,1523	0,023186
16	6,495	6,4419	-0,0531	0,002824
17	6,72	6,5111	-0,2089	0,043635
18	6,92	6,7068	-0,2132	0,045453
19	6,47	6,5496	0,0796	0,006342
20	6,395	6,5348	0,1398	0,019545
21	6,555	6,5760	0,0210	0,000442
22	6,755	6,6862	-0,0688	0,00473
сумма			≈ 0	0,433155

По формулам (2.19) и (2.20) рассчитаем необъясненную дисперсию и стандартные отклонения случайных величин:

$$S^2 = \frac{\sum e_i^2}{n-m-1} = \frac{0,43155}{22-2-1} = 0,0227.$$

$$S_{b0}^2 = \left[\frac{1}{22} + \frac{(10,79914)^2 \cdot 5,2208 + (5,98067)^2 \cdot 34,1 - 2 \cdot 10,79914 \cdot 5,98067 \cdot 12,5953}{34,1 \cdot 5,2208 - (12,5953)^2}\right] \cdot 0,0227.$$

$$S_{b0} = \sqrt{S_{b0}^2} = \sqrt{0,238} = 0,4879.$$

$$S_{b1}^2 = \frac{5,2208}{34,1 \cdot 5,2208 - (12,5953)^2} \cdot 0,0227 = 0,0061.$$

$$S_{b1} = \sqrt{S_{b1}^2} = \sqrt{0,0061} = 0,0781.$$

$$S_{\gamma}^2 = \frac{34,10}{34,1 \cdot 5,2208 - (12,5953)^2} \cdot 0,0227 = 0,0399.$$

$$S_{\gamma} = \sqrt{S_{\gamma}^2} = \sqrt{0,0399} = 0,1997.$$

Определим значение t-статистики для каждого из коэффициентов:

$$t_{b0} = \frac{1,8}{0,4879} = 3,689, \quad t_{b1} = \frac{0,36}{0,0781} = 4,609, \quad t_{\gamma} = \frac{0,06}{0,1997} = 0,300.$$

Критическое значение определим из приложения 1 для уровня значимости 0,1 и числа степеней свободы $v = 22 - 2 - 1 = 19$: $t_{кр} = t_{\frac{0,1}{2},19} = 1{,}729$.

Очевидно, что коэффициенты b_0, b_1 являются статистически значимыми, а коэффициент γ является статистически незначимым с уровнем значимости 0,1.

Определим для рассчитанного уравнения коэффициент детерминации по формуле $R^2 = 1 - \frac{\sum e_i^2}{\sum (y_i - \bar{y})^2}$. $\qquad R^2 = 1 - \frac{0{,}433155}{5{,}3998} = 0{,}92$.

Столь высокое значение коэффициента детерминации свидетельствует о высоком качестве модели. Поэтому не следует удалять переменную y_{t-1} из уравнения.

Представим графически зависимость фактической переменной y_t и переменной $\hat{y}_t$ от *t* (рис. 7.1).

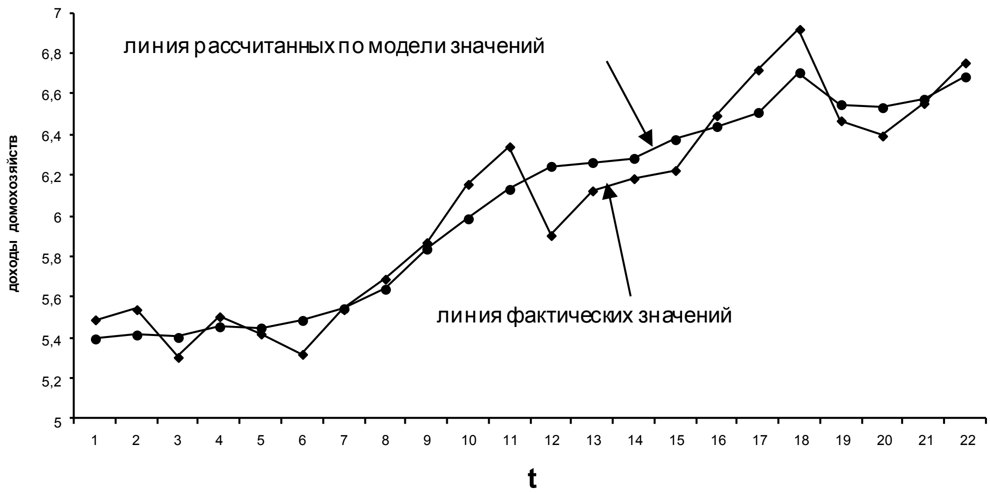


Рис. 7.1. Графики реальных и модельных значений дохода

Тема 8. Экономическая оценка инвестиций

Для определения экономической эффективности инвестиций рассчитывают следующие показатели:

- Чистая текущая стоимость (NPV)

$$NPV = \sum_{t=1}^{n} \frac{D_t}{(1+i)^t} - I$$

где D_t - чистые денежные поступления в t-ом году;

I – сумма инвестиций;

n – продолжительность проекта.

Показатель *NPV* позволяет определить стоимость, которую имеют капитальные вложения для инвестора, он дает вероятностную оценку прироста стоимости предприятия, в полной мере отвечает основной цели предприятия (наращивание экономического потенциала и рост благосостояния предприятия и его акционеров).

▪ Срок окупаемости инвестиций (к)

Определяется из условия: $\sum_{t=1}^{k} \frac{D_t}{(1+i)^t} = I$

▪ Индекс доходности (PI)

$$PI = \sum_{t=1}^{n} \frac{D_t}{(1+i)^t} \ / \ I$$

Взаимосвязь *NPV* и *PI* может свидетельствовать и о стоимости капитальных вложений и о привлекательности дохода, полученного от конкретной суммы вложенных средств

▪ Внутренняя норма доходности (прибыли) (IRR) – представляет собой ту ставку дисконтирования, при которой чистая текущая стоимость проекта равна нулю (все затраты, с учетом временной стоимости денег, окупаются). Определяется из условия:

$$\sum_{t=1}^{n} \frac{D_t}{(1+i)^t} = I \ .$$

Для определения IRR графическим методом нужно:

- Задать некую ставку дисконтирования и определить NPV проекта. Отметить соответствующую точку на графике (по оси ординат – ставки доходности, дисконтирования IRR, по оси абсцисс – NPV).
- Задать гораздо большую ставку дисконтирования (тогда NPV резко уменьшится), вычислить NPV и отметить соответствующую точку на графике.
- Соединить данные две точки и, если необходимо, продлить кривую NPV до пересечения с осью IRR. В точке пересечения кривой NPV с осью IRR чистая настоящая стоимость проекта равна нулю.

Внутренняя норма доходности *IRR* сравнивается с показателем *WACC:* если *IRR* > *WACC* – проект следует принять; если *IRR* < *WACC* - отвергнуть; при *IRR* = *CC* – проект не является ни прибыльным, ни убыточным.

Результаты проведения финансового анализа являются основой принятия окончательных решений о включении инвестиционных проектов в портфель. Сравнивая полученные в ходе финансового анализа результаты с заранее определенными критериями отбора проектов в портфель, принимается инвестиционное решение. Критерии отбора каждое предприятие устанавливает для себя произвольно, исходя из конкретных условий производственной и инвестиционной деятельности. В основе выбора критерия могут лежать такие признаки, как классификационная категория, к которой отнесены проекты; тип используемого финансового анализа; длительность проектов; эффективность (доходность) и ликвидность проектов; степень их важности для предприятия; стоимость капитала для финансирования проектов и уровень их риска и т.д.

Большое значение для установления критерия имеет правильный выбор нужного показателя и точное определение его порогового значения, являюще-

гося для предприятия мерой оценки капитальных вложений. Основная проблема при отборе инвестиционных проектов в портфель заключается в противоречивости этих показателей. Даже при принятии решения в отношении единичного проекта могут возникнуть диаметрально противоположные выводы о приемлемости инвестиций, основанные на том, какой показатель взят за основу.
Если установление критериев приемлемости и финансовый анализ производятся одним и тем же методом, инвестиционные решения, принятые на их основе, различаться не будут. Тем не менее, предприятия могут устанавливать различные пороговые значения этих критериев, поэтому возможность возникновения противоречий не исключена. Взаимосвязи между критериями, основанными на экономическом (финансовом) подходе к анализу и оценке капитальных вложений, намного сложнее. Если в портфель объединяются независимые проекты, то противоречий между рассчитываемыми показателями *NPV, IP IRR,* как правило, не возникает. Все они дают одинаковые результаты относительно принятия инвестиционного решения. Взаимосвязь между критериями:

если $NPV > 0$, одновременно $IRR > WACC$ и $PI > 1$;
если $NPV < 0$, одновременно $IRR < WACC$ и $PI < 1$;
если $NPV = 0$, одновременно $IRR = WACC$ и $PI = 1$.

Инвестиционные проекты могут конкурировать между собой в силу того, что они являются взаимоисключающими. Обычно взаимоисключающими оказываются инвестиции, которые обеспечивают альтернативные способы достижения одного и того же результата или использование какого-либо ограниченного ресурса (только не денежных средств). Ограниченность финансовых ресурсов также делает невозможным включение в портфель одновременно всех приемлемых инвестиционных проектов и некоторые из них приходится либо отвергать, либо откладывать на будущее. В подобных ситуациях формирование портфеля требует ранжирования проектов по степени их приоритетности (независимо от того, являются они независимыми или взаимоисключающими), а затем отбора в зависимости от используемого критерия.

Каждый из существующих критериев отбора проектов в портфель имеет свои преимущества и недостатки.

В случае противоречия критериев предпочтение следует отдавать критерию, основанному на значениях показателя *NPV*. При определении *NPV* не приходится сталкиваться с вычислительными проблемами, как при расчете *IRR*. Однако, несмотря на то, что по многим сравнительным параметрам приоритетотдается *NPV,* практики часто предпочитают критерий, основанный на значениях показателя *IRR*. Это объясняется относительностью показателя *IRR,* на основе которого легче принимать решение, тогда как *NPV* является абсолютным показателем.

В сравнительном анализе альтернативных, взаимоисключающих проектов критерий *IRR* дает оценку, которая не всегда совпадает с результатами анализа *NPV,* поэтому может использоваться достаточно условно. В подобных ситуациях метод *IRR* может быть скорректирован. При анализе исходят из показателя *IRR,* рассчитанного из разницы денежных потоков каждого рассматриваемого проекта. Расчет *IRR* ведется для приростных показателей капитальных вложений и доходов по проекту. Если при этом *IRR >WA CC,* приростные показатели оправданы и целесообразно принять проект с более низким значением *IRR.*

Для стандартных, единичных проектов критерий *IRR* показывает лишь максимальный уровень затрат, допустимый для оцениваемого проекта. Если цена капитала, привлекаемого для финансирования альтернативных проектов, меньше значения *IRR* для них, выбор может быть сделан лишь с помощью дополнительных критериев.

Недостатком критерия *IRR* является невозможность с его помощью различать ситуации, когда цена капитала меняется. Расчет показателя *IRR* предполагает, что ставка дисконтирования будет постоянной во время всего срока жизни проекта. Критерий *IRR* совершенно непригоден для анализа инвестиционных проектов с неординарными денежными потоками. В этих случаях возникает множественность (или отсутствие вовсе) значений *IRR* и неочевидность экономической интерпретации возникающих соотношений между *IRR* и ценой капитала.

Еще одна проблема метода *IRR* обусловлена допущением о реинвестициях. Модель расчета *IRR* предполагает, что все денежные потоки от осуществления проекта могут быть реинвестированы по ставке *IRR.* Однако это не реально. Реинвестирование потоков от капитальных вложений будет происходить по сложившейся на рынке инвестиционной процентной ставке, или по ставке дисконтирования, или в соответствии со стоимостью капитала. Поэтому метод *IRR* преувеличивает доход, который может быть получен от инвестиций.

При методе *NPV* подобных допущений не бывает. Возможность изменять ставку дисконтирования при расчетах, чтобы учесть изменяющиеся инвестиционные условия, делает метод *NPV* предпочтительнее. Однако ориентация на единственный критерий *NPV* также не всегда оправдана. Основной его недостаток в том, что это абсолютный показатель, который не дает информации о резерве безопасности проекта. Любая серьезная погрешность или ошибка в прогнозе денежного потока, в выборе ставки дисконтирования может привести к тому, что проект, который ранее рассматривался как прибыльный, окажется убыточным.

Информацию о резерве безопасности проекта дают критерии *IRR* и *PI.* При прочих равных условиях, чем больше *IRR* по сравнению с ценой капитала, тем больше резерв безопасности. Чем больше значение *PI* превосходит единицу, тем больше резерв безопасности.

С позиции риска можно сравнивать проекты по критериям *IRR* и *PI,* но нельзя — по критерию *NPV.* Высокое значение *NPV* также не может служить решающим аргументом при принятии решений, так как оно, во-первых, определяется масштабом проекта, а во-вторых, может быть сопряжено с достаточно высоким риском и степень риска, присущая этому значению *NPV,* не ясна.

При отборе инвестиционных проектов в портфель существует ряд практических аспектов, которые обязательно должны учитываться: следует принимать во внимание действие налогов, норм амортизации, а также необходимо учитывать инфляционные процессы. Еще на стадии первичного рассмотрения проектов заведомо негодными обычно признаются проекты, рентабельность (норма прибыли) которых ниже уровня инфляции. Такие проекты не обеспечивают предприятию противоинфляционной защиты.

Универсальная формула вычисления чистой текущей стоимости (NPV) проекта, позволяющая оценить эту величину в случае неодинакового инфляционного искажения доходов и затрат. Формула удобна тем, что позволяет одновременно производить и инфляционную коррекцию денежных потоков, и дисконтирование на основе средневзвешенной стоимости капитала, включающей инфляционную премию.

$$NPV = \sum_{i=1}^{n} \frac{\left[R_t / \prod_{r=1}^{t} (1 + i_r) - C_t / \prod_{r=1}^{t} \left(1 + i_r^{/}\right) \right] (1 - T) + D_t \cdot T}{(1 + k)^t} - I_0,$$

где

R_t – номинальная выручка t-го года, оцененная для безинфляционной ситуации, т. е. в ценах базового периода;

i_r – темпы инфляции доходов r-го года;

C_t – номинальный денежные затраты t-го года в ценах базового периода;

$i_r^{/}$ – темпы инфляции издержек r-го года;

T – ставка налогообложения прибыли;

I_0 – первоначальные затраты на приобретение основных средств;

к – средневзвешенная стоимость капитала, включающая инфляционную премию;

D_t – амортизационные отчисления t-го года

В выборе того или иного проекта на практике не всегда руководствуются критерием «внутренняя ставка рентабельности должна быть выше средневзвешенной стоимости капитала». Существует целый ряд проектов, осуществление которых диктуется экологической необходимостью или мотивировано повышением безопасности труда. От подобных проектов трудно, да и не следует ожидать значительных чистых денежных потоков. Но тогда доходы от остальных проектов предприятия должны обеспечить такую IRR, чтобы компенсировать пониженные денежные потоки или даже убытки от нерентабельных проектов.

Например: предприятие инвестирует 10 млрд. руб., из них 2 млрд. руб. – в необходимые, но нерентабельные проекты. Если средневзвешенная стоимость капитала равна 15%, то 8 млн. руб. инвестиций должны обеспечить не менее 1,5 млн. руб. чистых денежных потоков в год (15% на все 10 млн. руб. инвестиций), т. е. использоваться с рентабельностью не менее 18,75%.

Тема 9. Риск-менеджмент

Финансовый менеджмент всегда ставит получение дохода в зависимость от риска. Риск и доход представляют собой две взаимосвязанные и взаимообусловленные финансовые категории.

Под риском понимается возможная опасность потерь, вытекающая из специфики тех или иных явлений природы и видов деятельности человека.

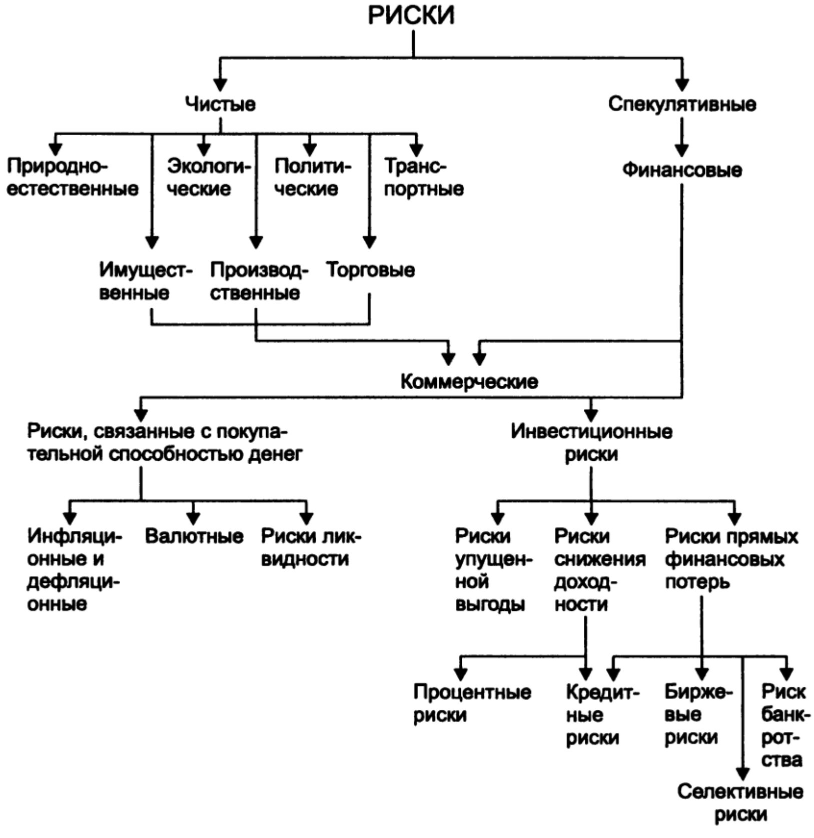
Для финансового менеджера <u>*риск*</u> — это вероятность неблагоприятного исхода. Различные инвестиционные проекты имеют различную степень риска, самый высокодоходный вариант вложения капитала может оказаться настолько рискованным, что, как говорится, «игра не стоит свеч».

Конечно, риска можно избежать, т. е. просто уклониться от мероприятия, связанного с риском. Однако для предпринимателя избежание риска зачастую означает отказ от возможной прибыли.

Риском можно управлять, т. е. использовать различные меры, позволяющие в определенной степени прогнозировать наступление рискового события и принимать меры к снижению степени риска. Эффективность организации управления риском во многом определяется классификацией риска.

Под классификацией рисков следует понимать их распределение на отдельные группы по определенным признакам для достижения определенных целей. Научно обоснованная классификация рисков позволяет четко определить место каждого риска в их общей системе. Она создает возможности для эффективного применения соответствующих методов и приемов управления риском. Каждому риску соответствует свой прием управления риском.

Квалификационная система рисков включает в себя категории, группы, виды, подвиды и разновидности рисков.



<u>*В зависимости от возможного результата*</u> (рискового события) риски можно подразделить на две большие группы: чистые и спекулятивные.

Чистые риски означают возможность получения отрицательного или нулевого результата. К этим рискам относятся: природно-естественные, экологические, политические, транспортные и часть коммерческих рисков (имущественные, производственные, торговые).

Спекулятивные риски выражаются в возможности получения как положительного, так и отрицательного результата. К ним относятся финансовые риски, являющиеся частью коммерческих рисков.

В зависимости от основной причины возникновения (базисный или природный признак), риски делятся на следующие категории: природно-естественные, экологические, политические, транспортные и коммерческие.

К *природно-естественным* относятся риски, связанные с проявлением стихийных сил природы: землетрясение, наводнение, буря, пожар, эпидемия и т. п.

Экологические риски — это риски, связанные с загрязнением окружающей среды.

Политические риски связаны с политической ситуацией в стране и деятельностью государства. Политические риски возникают при нарушении условий производственно-торгового процесса по причинам, непосредственно не зависящим от хозяйствующего субъекта. К политическим рискам относятся: невозможность осуществления хозяйственной деятельности вследствие военных действий, революции, обострения внутриполитической ситуации в стране, национализации, конфискации товаров и предприятий, введения эмбарго, из-за отказа нового правительства выполнять принятые предшественниками обязательства и т. п.; введение отсрочки (моратория) на внешние платежи на определенный срок ввиду наступления чрезвычайных обстоятельств (забастовка, война и т. д.); неблагоприятное изменение налогового законодательства; запрет или ограничение конверсии национальной валюты в валюту платежа. В этом случае обязательство перед экспортерами может быть выполнено в национальной валюте, имеющей ограниченную сферу применения.

Транспортные риски — это риски, связанные с перевозками грузов транспортом: автомобильным, морским, речным, железнодорожным, самолетами и т. д.

Коммерческие риски представляют собой опасность потерь в процессе финансово-хозяйственной деятельности. Они означают неопределенность результата от данной коммерческой сделки.

По структурному признаку коммерческие риски делятся на имущественные, производственные, торговые, финансовые.

Имущественные риски — это риски, связанные с вероятностью потерь имущества гражданина-предпринимателя по причине кражи, диверсии, халатности, перенапряжения технической и технологической систем и т. п.

Производственные риски — это риски, связанные с убытком от остановки производства вследствие воздействия различных факторов и, прежде всего, с гибелью или повреждением основных и оборотных фондов (оборудование, сырье, транспорт и т. п.), а также риски, связанные с внедрением в производство новой техники и технологии.

Торговые риски представляют собой риски, связанные с убытком по причине задержки платежей, отказа от платежа в период транспортировки товара, непоставки товара и т. п.

Финансовые риски связаны с вероятностью потерь финансовых ресурсов (т. е. денежных средств).

Финансовые риски подразделяются на два вида:

1) риски, связанные с покупательной способностью денег;

2) риски, связанные с вложением капитала (инвестиционные риски).

К рискам, связанным с покупательной способностью денег, относятся следующие разновидности рисков: инфляционные и дефляционные риски, валютные риски, риск ликвидности.

Инфляционный риск — это риск того, что при росте инфляции, получаемые денежные доходы обесцениваются с точки зрения реальной покупательной способности быстрее, чем растут. В таких условиях предприниматель несет реальные потери.

Дефляционный риск — это риск того, что при росте дефляции происходит падение уровня цен, ухудшение экономических условий предпринимательства и снижение доходов.

Валютные риски представляют собой опасность валютных потерь, связанных с изменением курса одной иностранной валюты по отношению, к другой при проведении внешнеэкономических, кредитных и других валютных операций.

Риски ликвидности — это риски, связанные с возможностью потерь при реализации ценных бумаг или других товаров из-за изменения оценки их качества и потребительной стоимости.

Инвестиционные риски включают в себя следующие подвиды рисков:

1) риск упущенной выгоды;

2) риск снижения доходности;

3) риск прямых финансовых потерь.

Риск упущенной выгоды – это риск наступления косвенного (побочного) финансового ущерба (неполученная прибыль) в результате неосуществления какого-либо мероприятия (например, страхование, хеджирование, инвестирование т. п.).

Риск снижения доходности может возникнуть в результате уменьшения размера процентов и дивидендов по портфельным инвестициям, по вкладам и кредитам. Риск снижения доходности включает в себя следующие разновидности: процентные риски и кредитные риски.

К *процентным рискам* относится опасность потерь коммерческими банками, кредитными учреждениями, инвестиционными институтами в результате превышения процентных ставок, выплачиваемых ими по привлеченным средствам, над ставками по предоставленным кредитам. К процентным рискам относятся также риски потерь, которые могут понести инвесторы в связи с изменением дивидендов по акциям, процентных ставок на рынке по облигациям, сертификатам и другим ценным бумагам.

Кредитный риск – опасность неуплаты заемщиком основного долга и процентов, причитающихся кредитору. К кредитному риску относится также риск такого события, при котором эмитент, выпустивший долговые ценные бумаги, окажется не в состоянии выплачивать проценты по ним или основную сумму долга.

Кредитный риск может быть также разновидностью рисков прямых финансовых потерь.

Риски прямых финансовых потерь включают в себя следующие разновидности: биржевой риск, селективный риск, риск банкротства, а также кредитный риск.

Биржевые риски представляют собой опасность потерь от биржевых сделок. К этим рискам относятся: риск неплатежа по коммерческим сделкам, риск неплатежа комиссионного вознаграждения брокерской фирмы и т. п.

Селективные риски – это риски неправильного выбора способа вложения капитала, вида ценных бумаг для инвестирования в сравнении с другими видами ценных бумаг при формировании инвестиционного портфеля.

Риск банкротства представляет собой опасность в результате неправильного выбора способа вложения капитала, полной потери предпринимателем собственного капитала и неспособности его рассчитываться по взятым на себя обязательствам. В результате предприниматель становится банкротом.

Финансовый риск представляет собой функцию времени. Как правило, степень риска для данного финансового актива или варианта вложения капитала увеличивается во времени.

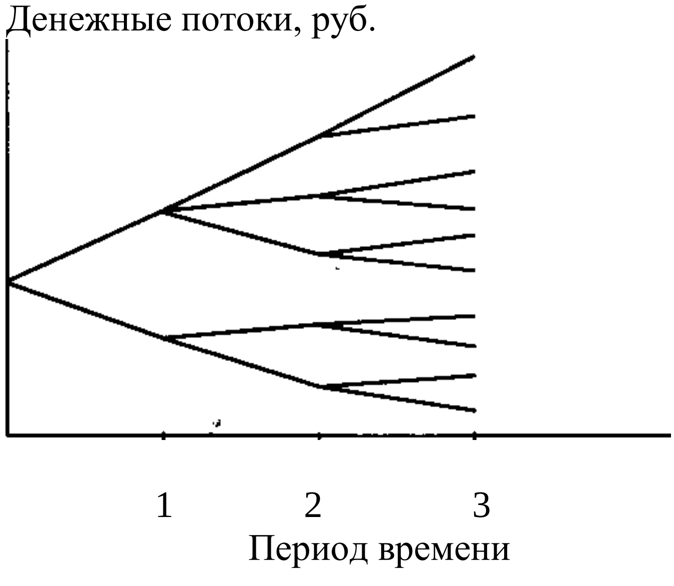
В зарубежной практике в качестве метода количественного определения риска вложения капитала предлагается использовать дерево вероятностей.

Этот метод позволяет точно определить вероятные будущие денежные потоки инвестиционного проекта в их связи с результатами предыдущих периодов времени. Если проект вложения капитала приемлем в первом периоде времени, то он может быть также приемлем и в последующих периодах времени.

Если же предполагается, что денежные потоки в разных периодах времени являются независимыми друг от друга, то необходимо определить вероятное распределение результатов денежных потоков для каждого периода времени.

В случае, когда связь между денежными потоками в разных периодах времени существует, необходимо принять данную зависимость и на ее основе представить будущие события так, как они могут произойти.

В качестве примера приведем дерево вероятностей для трех периодов времени.



Дерево вероятностей показывает, что если в периоде 1 результатом будет верхняя ветвь, то она приведет в периоде 2 к другому множеству возможных результатов, чем это было бы, если бы результат в периоде 1 выражался нижней ветвью. Аналогичная картина наблюдается и при переходе от периода времени 2 к периоду 3. Поэтому в момент временного периода 0 дерево вероятностей представляет наилучшую оценку того результата, который, вероятно, будет иметь место в будущем, в зависимости от того, что происходило прежде. Каждой ветви сопоставлена вероятность получения соответствующего результата.

Пример:

Фирма оценивает возможность производства нового товара А со сроком использования два года. Стоимость товара 1000 тыс. руб. Величина денежного потока зависит от спроса на данный товар. Дерево вероятностей возможных будущих денежных потоков, связанных с новым товаром, имеет вид:

Иллюстрация дерева вероятностей денежных потоков при производстве и реализации нового товара А

I год		II год		III год	
Исходная вероятность	Денежный поток (ожидаемый прогноз), тыс. руб.	Условная вероятность	Денежный поток (ожидаемый прогноз), тыс. руб.	№ ветви	Совместная вероятность
0,4	300	0,3	200	1	0,4 * 0,3 = 0,12
		0,7	400	2	0,4 * 0,7 = 0,28
0,6	600	0,4	600	3	0,6 * 0,4 = 0,24
		0,6	1000	4	0,6 * 0,6 = 0,36
					Итого 1,0

Основные методы оценки риска

Эффективность любой финансовой или хозяйственной операции и величина сопутствующего ей риска взаимосвязаны («за риск приплачивают»). Не учитывая фактора риска, невозможно провести полноценный инвестиционный анализ.

Независимо от происхождения и сущности риска, главнейшей цели бизнеса – получению дохода на вложенный капитал – соответствует следующее определение риска.

Риск – это возможность неблагоприятного исхода, т.е. неполучения инвестором ожидаемой прибыли.

Чем выше вероятность получения низкого дохода или даже убытков, тем рискованнее проект и тем выше должна быть норма его доходности.

Инструментом для проведения необходимых вычислений является математическая теория вероятностей. Каждому событию ставится в соответствие некоторая величина, характеризующую возможность того, что оно (событие) произойдет – *вероятность* данного события – р. Если в результате проведения эксперимента или наблюдения установлено, что некоторое событие происходит в n случаях из N, то ему приписывается вероятность p = n / N. Сумма вероятностей всех событий, которые могут произойти в результате некоторого эксперимента, должна быть равна единице. Перечисление всех возможных событий с соответствующими им вероятностями называется *распределением вероятностей* в данном эксперименте.

Вероятность может быть выражена в процентах: *p = (n/N)*100%*, тогда значение p может находиться в пределах от 0 до 100%.

<u>Пример:</u> Рассмотрим два финансовых проекта А и В, для которых возможные нормы доходности (IRR) находятся в зависимости от будущего состояния экономики. Данная зависимость отражена в следующей таблице:

Состояние экономики	Вероятность данного состояния	Проект А, IRR	Проект В, IRR
Подъем	P1 = 0,25	90%	25%
Норма	P2 = 0,5	20%	20%
Спад	P3 = 0,25	-50%	15%

Для каждого из проектов А и В может быть рассчитана ожидаемая норма доходности ERR – средневзвешенное (где в качестве весов берутся вероятности) возможных IRR.

$ERR = \sum p_i \, IRR_i$

Для проекта А получаем:

$ERR_A = 0{,}25 * 90\% + 0{,}5 * 20\% + 0{,}25 * (-50\%) = 20\%$

Для проекта В:

$ERR_B = 0{,}25 * 25\% + 0{,}5 * 20\% + 0{,}25 * 15\% = 20\%$

Таким образом, для двух рассматриваемых проектов ожидаемые нормы доходности совпадают, несмотря на то, что диапазон возможных значений IRR сильно различается: у проекта А от -50% до 90%, у проекта В – от 15% до 25%.

Также за достаточно корректную меру риска можно принять показатель *среднее квадратическое отклонение*

$$\sigma = \sqrt{\sum_i (IRR - ARR)^2 / n},$$

где $ARR = \sum_i^n IRR_i / n$ – средняя арифметическая норма доходности.

Чем меньше величина, тем менее рискован проект. Считается, что среднерискованной операции соответствует значение σ около 30%.

Еще одной величиной, характеризующей степень риска, является коэффициент CV:

$CV = \sigma / ERR$

<u>*Практический пример:*</u> Какой проект предпочтительнее?

Распределение вероятности для проектов C и D:

Состояние экономики	Вероятность состояния	Проект C, IRR	Проект D, IRR
Подъем	0.2	30%	25%
Норма	0.6	20%	32%
Спад	0.2	10%	8%

<u>*Практический пример:*</u> Какой проект предпочтительнее?

Доходность проектов А и В в динамике:

Год	Доходность проекта А	Доходность проекта В
2002	20%	40%
2003	15%	24%
2004	18%	30%
2005	23%	50%

Сущность, содержание и организация риск-менеджмента

Риск – это финансовая категория. Поэтому на степень и величину риска можно воздействовать через финансовый механизм. Такое воздействие осуществляется с помощью приемов финансового менеджмента и особой стратегии. В совокупности стратегия и приемы образуют своеобразный механизм управления риском, т. е. риск-менеджмент. Таким образом, риск-менеджмент представляет собой часть финансового менеджмента.

В основе риск-менеджмента лежат целенаправленный поиск и организация работы по снижению степени риска, искусство получения и увеличения дохода (выигрыша, прибыли) в неопределенной хозяйственной ситуации.

Конечная цель риск-менеджмента соответствует целевой функции предпринимательства. Она заключается в получении наибольшей прибыли при оптимальном, приемлемом для предпринимателя соотношении прибыли и риска.

Риск-менеджмент включает в себя стратегию и тактику управления.

Под *стратегией управления* понимаются направление и способ использования средств для достижения поставленной цели. Этому способу соответствует определенный набор правил и ограничений для принятия решения. Стратегия позволяет сконцентрировать усилия на вариантах решения, не противоречащих принятой стратегии, отбросив все другие варианты. После достижения поставленной цели стратегия как направление и средство ее достижения прекращает свое существование. Новые цели ставят задачу разработки новой стратегии.

Тактика – это конкретные методы и приемы для достижения поставленной цели в конкретных условиях. Задачей тактики управления является выбор оптимального решения и наиболее приемлемых в данной хозяйственной ситуации методов и приемов управления.

Риск-менеджмент выполняет следующие *функции:*

- прогнозирование;

- организация;
- регулирование;
- координация;
- стимулирование;
- контроль.

Прогнозирование в риск-менеджменте представляет собой разработку на перспективу изменений финансового состояния объекта в целом и его различных частей. Управление на основе предвидения изменений требует выработки у менеджера определенного чутья рыночного механизма и интуиции, а также применения гибких экстренных решений.

Организация в риск-менеджменте представляет собой объединение людей, совместно реализующих программу рискового вложения капитала на основе определенных правил и процедур

Регулирование в риск-менеджменте представляет собой воздействие на объект управления, посредством которого достигается состояние устойчивости этого объекта в случае возникновения отклонения от заданных параметров. Регулирование охватывает главным образом текущие мероприятия по устранению возникших отклонений.

Координация в риск-менеджменте представляет собой согласованность работы всех звеньев системы управления риском, аппарата управления и специалистов.

Стимулирование в риск-менеджменте представляет собой побуждение финансовых менеджеров и других специалистов к заинтересованности в результате своего труда.

Контроль в риск-менеджменте представляет собой проверку организации работы по снижению степени риска.

Сферой предпринимательской деятельности риск-менеджмента является страховой рынок.

Организация риск-менеджмента представляет собой систему мер, направленных на рациональное сочетание всех его элементов в единой технологии процесса управления риском.

Первым этапом организации риск-менеджмента является определение цели риска и цели рисковых вложений капитала. Цель риска — это результат, который необходимо получить. Им может быть выигрыш, прибыль, доход и т. п. Цель рисковых вложений капитала — получение максимальной прибыли.

Следующим важным моментом в организации риск-менеджмента является получение информации об окружающей обстановке, которая необходима для принятия решения в пользу того или иного действия.

Для предпринимателя важно знать действительную стоимость риска, которому подвергается его деятельность.

Под стоимостью риска следует понимать фактические убытки предпринимателя, затраты на снижение величины этих убытков или затраты по возмещению таких убытков и их последствий. Правильная оценка финансовым менеджером действительной стоимости риска позволяет ему объективно представлять объем возможных убытков и наметить пути к их предотвращению или уменьшению, а в случае невозможности предотвращения убытков обеспечить их возмещение. На основе имеющейся информации об окружающей среде, вероятности, степени и величине риска разрабатываются различные варианты рискового вложения капитала и проводится оценка их оптимальности путем сопоставления ожидаемой прибыли и величины риска.

Это позволяет правильно выбрать стратегию и приемы управления риском, а также способы снижения степени риска.

В риск-менеджменте готовых рецептов нет и быть не может. Он учит тому, как, зная методы, приемы, способы решения тех или иных хозяйственных задач, добиться ощутимого успеха в конкретной ситуации, сделав ее для себя более или менее определенной.

Основные правила риск-менеджмента:

- Нельзя рисковать больше, чем это может позволить собственный капитал.
- Надо думать о последствиях риска.
- Нельзя рисковать многим ради малого.
- Положительное решение принимается лишь при отсутствии сомнения.
- При наличии сомнений принимаются отрицательные решения.
- Нельзя думать, что всегда существует только одно решение. Возможно, есть и другие.
- Если сомневаешься, принимай отрицательное решение.

Соотношение максимально возможного объема убытка и объема собственных финансовых ресурсов инвестора представляет собой степень риска, ведущую к банкротству. Она измеряется с помощью коэффициента риска: Кр = У/С,

где Кр – коэффициент риска;

У – максимально возможная сумма убытка, руб.;

С – объем собственных финансовых ресурсов с учетом точно известных поступлений средств, руб.

При разработке программы действия по снижению риска необходимо учитывать психологическое восприятие рисковых решений. Принятие решений в условиях риска является психологическим процессом. Поэтому наряду с математической обоснованностью решений следует иметь в виду проявляющиеся при принятии и реализации рисковых решений психологические особенности человека: агрессивность, нерешительность, сомнения, самостоятельность, экстраверсию, интроверсию и др.

Важным этапом организации риск-менеджмента являются контроль за выполнением намеченной программы, анализ и оценка результатов выполнения выбранного варианта рискового решения.

Стратегия риск-менеджмента — это искусство управления риском в неопределенной хозяйственной ситуации, основанное на прогнозировании риска и приемов его снижения. Стратегия риск-менеджмента включает правила, на основе которых принимаются рисковое решение и способы выбора варианта решения.

Правила — это основополагающие принципы действия.

В стратегии риск-менеджмента применяются следующие *правила*:

1. Максимум выигрыша.
2. Оптимальная вероятность результата.
3. Оптимальная колеблемость результата.
4. Оптимальное сочетание выигрыша и величины риска.

Сущность правила максимума выигрыша заключается в том, что из возможных вариантов рисковых вложений выбирается вариант, дающий наибольшую эффективность результата (выигрыш, доход, прибыль) при минимальном или приемлемом для инвестора риске

Пример:

Владелец груза, который следует перевезти морским транспортом, знает, что в результате возможной гибели корабля он теряет груз стоимостью 100 млн. руб. Ему также известно, что вероятность кораблекрушения 0,05; страховой тариф при страховании груза составляет 3% от страховой суммы.

С учетом этих данных можно определить вероятность перевозки груза без кораблекрушения. Она равна (1 – 0,05) = 0,95.

Затраты владельца на страхование груза, т.е. его потери при перевозке груза без кораблекрушения, составляют (3 * 100 / 100) = 3 млн. руб.

Владелец груза стоит перед выбором: страховать или не страховать свой груз.

Для решения вопроса определим величину выигрыша владельца груза для двух вариантов его действия. При страховании владельцем груза его выигрыш составляет 97 млн. руб. = 0,05 * 100 + 0,95*100 + (–3). При отказе от страхования выигрыш составит 95 млн. руб. = 0,05(-100) + 0,95*100.

Владелец принимает решение страховать свой груз, так как оно обеспечивает ему наибольший выигрыш.

Сущность правила оптимальной вероятности результата состоит в том, что из возможных решений выбирается то, при котором вероятность результата является приемлемой для инвестора, т. е. удовлетворяет финансового менеджера.

Пример:

Имеются два варианта рискового вложения капитала. По первому варианту ожидается получить прибыль 1 млн. руб. при вероятности 0,9. По второму варианту ожидается получить прибыль 1,8 млн. руб. при вероятности 0,7. Сопоставление результатов двух вариантов показывает, что по второму варианту сумма больше на 80%, а вероятность ее получения ниже на 20%, чем в первом варианте. Менеджер выбирает второй вариант, считая, что он удовлетворяет его интересам.

Сущность правила оптимальной колеблемости результата заключается в том, что из возможных решений выбирается то, при котором вероятности выигрыша и проигрыша для одного и того же рискового вложения капитала имеют небольшой разрыв, т. е. наименьшую величину дисперсии, среднего квадратичного отклонения, вариации.

Пример:

Имеем два варианта рискового вложения капитала. В первом варианте с вероятностью 0,6 можно получить доход 1 млн. руб. и с вероятностью (1 – 0,6) = 0,4 получить убыток 0,6 млн. руб.

Во втором варианте с вероятностью 0,8 можно получить доход 2 млн. руб. и с вероятностью (1 – 0,8) = 0,2 получить убыток в 1,5 млн. руб.

Средний ожидаемый доход составляет по вариантам:

первый вариант

0,6 * 1 + 0,4*(-0,6) = 0,36 млн. руб.

второй вариант

0,8 * 2 + 0,2*(-1,5) = 1,3 млн. руб.

На первый взгляд, более доходным является второй вариант, так как при нем доходность на 261% выше, чем при первом варианте [(1,3-0,36)/0,36* 100].

Однако более глубокий анализ показывает, что первый вариант имеет определенные преимущества перед вторым вариантом, а именно:

1. Меньший разрыв показателей вероятности результатов. Этот разрыв в первом варианте составляет 0,2 или 50% = (0,6 – 0,4)/0,4 * 100,

во втором варианте – 0,6 или 300% = (0,8 – 0,2) / 0,2 * 100.

2. Незначительный разрыв вероятности выигрыша, т. е. получения дохода. Вероятность получения дохода во втором варианте – 0,8, а в первом варианте – 0,6, т. е. меньше всего на 25%= (100 – 0,6/0,8 * 100).

3. Более низкий темп изменения вероятности и суммы убытка по сравнению с темпом изменения вероятности и суммы дохода.

Так, во втором варианте по сравнению с первым вариантом при росте вероятности дохода с 0,6 до 0,8 (на 33%) сумма дохода возрастает на 100% (с 1 до 2 млн. руб.). В то же время при снижении вероятности убытка на 100% (с 0,4 до 0,2) сумма убытка увеличивается на 150% (с 0,6 до 1,5 млн. руб.).

Сущность правила оптимального сочетания выигрыша и величины риска заключается в том, что менеджер оценивает ожидаемые величины выигрыша и риска (проигрыша, убытка) и принимает решение вложить капитал в то мероприятие, которое позволяет получить ожидаемый выигрыш и одновременно избежать большого риска.

Пример:

Имеются два варианта рискового вложения капитала. При первом варианте доход составляет 10 млн. руб., а убыток – 3 млн. руб. По второму варианту доход составляет 15 млн. руб., а убыток – 5 млн. руб. Соотношение дохода и убытка позволяет сделать вывод в пользу принятия первого варианта вложения капитала, так как по первому варианту на 1 руб. убытка приходится 3,33 руб. дохода (10 / 3), по второму варианту – 3,0 руб. дохода (15 / 5).

Таким образом, если соотношение дохода и убытка по первому варианту составляет 3,3 : 1, то по второму варианту – 3 : 1. Делаем выбор в пользу первого варианта.

Правила принятия решения рискового вложения капитала дополняются способами выбора варианта решения. Существуют следующие способы выбора решения:

- Выбор варианта решения при условии, что известны вероятности возможных хозяйственных ситуаций.
- Выбор варианта решения при условии, что вероятности возможных хозяйственных ситуаций неизвестны, но имеются оценки их относительных значений.
- Выбор варианта решения при условии, что вероятности возможных хозяйственных ситуаций неизвестны, но существуют основные направления оценки результатов вложения капитала.