Федеральное государственное образовательное бюджетное учреждение высшего образования

«ФИНАНСОВЫЙ УНИВЕРСИТЕТ ПРИ ПРАВИТЕЛЬСТВЕ РОССИЙСКОЙ ФЕДЕРАЦИИ»

Кафедра системного анализа в экономике

И.В. Орлова, С.А. Рытиков, С.Е. Щепетова, Г.В. Росс, М.Г. Бич

ОСНОВЫ МАТЕМАТИЧЕСКОГО МОДЕЛИРОВАНИЯ СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ ПРОЦЕССОВ

Практикум. Часть 2

Учебное пособие

Под редакцией С.А. Рытикова

Москва 2016

УДК 519.86 (073)
ББК 65в631
О-66

Рецензенты:

д.э.н., профессор ***И.Н. Дрогобыцкий***; (Финансовый университет)

д.э.н. ***Н.Б. Кобелев,*** президент НП «Ремесленная палата России»

Орлова И.В., Рытиков С.А., Щепетова С.Е., Росс Г.В., Бич М.Г.

О-66 Основы математического моделирования социально-экономических процессов. Практикум. Часть 2: учебное пособие / под ред. С.А. Рытикова. — М.: Финансовый университет, 2016. — 132 с.

ISBN 978-5-7942-1310-2
ISBN 978-5-7942-1368-3 (ч. 2)

Пособие посвящено перспективным направлениям экономико-математического моделирования — применению эконометрических и эволюционно-симулятивных моделей и методов для решения наиболее важных и часто встречающихся на практике проблем экономики, маркетинга и менеджмента. Пособие имеет практическую направленность. Оно содержит все необходимое для изучения студентами данных типов моделей, а их применение разъясняется и иллюстрируется на практических примерах. Каждая глава завершается набором задач для самостоятельной работы, решение которых поможет овладеть практическими приемами, которые позволят в дальнейшем вести самостоятельные исследования, и способствует процессу подготовки студентов к экзаменам. В качестве инструментальных средств предлагается использовать пакет Microsoft Excel и диалоговую инструментальную систему «Decision».

Пособие ориентировано на студентов бакалавриата, обучающихся по направлениям «Государственное и муниципальное управление», «Экономика» и «Менеджмент», а также на специалистов в области экономико-математического моделирования.

УДК 519.86 (073)
ББК 65в631

ISBN 978-5-7942-1368-3 (ч. 2)
ISBN 978-5-7942-1310-2

**The Federal State-Funded Educational Institution
of Higher Education**

**«FINANCIAL UNIVERSITY UNDER THE GOVERNMENT
OF THE RUSSIAN FEDERATION»**

The Department of Systemic
Analysis and Economic

**I.V. Orlova, S.A. Rytikov, S.E. Schepetova,
G.V. Ross, M.G. Bich**

FUNDAMENTALS OF MATHEMATICAL MODELING OF SOCIO-ECONOMIC PROCESSES

Study Guide for Practical Work. Part 2

Edited by S.A. Rytikov

Moscow 2016

Referees:
Doctor of Economics, Professor ***I.N. Drogobytsky***; (Financial University)
Doctor of Economics ***N.B. Kobelev***, President of the Russian Chamber of Crafts

Orlova I.V., Rytikov S.A., Schepetova S.E., Ross G.V., Bich M.G.
Fundamentals of Mathematical Modeling of socio-economic processes. Study Guide for Practical Work. Part 2 / Edited by S.A. Rytikov. — M.: Financial University, 2016. — 132 p.

ISBN 978-5-7942-1310-2
ISBN 978-5-7942-1368-3 (p. 2)

The Study Guide is devoted to advanced direction of economic and mathematical methods — applications of econometric and evolutionarily-simulative modeling (ESM) techniques for solving the most important and frequently encountered in practice problems of economy, marketing and management. The Guide is practice-oriented and contains everything need for studying these types of models. Application of methods has explained and illustrated by practical examples. Each chapter ends with self-study exercises, it is solving help to inquire practice skills for future scientific research, and useful for exam preparation. Microsoft Excel and interactive tool system «Decision» are used as computer-based modeling tools.

The Study Guide is intended for students of Public and Municipal Administration, Economics and Management Bachelor Degree Programs as well as for research associates in the sphere of economic-mathematical methods.

ISBN 978-5-7942-1368-3 (p. 2)
ISBN 978-5-7942-1310-2

Введение

Учебное пособие «Основы математического моделирования социально-экономических процессов» посвящено изучению количественных и качественных экономических взаимосвязей с помощью перспективных методов экономико-математического моделирования социально-экономических систем. В нем содержится все необходимое для изучения студентами наиболее важных и часто встречающихся на практике проблем экономики, маркетинга и менеджмента. Отбор и изложение материала учебного пособия имеют некоторые особенности по сравнению с существующими учебниками и учебными пособиями, суть которых заключается в описании практических применений широкого спектра математических моделей социально-экономических систем.

Первая часть практикума была посвящена оптимизационным и балансовым моделям — задача линейного программирования, задача целочисленного программирования, транспортная задача, задача о назначениях, модели сетевого планирования и управления, матричные балансовые модели, модели одновременного инвестиционно-финансового планирования.

Вторая часть практикума рассматривает два других класса экономико-математических моделей: эконометрические модели, (глава 1 «Основы эконометрики») и имитационные (эволюционно-симулятивные) модели, представленные (глава 2 «Равновесные модели эволюционно-симулятивного метода»).

Глава 1 данного пособия включает описания типов данных, используемых в эконометрике, в ней подробно рассмотрены технологии эконометрического моделирования, количественного и графического анализа экономических данных в Microsoft Excel. Приведены конкретные примеры проведения корреляционно-регрессионного анализа с помощью различных функций Microsoft Excel и важнейших специальных инструментов — пакета **«Анализ данных»** и надстройки **«Поиск решения»**.

Особое внимание уделяется эконометрическим методам корректировки нарушений предпосылок Гаусса–Маркова и применению моделей с фиктивными переменными, которые позволяют учесть влияние качественных признаков при исследовании экономических объектов и являются эффективным инструментом при анализе структурных изменений, выборе альтернатив, построении рейтинговых систем.

Изложение материала сопровождается справочными материалами по основным формулам, используемым в корреляционном и регрессионном анализе и функциям Microsoft Excel. Примеры решения задач включают фрагмент или полный текст рабочего документа Microsoft Excel, снабженный комментариями и краткими указаниями, помогающими реализовать решение задачи на компьютере. Завершает главу 1 комплексный пример исследования экономических данных и предлагается набор задач для самостоятельной работы, решение которых помогает процессу подготовки студентов к экзаменам.

Глава 2 посвящена решению проблем эффективного применения экономико-математических методов в практическом бизнесе. В ней разъясняются и иллюстрируются на практических примерах методы, которые позволяют с новых позиций решать широкий круг прикладных экономических задач с учетом неопределенностей и рисков. Рассмотренный в главе 2 эволюционно-симулятивный метод позволяет за счет оптимизации снижать неопределенность на один–два порядка, что принципиально меняет ситуацию, делая теоретическую экономику эффективным инструментом практического бизнеса. В главе 2 дано описание компьютерной системы **«Decision»**, в рамках которой предлагается пример, иллюстрирующий процесс прогнозирования рынков товаров и услуг, и проведен расчет оценок рисков, надежности выполнения плана, емкости рынка, прибыли и других интегральных характеристик.

Пособие адресовано студентам, специализирующимся в области государственного и муниципального управления, финансов и кредита, экономики предприятия, менеджмента, маркетинга, инвестиций, математической статистики, управления проектами, управления в чрезвычайных ситуациях и т.д. Его использование не требует специальных математических или иных знаний, кроме полученных студентами на начальных курсах по указанным специальностям.

Глава 1

Эконометрические модели

1.1. Типы данных, используемых в эконометрике. Оценка тесноты линейной связи

Эконометрические модели отражают статистические закономерности, устанавливаемые экономической наукой, и могут применяться как на макро-, так и на микроуровне. Целью их использования являются количественный анализ и прогнозирование взаимосвязей показателей, описывающих экономический объект для подготовки и принятия обоснованных экономических решений.

При моделировании экономических процессов используют следующие типы данных.

Пространственные данные — характеризуют ситуацию по конкретной переменной (или набору переменных), относящейся к пространственно разделенным сходным объектам в один и тот же момент времени (табл. 1.1).

Таблица 1.1

Валовой региональный продукт в 2012 г.

Регион	**Валовой региональный продукт, млн руб.**	**Среднегодовая численность занятых в экономике, тыс. чел.**	**Инвестиции в основной капитал (в фактически действовавших ценах), млн руб.**
1. Тюменская область	4 618 711,00	1 963,00	1 439 576,00
3. Свердловская область	1 484 447,00	2 043,20	341 635,00
4. Республика Татарстан	1 436 933,00	1 821,80	464 745,00
5. Красноярский край	1 192 648,00	1 439,00	376 090,00
7. Республика Башкортостан	1 154 056,00	1 797,10	232 873,00
8. Самарская область	941 611,00	1 507,30	204 165,00
9. Пермская область	897 598,00	1 298,70	158 314,00
10. Челябинская область	843 340,00	1 672,90	179 728,00
11. Нижегородская область	838 559,00	1 703,20	258 176,00
12. Иркутская область	743 764,00	1 137,00	156 470,00
13. Кемеровская область	717 700,00	1 305,40	264 440,00

Временные ряды — отражают изменения (динамику) какой-либо переменой на промежутке времени (табл. 1.2).

Таблица 1.2

Численность населения России, млн чел.

Год	Численность населения (на конец года), млн чел.
2009	142,8
2010	142,9
2011	143,0
2012	143,3
2013	143,7
2014	146,3

Панельные данные — представляют собой прослеженные во времени пространственные выборки, которые состоят из наблюдений одних и тех же экономических объектов в последовательные периоды времени. Панельные данные состоят из трех измерений: признаки — объекты — время.

Переменные, участвующие в эконометрической модели любого типа, разделяются на следующие типы.

Результирующая (эндогенная, зависимая) переменная (*y*) — характеризует результат или эффективность функционирования экономической системы. Значения ее формируются в процессе и внутри функционирования этой системы под воздействием ряда других переменных и факторов, часть из которых поддается регистрации, управлению и планированию. В регрессионном анализе результирующая переменная играет роль функции, значение которой определяется значениями объясняющих переменных, выполняющих роль аргументов. По своей природе результирующая переменная всегда случайна (стохастична).

Объясняющие (экзогенные, независимые) переменные (*x*) — переменные, которые поддаются регистрации и описывают условия функционирования реальной экономической системы. Они в значительной мере определяют значения результирующих переменных. Обычно часть из них поддается регулированию и управлению. Значение этих переменных могут задаваться вне анализируемой системы, поэтому их называют экзогенными, или факторными признаками. По своей природе они могут быть как случайными, так и неслучайными.

Переменные, выступающие в системе в роли факторов-аргументов или объясняющих переменных, называют **предопределенными**. Множество предопределенных переменных формируется из всех экзогенных переменных и так называемых лаговых эндогенных переменных, то есть таких **эндогенных переменных**, значения которых входят в уравнения анализируемой эконометрической системы, измеренных в прошлые моменты времени, а следовательно, являются уже известными, заданными.

Можно сказать, что любая эконометрическая модель предназначена для объяснения значений текущих эндогенных переменных (одной или нескольких) в зависимости от значений заранее определенных переменных.

Изучая взаимосвязи между переменными, применяют ***ковариацию*** и ***корреляцию***.

Используемые обозначения:

$E(\xi)$ — математическое ожидание случайной величины ξ;

$Var(\xi)$ — дисперсия случайной величины ξ;

$Cov(\xi, \eta)$ — коэффициент ковариации между случайными величинами ξ и η;

$Corr(\xi, \eta) = r_{\xi,\eta}$ — коэффициент корреляции между случайными величинами ξ и η.

Если между случайными величинами x и y существует стохастическая связь, то одним из параметров, характеризующих меру этой связи, является ковариация $Cov(x, y)$. Ковариацию вычисляют по формуле

$$Cov(x, y) = E[(x - E_x)(y - E_y)] = E(xy) - E_x E_y. \qquad (1.1)$$

Если случайные величины x и y независимы, то $Cov(x, y) = 0$.

Если ковариация случайных величин отлична от нуля, то между ними существует стохастическая связь, мерой которой и является величина ковариации.

Следует отметить, что $Cov(x, x) = Var(x)$ и $Cov(y, y) = Var(y)$.

Выборочный коэффициент ковариации между двумя переменными x и y рассчитывается следующим образом:

$$Cov(x, y) = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) = \overline{xy} - \bar{x}\cdot\bar{y}, \qquad (1.2)$$

где (x_1, y_1), (x_2, y_2), ..., (x_n, y_n) — фактические значения случайных переменных x и y;

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i;$$

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i.$$

Ковариация зависит от единиц, в которых измеряются переменные x и y, поэтому для измерения силы связи между двумя переменными используется другая статистическая характеристика, называемая ***коэффициентом корреляции***

$$Corr(x, y) = r_{x,y} = \frac{Cov(x, y)}{S_x S_y} = \frac{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{S_x^2 S_y^2} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}}, \qquad (1.3)$$

где S_x и S_y — оценки среднеквадратических отклонений (СКО) величин x и y

$$S_x = \sqrt{\frac{1}{n-1}\sum(x_i - \bar{x})}; \quad S_y = \sqrt{\frac{1}{n-1}\sum(y_i - \bar{y})}. \qquad (1.4)$$

Коэффициент корреляции принимает значение на отрезке $[-1, 1]$.

Оценка значимости коэффициента корреляции при малых объемах выборки выполняется с использованием ***t-критерия Стьюдента***. При этом фактическое (наблюдаемое) значение этого критерия определяется по формуле

$$t_{\text{набл}} = \sqrt{\frac{r^2_{y,x}}{1 - r^2_{y,x}}(n - 2)}. \qquad (1.5)$$

Вычисленное по этой формуле значение $t_{\text{набл}}$ сравнивается с критическим значением t-критерия, которое берется из таблицы значений критерия Стьюдента с учетом заданного уровня значимости и числа степеней свободы $(n - 2)$.

Если $t_{\text{набл}} > t_{\text{кр}}$, то полученное значение коэффициента корреляции признается значимым (то есть нулевая гипотеза, утверждающая равенство нулю генерального коэффициента корреляции, отвергается). И таким образом делается вывод о том, что между исследуемыми переменными есть тесная статистическая взаимосвязь.

Значимость коэффициентов корреляции можно проверить, используя критическое значение коэффициента корреляции. При условии, что нулевая гипотеза $H_0: r_{ij} = 0$, критическое значение коэффициента корреляции определяется статистикой

$$r^* = \sqrt{\frac{t^2_{(\alpha,n-2)}}{n-2} \Bigg/ \left(1 + \frac{t^2_{(\alpha,n-2)}}{n-2}\right)}, \tag{1.6}$$

где $t^2_{(\alpha,n-2)}$ — критическое значение t-статистики Стьюдента для уровня значимости α и количества степеней свободы, равного $(n - 2)$.

При проведении эконометрического моделирования можно использовать надстройку Excel — пакет «Анализ данных». В таблицах 1.3 и 1.6—1.8 (см. с. 20, 21) приведены функции Excel или пакета «Анализ данных» для проведения корреляционного (табл. 1.3) и регрессионного (табл. 1.6—1.8) анализов.

Таблица 1.3

Использование средств Excel для проведения корреляционного анализа

Формула для вычислений	**Функция Excel или пакета «Анализ данных»**	**Результат вычислений**
$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$	**СРЗНАЧ(массив Х)**	Возвращает среднее значение (среднее арифметическое) аргументов
$S_x^2 = \frac{1}{n-1}\sum (x_i - \bar{x})^2$	**ДИСП(массив Х)**	Оценивает дисперсию по выборке
$\sum_{i=1}^{n} (x_i - \bar{x})^2$	**КВАДРОТКЛ (число1; число2; ...)**	Возвращает сумму квадратов отклонений точек данных от их среднего
$S_x = \sqrt{S_x^2}$	**СТАНДОТКЛОН(массив Х)** $S_x^2 = \frac{1}{n-1}\sum (x_i - \bar{x})^2$	Оценивает стандартное отклонение по выборке. Стандартное отклонение — это мера того, насколько широко разбросаны точки данных относительно их среднего
Коэффициент корреляции $r_{x,y} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}}$	**КОРРЕЛ(массив1; массив2)** Массив1 — это ячейки интервала значений. Массив2 — это второй интервал ячеек со значениями	Возвращает коэффициент корреляции меду интервалами ячеек «Массив1» и «Массив2»

Формула для вычислений	**Функция Excel или пакета «Анализ данных»**	**Результат вычислений**
Оценка значимости коэффициента парной корреляции с использованием *t*-критерия Стьюдента. $$t_{\text{набл}} = \sqrt{\frac{r_{y,x}^2}{1 - r_{y,x}^2}(n-2)}.$$ Вычисленное по этой формуле значение $t_{\text{набл}}$ сравнивается с критическим значением t-критерия, которое берется из таблицы значений критерия Стьюдента с учетом заданного уровня значимости и числа степеней свободы $(n - 2)$.	**СТЬЮДРАСПОБР (Вероятность; Степени_свободы)**, где **вероятность** — вероятность, соответствующая двустороннему распределению Стьюдента. **степени_свободы** — число степеней свободы, характеризующее распределение. При оценке значимости коэффициента парной корреляции число степеней свободы $(n - 2)$	Возвращает t-значение распределения Стьюдента как функцию вероятности и числа степеней свободы
Матрица коэффициентов парной корреляции $$R = \begin{pmatrix} 1 & r_{yx_1} & r_{yx_2} & \dots & r_{yx_m} \\ r_{yx_1} & 1 & r_{x_1x_2} & \dots & r_{x_1x_m} \\ r_{yx_2} & r_{x_1x_2} & 1 & \dots & r_{x_2x_m} \\ \dots & \dots & \dots & \dots & \dots \\ r_{yx_m} & r_{x_1x_m} & r_{x_2x_m} & \dots & 1 \end{pmatrix}$$	**Обращение к средствам анализа данных**. Для вычисления матрицы коэффициентов парной корреляции R следует воспользоваться инструментом **«Корреляция»** в надстройке **«Анализ данных»**	Инструмент **«Корреляция»** применяется, если имеется более двух переменных измерений для каждого объекта. В результате выдается корреляционная матрица, показывающая значение функции **КОРРЕЛ** для каждой возможной пары переменных измерений

Пример 1

На основании данных о десяти случаях ущерба (табл. 1.4), нанесенного пожаром в жилых помещениях, требуется:

1) построить диаграмму рассеяния (корреляционное поле) для переменных «Общая сумма ущерба (y), тыс. руб.» и «Расстояние до ближайшей станции (x), км»;

2) оценить тесноту связи общей суммы ущерба (тыс. руб.) с расстоянием до ближайшей пожарной станции (км);

3) проверить значимость вычисленного коэффициента парной корреляции, сделать выводы о направлении и тесноте связи.

Таблица 1.4

Ущерб от пожара

Показатель	**Случай ущерба**									
	1-й	**2-й**	**3-й**	**4-й**	**5-й**	**6-й**	**7-й**	**8-й**	**9-й**	**10-й**
Расстояние до ближайшей пожарной станции (x), км	3,4	1,8	4,6	2,3	3,1	5,5	0,7	3	2,6	4,3
Общая сумма ущерба (y), тыс. руб.	262	178	313	231	275	360	141	223	196	313

Решение

1) Создание парных диаграмм рассеяния. Важнейшим элементом эконометрического исследования является графический анализ исходных данных. Для создания диаграммы рассеяния нужно выделить два столбца данных со значениями показателей, включая их названия (метки) в первой строке матрицы данных, и выполнить следующие действия: на вкладке «Вставка» в группе «Диаграммы» выбрать тип диаграммы «Точечная», диаграмма добавится на лист.

Нужно помнить, что для того чтобы Microsoft Excel правильно определил переменные, объясняемая переменная y должна быть расположена в правом из двух выделенных столбцов, а объясняющая переменная x — в левом. В нашем примере диаграмма рассеяния имеет вид, приведенный на рис. 1.1.

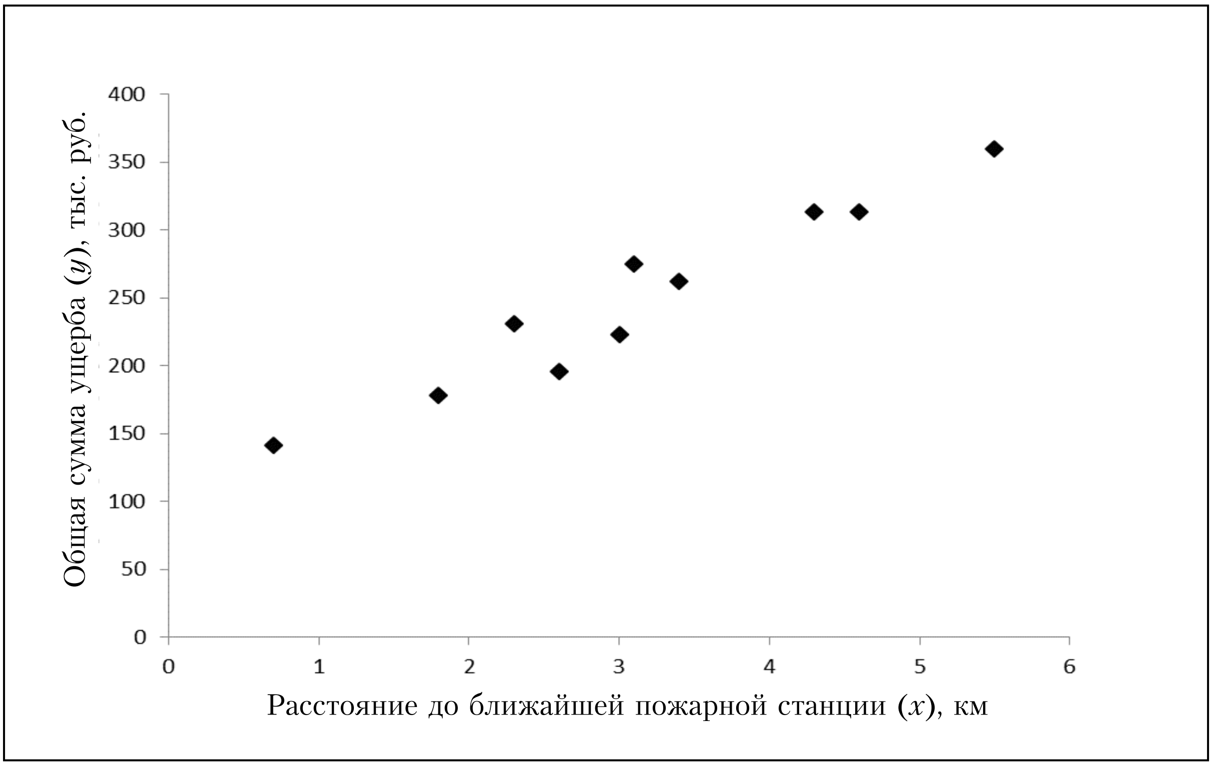
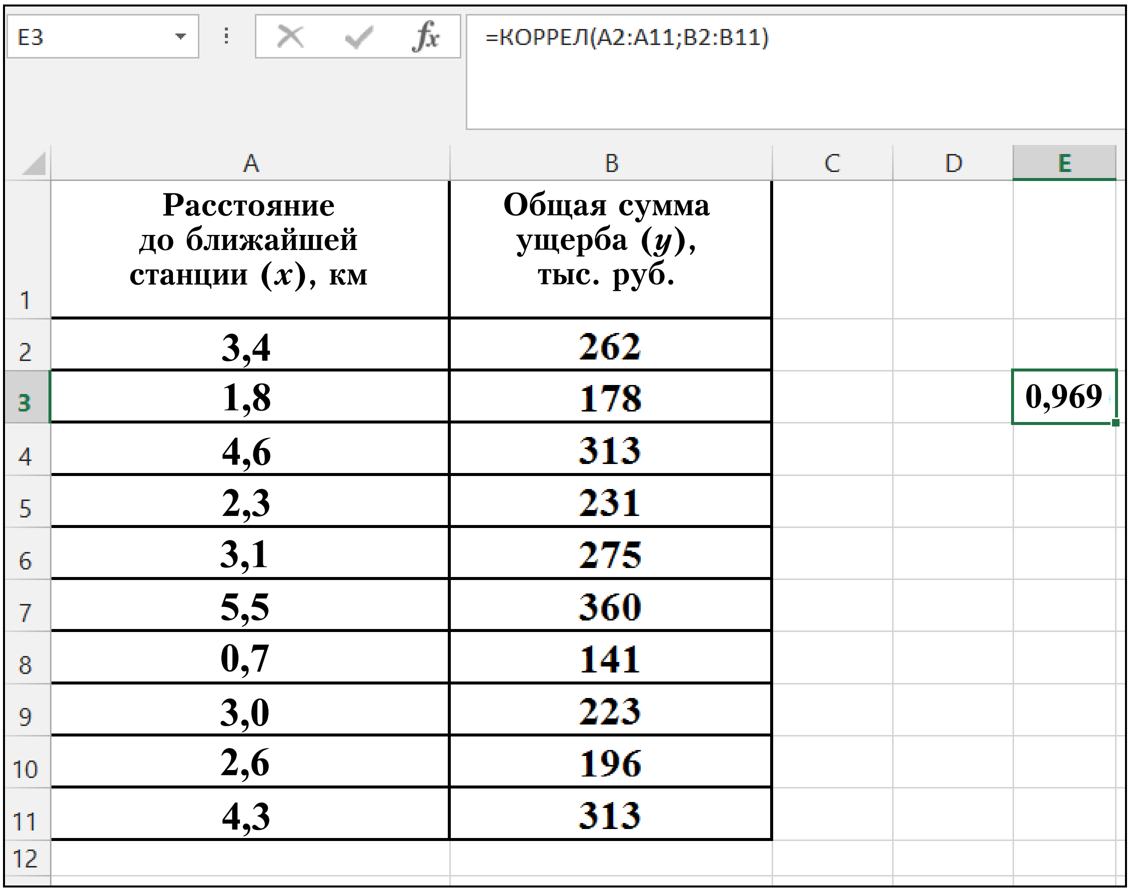


Рис. 1.1. Диаграмма рассеяния

Вывод

Вытянутость облака точек на диаграмме рассеяния вдоль наклонной прямой позволяет сделать предположение, что существует некоторая объективная тенденция прямой линейной связи между значениями переменных x и y, то есть при увеличении расстояния до пожарной станции возрастает общая сумма ущерба.

2) Оценка тесноты связи. Чтобы вычислить корреляцию средствами Microsoft Excel, можно воспользоваться функцией **«=КОРРЕЛ(·)»**, указав адреса двух столбцов чисел, как показано на рис. 1.2. Ответ помещен в ячейке Е3 и равен 0,969.



	A	B	C	D	E
1	Расстояние до ближайшей станции (x), км	Общая сумма ущерба (y), тыс. руб.			
2	3,4	262			
3	1,8	178			0,969
4	4,6	313			
5	2,3	231			
6	3,1	275			
7	5,5	360			
8	0,7	141			
9	3,0	223			
10	2,6	196			
11	4,3	313			
12					

Рис. 1.2. Вычисление коэффициента парной корреляции с помощью функции «=КОРРЕЛ(·)»

Вычисление «вручную» по формуле (1.3) дает такие же результаты (табл. 1.5).

Таблица 1.5

Промежуточные результаты для вычисления коэффициента корреляции

($\bar{x}$, $\bar{y}$ — средние значения x y, соответственно равные 3,13 и 249,2)

Случай ущерба	$x_i - \bar{x}$	$y_i - \bar{y}$	$(y_i - \bar{y})(x_i - \bar{x})$	$(x_i - \bar{x})^2$	$(y_i - \bar{y})^2$
1-й	0,27	12,8	3,456	0,0729	163,84
2-й	−1,33	−71,2	94,696	1,7689	5 069,44
3-й	1,47	63,8	93,786	2,1609	4 070,44
4-й	−0,83	−18,2	15,106	0,6889	331,24
5-й	−0,03	25,8	−0,774	0,0009	665,64
6-й	2,37	110,8	262,596	5,6169	12 276,64
7-й	−2,43	−108,2	262,926	5,9049	11 707,24
8-й	−0,13	−26,2	3,406	0,0169	686,44
9-й	−0,53	−53,2	28,196	0,2809	2 830,24
10-й	1,17	63,8	74,646	1,3689	4 070,44
Сумма	$-4{,}2 \cdot 10^{-15}$	$1{,}14 \cdot 10^{-13}$	838,04	17,8810	41 871,60

3) Оценка значимости коэффициента корреляции. Для этого рассчитаем значение t-статистики по формуле (1.5)

$$t_{\text{расч}} = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} = \frac{0{,}969\sqrt{8}}{\sqrt{1-0{,}938}} = 11{,}005.$$

Табличное значение критерия Стьюдента равно

$$t_{\text{табл }(\alpha = 0{,}1;\; k = n-2 = 8)} = 1{,}86.$$

Сравнивая числовые значения критериев, видим, что $t_{\text{расч}} > t_{\text{табл}}$, то есть полученное значение коэффициента корреляции значимо.

Таким образом, расстояние до ближайшей станции оказывает весьма большое влияние на общую сумму ущерба от пожара.

1.2. Линейные регрессионные модели

Используемые обозначения:

y	—	зависимая (эндогенная) переменная;
$x_1, x_2, ..., x_k$	—	объясняющие (экзогенные, независимые) переменные, регрессоры;
X	—	матрица объясняющих переменных;
X^T	—	транспонированная матрица объясняющих переменных;
ε_i	—	ошибка, ошибка модели, случайная составляющая;
$\hat{y}_i$	—	предсказанное по модели значение y, прогноз, прогнозное значение;
$\hat{\varepsilon}_i = y_i - \hat{y}_i$	—	остаток, ошибка прогноза.

Регрессионный анализ предназначен для исследования зависимости исследуемой переменной от различных факторов и отображения их взаимосвязи в форме регрессионной модели. В регрессионных моделях зависимая (объясняемая) переменная y может быть представлена в виде функции $f(x_1, x_2, ..., x_k)$, где $x_1, x_2, ..., x_k$ — независимые (объясняющие) переменные, или факторы.

В то время как зависимая переменная должна быть непрерывной (за исключением логистической регрессии), независимые переменные могут быть как прерывными, так и категориальными, то есть как «пол» или «тип применяемого препарата».

Функция $f(x_1, x_2, ..., x_k)$, описывающая зависимость показателя от параметров, называется уравнением (функцией) регрессии.

В зависимости от количества включенных в модель факторов x модели делятся на однофакторные (парная модель регрессии) и многофакторные (модель множественной регрессии), которые, в свою очередь, в зависимости от вида функции $f(x_1, x_2, ..., x_k)$ модели делятся на линейные и нелинейные.

Рассмотрим переменные x и y. Между ними существует регрессионная зависимость: если найдется функция $f(x)$ такая, когда имеет место равенство $y = f(x) + \varepsilon$, в котором ε случайная величина отражает тот факт, что изменение y будет неточно описываться изменением x — присутствуют другие факторы, неучтенные в данной модели.

Рассмотрим случай, для которого функция $f(x)$ линейна относительно описываемых параметров:

$$y = f(x) + \varepsilon = \alpha_0 + \alpha_1 x + \varepsilon. \quad (1.7)$$

Слагаемое $f(x)$ называется функцией, или уравнением регрессии y на x.

Предположим, что для оценки линейной регрессии (1.7) взята выборка, состоящая из n пар значений переменных (x_i, y_i), где $i = 1, 2, ..., n$. Тогда $f(x)$ можно представить в виде:

$$\hat{y}_i = \hat{\alpha}_0 + \hat{\alpha}_1 x_i, \quad (1.8)$$

где $\hat{\alpha}_0$ и $\hat{\alpha}_1$ — параметры регрессии, которые должны быть определены по выборочным данным с помощью **метода наименьших квадратов** (МНК).

Согласно принципу метода наименьших квадратов оценки $\hat{\alpha}_0$ и $\hat{\alpha}_1$ находятся путем минимизации суммы квадратов

$$Q(\hat{\alpha}_0, \hat{\alpha}_1) = \sum_{i=1}^{n} \hat{\varepsilon}_i^2 = \sum_{i=1}^{n} \left(y_i - \hat{y}_i \right)^2 = \sum_{i=1}^{n} (y_i - \alpha_0 - \alpha_1 \ x_i)^2 \quad (1.9)$$

по всем возможным значениям $\hat{\alpha}_0$ и $\hat{\alpha}_1$ при заданных (наблюдаемых) значениях $x_1, ..., x_n, y_1, ..., y_n$. Задача сводится к математической задаче поиска точки минимума функции двух переменных. Точка минимума находится путем приравнивания нулю частных производных функции $z = Q(\alpha_0, \alpha_1)$ по переменным $\hat{\alpha}_0$ и $\hat{\alpha}_1$.

Оценки МНК параметров имеют вид:

$$\hat{\alpha}_1 = \frac{\sum_{i=1}^{n} \left(y_i - \overline{y} \right)\left(x_i - \overline{x} \right)}{\sum_{i=1}^{n} \left(x_i - \overline{x} \right)^2}; \quad \hat{\alpha}_0 = \overline{y} - \hat{\alpha}_1 \overline{x}. \quad (1.10)$$

Оценки $\hat{\alpha}_0$ и $\hat{\alpha}_1$ называют оценками наименьших квадратов. Обратим внимание на полученное выражение для параметра $\hat{\alpha}_1$. Используя формулы для вычисления выборочной дисперсии и коэффициента парной корреляции (формулы 1.3—1.4), параметр $\hat{\alpha}_1$ можно получить следующим образом:

$$\hat{\alpha}_1 = \frac{\sum_{i=1}^{n} (x_i - \overline{x})(y_i - \overline{y}) \Big/ (n-1)}{S_x^2} = \frac{\sum_{i=1}^{n} (x_i - \overline{x})(y_i - \overline{y})}{\sum_{i=1}^{n} (x_i - \overline{x})^2} =$$

$$= r_{x,y} \frac{S_y}{S_x} = \frac{\overline{yx} - \overline{y}\,\overline{x}}{\overline{x^2} - \overline{x}^2} = \frac{\sum_{i=1}^{n} y_i x_i - n\overline{yx}}{\sum_{i=1}^{n} x_i^2 - n\overline{x}^2}. \quad (1.11)$$

Оценка параметров модели множественной регрессии с помощью метода наименьших квадратов

Формулу для вычисления параметров регрессионного уравнения приведем без вывода

$$A = (X^T X)^{-T} X^T y, \tag{1.12}$$

где

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ ... \\ y_n \end{bmatrix}, \quad X = \begin{bmatrix} 1 & x_{11} & ... & x_{1k} \\ 1 & x_{21} & ... & x_{2k} \\ ... & ... & ... & ... \\ 1 & x_{n1} & ... & x_{nk} \end{bmatrix}, \quad A = \begin{bmatrix} \hat{\alpha}_0 \\ \hat{\alpha}_1 \\ ... \\ \hat{\alpha}_k \end{bmatrix}$$

Используемые обозначения:

$RSS = \sum_{i=1}^{n} \hat{\varepsilon}_i^2$ – сумма квадратов остатков (Residual Sum of Squares);

$TSS = \sum_{i=1}^{n} (y_i - \overline{y}_i)^2$ — общая сумма квадратов (Total Sum of Squares);

$ESS = \sum_{i=1}^{n} (\hat{y}_i - \overline{y}_i)^2$ — объясненная сумма квадратов (Explained Sum of Squares).

Качество модели регрессии проверяется путем анализа остатков регрессии $\hat{\varepsilon}_i = y_i - \hat{y}_i$, который позволяет получить представление, насколько хорошо подобрана сама модель и насколько правильно выбран метод оценки коэффициентов.

При анализе качества модели регрессии, в первую очередь, используется ***коэффициент детерминации***, который определяется следующим образом:

$$R^2 = \frac{\sum_{i=1}^{n} (\hat{y}_i - \overline{y})^2}{\sum_{i=1}^{n} (y_i - \overline{y})^2} = 1 - \frac{\sum_{i=1}^{n} \hat{\varepsilon}_i^2}{\sum_{i=1}^{n} (y_i - \overline{y})^2} = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS}, \tag{1.13}$$

где $\overline{y}$ — среднее значение зависимой переменной;

$\hat{y}_i$ — предсказанное (рассчитанное по уравнению регрессии) значение зависимой переменной.

Коэффициент детерминации показывает долю вариации результативного признака, находящегося под воздействием изучаемых факторов, то есть определяет, какая доля вариации признака y учтена в модели и обусловлена влиянием на него факторов, включенных в модель.

Чем ближе R^2 к 1, тем выше качество модели. Для оценки качества регрессионных моделей целесообразно также использовать ***коэффициент множественной корреляции (индекс корреляции)*** R

$$R=\sqrt{1-\frac{\sum_{i=1}^{n}\hat{\varepsilon}_i^2}{\sum_{i=}^{n}(y_i-\overline{y})}}=\sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i-\overline{y})^2}{\sum_{i=1}^{n}(y_i-\overline{y})^2}}=Corr(\hat{y}_i,y_i). \quad (1.14)$$

Данный коэффициент является универсальным, так как он отражает тесноту связи и точность модели, а также может использоваться при любой форме связи переменных.

Проверка значимости построенного уравнения в целом и его отдельных параметров

Для проверки значимости модели регрессии используется ***F-критерий Фишера***. Если расчетное значение с $\nu_1 = k$ и $\nu_2 = (n - k - 1)$ степенями свободы, где k — количество факторов, включенных в модель, больше табличного при заданном уровне значимости, то модель считается значимой

$$F=\frac{R^2/k}{\left(1-R^2\right)/\left(n-k-1\right)}. \quad (1.15)$$

В качестве **меры точности** применяют несмещенную оценку дисперсии остаточной компоненты, которая представляет собой отношение суммы квадратов уровней остаточной компоненты к величине $(n - k - 1)$, где k — количество факторов, включенных в модель. Квадратный корень из этой величины $(\hat{\sigma})$ называется ***стандартной ошибкой***

$$\hat{\sigma}=\sqrt{\frac{RSS}{n-k-1}}. \quad (1.16)$$

Также для оценки качества регрессионных моделей целесообразно использовать ***среднюю ошибку аппроксимации***

$$E_{\text{отн}}=\frac{1}{n}\sum_{i=1}^{n}\frac{|\hat{\varepsilon}_i|}{y_i}100\%. \quad (1.17)$$

Чем меньше рассеяние эмпирических точек вокруг теоретической линии регрессии, тем меньше средняя ошибка аппроксимации. Ошибка аппроксимации меньше 7% свидетельствует о хорошем качестве модели.

Значимость отдельных коэффициентов регрессии проверяется по t-статистике путем проверки гипотезы о равенстве нулю j-го параметра уравнения (кроме свободного члена)

$$t_{a_j}=\hat{a}_j\big/\hat{\sigma}_{a_j}, \quad (1.18)$$

где $\hat{\sigma}_{aj}$ — стандартное (среднеквадратическое) отклонение коэффициента уравнения регрессии a_j.

Величина $\hat{\sigma}_{aj}$ представляет собой квадратный корень из произведения несмещенной оценки дисперсии $\hat{\sigma}^2$ и j-го диагонального элемента матрицы, обратной матрице системы нормальных уравнений

$$\hat{\sigma}_{aj} = \hat{\sigma}\sqrt{b_{jj}}, \tag{1.19}$$

где b_{jj} — диагональный элемент матрицы $(X^T X)^{-1}$.

Если расчетное значение t-критерия с $(n - k - 1)$ степенями свободы превосходит его табличное значение при заданном уровне значимости, коэффициент регрессии считается значимым. В противном случае фактор, соответствующий этому коэффициенту, следует исключить из модели (при этом ее качество не ухудшится).

Если модель адекватна и достаточно точна, то ее можно использовать для анализа и прогнозирования. Анализ на основе регрессионных моделей проводят, во-первых, для выявления факторов, наиболее сильно влияющих на зависимую переменную, а, во-вторых, с целью ранжирования объектов по степени их эффективности. Выбор факторов, наиболее сильно влияющих на зависимую переменную, важен для принятия решений по улучшению результатов деятельности исследуемой системы. Значительную роль при оценке влияния факторов играют коэффициенты регрессионной модели. Однако непосредственно с их помощью нельзя сопоставить факторы по степени их влияния на зависимую переменную из-за различия единиц измерения и разной степени колеблемости. Для устранения таких различий при интерпретации применяются средние частные коэффициенты эластичности $Э_j$ бета-коэффициенты β_j или коэффициенты регрессии в стандартизированном масштабе и дельта-коэффициенты ($\Delta(j)$).

Эластичность y по отношению к x_j определяется как процентное изменение y, отнесенное к соответствующему процентному изменению x. В общем случае эластичности не постоянны, они различаются, если измерены для различных точек на линии регрессии. По умолчанию стандартные программы, оценивающие эластичность, вычисляют ее в точках средних значений

$$Э_j = \hat{a}_j \frac{\bar{x}_j}{\bar{y}} \approx \frac{\delta y}{\bar{y}} \bigg/ \frac{\delta x}{\bar{x}}. \tag{1.20}$$

Коэффициент эластичности показывает, на сколько процентов изменится зависимая переменная при изменении j-го фактора на 1%.

Эластичность не нормирована и может изменяться от $-\infty$ до $+\infty$. Высокий уровень эластичности означает сильное влияние независимой переменной на объясняемую переменную.

Однако средний частный коэффициент эластичности не учитывает степени колеблемости факторов, которая может значительно различаться у отдельных факторов. Поэтому для устранения различий в измерении и степени колеблемости факторов используется другой показатель — коэффициент регрессии в стандартизированном масштабе **(*бета-коэффициент*)**

$$\beta_j = \hat{\alpha}_j \frac{S_{x_j}}{S_y}, \tag{1.21}$$

где S_{x_j} — среднеквадратическое отклонение фактора j:

$$S_{x_j} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})}, \; S_y = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(y_i - \overline{y})}.$$

Бэта-коэффициент показывает, на какую часть величины СКО изменяется среднее значение зависимой переменной с изменением соответствующей независимой переменной на одно среднее квадратическое отклонение при фиксированном на постоянном уровне значении остальных независимых переменных.

Долю влияния фактора в суммарном влиянии всех факторов можно оценить по величине ***дельта-коэффициентов*** Δ_j

$$\Delta_j = r_{y,x_j}\frac{\beta_j}{R^2}, \tag{1.22}$$

где r_{y,x_j} — коэффициент парной корреляции между фактором j и зависимой переменной.

В практических задачах при корректно проведенном анализе величины дельта-коэффициентов положительны, то есть все коэффициенты регрессии имеют тот же знак, что и соответствующие парные коэффициенты корреляции. Указанные характеристики позволяют упорядочить факторы по степени влияния факторов на зависимую переменную.

Уравнение регрессии применяют для расчета значений показателя в заданном диапазоне изменения параметров. Оно ограниченно пригодно для расчета вне этого диапазона, то есть его можно применять для решения задач интерполяции и в ограниченной степени для экстраполяции.

Прогноз, полученный подстановкой в уравнение регрессии ожидаемого значения параметра, является точечным. Вероятность реализации такого прогноза ничтожно мала. Целесообразно определить доверительный интервал прогноза.

Для линейной модели множественной регрессии при прогнозировании индивидуальных значений доверительный интервал рассчитывается по формуле (1.23). Для этого оценивается величина отклонения от линии регрессии (обозначим ее $U(X_{\text{прогн}})$)

$$U(X_{\text{прогн}}) = \hat{\sigma}t_\alpha \sqrt{1 + X_{\text{прогн}}^T (X^T X)^{-1} X_{\text{прогн}}}, \tag{1.23}$$

где $X_{\text{прогн}}^T = (1, x_{1\,\text{прогн}}, x_{2\,\text{прогн}}, ..., x_{k\,\text{прогн}})$, то есть $y_{\text{прогн}} \in [\hat{y}_{\text{прогн}} \pm U(X_{\text{прогн}})]$.

Для модели парной регрессии формула (1.23) может быть записана в следующем виде:

$$y_{\text{прогн}} \in \left[\hat{y}_{\text{прогн}} - \hat{\sigma}\,t_\alpha \sqrt{1 + \frac{1}{n} + \frac{(x_{\text{прогн}} - \overline{x})^2}{\sum_{i=1}^{n}(x_i - \overline{x})^2}}; \hat{y}_{\text{прогн}} + \hat{\sigma}\,t_\alpha \sqrt{1 + \frac{1}{n} + \frac{(x_{\text{прогн}} - \overline{x})^2}{\sum_{i=1}^{n}(x_i - \overline{x})^2}}\right] \tag{1.24}$$

Таблица 1.6

Использование средств Excel для проведения регрессионного анализа

Формула для вычислений	**Функция Excel или пакета «Анализ данных»**	**Результат вычислений**
Оценка параметров модели парной и множественной линейной регрессии $A = (X^T X)^{-T} X^T y$	Для вычисления параметров уравнения регрессии следует воспользоваться инструментом **«Регрессия»** из пакета **«Анализ данных»**	Возвращает подробную информацию о параметрах модели, качестве модели, расчетных значениях и остатках в виде четырех таблиц: *«Регрессионная статистика»*, *«Дисперсионный анализ»*, *«Коэффициенты»*, *«Вывод остатка»*. Также могут быть получены *график подбора* и *график остатков*
Оценка качества модели регрессии		
***F*-критерий Фишера** для проверки значимости модели регрессии $F = \dfrac{R^2/k}{\left(1-R^2\right)/\left(n-k-1\right)}$	**=FРАСПОБР(вероятность; степени_свободы1; степени_свободы2)**, где **вероятность** — это вероятность, связанная с F-распределением **степени_свободы 1** — это числитель степеней свободы ($\nu_1 = k$) **степени_свободы 2** — это знаменатель степеней свободы ($\nu_2 - (n - k - 1)$, где k — количество факторов, включенных в модель)	Возвращает обратное значение для *F*-распределения вероятностей. **FРАСПОБР(·)** можно использовать, чтобы определить критические значения *F*-распределения. Чтобы определить критическое значение *F*, нужно использовать уровень значимости α как аргумент **вероятность** для **FРАСПОБР(·)**
Коэффициент детерминации $R^2 = \dfrac{\sum(\hat{y}_i - \bar{y})^2}{\sum(y_i - \bar{y})^2} = \dfrac{ESS}{TSS} = 1 - \dfrac{RSS}{TSS}$	**Коэффициент детерминации** показывает долю вариации результативного признака, находящегося под воздействием изучаемых факторов, то есть определяет, какая доля вариации признака *Y* учтена в модели и обусловлена влиянием на него факторов. Чем ближе R^2 к 1, тем выше качество модели	
Коэффициент множественной корреляции (индекс корреляции) $R = \sqrt{\dfrac{\sum(\hat{y}_i - \bar{y})^2}{\sum(y_i - \bar{y})^2}} = \sqrt{1 - \dfrac{RSS}{TSS}}$	Данный коэффициент является универсальным, так как он отражает тесноту связи и точность модели, а также может использоваться при любой форме связи переменных. Чем ближе *R* к 1, тем выше качество модели	
***t*-критерий Стьюдента** для оценки значимости параметров модели линейной регрессии $t_{a_j} = \hat{a}_j / \sigma_{aj}$	Вычисленное значение t_{a_j} сравнивается с критическим значением *t*-критерия, которое берется из таблицы значений *t*-распределения Стьюдента с учетом заданного уровня значимости и числа степеней свободы ($n - k - 1$). В Excel критическое значение *t*-критерия можно получить с помощью функции **СТЬЮДРАСПОБР(вероятность; степени_свободы)**, где **вероятность** — вероятность, соответствующая двустороннему распределению Стьюдента; **степени_свободы** — число степеней свободы, характеризующее распределение	
Средняя относительная ошибка аппроксимации $E_{отн} = \dfrac{1}{n}\sum_{i=1}^{n}\dfrac{\lvert\hat{\varepsilon}_i\rvert}{y_i}100\%$	Средняя относительная ошибка аппроксимации — оценка точности модели	

Окончание табл. 1.6

Оценка влияния отдельных факторов на зависимую переменную на основе модели	
Коэффициент эластичности $Э_j = \hat{a}_j \dfrac{\overline{x}_j}{\overline{y}}$	Коэффициент эластичности показывает, на сколько процентов изменится значение исследуемой величины при изменении соответствующего фактора на 1%
Бета-коэффициент $\beta_j = \hat{\alpha}_j \dfrac{S_{x_j}}{S_y}$	Бета-коэффициент показывает, на какую часть своего СКО изменится значение исследуемой переменной при изменении соответствующего фактора на 1 СКО
Дельта-коэффициент $\Delta_j = r_{y,x_j} \dfrac{\beta_j}{R^2}$	Дельта-коэффициент показывает среднюю долю влияния соответствующего фактора в совокупном влиянии всех факторов, включенных в модель
Построение интервальных прогнозов по модели регрессии	
$U(X_{\text{прогн}}) = \hat{\sigma}\, t_\alpha \sqrt{1 + X^T_{\text{прогн}} (X^T X)^{-1} X_{\text{прогн}}}$ — ошибка прогнозирования, которая позволяет определить доверительный интервал прогноза, где $\hat{\sigma} = \sqrt{\dfrac{RSS}{n-k-1}}$ — стандартная ошибка модели. Для модели парной регрессии ошибка прогнозирования может быть записана в следующем виде: $U(X_{\text{прогн}}) = \hat{\sigma}\, t_\alpha \sqrt{1 + \dfrac{1}{n} + \dfrac{(x_{\text{прогн}} - \overline{x})^2}{\sum_{i=1}^{n} (x_i - \overline{x})^2}}$.	

Таблица 1.7

Регрессионная статистика в отчете Microsoft Excel

Наименование в отчете Excel	Принятое наименование	Формула
Множественный R	Коэффициент множественной корреляции, индекс корреляции	$R^2 = \sqrt{R^2}$
R-квадрат	Коэффициент детерминации, R^2	$R^2 = \dfrac{\sum (\hat{y}_i - \overline{y})^2}{\sum (y_i - \overline{y})^2} = \dfrac{ESS}{TSS} = 1 - \dfrac{RSS}{TSS}$
Нормированный R-квадрат	Скорректированный R^2	$\overline{R}^2 = 1 - (1 - R^2)\dfrac{n-1}{n-k-1}$
Стандартная ошибка	Среднеквадратическое отклонение от модели σ	$\sigma = \sqrt{\sum \hat{\varepsilon}_i^2 / (n-k-1)} = \sqrt{RSS / (n-k-1)}$

Таблица 1.8

Дисперсионный анализ в отчете Microsoft Excel

Наименование в отчете Excel	***Df*** **— число степеней свободы**	***SS*** **— сумма квадратов**	***MS*** **— дисперсия на одну степень свободы**	***F*****-критерий Фишера**
Регрессия	k	$ESS = \sum (\hat{y}_i - \overline{y})^2$	$\sum (\hat{y}_i - \overline{y})^2 / k = ESS / k$	$F = \dfrac{R^2 / k}{(1 - R^2)/(n-k-1)}$
Остаток	$n - k - 1$	$RSS = \sum \hat{\varepsilon}_i^2$	$\sum \hat{\varepsilon}_i^2 / (n-k-1) = RSS / (n-k-1)$	
Сумма	$n - 1$	$TSS = \sum (y_i - \overline{y})^2$		

Названия некоторых функций в Microsoft Excel 2010 были изменены по сравнению с более ранними версиями.

Чтобы повысить точность работы функций Microsoft Excel, обеспечить их согласованность и привести имена функций в соответствии с их назначением, корпорация Microsoft изменила, переименовала и добавила несколько функций в библиотеку Microsoft Excel 2010 (табл. 1.9).

Для обеспечения обратной совместимости переименованные функции доступны также и по их старым именам.

Таблица 1.9

Переименованные и добавленные функции в библиотеку Microsoft Excel 2010

Название функции в Excel более ранних версий	Название функции в Excel 2010	Примечания
ДИСП(число1, число2, ...)	**ДИСП.В(число1, число2, ...)**	Оценивает дисперсию по выборке
СТЬЮДРАСПОБР(вероятность; степени_свободы)	**СТЬДЕНТ.ОБР.2Х(вероятность, степени_свободы)**	Возвращает двустороннее обратное t-распределение Стьюдента
FРАСПОБР(вероятность; степени_свободы1;степени_свободы2)	**F.ОБР.ПХ(вероятность,степени_свободы1, степени_свободы2)**	Возвращает значение, обратное (правостороннему) F-распределению вероятностей
ХИ2ОБР(вероятность,степени_свободы)	**ХИ2.ОБР.ПХ(вероятность,степе-ни_свободы)**	Возвращает обратное значение односторонней вероятности распределения хи-квадрат

1.3. Примеры использования различных функций Excel для оценки параметров парной линейной регрессии

В Microsoft Excel есть несколько инструментов, которыми можно воспользоваться для построения и последующего анализа эконометрических моделей. Мы рассмотрим пакет **«Анализ данных»** и надстройку **«Поиск решения»**.

Пример 2

На основании данных (см. табл. 1.4) о десяти случаях ущерба, нанесенного пожаром в жилых помещениях, требуется:

1) построить модель парной регрессии зависимости стоимости ущерба, нанесенного пожаром, от расстояния до ближайшей пожарной станции с использованием: пакета **«Анализ данных»**; b) надстройки **«Поиска решений»**; c) матричных функций; d) функции **ЛИНЕЙН** и дать экономическую интерпретацию параметров модели регрессии;

2) оценить качество построенной модели;

3) оценить стоимость ущерба, нанесенного пожаром, если полагают, что расстояние до ближайшей пожарной станции уменьшится на 10% от своего среднего уровня;

4) изобразить на графике исходные данные, результаты моделирования и прогнозирования.

Решение

1а) Построение модели линейной регрессии зависимости стоимости ущерба, нанесенного пожаром, от расстояния до ближайшей пожарной станции с использованием инструмента «Регрессия» в пакете «*Анализ данных*» Excel:

- данные записываем в таблицу Excel, упорядочив по возрастанию x (процедура необязательная);
- выбраем команду **«Анализ данных»** на вкладке «Данные»;
- в диалоговом окне **«Анализ данных»** выбраем инструмент **«Регрессия»**;
- в диалоговом окне **«Регрессия»** в поле «*Входной интервал Y*» вводим адрес одного диапазона ячеек, который представляет зависимую переменную. В поле «*Входной интервал X*» вводим адрес диапазона, который содержит значения независимой переменной (рис. 1.3);
- устанавливаем флажок «*Метки*» в первой строке для отображения заголовков столбцов;
- выбираем параметры вывода (в данном примере «*Выходной интервал*» — $A$15);
- в поле «*Остатки*» поставим необходимые флажки;
- щелкаем «ОК».

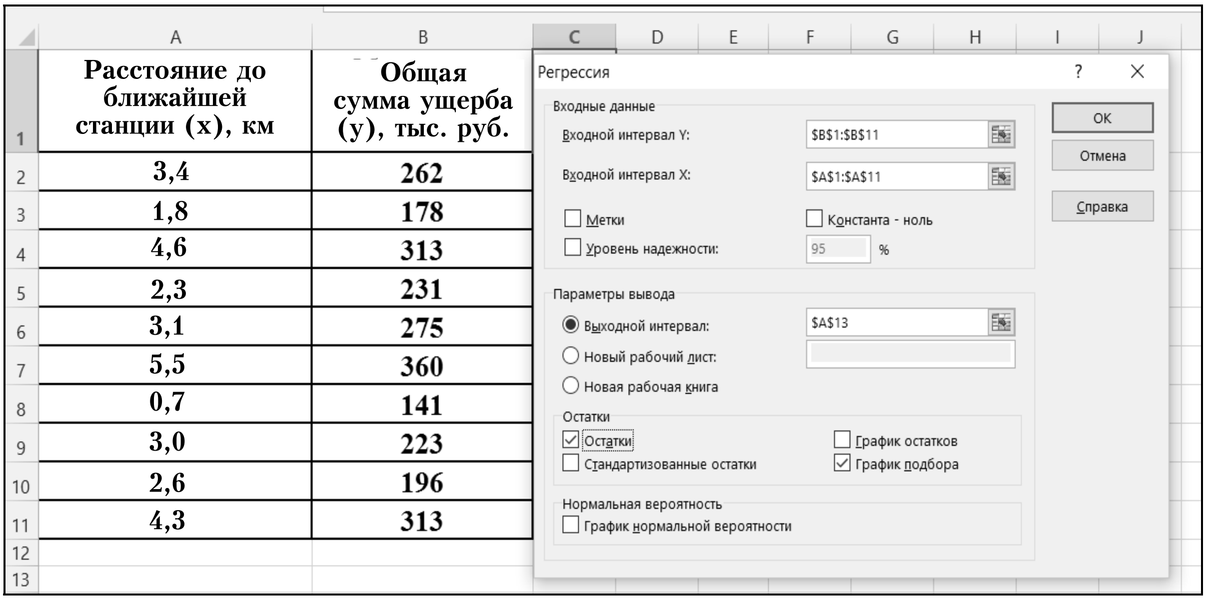


Рис. 1.3. Диалоговое **«Регрессия»**, подготовленное к построению модели регрессии

Рассмотрим содержание протокола регрессионного анализа примера 2. На рис. 1.4 показаны четыре таблицы протокола регрессионного анализа, в которых отражены основные итоги расчетов.

Регрессионная статистика						
Множественный R	0,969					
R-квадрат	0,938					
Нормированный R-квадрат	0,930					
Стандартная ошибка	18,009					
Наблюдения	10					
Дисперсионный анализ						
	df	SS	MS	F	Значимость F	
Регрессия	1	39276,94	39276,94	121,10	4,13578E-06	
Остаток	8	2594,66	324,33			
Итого	9	41871,6				
	Коэффициенты	Стандартная ошибка	t-статистика	P-Значение	Нижние 95%	Верхние 95%
Y-пересечение	102,504	14,496	7,071	0,000	69,077	135,932
Расстояние до ближайшей станции	46,868	4,259	11,005	0,000	37,047	56,689

Рис. 1.4. Фрагмент протокола выполнения регрессионного анализа

В табл. 1.10 приведены формулы, по которым выполнены расчеты в протоколе «Регрессионная статистика», и результаты вычислений. В табл. 1.11 приведены формулы, по которым выполнены расчеты в протоколе «Дисперсионный анализ», и результаты вычислений.

Таблица 1.10

Пояснения к таблице «Регрессионная статистика»

Наименование в отчете EXCEL	Принятое наименование	Формула
Множественный R	Коэффициент множественной корреляции, индекс корреляции	$R^2 = \sqrt{R^2} = 0{,}969$
R-квадрат	Коэффициент детерминации, R^2	$R^2 = 1 - \frac{\sum \hat{\varepsilon}_i^2}{\sum (y_i - \bar{y})^2} = \frac{\sum (\hat{y}_i - \bar{y})^2}{\sum (y_i - \bar{y})^2} = 0{,}938$
Нормированный R-квадрат	Скорректированный R^2	$\bar{R}^2 = 1 - (1 - R^2)\frac{n-1}{n-k-1} = 0{,}930$
Стандартная ошибка	Среднеквадратическое отклонение от модели $\hat{\sigma}$	$\hat{\sigma} = \sqrt{\frac{\sum \hat{\varepsilon}_i^2}{n-k-1}} = 18{,}009$
Наблюдения	Количество наблюдений, n	$n = 10$

Таблица 1.11

Пояснения к таблице «Дисперсионный анализ»

	Df — число степеней свободы	**SS — сумма квадратов отклонений**	**$MS = SS / Df$**	**F-критерий Фишера= MS(регрессия) / MS(остаток)**	**Значимость F**
Регрессия	$k = 1$	$\sum(\hat{y}_i - \overline{y})^2$ 392769,94	$\sum(\hat{y}_i - \overline{y})^2 / k$ 39276,94	$F = \dfrac{R^2/k}{(1-R^2)/(n-k-1)}$ 121,101	Значение уровня значимости, соответствующее вычисленной F-статистике $4{,}13578 \cdot 10^{-6}$
Остаток	$n-k-1=8$	$\sum\hat{\varepsilon}_i^2$ 2594,66	$\sum\hat{\varepsilon}_i^2 / n-k-1$ 3,243		
Сумма	$n - 1 = 9$	$\sum(y_i - \overline{y})^2$ 41871,6			

В табл. 1.12 приводятся значения следующих параметров:

«*Коэффициенты*» — значения коэффициентов уравнения регрессии $\hat{\alpha}_0$ и $\hat{\alpha}_1$.

«*Стандартная ошибка*» — стандартные ошибки коэффициентов уравнения регрессии.

«*t-статистика*» — расчетные значения t-критерия, используемые для проверки значимости коэффициентов уравнения регрессии.

«*P-значение*» — (значимость t) — это значение уровня значимости, соответствующее вычисленной t-статистике.

Если P-значение меньше стандартного уровня значимости, то соответствующий коэффициент статистически значим.

«*Нижние 95%*» и «*Верхние 95%*» — нижние и верхние границы 95%-ных доверительных интервалов для коэффициентов теоретического уравнения регрессии.

Таблица 1.12

Третья таблица отчета Microsoft Excel

	Коэффициенты	**Стандартная ошибка**	**t-статистика**	**P-значение**	**Нижние 95%**	**Верхние 95%**
Y-пересечение	102,504	14,496	7,07	0,00	69,077	135,932
Расстояние до ближайшей пожарной станции (x), км	46,868	4,259	11,00	0,00	37,047	56,689

В табл. 1.13 приведены вычисленные (предсказанные) по модели значения зависимой переменной $\hat{y}_i$ и значения остаточной компоненты $\hat{\varepsilon}_i$ — «*Остатки*».

Таблица 1.13

Вывод остатка

Наблюдение	Общая сумма ущерба, тыс. руб.	Остатки
1	135,31	5,69
2	186,87	−8,87
3	210,30	20,70
4	224,36	−28,36
5	243,11	−20,11
6	247,79	27,21
7	261,85	0,15
8	304,04	8,96
9	318,10	−5,10
10	360,28	−0,28

Уравнение регрессии зависимости стоимости ущерба, нанесенного пожаром, от расстояния до ближайшей пожарной станции можно записать в следующем виде (в скобках указаны стандартные ошибки параметров регрессии):

$$\hat{y}_i = 102{,}5 + 46{,}9x_i$$
$$(14{,}5) \quad (4{,}6).$$

Интерпретация параметров модели: 46,9 показывает, что при увеличении расстояния до ближайшей пожарной станции на 1 км общая сумма ущерба увеличится в среднем на 46,9 тыс. руб.; 102,5 показывает среднюю сумму ущерба, если объект возгорания находится рядом с пожарной станцией (0 км).

***1b) Построение модели парной линейной регрессии зависимости стоимости ущерба, нанесенного пожаром, от расстояния до ближайшей пожарной станции с использованием надстройки Excel* «Поиск решения»**.

Согласно принципу метода наименьших квадратов, оценки $\hat{\alpha}_0$ и $\hat{\alpha}_1$ находятся путем минимизации суммы квадратов отклонений RSS по всем возможным значениям $\hat{\alpha}_0$ и $\hat{\alpha}_1$ при заданных (наблюдаемых) значениях $x_1, \ldots, x_n, y_1, \ldots, y_n$. Задача сводится к математической задаче поиска точки минимума функции двух переменных. Задача может быть решена с использованием надстройки Excel **«Поиск решения»,** которая позволяет решать оптимизационные задачи.

В диалоговом окне Excel **«Поиск решения»** есть три основных параметра:

«*Оптимизировать целевую функцию*»;

«*Изменяя ячейки переменных*»;

«*В соответствии с ограничениями*».

Подробно применение надстройки Excel **«Поиск решения»** рассмотрено в учебном пособии [1].

Изложим технологию оценки параметров модели линейной регрессии зависимости стоимости ущерба, нанесенного пожаром, от расстояния до ближайшей пожарной станции на основании данных (табл. 1.2) с использованием надстройки Excel **«Поиск решения»** (рис. 1.5—1.7):

• изменяем ячейки переменных. Здесь указываются ячейки, значения в которых будут изменяться для того, чтобы оптимизировать результат в целевой ячейке. В примере 2 это ячейки *E*8 — *F*8 (см. рис. 1.7);

• заполняем поле «*Оптимизировать целевую функцию*». Целевая ячейка связана с другими ячейками этого рабочего листа с помощью формул. В примере 2 это ячейка *J*15, в которой в результате введенных формул получим сумму квадратов отклонений расчетных данных от фактических (рис. 1.5—1.7);

• для запуска надстройки **«Поиск решений»** надо на вкладке «Данные» выбрать команду «*Поиск решения»* и указать в появившемся меню, адреса целевой функции изменяемых ячеек и выполнить поиск наименьшего значения.

	A	B	C	D	E	F	G	H	I	J
1										
2					**МНК в Поиске решений**					
3										
4	**Расстояние до ближайшей станции (x), км**	**Общая сумма ущерба (y), тыс. руб.**			$\hat{y}_i = \hat{a}_0 + \hat{a}_1 \cdot x_i$ $\min\sum_i^n (y_i - \hat{y}_i)^2 = \min\sum_i^n e_i^2$				$\hat{y}_i = \hat{a}_0 + \hat{a}_1 \cdot x_i$	e_i^2
5	0,7	141							=$E$8+$F$8*A5	=(B5-I5)^2
6	1,8	178							=$E$8+$F$8*A6	=(B6-I6)^2
7	2,3	231			$\hat{a}_0$	$\hat{a}_1$			=$E$8+$F$8*A7	=(B7-I7)^2
8	2,6	196							=$E$8+$F$8*A8	=(B8-I8)^2
9	3,0	223							=$E$8+$F$8*A9	=(B9-I9)^2
10	3,1	275							=$E$8+$F$8*A10	=(B10-I10)^2
11	3,4	262							=$E$8+$F$8*A11	=(B11-I11)^2
12	4,3	313							=$E$8+$F$8*A12	=(B12-I12)^2
13	4,6	313							=$E$8+$F$8*A13	=(B13-I13)^2
14	5,5	360							=$E$8+$F$8*A14	=(B14-I14)^2
15										=СУММ(J5:J14)

Рис. 1.5. Введены формулы для вычисления значения целевой функции (режим **«Показать формулы»**)

	A	B	C	D	E	F	G	H	I	J	K
1					**МНК в Поиске решений**						
2											
3	**Расстояние до ближайшей станции (x), км**	**Общая сумма ущерба (y), тыс. руб.**			$\hat{y}_i = \hat{a}_0 + \hat{a}_1 \cdot x_i$ $\min\sum_i^n (y_i - \hat{y}_i)^2 = \min\sum_i^n e_i^2$				$\hat{y}_i = \hat{a}_0 + \hat{a}_1 \cdot x_i$	e_i^2	
4	0,7	141							0.00	19881.00	
5	1,8	178							0.00	31684.00	
6	2,3	231			$\hat{a}_0$	$\hat{a}_1$			0.00	53361.00	
7	2,6	196							0.00	38416.00	
8	3,0	223							0.00	49729.00	
9	3,1	275							0.00	75625.00	
10	3,4	262							0.00	68644.00	
11	4,3	313							0.00	97969.00	
12	4,6	313							0.00	97969.00	Значение целевой функции
13	5,5	360							0.00	129600.00	
14										662878.00	

Рис. 1.6. Получено начальное значение целевой функции

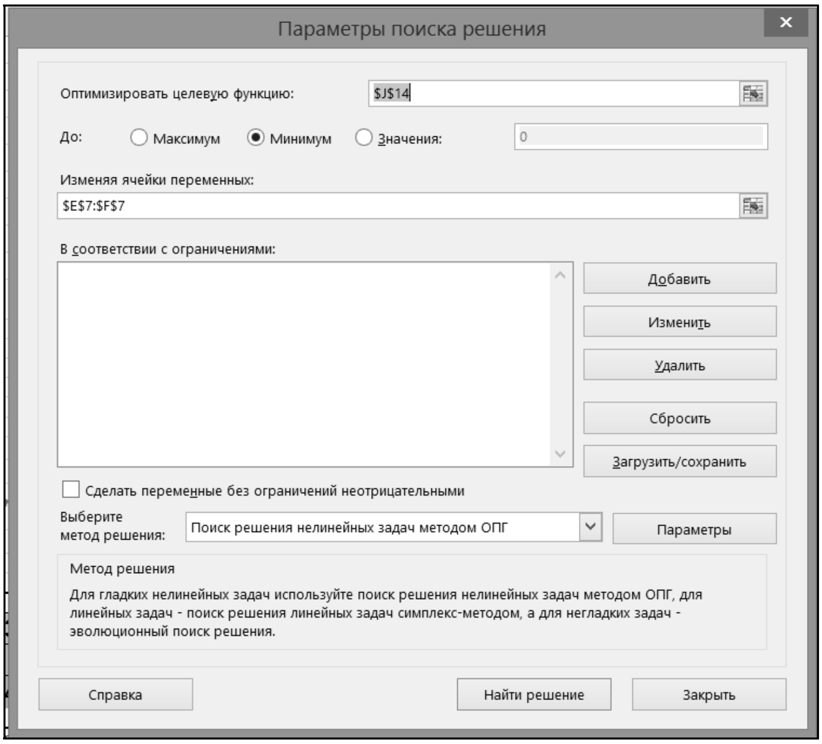


Рис. 1.7. Диалоговое окно **«Поиск решения»**

На рис. 1.8 показано, что оптимальное решение найдено и можно записать полученную модель: $\hat{y}_i = 102,5 + 46,9x_i$. Результаты, полученные с помощью разных инструментов (**«Поиск решения»** и **«Анализ данных»**), совпадают (рис. 1.9).

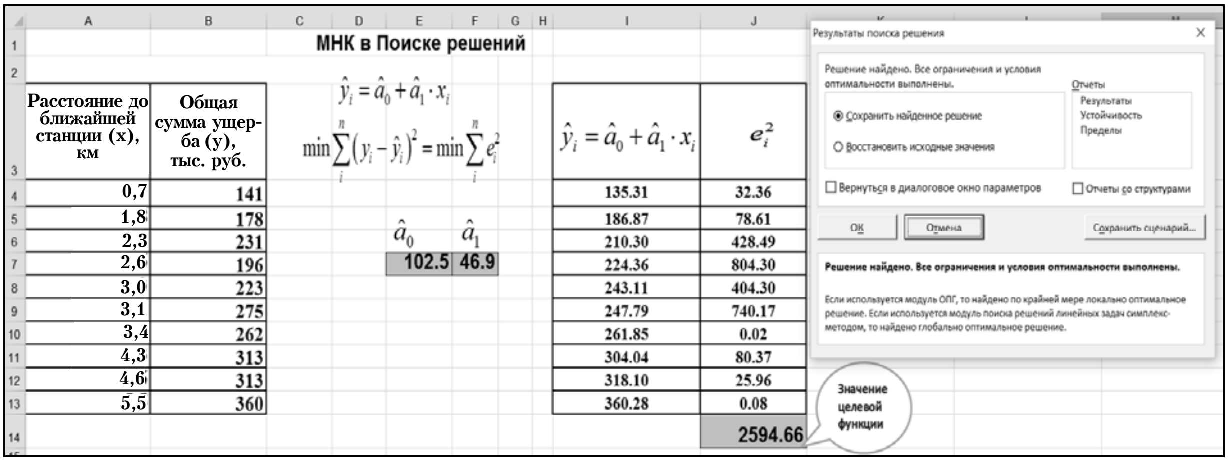


Рис. 1.8. Оптимальное решение найдено

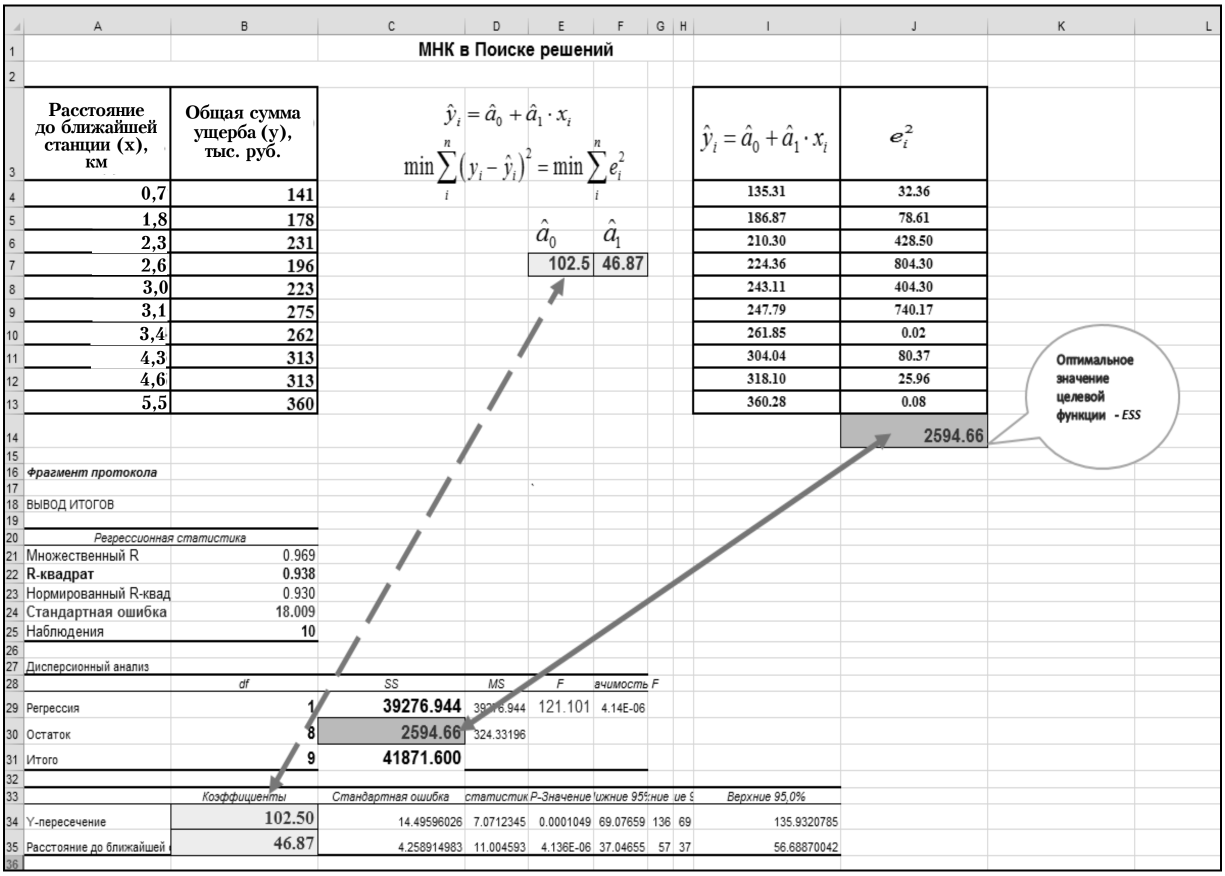


Рис. 1.9. Сравнение результатов, полученных с помощью надстройки **«Поиск решения»** и пакета **«Анализ данных»**

1c) Построение модели линейной регрессии зависимости стоимости ущерба, нанесенного пожаром от расстояния до ближайшей пожарной станции с использованием матричных функций.

В матричной форме расчет параметров линейной модели парной регрессии по приведенной формуле $(A = (X^T X)^{-T} X^T y)$ может быть выполнен с помощью матричных функций **МУМНОЖ** и **МОБР.**

При подготовке данных, формируя матрицу X для вычисления свободного члена a_0, необходимо добавить столбец x_0, состоящий из единиц.

Рассмотрим технологию оценки параметров модели линейной регрессии зависимости стоимости ущерба, нанесенного пожаром от расстояния до ближайшей пожарной станции на основании данных (табл. 1.2) с использованием матричных функций:

• транспонируем матрицу X с использованием функции **ТРАНСП** или путем последовательного копирования и специальной вставки транспонирования. Итак, копируем матрицу X, а затем, используя специальную вставку, получаем транспонированную матрицу;

• умножаем транспонированную матрицу X^T на матрицу X. Выделяем диапазон ячеек для результата умножения матриц — (H8:I9) размером 2×2. Результатом является массив с таким же числом строк, как массив транспонированная матрица X^T, то есть 2 и с таким же числом столбцов, как матрица X, то есть тоже 2. Затем вводим формулу умножения матриц = **МУМНОЖ**(H3:Q4, D4: E13). Нажимаем клавишу **CTRL+SHIFT+ENTER**.

Вычисляем обратную матрицу $(X^T X)^{-1}$, для чего:

• выделяем диапазон для размещения обратной матрицы (H12:I13), в категории «Математические» выбираем функцию вычисления обратной матрицы =**МОБР**(H8:I9), в качестве массива указываем диапазон ячеек H8:I9, где размещена матрица, к которой надо вычислить обратную, и нажимаем клавишу **CTRL+SHIFT+ENTER**;

• умножаем обратную матрицу $(X^T X)^{-1}$ на транспонированную матрицу X^T, для чего выделяем диапазон ячеек для результата умножения матриц (H16:Q17) размером 2×10. Результатом является массив с таким же числом строк, как обратная матрица $(X^T X)^{-1}$, то есть 2, и с таким же числом столбцов, как матрица X^T, то есть 10, вводим формулу умножения матриц = **МУМНОЖ**(H12:I13, D4: E13) и нажамаем клавишу **CTRL+SHIFT+ENTER.**

• умножаем матрицу $(X^T X)^{-1} X^T$ на *y,* для чего выделяем диапазон ячеек для результата умножения матриц (H19:H20) размером 2×1, затем вводим формулу умножения матриц = **МУМНОЖ**(H16:Q17, B4:B13) и нажимаем клавишу **CTRL+SHIFT+ENTER.**

В ячейках H19:H20 будут находиться параметры модели линейной регрессии a_0 = 102,5 и a_1 = 46,87.

В матричной форме расчет параметров модели представлен на рисунках 1.10—1.11.

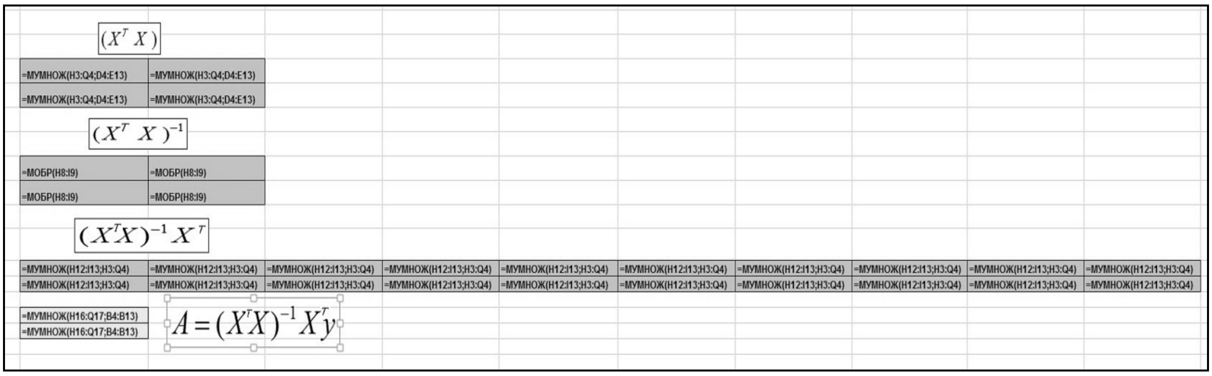


Рис. 1.10. Фрагмент рис. 1.11 в режиме **«Показать формулы»**

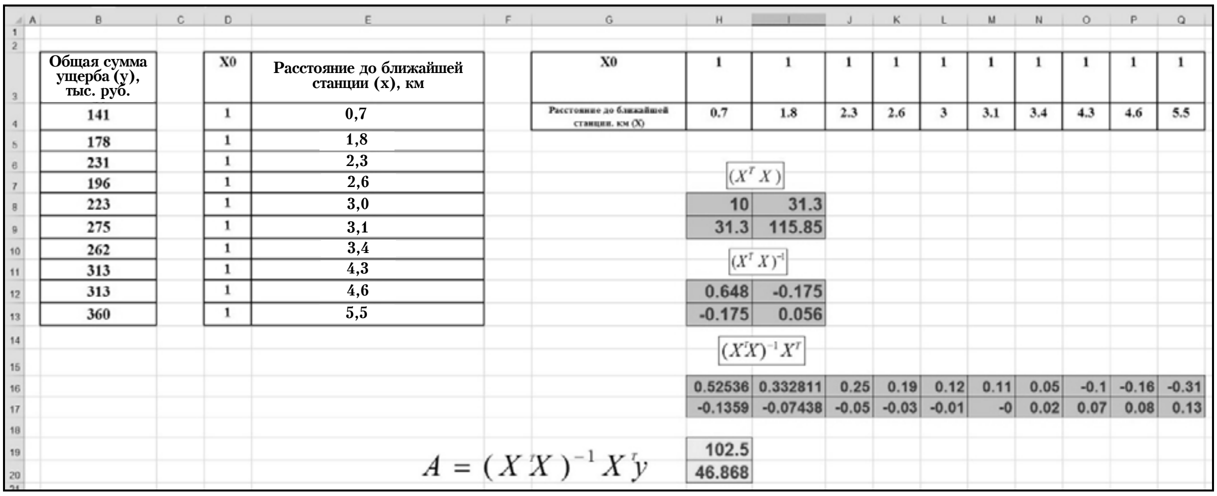


Рис. 1.11. Вычисления параметров модели с помощью матричных функций **МУМНОЖ** и **МОБР**

Результаты, полученные с помощью разных инструментов (надстройка **«Поиск решения»,** пакет **«Анализ данных»** и матричные функции), совпадают (см. рисунок 1.11).

1d) Построение модели линейной регрессии зависимости стоимости ущерба, нанесенного пожаром, от расстояния до ближайшей пожарной станции с использованием функции ЛИНЕЙН табличного процессора Excel.

Рассмотрим технологию оценки параметров модели линейной регрессии зависимости стоимости ущерба, нанесенного пожаром, от расстояния до ближайшей пожарной станции на основании данных (см. таблицу 1.2) с использованием функции **ЛИНЕЙН**:

• вначале выделяем диапазон для размещения результатов выполнения функции **ЛИНЕЙН**. На листе выделяется область высотой пять строк и шириной равной количеству столбцов с данными, в примере 2 их два (F5:G9). Затем вызываем функцию **ЛИНЕЙН**;

• при применении функции **ЛИНЕЙН** табличного процессора Excel на экран дисплея выдается диалоговое окно (рис. 1.12) для задания значений y и x;

• в данном окне в строке «*Известные_значения_y*» следует внести все 10 значений из столбца y, а в строке «*Известные_значения_x*» — соответственно все 10 значений из столбца x. Если необходимо выполнить оценку двух параметров — постоянного и регрессионного, то в строке «*Конст*» следует указать значение 1. Так как при оценке параметров модели интересны, в первую очередь, статистические сведения, то в окне строки «*Статистика*» следует указать значение 1. После ввода всех значений, как показано на рис. 1.13, следует нажать клавиши **CTRL+SHIFT+ENTER.**

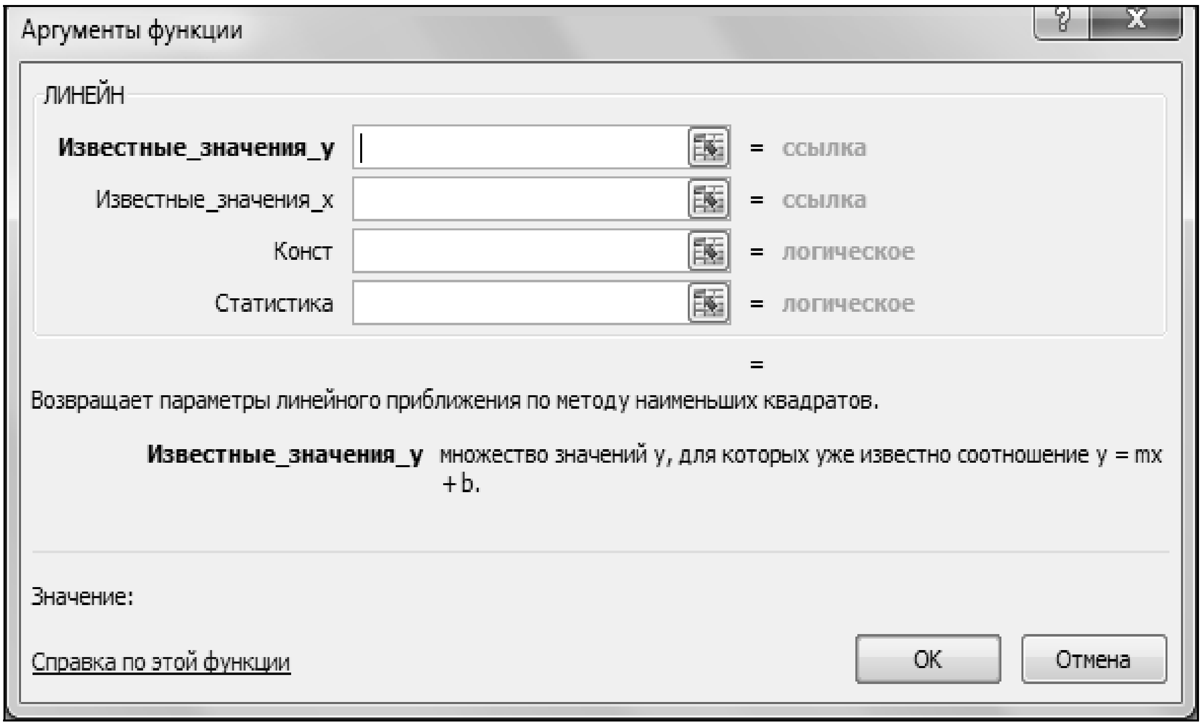


Рис. 1.12. Диалоговое окно функции **ЛИНЕЙН**

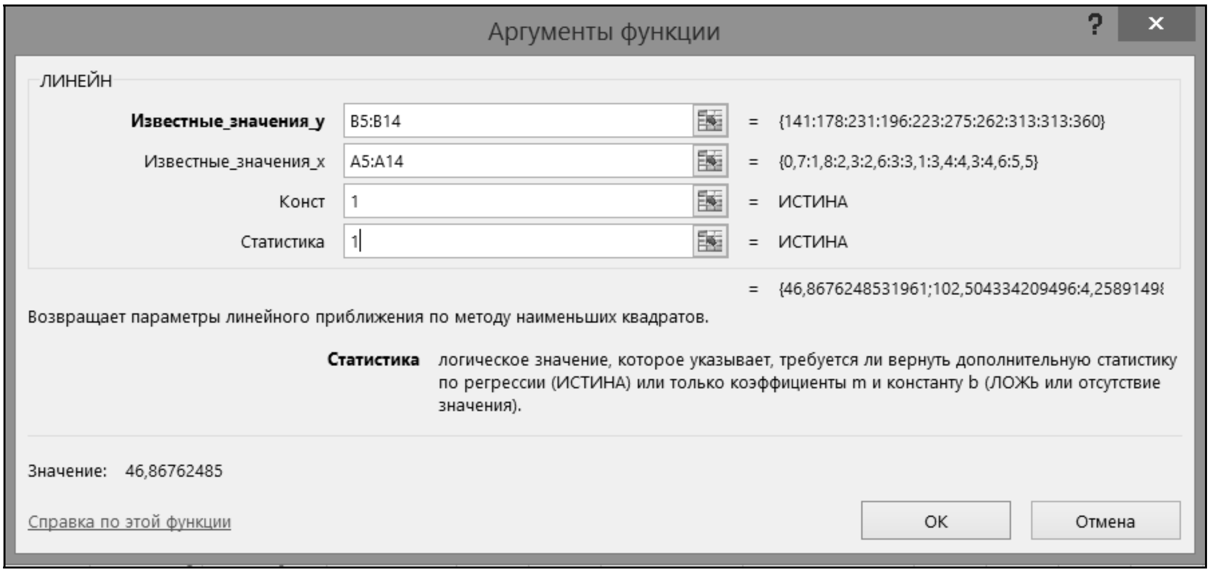


Рис. 1.13. Диалоговое окно функции **ЛИНЕЙН** после задания необходимых данных

После этого будет выдан результат, как показано на рис. 1.14.

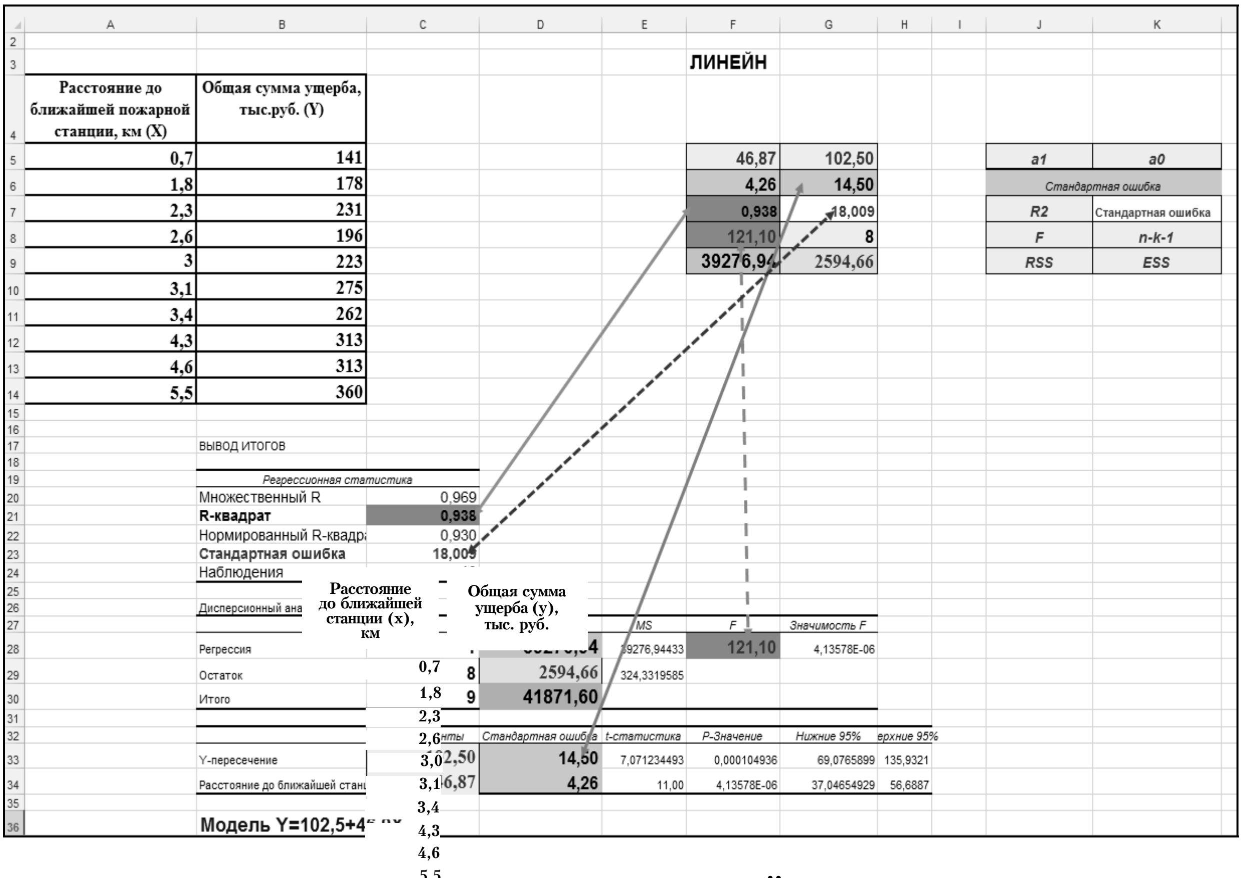


Рис. 1.14. Результат выполнения функции **ЛИНЕЙН** совпадает с результатом, полученным с использованием пакета **«Анализ данных»**

На рис. 15 приведен фрагмент рис. 1.14.

$\hat{a}_0$	46,87	102,50	$\hat{a}_1$
$\hat{\sigma}_{a_0}$	4,26	14,50	$\hat{\sigma}_{a1}$
R^2	0,94	18,01	$\hat{\sigma}$
F	121,10	8	ν
ESS	39276,94	2594,66	*RSS*

Рис. 15. Фрагмент рис. 14 со значениями параметров

Слева и справа от полученных значений указаны их обозначения (наименования), которые поясняются ниже:

$\hat{a}_0$ и $\hat{a}_1$ — оценки параметров полученной парной регрессии (1-я строка);

$\hat{\sigma}_{a_0}$ и $\hat{\sigma}_{a_1}$ — стандартные ошибки оцененных параметров (2-я строка);

R^2 — коэффициент детерминации (3-я строка, 1-й столбец);

$\hat{\sigma}$ — стандартная ошибка полученной регрессии (3-я строка, 2-й столбец);

F — статистика Фишера полученной парной регрессии (4-я строка, 1-й столбец);

ν — степень свободы (3-я строка, 2-й столбец);

RSS — (5-я строка, 1-й столбец);

ESS — квадрат остатков для полученной парной регрессии (5-я строка, 2-й столбец).

Итак, в первой строке выдаются значения параметров $\hat{a}_0$ и $\hat{a}_1$, а во второй строке соответственно их стандартные ошибки $\hat{\sigma}_{a_0}$ и $\hat{\sigma}_{a_1}$.

В ячейках F8 и G8 будут находиться параметры модели линейной регрессии $\hat{a}_0$ = 102,5 и $\hat{a}_1$ = 46,87.

2) Проверка качество уравнения регрессии. Качество модели оценивается коэффициентом детерминации R^2

$$R^2 = 1 - \frac{\sum \hat{\varepsilon}_i^2}{\sum (y_i - \bar{y})^2} = 1 - 25{,}947/418{,}716 = 0{,}938.$$

Величина R^2 = 0,938 означает, что фактором расстояния до ближайшей пожарной станции можно объяснить 93,8% вариации (разброса) стоимости ущерба, нанесенного пожаром.

Точность модели оценим с помощью средней ошибки аппроксимации (1.17) $E_{\text{отн}}$ = 5,6%. Точность модели хорошая.

Оценим *значимость уравнения регрессии* с помощью критерия Фишера. Расчетное значение F-критерия вычислим по формуле (1.15)

$$F = \frac{R^2/k}{\left(1 - R^2\right) / \left(n - k - 1\right)} = \frac{0{,}938 / 1}{(1 - 0{,}938) / (10 - 2)} = 121{,}101.$$

Расчетное значение F-критерия Фишера можно найти в табл. 1.11 протокола Excel (см. с. 25).

Уравнение регрессии значимо на уровне α, если расчетное значение $F > F_{табл}$, где $F_{табл}$ — табличное значение F-критерия Фишера.

Табличное значение F-критерия можно найти с помощью функции **FРАСПОБР***. Табличное значение F-критерия при $\alpha = 0{,}05$ при $\nu_1 = k = 1$ и $\nu_2 = n - k - 1 = 10 - 1 - 1 = 8$ составляет 5,318.

Поскольку $F_{рас} > F_{табл}$, уравнение регрессии следует признать значимым.

Оценим *значимость коэффициентов полученной модели*, используя результаты отчета Excel. Это можно сделать тремя способами.

Коэффициент уравнения регрессии признается значимым в том случае, если:

• наблюдаемое значение t-статистики Стьюдента для этого коэффициента больше, чем критическое (табличное) значение статистики Стьюдента (для заданного уровня значимости, например, $\alpha = 0{,}05$ и числа степеней свободы $df = n - k - 1$, где n — это число наблюдений, k — число факторов в модели);

• P-значение t-статистики Стьюдента для этого коэффициента меньше, чем уровень значимости, например, $\alpha = 0{,}05$;

• доверительный интервал для этого коэффициента (вычисленный с некоторой доверительной вероятностью, например, 95%) не содержит ноль внутри себя, то есть если нижняя 95% и верхняя 95% границы доверительного интервала имеют одинаковые знаки.

Значимость коэффициента $\hat{a}_1$ проверим по второму и третьему способам, используя данные табл. 1.12.

P-значение $(\hat{a}_1) = 0{,}00 < 0{,}01 < 0{,}05$.

Следовательно, коэффициент $(\hat{a}_1)$ зна́чим при 1%-ном уровне, а тем более при 5%-ном уровне значимости.

Нижняя 95%-ная граница равна 3,70 и верхняя 95%-ная граница равна 5,67, границы доверительного интервала имеют одинаковые знаки, следовательно, коэффициент $(\hat{a}_1)$ значим.

3) Построение прогноза. Для того чтобы оценить стоимость ущерба, нанесенного пожаром, если расстояние до ближайшей пожарной станции уменьшится на 10% от своего среднего уровня, необходимо найти ожидаемое значение фактора x, затем следует подставить значение $x_{прогн} = x_{средн} \cdot 0{,}9 = 2{,}82$ в полученную модель

$$y_{прогн} = 102{,}5 + 0{,}469 \cdot 2{,}82 = 234{,}5.$$

* В Excel 2010 название функции **FРАСПОБР** изменено на **F.ОБР.ПХ**.

Доверительный интервал для прогнозов индивидуальных значений y_i определяется из соотношения (1.24)

$$y_i \in \left[\hat{y}_i \pm U(X_{\text{прогн}})\right] = \left[\hat{y}_i \pm t_\alpha \hat{\sigma}_\varepsilon \sqrt{1 + \frac{1}{n} + \frac{(x_{прогн} - \bar{x})^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}}\right] =$$

$$= 234{,}5 \pm 1{,}801 \times 1{,}86\sqrt{1 + \frac{1}{10} + \frac{(2{,}82 - 3{,}13)^2}{17{,}881}} = 234{,}53 \pm 3{,}521.$$

Коэффициент Стьюдента $t_{(0{,}1;8)}$ для m = 8 степеней свободы и уровня значимости 0,1 равен 1,86; стандартная ошибка равна 1,801 (рис. 1.16).

Таким образом, прогнозное значение $y_{\text{прогн}}$ = 234,53 с вероятностью 90% будет находиться между верхней границей, равной 269,74, и нижней границей, равной 199,32.

	A	B	C	D	E	F	G	H
1	**Общая сумма ущерба (y), тыс. руб.**	**Расстояние до ближайшей станции (x), км**						
2	**262**	**3,4**		$y_i \in \left[\hat{y}_i \pm U(X_{\text{прогн}})\right] = \left[\hat{y}_i \pm t_\alpha \hat{\sigma}_\varepsilon \sqrt{1+\frac{1}{n}+\frac{(x_{прогн}-\bar{x})^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2}}\right] =$				
3	**178**	**1,8**						
4	**313**	**4,6**						
5	**231**	**2,3**				**хпрогн**	**ВГ**	**НГ**
6	**275**	**3,1**		$234{,}5 \pm 1{,}801 \times 1{,}86\sqrt{1+\frac{1}{10}+\frac{(2{,}82-3{,}13)^2}{17{,}881}} =$		**2.817**	**=E18+E16**	**=E18-E16**
7	**360**	**5,5**				**2.817**	**269.74**	**199.32**
8	**141**	**0,7**						
9	**223**	**3,0**		$= 234{,}53 \pm 3{,}521.$				
10	**196**	**2,6**						
11	**313**	**4,3**		$x_{прогн} = 0{,}9 \cdot \bar{x}$	**=0.9*B12**	**2.82**		
12	$\bar{x}$	***=СРЗНАЧ(B2:B11)***	**3.13**	$\hat{\sigma}_\varepsilon =$	**стандартная ошибка**	**18.009**		
13	$\sum_{i=1}^{n}(x_i-\bar{x})^2$	***=КВАДРОТКЛ(B2:B11)***	**17.881**	t_α	**=СТЬЮДЕНТ.ОБР.2Х(0.1;8)**	**1.86**		
14				$(x_{прогн}-\bar{x})^2$	**=(E11-B12)^2**	**=(2.82-3.13)^2**		
15								
16	Из протокола			$U(X_{\text{прогн}})$	**=F12*F13*КОРЕНЬ(1+1/10+F14/C13)**	**3.521**		
17								
18		*Коэффициенты*		$\hat{y}_i = y_{прогн}$	**=B19+B20*E11**	**234.53**		
19	Y-пересечение	102.504334209496						

Рис. 1.16. Результат выполнения пункта 3 в Microsoft Excel
(в режиме **«Показать формулы»**)

4) Построение графика. Выполнение пункта 4 — построение графика прогноза приведено на рис. 1.17. На график подбора нанесем прогнозное значение y, нижнюю и верхнюю границы.

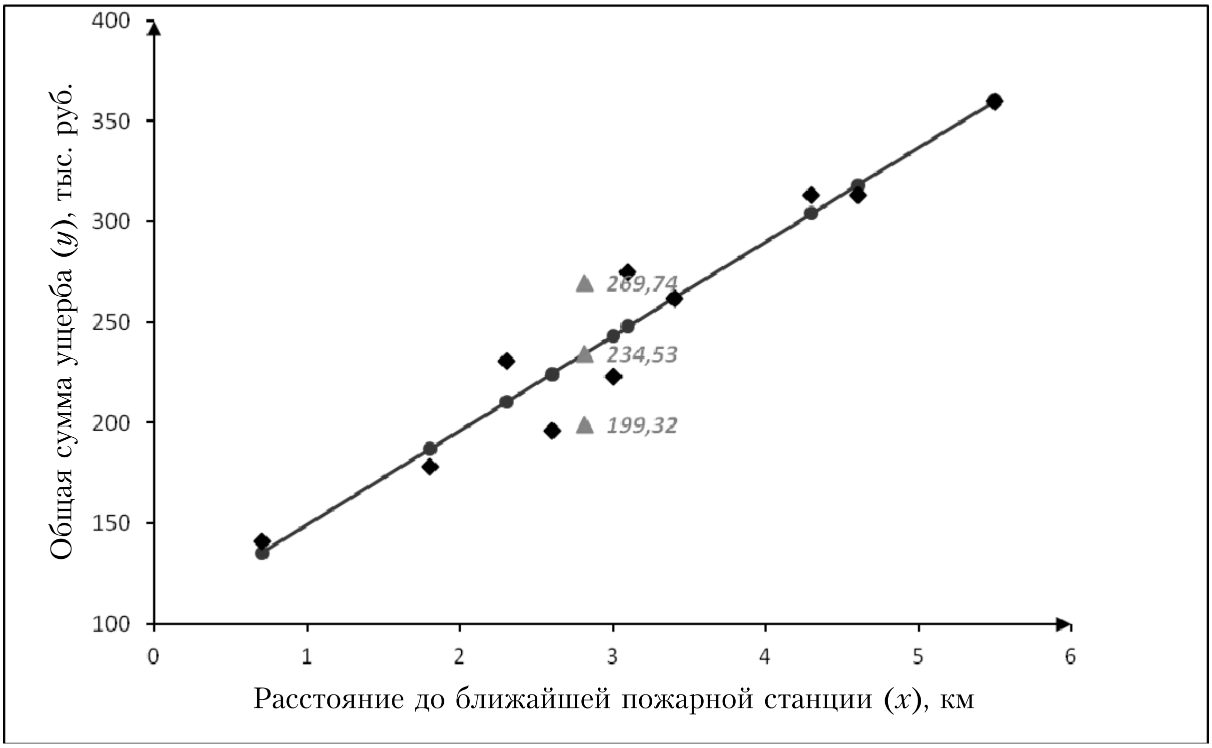


Рис. 1.17. График модели парной регрессии зависимости стоимости ущерба, нанесенного пожаром, от расстояния до ближайшей пожарной станции (точечный и интервальный прогноз)

1.4. Некоторые вопросы применения моделей множественной регрессии

Линейная модель множественной регрессии имеет вид

$$y_i = \alpha_0 + \alpha_1 x_{i1} + \alpha_2 x_{i2} + \ldots + \alpha_k x_{ik} + \varepsilon_i, \quad i = 1, \ldots, n, \tag{1.25}$$

где n — количество наблюдений;

k — количество факторов, включенных в модель.

Для того чтобы регрессионный анализ, основанный на обычном методе наименьших квадратов, давал наилучшие из всех возможных результаты, должны выполняться следующие условия (предпосылки), известные как условия Гаусса–Маркова.

Основные предпосылки метода наименьших квадратов

Предполагается, что истинная зависимость (1.25) имеет ***линейный вид***. Линейность по переменным и линейность по оценкам. В матричном виде это можно записать в следующем виде:

$$y = XA + \varepsilon.$$

С помощью МНК оценивается регрессия y

$$A = (X^T X)^{-1} X^T y = \begin{pmatrix} \hat{\alpha}_0 \\ \hat{\alpha}_1 \\ \dots \\ \hat{\alpha}_k \end{pmatrix} \tag{1.26}$$

Наблюдений больше, чем оцениваемых коэффициентов, то есть $n > k$.

Предпосылки на остаточную компоненту ε_i

Первое условие. Математическое ожидание случайной составляющей в любом наблюдении должно быть равно нулю

$$E(\varepsilon_i \mid X) = 0. \tag{1.27}$$

Фактически если уравнение регрессии включает постоянный член, то обычно это условие выполняется автоматически, так как роль константы состоит в определении любой систематической тенденции y, которую не учитывают объясняющие переменные, включенные в уравнение регрессии.

Второе условие. Дисперсия случайной составляющей должна быть постоянна для всех наблюдений.

Это условие гомоскедастичности составляющей (возмущения):

$Var(\varepsilon_i \mid \textit{все регрессоры}) = \sigma^2$, или в матричном виде

$$E(\varepsilon_i^2 \mid X) = \sigma^2. \tag{1.28}$$

Третье условие. Предполагает отсутствие систематической связи между значениями случайной составляющей в любых двух наблюдениях.

Условная некоррелированность

$$Cov(\varepsilon_i, \varepsilon_j \mid X) = 0 \text{ при } i \neq j. \tag{1.29}$$

Возмущения ε_i и ε_j не коррелированы (условие независимости случайных составляющих в различных наблюдениях).

К предпосылкам на регрессоры относятся следующие:

— среди регрессоров нет линейно зависимых, или, что то же самое, матрица $(X^T X)^{-1}$ существует, или определитель $(X^T X)$ не равен нулю $\det(X^T X) \neq 0$

$$Cov(\varepsilon_i, \varepsilon_j \mid X) = 0 \text{ при } i \neq j; \tag{1.30}$$

— векторы отдельных наблюдений независимы и одинаково распределены.

Предположение о нормальности

Наряду с перечисленными условиями обычно также предполагается нормальность распределения случайного члена. Дело в том, что если случайный член нормально распределен, то так же будут распределены и коэффициенты регрессии.

Свойства оценок МНК

В тех случаях, когда предпосылки выполняются, оценки, полученные по МНК, будут обладать свойствами несмещенности, состоятельности и эффективности.

Несмещенность оценки означает, что математическое ожидание остатков равно нулю. Если оценки обладают свойством несмещенности, то их можно сравнивать по разным исследованиям.

$$E(\hat{a}_j \mid X) = a_j.$$

Для практических целей важна не только несмещенность, но и эффективность оценок.

Оценки считаются эффективными среди линейных и несмещенных, если они характеризуются наименьшей дисперсией. Поэтому несмещенность оценки должна дополняться минимальной дисперсией.

Степень достоверности доверительных интервалов параметров регрессии обеспечивается, если оценки будут не только несмещенными и эффективными, но и состоятельными. ***Состоятельность оценок*** характеризует увеличение их точности с увеличением объема выборки. Метод наименьших квадратов дает оценки, имеющие наименьшую дисперсию в классе всех линейных оценок, если выполняются предпосылки нормальной линейной регрессионной модели.

Наиболее сложным и важным при построении модели множественной регрессии является отбор факторов для включения в модель. С возникающими при этом проблемами мультиколлинеарности факторов, гетероскедастичности остатков и автокорреляции остатков можно ознакомиться в работах [1–5; 9] и др. С построением моделей с фиктивными переменными можно ознакомиться в работах [4; 5; 8; 9].

1.4.1. Проверка условия гомоскедастичности случайной составляющей (возмущения)

Гетероскедастичностью называется нарушение условия постоянства дисперсии ошибок в линейной модели регрессии (1.28), но при этом предполагается, что остальные условия выполнены.

Нарушение условия гомоскедастичности возмущений означает, что дисперсия возмущения зависит от значений факторов. Такие регрессионные модели называются моделями с гетероскедастичностью возмущений. Часто гетероскедастичность ошибок возникает при построении регрессионных моделей для неоднородных.

Обнаружение гетероскедастичности

Для обнаружения гетероскедастичности обычно используют тесты, в которых делаются различные предположения о зависимости между дисперсией случайного члена и объясняющей переменной: тест ранговой корреляции Спирмена, тест Голдфельда–Квандта, тест Уайта, тест Глейзера, двусторонний критерий Фишера и др. [1].

Последствия гетероскедастичности

При гетероскедастичности последствия применения МНК будут следующими.

1. Оценки коэффициентов по-прежнему останутся несмещенными и линейными.

2. Оценки не будут эффективными (не будут иметь наименьшую дисперсию по сравнению с другими оценками такого же параметра). При увеличении дисперсии оценок снижается вероятность получения максимально точных оценок.

3. Дисперсии оценок будут рассчитываться со смещением.

4. Вследствие того, что было сказано выше, все выводы, получаемые на основе соответствующих t- и F- статистик (критериев Стьюдента и Фишера), а также ин-

тервальные оценки будут ненадежными. Значит, статистические выводы, которые получаются при стандартных проверках качества оценок, могут быть ошибочными и приводить к неверным выводам по построенной модели. Вполне вероятно, что стандартные ошибки коэффициентов будут занижены, следовательно, t-статистики будут завышены. Это может приводить к признанию статистически значимыми коэффициентов, таковыми на самом деле не являющихся.

Обнаружение гетероскедастичности

Для обнаружения гетероскедастичности обычно используют тесты, в которых делаются различные предположения о зависимости между дисперсией случайного члена и объясняющей переменной: тест ранговой корреляции Спирмена, тест Голдфельда–Квандта, тест Глейзера, двусторонний критерий Фишера и другие [1; 4; 9].

Следует отметить, что условие (1.28) означает, что дисперсия истинной ошибки является постоянной величиной на любом из отрезков рассматриваемого интервала. В связи с этим проверка условия (1.28) может быть идентична проверке гипотезы о постоянстве дисперсии фактической ошибки на различных отрезках интервала.

Графический анализ является простым и эффективным методом предварительного анализа однородности ошибок регрессии.

Для этого строят следующие графики:

- остатки в зависимости от оцененных значений — $\hat{\varepsilon}$ против $\hat{y}$;
- остатки в зависимости от отдельных объясняющих переменных — $\hat{\varepsilon}$ против x;
- модуль остатков в зависимости от отдельных объясняющих переменных — $|\hat{\varepsilon}|$ против x;
- остатки в зависимости от номера наблюдений $\hat{\varepsilon}$ против t (только для временных рядов).

Тест Голдфельда–Квандта (Goldfeld–Quandt)

Данный тест используется для проверки такого типа гетероскедастичности, когда дисперсия остатков возрастает пропорционально квадрату фактора. При этом делается предположение, что случайная составляющая ε распределена нормально.

Чтобы оценить нарушение гомоскедастичности по тесту Голдфельда–Квандта, необходимо выполнить следующие шаги.

1. Упорядочение n наблюдений по мере возрастания переменной x.
2. Исключение средних наблюдений (например, 20%).
3. Разделение совокупности на две группы (соответственно с малыми и большими значениями фактора x) и определение по каждой из групп уравнений регрессии.
4. Определение остаточной суммы квадратов для первой регрессии $RSS_1 = S_{1\hat{y}} = \sum_{i=1}^{n1}(y_i - \hat{y}_{1i})^2$ и второй регрессии $RSS_2 = S_{2\hat{y}} = \sum_{i=n-n1+1}^{n}(y_i - \hat{y}_{2i})^2$.
5. Вычисление отношений $S_{2\hat{y}} / S_{1\hat{y}}$ (или $S_{1\hat{y}} / S_{2\hat{y}}$). В числителе должна быть большая сумма квадратов.

Полученное отношение имеет F-распределение со степенями свободы $k_1 = n_1 - k$ и $k_2 = n - n_1 - k$, (k — число оцениваемых параметров в уравнении регрессии).

Если $S_{2\hat{y}} / S_{1\hat{y}}$ $F_{\text{набл}} = \dfrac{S_{1\hat{y}}}{S_{2\hat{y}}} > F_{\text{кр}\,(\alpha;\nu_1;\nu_2)}$, то гетероскедастичность имеет место.

Чем больше величина F превышает табличное значение F-критерия, тем более нарушена предпосылка о равенстве дисперсий остаточных величин.

Пример 3

По данным 39 стран построено уравнение регрессии зависимости числа банков на 100 000 человек (y) от активов банков к ВВП (%) — x

$$\hat{y}_i = -4,77 + 0,07x_i.$$

Для предварительного анализа однородности ошибок регрессии построен график: остатки в зависимости от оцененных значений — $\hat{\varepsilon}$ против $\hat{y}$.

График, приведенный на рис. 1.18, свидетельствует о неоднородности ошибок регрессии.

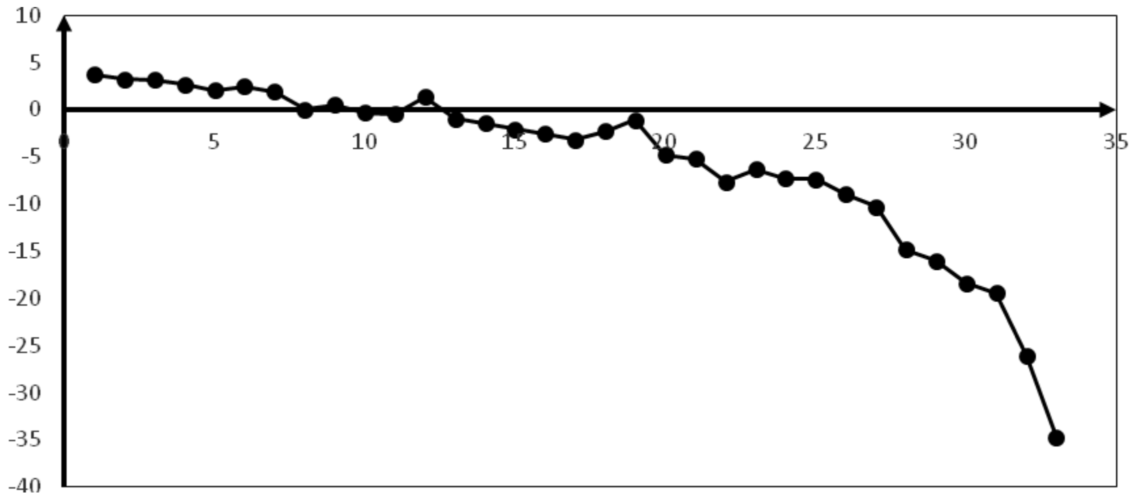


Рис. 1.18. График остатков (ошибок регрессии)

Чтобы подтвердить наличие гетероскедастичности с помощью теста Голдфельда–Квандта, оценим эту модель по 16 странам с наименьшим значением x и по 16 странам с наибольшим значением x. В первом случае RSS_1 = 4,12, а во втором RSS_2 = 165 700,2.

Решение

Проверим гипотезу H_0: $E(\varepsilon_i^2 \mid X) = \sigma^2$ с помощью критерия Фишера

$$F_{\text{набл}} = \frac{RSS_2}{RSS_1} = \frac{165\ 700,2}{4,12} = 40\ 232 > F_{\text{кр}\,(0,05;14;14)} = 2,48.$$

Так как наблюдаемое значение F-статистики больше критического, то H_0 отвергается.

Есть разные способы борьбы с гетероскедастичностью [4; 9].

В примере 3 эту проблему удалось решить, построив уравнение регрессии по логарифмам: $\ln\hat{y}_i = -7,54 + 1,35\ln x_i$. Анализ графика остатков (рис. 1.19) позволяет предположить, что условие гомоскедастичности выполняется. Проверим предположение с помощью теста Голдфельда–Квандта.

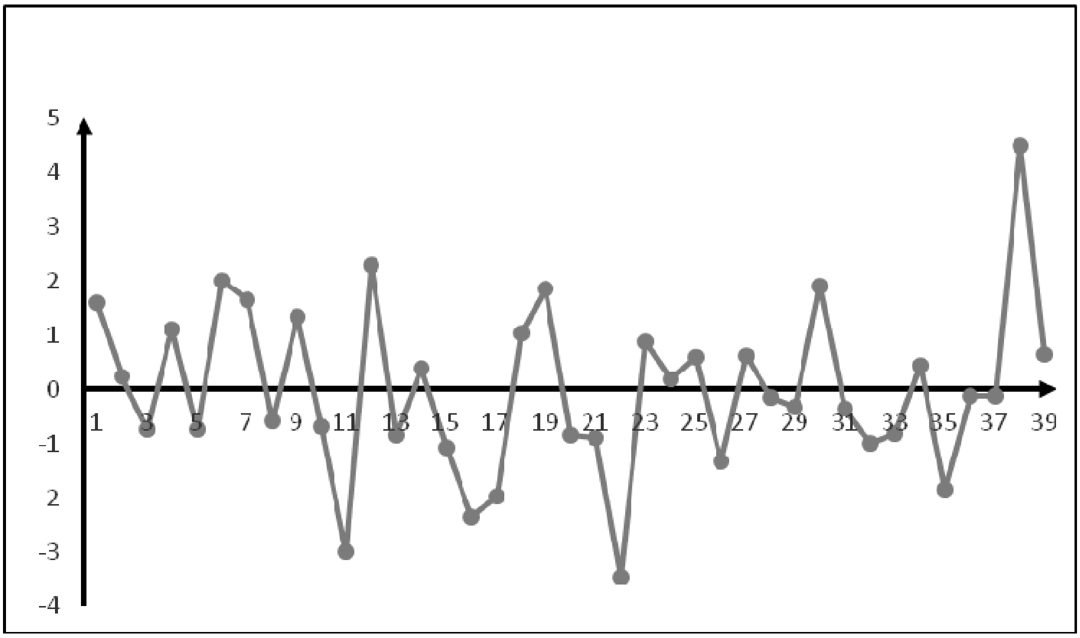


Рис. 1.19. График остатков модели $\ln\hat{y}_i = -7,54 + 1,351\mathrm{n}x_i$

Проведем тест Голдфельда–Квандта для модели $\ln\hat{y}_i = -7,54 + 1,351\mathrm{n}x_i$. Оценим эту модель по 16 странам с наименьшим значением $\ln x$ и по 16 странам с наибольшим значением $\ln x$ и проверим гипотезу H_0: $E(\varepsilon_i^2 \mid x) = \sigma^2$ с помощью критерия Фишера

$$F_{\text{набл}} = \frac{32,26}{30,06} = 1,04 < F_{\text{кр }(0,05;14;14)} = 2,48.$$

Так как наблюдаемое значение F-статистики меньше критического, то H_0 не отвергается.

1.4.2. Проверка условия независимости случайных составляющих в различных наблюдениях

Автокорреляция — статистическая взаимосвязь между случайными величинами из одного ряда, но взятых со сдвигом.

Автокорреляция случайной составляющей нарушает третьею предпосылку нормальной линейной модели регрессии, которая предполагает отсутствие систематической связи между значениями случайной составляющей в любых двух наблюдениях.

Автокорреляция отклонений чаще всего наблюдается тогда, когда эконометрическая модель строится на основе временных рядов. Если существует корреляция между последовательными значениями некоторой независимой переменной, то будет наблюдаться и корреляция последовательных значений остатков.

Автокорреляция может быть также следствием ошибочной спецификации эконометрической модели. Кроме того, наличие автокорреляции остатков может означать, что необходимо ввести в модель новую независимую переменную.

Последствия автокорреляции

Последствия автокорреляции в определенной степени сходны с последствиями гетероскедастичности. Среди них при использовании МНК обычно выделяют следующие.

1. Оценки параметров, оставаясь линейными и несмещенными, перестают быть эффективными. Следовательно, они перестают обладать свойствами наилучших линейных несмещенных оценок.

2. Дисперсии оценок являются смещенными. Зачастую дисперсии, вычисляемые по стандартным формулам, являются заниженными, что приводит к увеличению t-статистик. Это может привести к признанию статистически значимыми объясняющие переменные, которые в действительности таковыми могут и не являться.

3. Оценка дисперсии регрессии является смещенной оценкой истинного значения, во многих случаях занижая его.

4. В силу вышесказанного выводы по t- и F-статистикам, определяющим значимость коэффициентов регрессии и коэффициента детерминации, возможно, будут неверными. Вследствие этого ухудшаются прогнозные качества модели.

В линейной регрессионной модели либо в моделях, сводящихся к линейной, наиболее целесообразным и простым преобразованием, устраняющим автокорреляцию, является авторегрессионная схема первого порядка AR(1).

Обнаружение автокорреляции

Наличие (отсутствие) автокорреляции в отклонениях проверяют с помощью критерия Дарбина–Уотсона. Численное значение коэффициента равно

$$dw = \frac{\sum_{i=2}^{n} (\hat{\varepsilon}_i - \hat{\varepsilon}_{i-1})^2}{\sum_{i=1}^{n} \hat{\varepsilon}_i^2}, \tag{1.31}$$

где $\hat{\varepsilon}_i = y_i - \hat{y}_i$.

Значение dw статистики близко к величине $2(1 - r(1))$, где $r(1)$ — выборочная автокорреляционная функция остатков первого порядка. Таким образом, значение статистики Дарбина–Уотсона распределено в интервале от 0 до 4. Соответственно, идеальное значение статистики — 2 (автокорреляция отсутствует). Меньшие значения критерия соответствуют положительной автокорреляции остатков, большие значения — отрицательной. Статистика учитывает только автокорреляцию первого порядка. Верхние (d_2) и нижние (d_1) критические значения, позволяющие принять или отвергнуть гипотезу об отсутствии автокорреляции, зависят от количества уровней динамического ряда и числа независимых переменных модели. Значения этих границ для уровня значимости $\alpha = 0{,}05$ даны в специальных таблицах (см. Приложение 1).

При сравнении расчетного значения dw статистики (1.31) с табличным могут возникнуть такие ситуации: $d_2 < dw < 2$ — ряд остатков не коррелирован; $dw < d_1$ — остатки содержат автокорреляцию; $d_1 < dw < d_2$ — область неопределенности, когда нет оснований ни принять, ни отвергнуть гипотезу о существовании автокорреляции. Если d превышает 2, то это свидетельствует о наличии отрицательной корреляции. Перед сравнением с табличными значениями dw критерий следует преобразовать по формуле $dwr = 4 - dw$.

Следует иметь в виду, что тест Дарбина–Уотсона можно применять только в случае выполнения следующих условий:

1) в регрессионном уравнении присутствует свободный член;

2) регрессоры являются не стохастическими;

3) в регрессионном уравнении нет лаговых значений зависимой переменной.

Установив наличие автокорреляции остатков, переходят к улучшению модели. Если же ситуация оказалась неопределенной ($d_1 < dw < d_2$), то применяют другие критерии. В частности, можно воспользоваться **первым коэффициентом автокорреляции**

$$r(1) = \left(\sum_{i=2}^{n} \hat{\varepsilon}_i \hat{\varepsilon}_{i-1} \right) \Big/ \sum_{i=1}^{n} \hat{\varepsilon}_i^2 . \tag{1.32}$$

Для принятия решения о наличии или отсутствии автокорреляции в исследуемом ряду фактическое значение коэффициента автокорреляции $r(1)$ сопоставляется с табличным (критическим) значением для 5%-ного уровня значимости (вероятности допустить ошибку при принятии нулевой гипотезы о независимости уровней ряда). Если фактическое значение коэффициента автокорреляции меньше табличного, то гипотеза об отсутствии автокорреляции в ряду может быть принята, а если фактическое значение больше табличного — делают вывод о наличии автокорреляции в ряду динамики.

В связи с тем, что наличие в модели регрессии автокорреляции между остатками модели может привести к негативным результатам всего процесса построения модели, автокорреляция остатков должна быть устранена. Можно попытаться исправить неверную спецификацию модели или включить неучтенные факторы.

Авторегрессионной схемой первого порядка называется метод устранения автокорреляции первого порядка между соседними членами остаточного ряда в линейных моделях регрессии либо моделях регрессии, которые можно привести к линейному виду.

Для устранения автокорреляции используют процедуру Кохрейна–Оркатта и процедуру Хильдрата–Лу.

Пример 4

На основании данных о численности населения Российской Федерации с 1991 по 2016 гг.* (табл. 1.14) была построена модель $\hat{y}_i = 1485\,21{,}7 - 232{,}12$, где $t = 1, 2, ..., 26$. График исходных данных и результатов моделирования приведен на рис. 1.20.

Таблица 1.14

Численность населения, Российской Федерации с 1991 по 2016 гг., млн чел.

Год	Численность населения Российской Федерации
1991	148543
1992	148704
1993	148673
1994	148366

* Текущие оценки численности населения на 1 января рассчитываются на основании итогов последней переписи населения, к которым ежегодно прибавляются числа родившихся и прибывших на данную территорию и из которых вычитаются числа умерших и выбывших с данной территории. Текущие оценки численности населения за прошедшие годы уточняются на основании итогов очередной переписи.

Год	Численность населения Российской Федерации
1995	148306
1996	147976
1997	147502
1998	147105
1999	146693
2000	145925
2001	146304
2002	145649
2003	144964
2004	144168
2005	143474
2006	142754
2007	142220
2008	141980
2009	141900
2010	142962
2011	142914
2012	143103
2013	143395
2014	143700
2015	146267
2016	146546

Для проверки выполнения третьей предпосылки МНК — отсутствия систематической связи между значениями случайной составляющей в любых двух наблюдениях $Cov(\varepsilon_i, \varepsilon_j \mid X)$ при $i \neq j$) — вычислим критерий Дарбина–Уотсона. Промежуточные расчеты приведены в табл. 1.15.

$$dw = \frac{\sum_{i=2}^{n}(\hat{\varepsilon}_i - \hat{\varepsilon}_{i-1})^2}{\sum_{i=1}^{n}\hat{\varepsilon}_i^2} = \frac{1\,720\,390}{5\,742\,399{,}3} = 0{,}22.$$

Из данных Приложения 1 найдем значения статистик Дарбина–Уотсона при 5%-ном уровне значимости для $n = 26$, $d_1 = 1{,}30$ и $d_2 = 1{,}46$. Так как $dw < d_1$, то остатки содержат автокорреляцию.

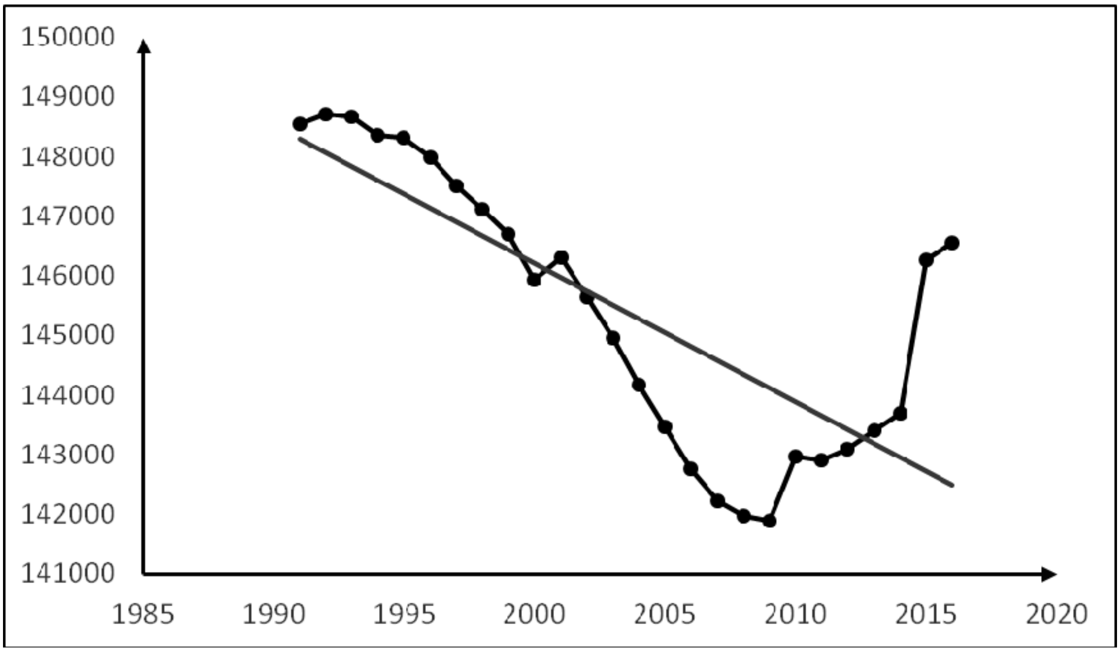


Рис. 1.20. График численности населения с 1991 по 2016 год и линейной модели

Таблица 1.15

Промежуточные расчеты для вычисления критерий Дарбина–Уотсона

t	$\hat{\varepsilon}_t$	$\hat{\varepsilon}_t^2$	$\left(\hat{\varepsilon}_t - \hat{\varepsilon}_{t-1}\right)^2$
1	253,37	64 195,10	
2	646,48	417 939,93	154 539,57
3	847,60	718 422,28	40 447,33
4	772,71	597 085,63	5 607,73
5	944,83	892 700,66	29 623,65
6	846,94	717 313,44	9 581,43
7	605,06	366 096,16	58 508,25
8	440,17	193 753,17	27 186,99
9	260,29	67 750,48	32 358,54
10	−275,60	75 952,91	287 172,50
11	335,52	112 573,44	373 461,80
12	−87,37	7 632,67	178 831,54
13	−540,25	291 869,97	205 104,63
14	−1104,13	1 219 113,44	317 966,05
15	−1 566,02	2 452 417,03	213 337,56
16	−2 053,90	4 218 522,76	238 031,56
17	−2 355,79	5 549 742,09	91 134,42
18	−2 363,67	5 586 954,05	62,17
19	−2 211,56	4 890 991,58	23 139,04
20	−917,44	841 702,43	1 674 734,19
21	−733,33	537 770,26	33 898,41

t	$\hat{\varepsilon}_t$	$\hat{\varepsilon}_t^2$	$\left(\hat{\varepsilon}_t - \hat{\varepsilon}_{t-1}\right)^2$
22	−312,21	97 476,95	177 338,02
23	211,90	44 902,55	274 696,76
24	749,02	561 027,12	288 492,75
25	3 548,13	12 589 245,30	7 835 045,98
26	4 059,25	16 477 493,22	261 238,76
Сумма		5 742 399,3	1 720 390

Значение коэффициента автокорреляции первого порядка свидетельствует о тесной связи между соседними уровнями ряда остатков

$$r(1) = \left(\sum_{i=2}^{n} (\hat{\varepsilon}_i \hat{\varepsilon}_{i-1})^2 \right) / \sum_{i=1}^{n} \hat{\varepsilon}_i^2 = 0,75.$$

Попытаемся скорректировать модель путем включения в модель переменной y_{t-1}. Применяя этот способ к новым данным МНК, получим модель $\hat{y}_i = -18\,017,36 + 74,82t + 1,12 y_{t-1}$, где t = 2, ..., 26 (рис. 1.21).

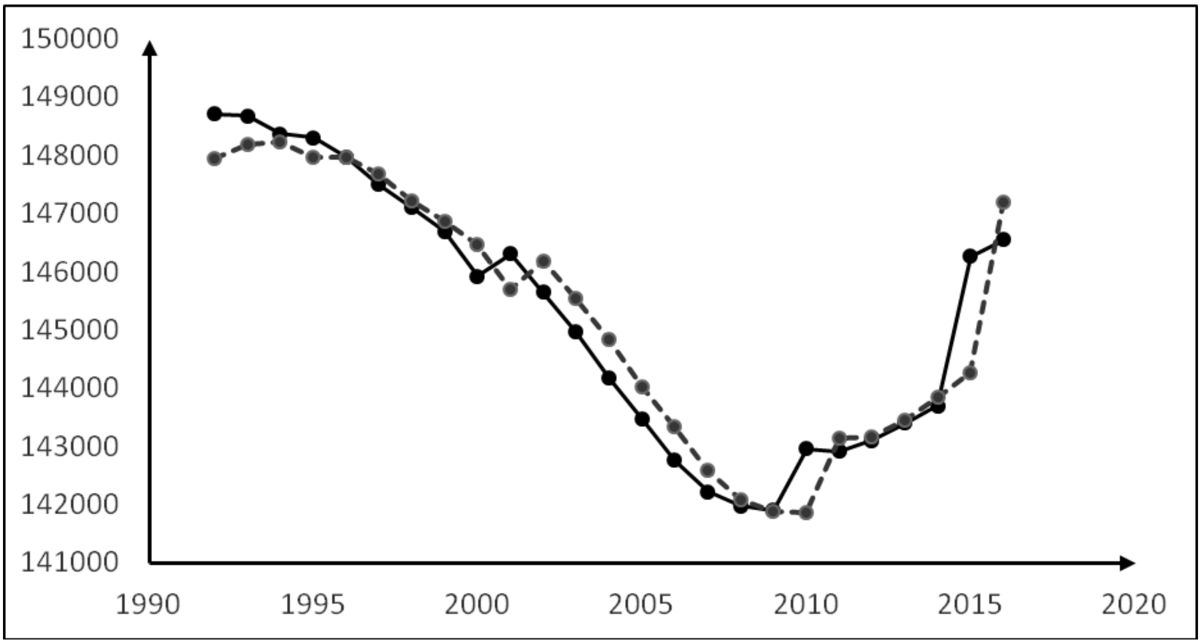


Рис. 1.21. График численности населения Российской Федерации с 1991 по 2016 гг. и скорректированной модели

Вычисленная статистика Дарбина–Уотсона dw = 1,93734 свидетельствует об отсутствии автокорреляции.

1.4.3. Проверка условия независимости объясняющих факторов

Виды мультиколлинеарности: ***строгая (perfect) мультиколлинеарность*** — наличие линейной функциональной связи между независимыми переменными и ***нестрогая (imperfect) мультиколлинеарность*** — наличие сильной линейной корреляционной связи между независимыми переменными. Заметим, что корреляционные связи есть всегда. Проблема мультиколлинеарности — проблема силы проявления корреляционных связей.

Строгая мультиколлинеарность нарушает одно из основных условий теоремы Гаусса–Маркова [2] и делает построение регрессии невозможным. Нестрогая мультиколлинеарность затрудняет работу, но не препятствует получению правильных выводов.

Основные причины возникновения мультиколлинеарности?

1. Ошибочное включение в уравнение двух или более линейно зависимых переменных.

2. В модели использованы факторные признаки, являющиеся составными элементами друг друга.

3. Исходные данные представлены временными рядами, имеющими одинаковые тенденции.

Мультиколлинеарность может проявляться и при отсутствии явных парных корреляционных зависимостей между переменными, так как мультиколлинеарность — ситуация линейной зависимости между объясняющими переменными. Однако вовсе необязательно эта зависимость должна быть парной.

При наличии мультиколлинеарности значимость отдельных коэффициентов регрессии уменьшается, так как стандартные ошибки становятся больше, что приводит к меньшей надежности полученных оценок.

Уменьшение значений t-статистики может выражаться в неверном с содержательной точки зрения знаке коэффициента регрессии. При мультиколлинеарности коэффициенты становятся неустойчивыми, поскольку в этом случае сложно отделить влияние одной переменной от влияния другой. Наличие доминантной переменной (коррелированной с зависимой переменной) делает коэффициенты при остальных объясняющих переменных незначимыми.

Наиболее характерные признаки мультиколлинеарности

1. Оценки становятся очень чувствительными к незначительному изменению результатов наблюдений и объема выборки.

2. Оценки имеют большие стандартные ошибки, малую значимость, в то время как модель в целом является значимой и обладает хорошей объясняющей способностью (хорошие значения F-статистики и коэффициента детерминации R^2).

3. Оценки коэффициентов имеют неправильные с точки зрения теории (и логики) знаки или неоправданно большие значения. Коэффициенты, которые по логике должны быть значимы, оказываются незначимыми.

Методы устранения или уменьшения мультиколлинеарности

Самый простой метод — исключение одной или нескольких переменных. При этом, какую переменную оставить, а какую удалить из анализа, решают в первую очередь на основании содержательных соображений. Если с экономической точки зрения ни одной из переменных нельзя отдать предпочтение, то оставляют ту из двух переменных, которая имеет больший коэффициент корреляции с зависимой переменной.

Для устранения мультиколлинеарности может быть использован переход от исходных объясняющих переменных x_1, x_2, ..., x_n, связанных между собой достаточно тесной корреляционной зависимостью, к новым переменным, представляющим линейные комбинации исходных. В качестве таких переменных берут, например, так называемые главные компоненты вектора исходных объясняющих переменных и рассматривают регрессию на главных компонентах, в которой последние выступают в качестве обобщенных объясняющих переменных, подлежащих в дальнейшем содержательной (экономической) интерпретации.

Пример 5

В примере 5 в качестве исходных данных будем использовать основные показатели баланса компаний-эмитентов, относящихся к сфере деятельности «Связь»* [6].

В качестве результирующей (эндогенной) переменной выбираем показатель *«Чистая прибыль (убыток)»* компаний.

Краткая экономическая характеристика основных показателей баланса, используемых в качестве экзогенных переменных:

- валюта баланса (ВБ) — это итог по всем счетам бухгалтерского баланса, сумма всех активов или всех пассивов компании;
- выручка (нетто) от продаж (ВП), дебиторская задолженность (краткосрочная) (ДЗ), запасы готовой продукции и товаров для перепродажи (ЗП), оборотные активы (ОА), основные средства (ОС), прибыль (убыток) от продаж (ПП), чистая прибыль (убыток) (ЧП) — это представители группы активов компании (здесь присутствуют как оборотные, так и внеоборотные активы);
- долгосрочные обязательства (ДО), краткосрочные обязательства (КО) — представители группы пассивов.

В результате пошагового отбора получено трехфакторное уравнение регрессии, все коэффициенты которого значимы при 5%-ном уровне значимости, вида

$$y = -4\,456{,}7 - 0{,}038x_1 + 0{,}647x_2 + 0{,}072x_3,$$

где y – *ЧП*, x_1 – *ОС*, x_2 – *ПП*, x_3 – *КО*.

Решение

Протестируем полученную модель на мультиколлинеарность с помощью метода дополнительных регрессий (тест *VIF* или метод инфляционных факторов).

Для измерения эффекта мультиколлинеарности используется показатель *VIF* — «фактор инфляции вариации»

$$VIF_{x_j} = \frac{1}{\left(1 - R^2_{x_j x_1 \ldots x_{j-1} x_{j+1} \ldots x_p}\right)},$$

где $R^2_{x_j x_1 \ldots x_{j-1} x_{j+1} \ldots x_k}$ — значение коэффициента множественной детерминации, полученное для регрессора x_j как зависимой переменной и остальных переменных.

Степень мультиколлинеарности, представляемая в регрессии переменной x_j, когда все переменные x включены в регрессию, есть функция множественной корреляции между x_j и другими переменными x.

* Источник данных — http://www.fira.ru/.

Если $VIF_{x_j} > 10,$ то считается, что данный регрессор приводит к мультиколлинеарности.

Расчеты были выполнены в Excel с помощью функции **ЛИНЕЙН**, построены три модели.

В первой модели $x_1 = f(x_2, x_3)$

0,3914	2,3907	100 014,7
0,1354	0,3008	93 120,34
0,7357	884 745,2730	#Н/Д
147,4948	106	#Н/Д
2,3091E+14	8,29741E+13	#Н/Д

$$VIF_{x_1} = \frac{1}{(1 - R^2_{x_2 x_3})} = \frac{1}{1 - 0,7357} = 3,783.$$

Во второй модели $x_2 = f(x_1, x_3)$

0,1562	−13 626,8936	0,1634
0,0197	23 893,0933	0,0323
0,7704	226 131,4928	#Н/Д
177,8257	106	#Н/Д
1,81864E+13	5,42036E+12	#Н/Д

$$VIF_{x_2} = \frac{1}{(1 - R^2_{x_1 x_3})} = \frac{1}{1 - 0,7704} = 4,355.$$

В третьей модели $x_3 = f(x_1, x_2)$

1,1929	0,1867	136 021,9
0,2355	0,0646	63 297,57
0,6604	611 040,7296	#Н/Д
103,0479	106	#Н/Д
7,69501E+13	3,95773E+13	#Н/Д

$$VIF_{x_3} = \frac{1}{(1 - R^2_{x_1 x_2})} = \frac{1}{1 - 0,6604} = 2,944.$$

Так как все значения *VIF* меньше 10, то можно сделать вывод, что построенная модель $Y = -4\ 456,7 - 0,038X_1 + 0,647X_2 + 0,072X_3$ не содержит коллинеарных факторов и может быть использована для анализа и прогнозирования.

1.4.4. Регрессионные модели с переменной структурой (фиктивные переменные)

При построении регрессионных моделей иногда приходится учитывать взаимосвязи не только между количественными характеристиками рассматриваемых явлений, объектов, но и принимать во внимание различия в их качестве. Это могут быть признаки, отражающие принадлежность объекта к какой-либо группе, наличие или отсутствие у него определенных свойств, например, пол, образование, сезон года и т.п.

Чтобы ввести такие переменные в регрессионную модель, их необходимо оцифровать, то есть качественные переменные должны быть преобразованы в количественные. Такого вида сконструированные переменные в эконометрике принято называть фиктивными переменными. В литературе можно встретить термины «структурные переменные» или «искусственные переменные»

Использование фиктивных переменных в моделях с временными рядами

В регрессионных моделях с временными рядами используется три основных вида фиктивных переменных.

1. Переменные-индикаторы принадлежности наблюдения к определенному периоду — для моделирования скачкообразных структурных сдвигов. Границы периода (моменты «скачков») должны быть установлены из априорных соображений. Например, 0, если наблюдение принадлежит периоду до кризиса 2008 г., и 1 для наблюдений последующий период. Это пример использования для моделирования временного структурного сдвига. Постоянный структурный сдвиг моделируется переменной, равной 0, до определенного момента времени и 1 для всех наблюдений после этого момента времени.

2. Сезонные переменные — для моделирования сезонности.

Например, модель потребления, учитывающая сезонные колебания.

$$y = b_0 + b_1x_1 + b_2x_2 + b_3x_3,$$

$$x_1 = \begin{cases} 1 & \text{для зимних месяцев} \\ 0 & \text{иначе} \end{cases}$$

$$x_2 = \begin{cases} 1 & \text{для весенних месяцев} \\ 0 & \text{иначе} \end{cases}$$

$$x_3 = \begin{cases} 1 & \text{для летних месяцев} \\ 0 & \text{иначе} \end{cases}$$

Следует отметить, что вводить четвертую переменную x_4 для осенних месяцев не требуется, так как в этом случае все переменные оказались бы связанными тождеством

$$x_1 + x_2 + x_3 + x_4 = 1,$$

что привело бы их к полной коллинеарности и вырожденности информационной матрицы (X^TX).

3. Линейный временной тренд — для моделирования постепенных плавных структурных сдвигов. Эта фиктивная переменная показывает, какой промежуток времени прошел от некоторого «нулевого» момента времени до того момента, к ко-

торому относится данное наблюдение (координаты данного наблюдения на временной шкале). Если промежутки времени между последовательными наблюдениями одинаковы, то временной тренд можно составить из номеров наблюдений.

Пример 6

Требуется построить регрессионную модель зависимости заработной платы работника (y) от возраста (x) с использованием фиктивной переменной по фактору пол по 20 работникам одного предприятия (табл. 1.16).

Таблица 1.16

Месячная заработная плата работников разного возраста и пола

Работник, номер	**Заработная плата в месяц (y), долл.**	**Пол (z), М/Ж**	**Возраст (x), лет**
1	300	Ж	29
2	400	М	40
3	300	Ж	36
4	320	Ж	32
5	200	М	23
6	350	Ж	45
7	350	Ж	38
8	400	М	40
9	380	М	50
10	400	М	47
11	250	Ж	28
12	350	М	30
13	200	М	25
14	400	М	48
15	220	Ж	30
16	320	М	40
17	390	М	40
18	360	М	38
19	260	Ж	29
20	250	М	25

Решение

Введем в модель фиктивную переменную Z, которая принимает два значения: 1 — если пол мужской; 0 — если пол женский. Оценим параметры модели $y = a_0 + a_1 x + a_2 z$ методом наименьших квадратов. Для вычислений воспользуемся пакетом **«Анализ данных»** в Excel (табл. 1.17).

Таблица 1.17

Фрагмент протокола регрессионного анализа

	Коэффициенты	**Стандартная ошибка**	***t*-статистика**	***P*-значение**	**Нижняя граница 95%**	**Верхняя граница 95%**
Y-пересечение	60,708	38,134	1,592	0,130	−19,748	141,165
X — возраст (лет)	6,983	1,072	6,511	0,000	4,720	9,245
Z — пол (1-М, 0-Ж)	17,275	17,462	0,989	0,336	−19,568	54,117

Уравнение множественной регрессии примет вид:

$$y = 60,71 + 6,98x + 17,27z.$$

Коэффициент детерминации R^2 равен 0,74.

Уравнение регрессии значимо по F-критерию на 5%-ном уровне, так как $F_{\text{расч}} = 24,51 > F_{(0,05;\,2;\,17)} = 3,59$.

Из полученного уравнения регрессии следует, что при одном и том же возрасте заработная плата у работников мужчин на 17,27 долл. в месяц выше, чем у женщин. Коэффициент при z незначим. Несмотря на существенную величину 17,27 долл., гипотеза о равенстве нулю не отвергается. Возможно, это связано с малым объемом выборки.

Из модели, включающей фиктивную переменную, можно получить частные уравнения регрессии для работников мужчин (z = 1) и женщин (z = 0):

$$y = 77,98 + 6,98x \quad (z = 1);$$
$$y = 60,71 + 6,98x \quad (z = 0).$$

Сопоставляя частные уравнения регрессии, видим, что эти уравнения регрессии отличаются значениями свободного члена, а соответствующие линии регрессии параллельны (рис. 1.22). График частного уравнения регрессии для мужчин будет располагаться выше, чем график частного уравнения регрессии для женщин.

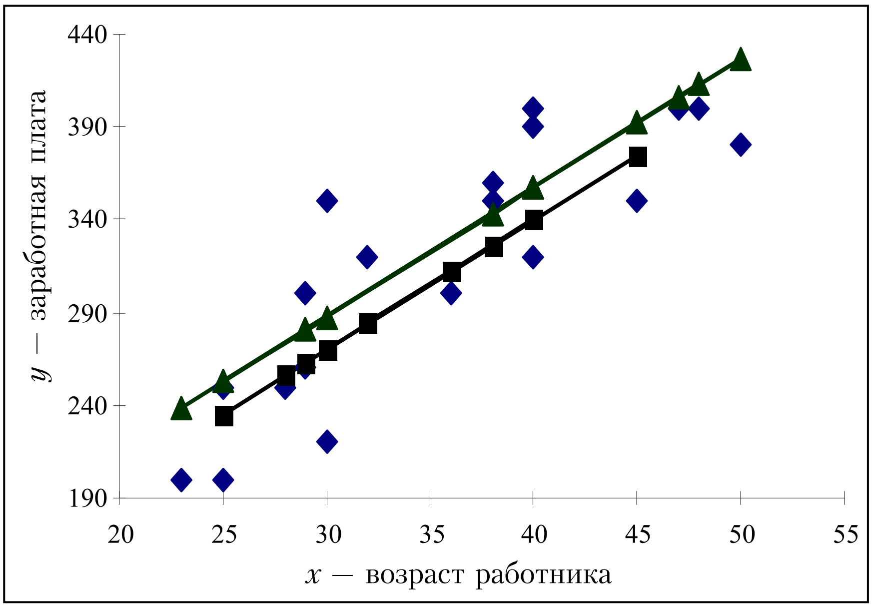


Рис. 1.22. Графики частных уравнений регрессии

1.5. Прогнозирование объема реализации продукции фирмы

Рассмотрим пример исследования экономических данных с использованием корреляционно-регрессионного анализа на два месяца вперед.

Пример 7

На основе статистических данных за 16 месяцев, приведенных в табл. 1.18, выполните корреляционно-регрессионный анализ с целью прогнозирования объема реализации продукции фирмы на два месяца вперед требуется:

1) построить матрицу парных коэффициентов линейной корреляции, проанализировать тесноту и направление связи между переменными, проверить значимость коэффициентов парной корреляции;

2) осуществить двумя способами выбор факторных признаков для построения регрессионной модели a) на основе анализа матрицы коэффициентов парной корреляции и b) с помощью пошагового отбора методом исключения;

3) для оценки качества модели определить a) коэффициент детерминации, b) коэффициент множественной корреляции и c) среднюю относительную ошибку аппроксимации;

4) провести оценку значимости уравнения регрессии и его коэффициентов;

5) проверить условие независимости остатков;

6) оценить по модели влияние факторов на зависимую переменную;

7) построить точечный и интервальный прогнозы результирующего показателя на два месяца вперед (α = 0,1).

Таблица 1.18

Исходные данные

Объем продаж (y)	**Время (x_1)**	**Затраты на рекламу (x_2)**	**Цена товара (x_3)**	**Средняя цена товара у конкурентов (x_4)**	**Индекс потребительских расходов (x_5)**
126	1	4,0	15,0	17,0	100,0
137	2	4,8	14,8	17,3	98,4
148	3	3,8	15,2	16,8	101,2
191	4	8,7	15,5	16,2	103,5
274	5	8,2	15,5	16,0	104,1
370	6	9,7	16,0	18,0	107,0
432	7	14,7	18,1	20,2	107,4
445	8	18,7	13,0	15,8	108,5
367	9	19,8	15,8	18,2	108,3
367	10	10,6	16,9	16,8	109,2
321	11	8,6	16,3	17,0	110,1
307	12	6,5	16,1	18,3	110,7
331	13	12,6	15,4	16,4	110,3
345	14	6,5	15,7	16,2	111,8
364	15	5,8	16,0	17,7	112,3
384	16	5,7	15,1	16,2	112,9

Содержательная интерпретация конечной цели задачи — прогнозирование объема продаж: **прогноз объема продаж** — предсказание будущего спроса, выраженное в денежных единицах или единицах продаваемого товара; в более узком смысле — это процесс определения объема реализации товара или группы товаров на несколько ближайших периодов времени.

Решение

1) Вычисление матрицы коэффициентов парной корреляции. Для проведения корреляционного анализа используем инструмент **«Корреляция»** (надстройка Excel пакет **«Анализ данных»**, рис. 1.23).

Объем реализации — это зависимая переменная y (тыс. руб.);

В качестве независимых, объясняющих переменных выбраны:

x_1 — время, месяцы;

x_2 — затраты на рекламу (тыс. руб.);

x_3 — цена товара, руб.;

x_4 — средняя цена товара у конкурентов (руб.);

x_5 — индекс потребительских расходов (%).

В примере 7 количество наблюдений n = 16, количество объясняющих переменных m = 5.

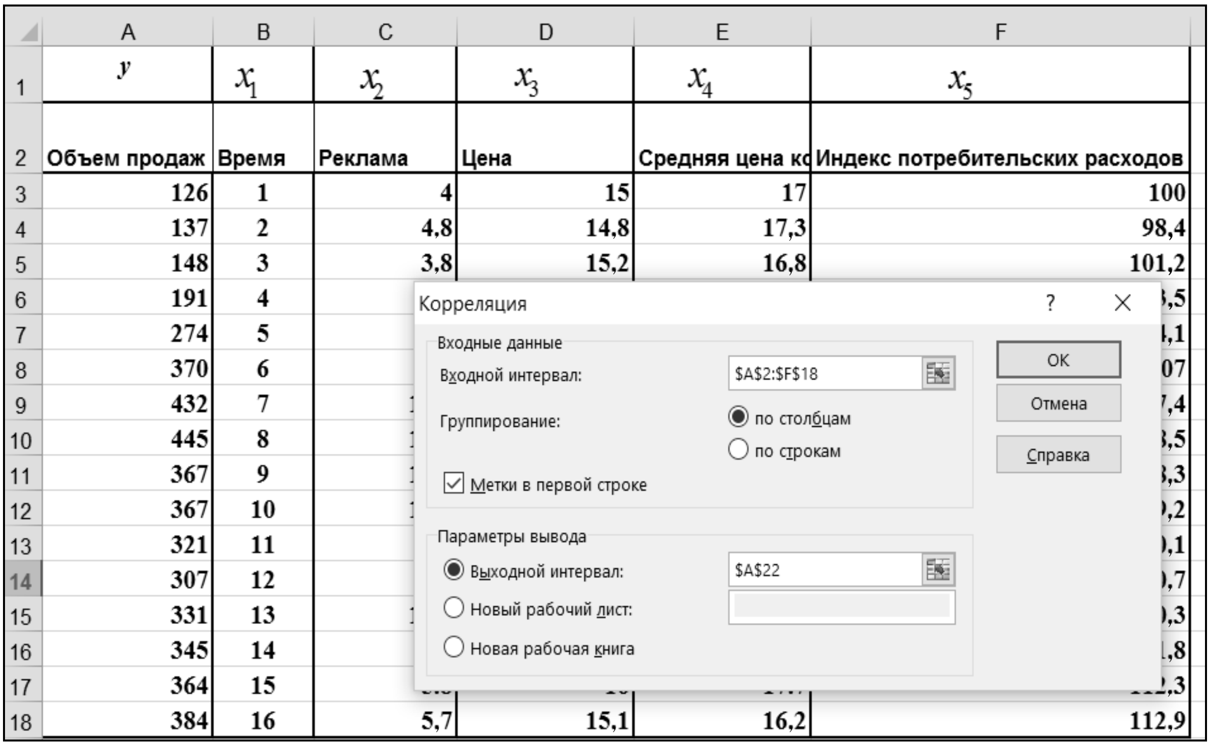


Рис. 1.23. Заполие диалогового окна инструмента **«Корреляция»**

В результате будет получена матрица коэффициентов парной корреляции (табл. 1.19).

Таблица 1.19

Результат корреляционного анализа

	Объем продаж, руб.	**Время, месяц**	**Затраты на рекламу, тыс. руб.**	**Цена товара, руб.**	**Средняя цена товара у конкурентов, руб.**	**Индекс потребительских расходов, %**
Объем продаж, тыс. руб.	1					
Время, месяц	0,678	1				
Затраты на рекламу, тыс. руб.	0,646	0,106	1			
Цена товара, руб.	0,233	0,174	−0,003	1		
Средняя цена товара у конкурентов, руб.	0,226	−0,051	0,204	0,698	1	
Индекс потребительских расходов, %	0,816	0,960	0,273	0,235	0,03	1

Анализ матрицы коэффициентов парной корреляции начнем с анализа первого столбца матрицы, в котором расположены коэффициенты корреляции, отражающие тесноту связи, зависимой переменной «*Объем продаж*» с включенными в анализ факторами. Анализ показывает, что зависимая переменная, то есть переменная «*Объем продаж*», имеет прямую тесную связь с факторами «*Индекс потребительских расходов*» ($r_{yx_5} = 0,816$), факторами «*Затраты на рекламу*» ($r_{yx_2} = 0,646$) и «*Время*» ($r_{yx_1} = 0,678$). Факторы x_3 и x_4 имеют слабую прямую связь с зависимой переменной и их не рекомендуется включать в модель регрессии.

Оценим значимость коэффициентов корреляции первого столбца матрицы. Для этого рассчитаем значение t-статистики для всех элементов первого столбца:

$$t_{\text{расч}}(r_{yx_1}) = \frac{r_{2,1}\sqrt{n-2}}{\sqrt{1-r_{2,1}^2}} = \frac{0,678\sqrt{14}}{\sqrt{1-0,460}} = 3,451;$$

$$t_{\text{расч}}(r_{yx_2}) = \frac{r_{3,1}\sqrt{n-2}}{\sqrt{1-r_{3,1}^2}} = \frac{0,646\sqrt{14}}{\sqrt{1-0,417}} = 3,166;$$

$$t_{\text{расч}}(r_{yx_3}) = \frac{r_{4,1}\sqrt{n-2}}{\sqrt{1-r_{4,1}^2}} = \frac{0,233\sqrt{14}}{\sqrt{1-0,054}} = 0,896;$$

$$t_{\text{расч}}(r_{yx_4}) = \frac{r_{5,1}\sqrt{n-2}}{\sqrt{1-r_{5,1}^2}} = \frac{0,226\sqrt{14}}{\sqrt{1-0,052}} = 0,869;$$

$$t_{\text{расч}}(r_{yx_5}) = \frac{r_{6,1}\sqrt{n-2}}{\sqrt{1-r_{6,1}^2}} = \frac{0,816\sqrt{14}}{\sqrt{1-0,666}} = 5,282.$$

Табличное значение критерия Стьюдента равно $t_{\text{табл}\ (\alpha\ =\ 0,05;\ k\ =\ n\ -\ 2\ =\ 14)} = 2,145$. Сравним числовые значения критериев с табличным. Сделаем вывод, что $t_{\text{расч}} > t_{\text{табл}}$, то есть полученные значения коэффициентов корреляции значимы для «*Индекса потребительских расходов*» ($r_{yx_5} = 0,816$), «*Затрат на рекламу*» ($r_{yx_2} = 0,646$) и «*Времени*» ($r_{yx_1} = 0,678$).

Значимость коэффициентов корреляции можно проверить, используя критическое значение коэффициента корреляции. При условии, что нулевая гипотеза H_0: $r_{ij} = 0$, критическое значение коэффициента корреляции определяется статистикой

$$r^* = \sqrt{\frac{t^2_{(\alpha,n-2)}}{n-2} \Big/ 1 + \frac{t^2_{(\alpha,n-2)}}{n-2}},$$

где $t_{(\alpha,n-2)}$ критическое значение t-статистики Стьюдента для уровня значимости α и количества степеней свободы, равного $n - 2$.

Так как в нашем примере критическое (табличное) значение критерия Стьюдента ($\alpha = 0,05$; $k = n - 2 = 14$) равно 2,145, то критическое значение коэффициента корреляции будет равно 0,497. Все коэффициенты парной корреляции в анализируемой матрице, превышающие значение 0,497, по абсолютной величине будут значимы.

2а) Выбор факторных признаков для построения регрессионной модели на основе анализа матрицы коэффициентов парной корреляции. Анализ первого столбца матрицы коэффициентов парной корреляции показал, что факторы «*Индекс потребительских расходов*» ($r_{yx_5} = 0,816$), «*Затраты на рекламу*» ($r_{yx_2} = 0,646$) и «*Время*» ($r_{yx_1} = 0,678$) имеют прямую тесную связь с зависимой переменной и их следует включить в модель, факторы x_3 и x_4 имеют слабую прямую связь с зависимой переменной и их можно не включать в модель регрессии.

Затем перейдем к анализу остальных столбцов матрицы с целью выявления коллинеарности [1]. Одним из условий регрессионной модели является предположение о линейной независимости объясняющих переменных, то есть решение задачи возможно лишь тогда, когда столбцы и строки матрицы исходных данных линейно независимы. Для экономических показателей это условие выполняется не всегда. Под ***мультиколлинеарностью*** понимается высокая взаимная коррелированность объясняющих переменных, которая приводит к линейной зависимости нормальных уравнений. Один из подходов определения наличия или отсутствия мультиколлинеарности заключается в анализе матрицы коэффициентов парной корреляции. Считают явление мультиколлинеарности в исходных данных установленным, если коэффициент парной корреляции между двумя переменными больше 0,8.

Факторы x_1 и x_5 тесно связаны между собой ($r_{x_1x_5} = 0,960$), что свидетельствует о наличии коллинеарности. Из этих двух переменных оставим x_5 — индекс потребительских расходов, так как $r_{x_1y} = 0,678 < r_{x_5y} = 0,816$. В нашем примере из двух тесно связанных друг с другом факторов x_1 и x_5 ($r_{x_1x_5} = 0,960$) исключаем x_1.

Таким образом, на основе анализа корреляционной матрицы для включения в модель регрессии остаются два фактора — «*Затраты на рекламу*» и «*Индекс потребительских расходов*» ($n = 16$, $k = 2$)

$$y = f(x_2, x_5).$$

Для проведения регрессионного анализа используем инструмент **«Регрессия»** (надстройка Excel — пакет **«Анализ данных»**, рис. 1.24). Результат вычислений приведен на рис. 1.25.

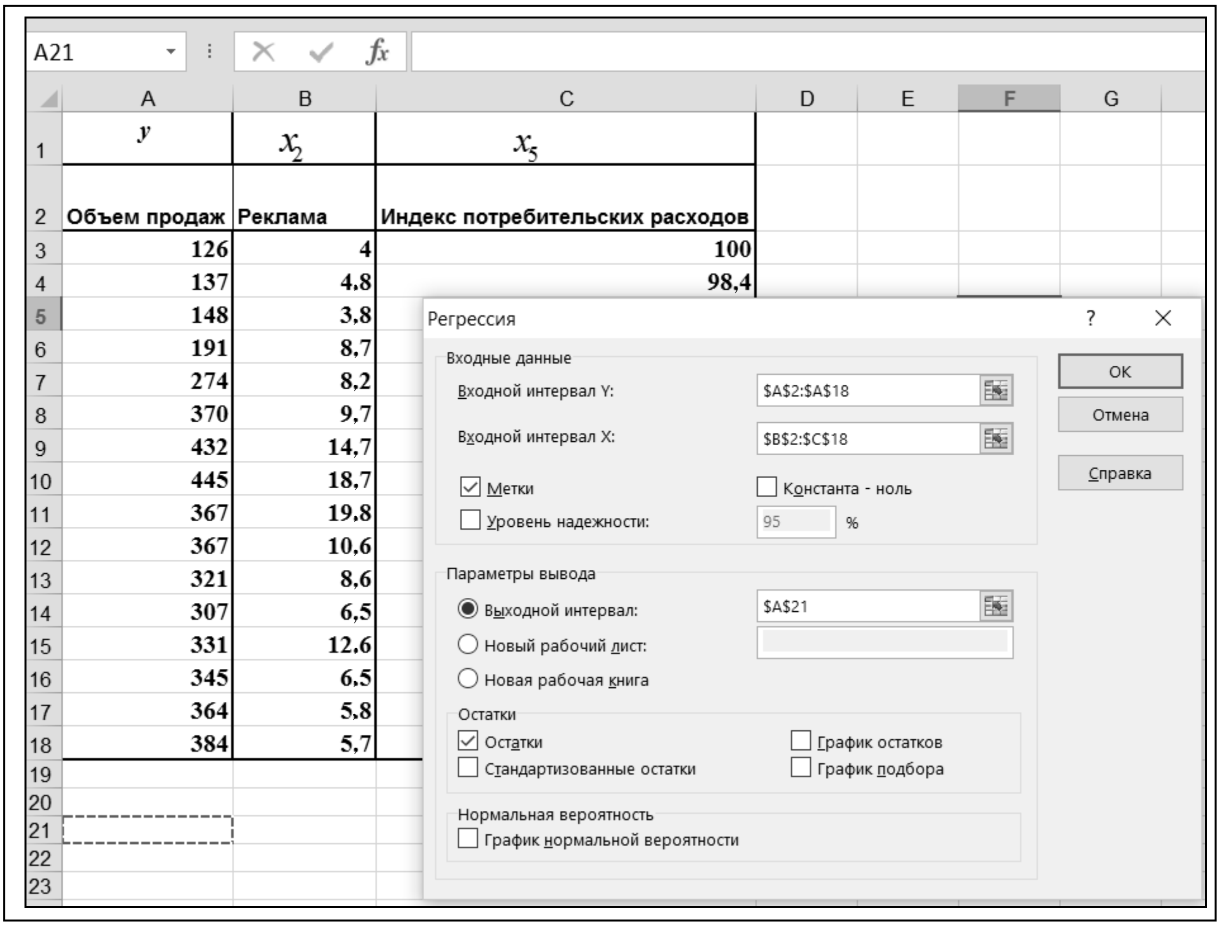


Рис. 1.24. Заполнение диалогового окна инструмента **«Регрессия»**

	A	B	C	D	E
22					
23	*Регрессионная статистика*				
24	Множественный R	0,927			
25	R-квадрат	**0,859**			
26	Нормированный R-квадрат	0,837			
27	Стандартная ошибка	41,473			
28	Наблюдения	16			
29					
30	Дисперсионный анализ				
31		*df*	*SS*	*MS*	*F*
32	Регрессия	2	136358,33	68179,17	**39,64**
33	Остаток	13	22360,10	1720,01	
34	Итого	15	158718,44		
35					
36		*Коэффициенты*	*Стандартная ошибка*	*t-статистика*	*P-Значение*
37	Y-пересечение	**-1471,31**	259,77	-5,66	0,00
38	Затраты на рекламу	**9,57**	2,27	4,22	0,00
39	Индекс потребительских расходов	**15,75**	2,47	6,39	0,00

Рис. 1.25. Результат вычислений

Используя протокол регрессионного анализа, уравнение зависимости объема реализации от затрат на рекламу и индекса потребительских расходов можно записать в следующем виде:

$$\hat{y}_i = -1\,471,31 + 9,57x_2 + 15,75x_5.$$

2б) Выбор факторных признаков для построения регрессионной модели методом исключения. Рассмотрим схему пошаговой регрессии, основанную на последовательном исключении факторов с помощью t-критерия. Она заключается в том, что после построения уравнения регрессии и оценки значимости всех коэффициентов регрессии из модели исключают тот фактор, коэффициент при котором незначим и имеет наименьший коэффициент t. После этого получают новое уравнение множественной регрессии и снова производят оценку значимости всех оставшихся коэффициентов регрессии. Если среди них опять окажутся незначимые, то опять исключают фактор с наименьшим значением t-критерия. Процесс исключения факторов останавливается на том шаге, при котором все регрессионные коэффициенты значимы.

На первом шаге строится модель регрессии по всем факторам (в скобках указаны значения стандартных ошибок коэффициентов регрессии)

$$\hat{y}_i = -3\,017,40 - \underset{(10,38)}{13,42x_1} + \underset{(3,01)}{6,67x_2} - \underset{(15,78)}{6,48x_3} + \underset{(14,41)}{12,24x_4} + \underset{(11,52)}{30,48x_5}$$

Фрагмент протокола регрессионного анализа приведен в табл. 1.20.

Таблица 1.20

Модель регрессии по пяти факторам

	Коэффициенты	**Стандартная ошибка**	***t*-статистика**	***P*-значение**	**Нижние 95%**	**Верхние 95%**
Y-пересечение	−3017,40	1094,49	−2,76	0,02	−5456,06	−578,73
Время (x_1)	−13,42	10,38	−1,29	0,23	−36,54	9,71
Затраты на рекламу (x_2)	6,67	3,01	2,22	0,05	−0,03	13,38
Цена товара (x_3)	−6,48	15,78	−0,41	0,69	−41,63	28,68
Средняя цена товара у конкурентов (x_4)	12,24	14,41	0,85	0,42	−19,87	44,34
Индекс потребительских расходов (x_5)	30,48	11,52	2,64	0,02	4,80	56,15

В данном случае коэффициенты уравнения регрессии при x_1, x_3, x_4 незначимы при 5%-ном уровне значимости. После построения уравнения регрессии и оценки значимости всех коэффициентов регрессии из модели исключают тот фактор, коэффициент при котором незначим и имеет наименьший по абсолютной величине коэффициент t, а именно x_3.

После этого получают новое уравнение множественной регрессии (в скобках указаны значения стандартных ошибок коэффициентов регрессии)

$$\hat{y}_i = -2\,914,33 - \underset{(9,78)}{12,57x_1} + \underset{(2,69)}{7,13x_2} + \underset{(9,49)}{7,93x_4} + \underset{(10,64)}{29,15x_5}$$

и снова производят оценку значимости всех оставшихся коэффициентов регрессии (табл. 1.21).

Таблица 1.21

Модель регрессии по четырем факторам

	Коэффициенты	**Стандартная ошибка**	***t*-статистика**	***P*-значение**	**Нижние 95%**	**Верхние 95%**
Y-пересечение	−2914,33	1024,23	−2,85	0,02	−5168,65	−66,00
Время (x_1)	−12,57	9,78	−1,29	0,23	−34,09	8,95
Затраты на рекламу (x_2)	7,13	2,69	2,65	0,02	1,20	1,20
Средняя цена товара у конкурентов (x_4)	7,93	9,49	0,84	0,42	−12,96	28,82
Индекс потребительских расходов (x_5)	29,15	10,64	2,74	0,02	5,74	52,56

Так как среди них есть незначимые (x_1 и x_4), то исключают фактор с наименьшим значением t-критерия — x_4. В табл. 1.22 представлены результаты, полученные после исключения фактора x_4. На следующем шаге исключаем незначимый фактор x_1.

Таблица 1.22

Модель регрессии по трем факторам

	Коэффициенты	**Стандартная ошибка**	***t*-статистика**	***P*-значение**	**Нижние 95%**	**Верхние 95%**
Y-пересечение	−2957,61	1009,97	−2,93	0,01	−5158,15	−2957,61
Время (x_1)	−14,32	9,43	−1,52	0,15	−34,86	−14,32
Затраты на рекламу (x_2)	7,23	2,65	2,72	0,02	1,45	7,23
Индекс потребительских расходов (x_5)	30,95	10,28	3,01	0,01	8,54	30,95

Процесс исключения факторов останавливается на том шаге, при котором все регрессионные коэффициенты значимы (табл. 1.23).

Таблица 1.23

Модель регрессии со значимыми факторами

	Коэффициенты	**Стандартная ошибка**	***t*-статистика**	***P*-значение**	**Нижние 95%**	**Верхние 95%**
Y-пересечение	−1471,31	259,77	−5,66	0,00	−2032,50	−910,12
Затраты на рекламу (x_2)	9,57	2,27	4,22	0,00	4,67	14,46
Индекс потребительских расходов (x_5)	15,75	2,47	6,39	0,00	10,42	21,08

Получено уравнение регрессии, все коэффициенты которого значимы не только при 5%-ном уровне значимости, но и при 1%-ном уровне значимости

$$\hat{y}_i = -1\,471{,}31 + 9{,}57x_2 + 15{,}75x_5.$$

Результаты проведенного теста не опровергают выводы, сделанные ранее только на основе корреляционной матрицы.

Коэффициент регрессии α_j показывает, на какую величину в среднем изменится результативный признак Y, если переменную x_j увеличить на единицу измерения, то есть α_j является нормативным коэффициентом.

В примере 7 величина, равная 9,57 (коэффициент при x_2), показывает, что при увеличении затрат на рекламу на 1 000 руб. объем реализации увеличится на 9,57 тыс. руб., а если на 1% увеличится индекс потребительских расходов, то объем реализации увеличится на 15,75 тыс. руб.

Расчетные значения y определяются путем последовательной подстановки в эту модель значений факторов, взятых для каждого наблюдения, или из последней таблицы регрессионного анализа «Вывод остатка» (столбец «Предсказанное y»).

3) Оценка качества модели регрессии. Для оценки качества модели множественной регрессии вычисляют коэффициент детерминации R^2 и коэффициент множественной корреляции (индекс корреляции) R. Чем ближе к 1 значение этих характеристик, тем выше качество модели.

3а) Значение коэффициентов детерминации и множественной корреляции можно найти в таблице «Регрессионная статистика» (см. рис. 1.25) или вычислить по формуле

$$R^2 = 1 - \frac{\sum_i \hat{\varepsilon}_i^2}{\sum_i (y_i - \bar{y})^2} = 1 - \frac{22\,360{,}104}{158\,718{,}438} = 0{,}859.$$

Коэффициент детерминации показывает долю вариации результативного признака под воздействием изучаемых факторов. Следовательно, около 86% вариации зависимой переменной учтено в модели и обусловлено влиянием факторов, включенных в модель;

3б) Коэффициент множественной корреляции

$$R = \sqrt{R^2} = 0{,}927.$$

Коэффициент множественной корреляции показывает высокую тесноту связи зависимой переменной Y с двумя включенными в модель объясняющими факторами.

3в) Точность модели оценим с помощью средней ошибки аппроксимации

$$E_{отн} = \frac{1}{n}\sum_{i=1}^{n}\frac{|\hat{\varepsilon}_i|}{y_i}100\% = 10{,}65\%.$$

Модель неточная. Фактические значения объема реализации отличаются от расчетных в среднем на 10,65%.

4) Оценка значимости уравнения регрессии и его коэффициентов.

Проверку значимости уравнения регрессии произведем на основе F-критерия Фишера

$$F = \frac{R^2/k}{\left(1-R^2\right)/\left(n-k-1\right)} = \frac{0,859/2}{(1-0,859)/(16-2-1)} = 39,6.$$

Значение F-критерия Фишера можно найти в таблице *«Дисперсионный анализ»* протокола Excel (см. рис. 1.24).

Табличное значение F-критерия при доверительной вероятности $\alpha = 0{,}95$ и числе степеней свободы, равном $\nu_1 = k = 2$ и $\nu_2 = n - k - 1 = 16 - 2 - 1 = 13$ составляет 3,81.

Поскольку $F_{\text{расч}} > F_{\text{табл}}$, уравнение регрессии следует признать значимым, то есть его можно использовать для анализа и прогнозирования.

Оценку значимости коэффициентов полученной модели, используя результаты отчета Excel, можно осуществить тремя способами.

Коэффициент уравнения регрессии признается значимым в том случае, если:

- наблюдаемое значение t-статистики Стьюдента для этого коэффициента больше, чем критическое (табличное) значение статистики Стьюдента (для заданного уровня значимости, например, $\alpha = 0{,}05$ и числа степеней свободы $df = n - k - 1$, где n — число наблюдений, а k — число факторов в модели);
- P-значение t-статистики Стьюдента для этого коэффициента меньше, чем уровень значимости, например, $\alpha = 0{,}05$;
- доверительный интервал для этого коэффициента, вычисленный с некоторой доверительной вероятностью (например, 95%), не содержит ноль внутри себя, то есть если «нижняя 95%» и «верхняя 95%» границы доверительного интервала имеют одинаковые знаки.

Значимость коэффициентов $\hat{a}_1$ и $\hat{a}_2$ проверим по второму и третьему способам, используя данные табл. 1.23:

$$P\text{-значение } (\hat{a}_1) = 0{,}00 < 0{,}01 < 0{,}05;$$

$$P\text{-значение } (\hat{a}_2) = 0{,}00 < 0{,}01 < 0{,}05.$$

Следовательно, коэффициенты $\hat{a}_1$ и $\hat{a}_2$ значимы при 1%-ном уровне, а тем более при 5%-ном уровне значимости.

Нижние и верхние 95% границы доверительного интервала имеют одинаковые знаки (см. рис. 1.25), следовательно, коэффициенты $\hat{a}_1$ и $\hat{a}_2$ значимы.

5) Проверка условия независимости остатков. При проверке независимости (отсутствие автокорреляции) определяется отсутствие в ряду остатков систематической составляющей с помощью dw-критерия Дарбина–Уотсона по формуле (промежуточные расчеты в табл. 1.24):

$$dw = \frac{\sum_{t=2}^{n}(\hat{\varepsilon}_t - \hat{\varepsilon}_{t-1})^2}{\sum_{t=1}^{n}\hat{\varepsilon}_t^2} = \frac{30\,336,23}{22\,360,10} \approx 1,37.$$

Таблица 1.24

Промежуточные расчеты для вычисления критерия Дарбина–Уотсона и коэффициента автокорреляции первого порядка

Номер	$\hat{\varepsilon}_t$	$\hat{\varepsilon}_t^2$	$(\hat{\varepsilon}_t - \hat{\varepsilon}_{t-1})^2$	$\hat{\varepsilon}_t \cdot \hat{\varepsilon}_{t-1}$
1	−16,25	263,96		
2	12,30	151,37	815,09	−199,89
3	−11,24	126,26	554,11	−138,24
4	−51,35	2 637,17	1 609,36	577,03
5	26,98	727,87	6 135,98	−1 385,47
6	62,94	3 961,84	1 293,41	1 698,15
7	70,80	5 012,64	61,73	4 456,38
8	28,20	795,14	1 814,91	1 996,43
9	−57,18	3 269,16	7 288,84	−1 612,27
10	16,68	278,07	5 454,09	−953,44
11	−24,37	593,68	1 684,34	−406,30
12	−27,72	768,59	11,28	675,50
13	−55,79	3 112,49	787,71	1 546,69
14	−7,05	49,73	2 375,39	393,41
15	10,77	115,99	317,60	−75,95
16	22,27	496,17	132,37	239,90
Сумма	0	22 360,10	30 336,23	6 811,93

Так как dw попадает в область неопределенности ($d_1 = 0{,}98 < dw < d_2 = 1{,}54$), когда нет оснований ни принять, ни отвергнуть гипотезу о существовании автокорреляции, то следует применить другой критерий. Воспользуемся первым коэффициентом автокорреляции

$$r(1) = \left(\sum_{i=2}^{n} \hat{\varepsilon}_i \hat{\varepsilon}_{i-1} \right) \Big/ \sum_{i=1}^{n} \hat{\varepsilon}_i^2 = \frac{6\,811{,}93}{22\,360{,}10} = 0{,}34.$$

Коэффициенты автокорреляции случайных данных должны обладать выборочным распределением, приближающимся к нормальному с нулевым математическим ожиданием и средним квадратическим отклонением, равным $\frac{1}{\sqrt{n}} = \frac{1}{\sqrt{16}} = 0{,}25.$ Если коэффициент автокорреляции первого порядка $r(1)$ находится в интервале $-1{,}96 \times 0{,}25 < r(1) < 1{,}96 \times 0{,}25,$ то можно считать, что данные не показывают наличия автокорреляции первого порядка. Так как $-0{,}49 < r(1) = 0{,}34 < 0{,}49,$ то свойство независимости остатков выполняется.

6) Оценка влияния факторов, включенных в модель, на объем реализации. Учитывая, что коэффициент регрессии невозможно использовать для непосредственной оценки влияния факторов на зависимую переменную из-за различия единиц измерения и разной колеблемости факторов, используем коэффициенты эластичности и бета-коэффициенты:

$$Э_j = a_j \overline{x}_j \,/\, \overline{y};$$

$$Э_2 = 9{,}568 \cdot 9{,}294/306{,}813 = 0{,}2898;$$

$$Э_5 = 15{,}7529 \cdot 107{,}231/306{,}813 = 5{,}506.$$

Коэффициент эластичности показывает, на сколько процентов изменяется зависимая переменная при изменении фактора на 1%:

$$\beta_j = \hat{a}_j S_{x_j} / S_y;$$

$$\hat{\beta}_2 = 9{,}568 \cdot 4{,}913/102{,}865 = 0{,}457;$$

$$\hat{\beta}_5 = 15{,}7529 \cdot 4{,}5128/102{,}865 = 0{,}691.$$

Бета-коэффициент с математической точки зрения показывает, на какую часть величины среднеквадратического отклонения меняется среднее значение зависимой переменной с изменением независимой переменной на одно среднеквадратическое отклонение при фиксированных на постоянном уровне значениях остальных независимых переменных. Это означает, что при увеличении затрат на рекламу на 4,91 тыс. руб. объем, реализации увеличится на 47 тыс. руб. (0,457 × 102,865).

Среднеквадратическое отклонение затрат на рекламу, равное 4,91, можно вычислить с помощью функции **СТАНДОТКЛОН**.

Долю влияния фактора в суммарном влиянии всех факторов можно оценить по величине дельта-коэффициентов Δ_j:

$$\Delta_j = r_{y,x_j} \hat{\beta}_j / R^2$$

$$\Delta_2 = 0{,}646 \cdot 0{,}457/0{,}859 = 0{,}344;$$

$$\Delta_5 = 0{,}816 \cdot 0{,}691/0{,}859 = 0{,}656.$$

Вывод

На объем реализации более сильное влияние оказывает фактор «*Индекс потребительских расходов*».

7) Определение точечных и интервальных прогнозных оценок объема реализации на два месяца вперед. При построении прогнозов введем обозначения x_1 «*Затраты на рекламу*» и x_2 «*Индекс потребительских расходов*». Прогнозные значения фактора x_1 $x_{1,17}$, $x_{1,18}$ и фактора x_2 $x_{2,17}$, $x_{1,18}$ на два месяца вперед можно определить с помощью методов экспертных оценок, с помощью средних абсолютных приростов или вычислить на основе экстраполяционных методов [1]. Воспользуемся инструментом **«Мастер диаграмм»** для подбора экстраполяционных моделей для объясняющих факторов.

Для фактора x_1 «*Затраты на рекламу*» выбрана модель

$$x_1 = 12{,}83 - 11{,}616\, t + 4{,}319\, t^2 - 0{,}552\, t^3 + 0{,}0292\, t^4 - 0{,}0006\, t^5,$$

по которой получен прогноз на два месяца вперед*.

Упреждение	Прогноз
1	5,75
2	4,85

График модели временного ряда «*Затраты на рекламу*» приведен на рис. 1.25.

* Следует обратить внимание, что полиномы таких высоких порядков редко используются при прогнозировании экономических показателей.

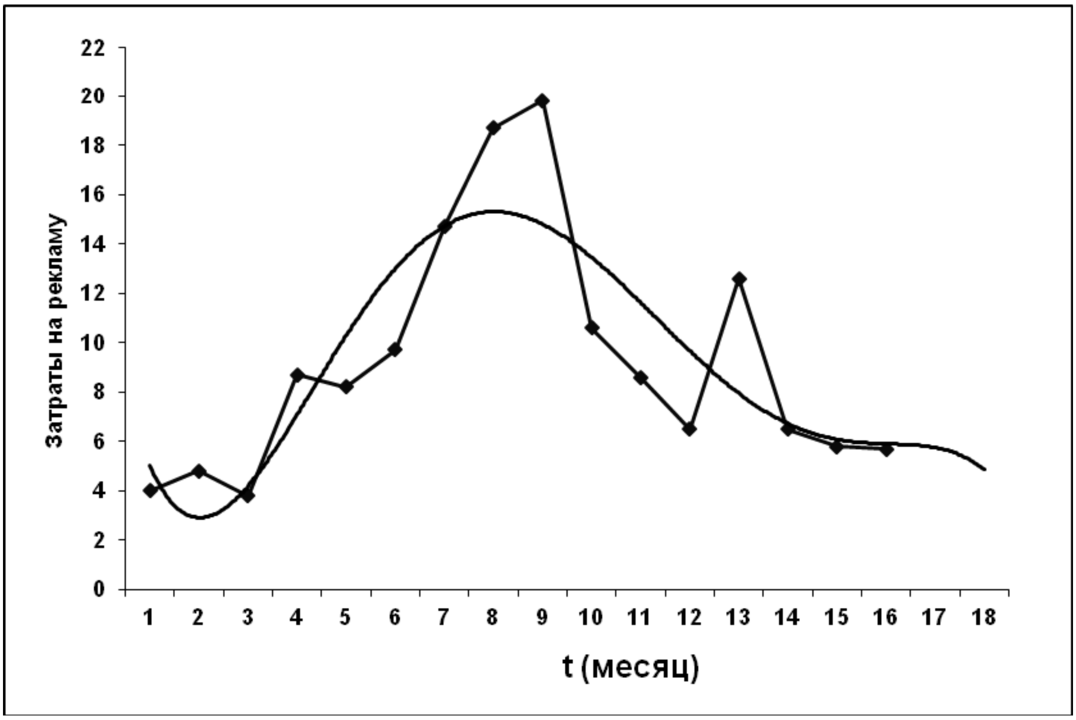


Рис. 1.26. Прогноз показателя *«Затраты на рекламу»*

Для временного ряда показателя *«Индекс потребительских расходов»* в качестве аппроксимирующей функции выбран полином второй степени (парабола), по которой построен прогноз на два шага вперед. На рис. 1.26 приведен результат построения тренда для временного ряда показателя *«Индекс потребительских расходов»*

$$x_2 = 97{,}008 + 1{,}739\ t - 0{,}0488\ t^2.$$

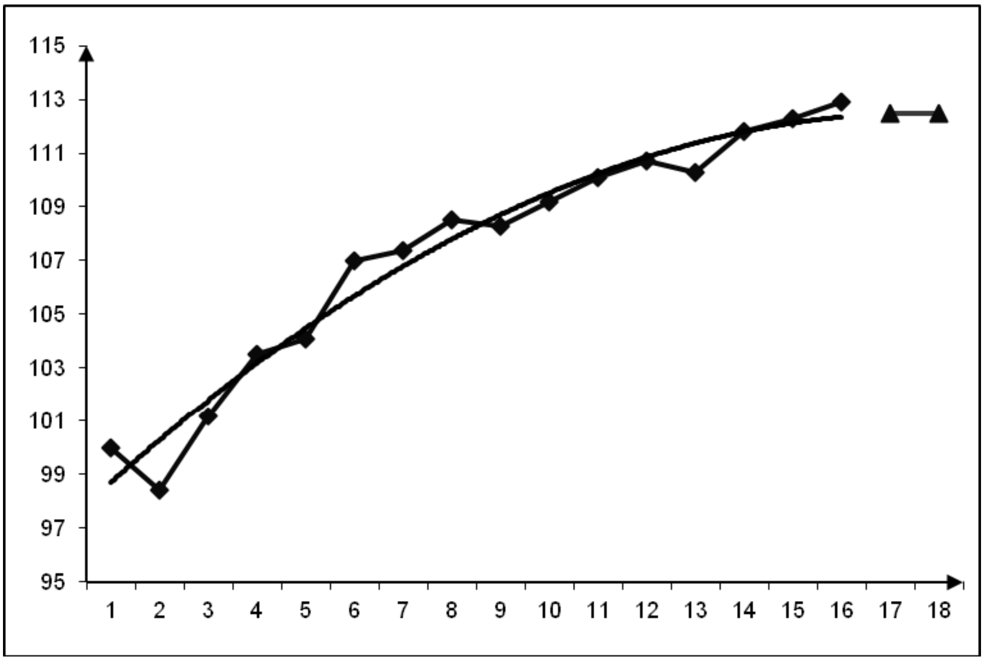


Рис. 1.27. Прогноз показателя *«Индекс потребительских расходов»*

Упреждение	Прогноз
1	112,468
2	112,488

Для получения прогнозных оценок зависимостей переменной по модели

$$y = -1\ 471{,}438 + 9{,}568\ x_1 + 15{,}754\ x_2$$

подставим в нее найденные прогнозные значения факторов x_1 и x_2:

$$y_{(t\,=\,17)} = -1\ 471{,}438 + 9{,}568 \times 5{,}75 + 15{,}754 \times 112{,}468 = 355{,}399;$$

$$y_{(t\,=\,18)} = -1\ 471{,}438 + 9{,}568 \times 4{,}85 + 15{,}754 \times 112{,}488 = 344{,}179.$$

Доверительный интервал прогноза будет иметь следующие границы:

Верхняя граница прогноза: $Y_{\text{прогн}\ (n\,+\,l)} + U(l)$;

Нижняя граница прогноза: $Y_{\text{прогн}\ (n\,+\,l)} - U(l)$.

$$U(l) = S_\varepsilon t_{\text{кр}} \sqrt{V_{\text{прогн}}} = S_\varepsilon t_{\text{кр}} \sqrt{1 + X^T_{\text{прогн}} (X^T X)^{-1} X_{\text{прогн}}};$$

$S_\varepsilon = 41{,}473$;

$t_{\text{кр}} = 1{,}77$ (значение $t_{\text{кр}}$ получено с помощью функции **СТЬЮДРАСПОБР** (0,1; 13) для выбранной вероятности 90% с числом степеней свободы 13).

На первый шаг ($l = 1$):

$$X^T_{\text{прогн}} = (1;\ 5{,}75;\ 112{,}48);$$

$$(X^T X)^{-1} = \begin{pmatrix} 39{,}23140 & 0{,}06752 & -0{,}37110 \\ 0{,}06752 & 0{,}00299 & -0{,}00088 \\ -0{,}37110 & -0{,}00088 & 0{,}00354 \end{pmatrix};$$

$U(1) = 83{,}39$.

На второй шаг ($l = 2$):

$$X^T_{\text{прогн}} = (1;\ 4{,}85;\ 112{,}51);$$

$U(2) = 84{,}44$.

Результаты прогнозирования на два месяца вперед представлены на рис. 1.27 и табл. 1.25.

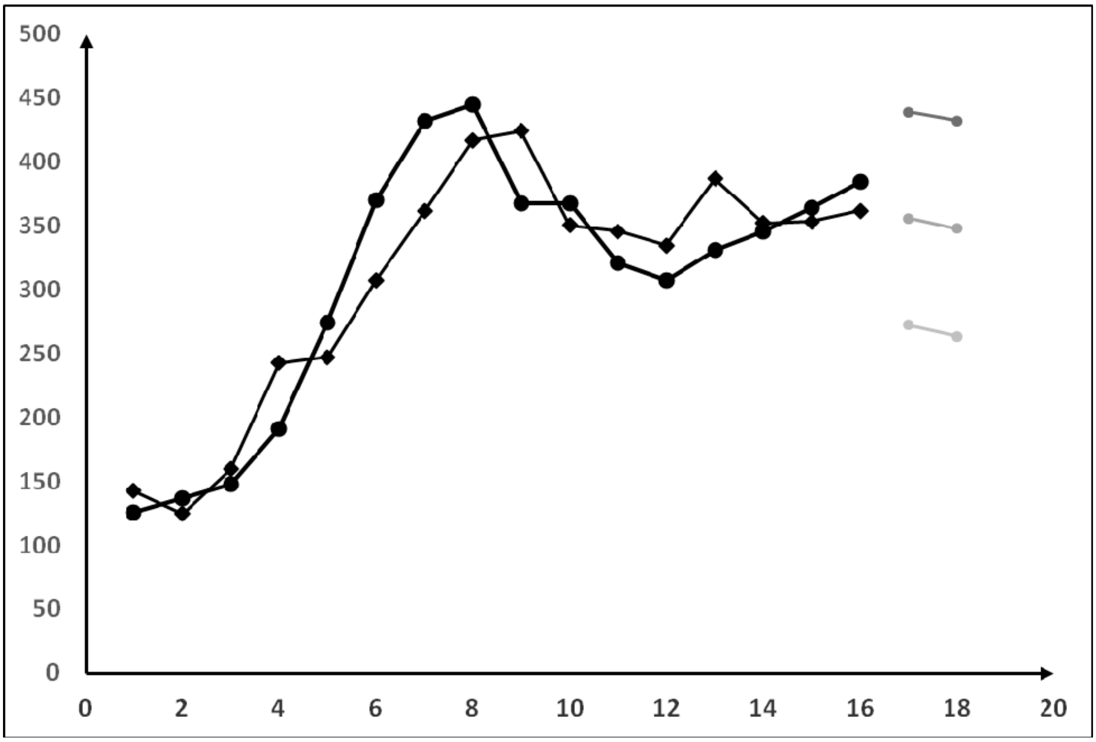


Рис. 1.28. Результаты прогнозирования объема реализации на два месяца

Таблица 1.25

Результаты точечного и интервального прогнозов (p = 90%)

Упреждение		Прогноз	Нижняя граница	Верхняя граница
1	(t = 17)	355,59	272,21	438,98
2	(t = 18)	347,49	263,05	431,93

1.6. Задачи для самостоятельного решения

Задача 1

Имеются следующие сведения о количестве пучков салата, предаваемого ежедневно в розницу и цене:

Количество, шт/день	28	29	34	35	37	37	41	46
Цена, руб. за единицу	30	31	25	26	22	24	16	12

Торговцу нужно выяснить, как изменяется количество продаваемого салата при изменении цены.

Требуется:

1) построить модель линейной регрессии с помощью надстройки Excel «Поиск решения»; дать интерпретацию построенной модели и отдельных ее коэффициентов;

2) оценить точность модели с помощью средней относительной ошибки аппроксимации;

3) если бы цена равнялась 35 руб. за каждый пучок, то сколько было бы продано салата; доверительные интервалы прогнозов; рассчитать при уровне значимости α = 0,2; отобразить на графике фактические данные, результаты моделирования и прогнозирования;

4) определить, при какой цене продавец не продаст ни одного пучка салата.

Задача 2

Для семи летних площадок предприятия общественного питания известны средние значения объема выручки за день (Y, руб.) и количества посетителей (X, человек):

Y	133000	150000	156000	195000	345000	365000	433000
X	95	145	165	233	333	350	456

Требуется:

1) построить линейную модель парной регрессии с помощью матричных функций Excel, отобразить ее вместе с исходными данными на графике, дать интерпретацию построенной модели и отдельных ее коэффициентов;

2) оценить точность модели с помощью средней относительной ошибки аппроксимации;

3) построить нелинейные модели парной регрессии: степенную, показательную и гиперболическую; пояснить, как выполнялась линеаризация данных. Привести графики моделей;

4) с вероятностью 0,9 (α = 0,1) дать прогноз относительно выручки новой площадки, если число посетителей составит 80% от максимального значения; результаты моделирования прогнозирования отобразить на графике.

Задача 3

Следует выявить зависимость объема реализации одного из продуктов кондитерской фабрики от влияющих факторов на основе ежеквартальных данных и построить соответствующую модель регрессии.

Используемые обозначения:

y – объем реализации (тыс. руб.);
x_1 – цена товара (руб.);
x_2 – средняя цена конкурентов (руб.);
x_3 – индекс потребительских расходов (%);
x_4 – затраты на рекламу (тыс. руб.);
x_5 – количество торговых точек (специализированных отделов), шт.;
x_6 – количество сотрудников в отделе маркетинга (чел.).

Требуется:

1) осуществить двумя способами выбор факторных признаков для построения регрессионной модели: на основе визуального анализа матрицы коэффициентов парной корреляции и с помощью пошагового отбора методом исключения и построить уравнение множественной регрессии в линейной форме с выбранными факторами (укажите, какая модель лучше и почему); дать экономическую интерпретацию коэффициентов модели регрессии;

2) протестировать данные на мультиколлинеарность (метод дополнительных регрессий);

3) проверить выполнение свойства независимости остатков;

4) дать сравнительную оценку силы связи факторов с результатом с помощью коэффициентов эластичности, β- и Δ-коэффициентов;

5) построить точечный и интервальный прогноз результирующего показателя на два квартала вперед $(\alpha = 0,2)$; прогнозные значения экзогенных переменных определить на основе экстраполяционных методов, представленных ниже:

Объем реализации (y), тыс. руб.	**Цена (x_1), тыс. руб.**	**Средняя цена конкурентов (x_2), тыс. руб.**	**Индекс потребительских расходов (x_3), тыс. руб.**	**Затраты на рекламу (x_4), тыс. руб.**	**Кол-во торговых точек (x_5), тыс. руб.**	**Кол-во сотрудников в отделе маркетинга (x_6), чел.**
1	2	3	4	5	6	7
126	15,9	16	100	6,1	2	10
130	15,8	16	99,6	6,9	2	8
135	15,9	15,9	98,9	7,2	2	8
137	15,6	15,5	98,4	7,7	2	7
141	15,7	16,2	98,9	6,5	2	11
146	15,8	16,1	100,1	6,8	2	11
157	15,7	15,7	101,2	7,7	2	10
166	15,6	15,5	102,1	8,3	2	10
184	15,5	15,8	102,9	13,6	2	3
191	15,7	15,9	103,5	15,5	2	2
216	15,6	15,8	103,7	14,7	3	4
246	15,4	15,9	103,8	15	3	6
274	15,5	15,8	104,1	14,5	3	8
317	15,4	15,5	105,2	16,1	3	9
341	14,9	15,4	105,8	16,4	3	10
370	14,8	15,2	107	26,8	3	3
392	15	15,1	107,1	29,2	3	3
370	14,9	15,1	107,3	25,6	3	4
432	14,7	14,7	107,4	27,5	4	5
437	14,8	14,7	107,9	29,1	4	5
442	14,9	15	108,1	32,4	4	3
425	15,2	15,4	108,5	35,5	4	1
403	15,3	15,3	108,4	34,1	4	1
464	14,8	14,7	108,5	37,4	4	2
367	15,4	15,5	108,3	37,7	4	4
362	15,3	15,5	108,4	31,3	4	1

Окончание таблицы

1	2	3	4	5	6	7
369	15	15	108,8	25,8	3	4
367	15,2	15,1	109,2	19,3	3	9
352	15,5	15,6	109,6	17,1	3	10
330	15,4	15,6	109,8	16,4	3	10
324	15,3	15,3	110,1	15,3	3	11
319	15,4	15,3	110,4	13,7	3	13
314	15,5	15,6	110,6	12,2	3	15
307	15,3	15,5	110,7	11,1	3	17
315	15,1	15,1	110,5	15,1	3	10
325	15,2	15,1	110,1	19,6	3	6
331	15,3	15,4	110,3	23,3	3	4
335	15,1	15,3	110,6	18,9	3	7
340	15	15	111,2	14,8	3	12
345	14,9	14,8	111,8	11,1	3	21
353	15	15,1	112	10,3	3	24
359	15,1	15,3	112,1	9,9	3	26
364	14,8	14,8	112,3	9,7	3	27
372	15	14,9	112,6	9,7	3	28
379	14,7	14,8	112,7	9	3	32
384	14,9	15,1	112,9	9,5	3	30
391	15,1	15,1	113	9,5	3	31
395	15,2	15,1	113,2	10,1	4	29
399	15,3	15,4	113,5	9,6	4	31
402	15,2	15,4	113,8	10,1	4	29

Задача 4

По тринадцати супермаркетам исследуется зависимость квартального торгового оборота от размера торговых площадей, района расположения (центральный или периферийные) и формы собственности (муниципальный или частный). Имеются следующие данные:

Номер магазина	Торговый оборот, млн руб.	Торговые площади, м^2	Район расположения	Форма собственности
1	59	2500	Периферийный	Муниципальная
2	85	2172	Периферийный	Частная
3	127	2928	Центральный	Муниципальная
4	178	3943	Центральный	Муниципальная
5	156	2819	Центральный	Частная
6	122	4902	Периферийный	Муниципальная
7	89	4236	Центральный	Муниципальная
8	159	5486	Периферийный	Муниципальная
9	256	7186	Центральный	Частная
10	156	4501	Центральный	Частная
11	149	3495	Центральный	Муниципальная
12	122	4562	Периферийный	Частная
13	178	2706	Центральный	Частная

Требуется:

1) построить линейную регрессионную модель торгового оборота магазина; оценить параметры модели; дать экономическую интерпретацию коэффициентам уравнения регрессии;

2) указать, существенна ли разница в торговом обороте магазинов; расположенных в центральном и периферийных районах города; и являющихся частной и муниципальной собственностью;

3) используя результаты регрессионного анализа, ранжировать компании по степени эффективности.

Литература к главе 1

1. **Орлова И.В., Половников В.А.** Экономико-математические методы и модели: компьютерное моделирование: учебное пособие. — М.: Вузовский учебник, 2007, 2011 (любое издание, но лучше третье)*.

2. **Костюнин В.И.** Эконометрика. Учебник и практикум для прикладного бакалавриата — М.: Юрайт, 2015.

3. **Кремер Н.Ш., Путко Б.А.** Эконометрика: учебник для вузов / под ред. проф. Н.Ш. Кремера. — М.: ЮНИТИ-ДАНА, 2003—2008.

4. **Магнус Я.Р., Катышев П.К., Пересецкий А.А.** Эконометрика. Начальный курс: Учебник. — 3-е изд., перераб. и доп. — М.: Дело, 2008.

5. Эконометрика: учебник / под ред. И.И. Елисеевой. — 2-е изд., перераб. и доп. — М.: Финансы и статистика, 2005—2008.

6. **Орлова И.В., Филонова Е.С.** Эконометрическое моделирование финансовой эффективности предприятий, относящихся к виду экономической деятельности «связь» // Международный бухгалтерский учет. — 2012. — № 43. С. 22—24.

7. **Орлова И.В., Филонова Е.С.** Выбор экзогенных факторов в модель регрессии при мультиколлинеарности данных // Международный журнал прикладных и фундаментальных исследований. — 2015. — № 5—1. С. 108—116.

8. **Орлова И.В., Турундаевский В.Б.** Многомерный статистический анализ при исследовании экономических процессов. Монография. — М.: МЭСИ, 2014. — С. 190.

9. Эконометрика: учебник для магистров / И.И. Елисеева [и др.]; под ред. И.И. Елисеевой. — М.: Издательство Юрайт, 2012; — Серия: Магистр; Информатизация и связь. — 2014. — № 3. С. 69—73.

* Учебное пособие есть на портале.

Глава 2

Равновесные модели эволюционно-симулятивного метода

2.1. Метод аналитического моделирования в задачах планирования

Каждой фирме вне зависимости от ее масштабов, направления деятельности или формы собственности необходимо составлять план продаж. План продаж является стратегическим документом фирмы, основой разработки производственного и финансового плана.

Основной задачей планирования является оценка ожидаемых объемов продаж и цен на товары и услуги, которые фирма сумеет реализовать на конкретном секторе рынка в некоторый намечаемый период времени, называемый плановым периодом.

Существуют различные методы планирования продаж. Среди них важное место занимает **метод аналитического моделирования**, позволяющий разработать план продаж, близкий к оптимальному.

Особую важность метод аналитического моделирования приобретает в тех случаях, когда планируется производство нового товара или предоставление нового спектра услуг. В этих случаях возможность получения необходимой информации о конъюнктуре на данном секторе рынка резко снижается ввиду целого ряда ограничений:

- частично или полностью исключены пробные продажи, поскольку товар (услуга) еще не производится;
- ограничены возможности опросов, так как потенциальные потребители недостаточно осведомлены о намечаемых к производству товарах (услугах);
- ограничены возможности получение информации о реакции конкурентов на появление нового товара (услуги), в частности, о возможных изменениях в ценовой политике конкурентов, о намечаемых новациях в способах продаж.

В условиях недостаточности информации возникает необходимость применения специальных методов планирования и прежде всего метода аналитического моделирования.

Разработки плана продаж методом аналитического моделирования включает несколько этапов.

Этап I. Составить экономико-математическую модель, выражающую зависимость ожидаемого объема продаж товаров (услуг) от факторов, определяющих спрос на данный товар (услугу)*. При выборе факторов для включения в модель следует исходить из того, что наиболее существенное влияние на объем продаж оказывают следующие факторы:

---

* Здесь и далее под спросом на товар понимается количество товара, которое можно реализовать на данном секторе рынка по приемлемой цене.

• количество потенциальных покупателей, их предпочтения и намерения;
• уровень денежных доходов покупателей;
• структура потребления населением товаров (услуг);
• позиции и ценовая политика конкурентов на данном секторе рынка.

Этап II. Оценить факторы, включенные в модель, то есть присвоить им числовые значения, отражающие сложившуюся конъюнктуру на данном секторе рынка. Как правило, для этой цели используется метод экспертных оценок.

Этап III. Выполнить на основе составленной экономико-математической модели и числовых оценок факторов расчет оптимального плана продаж, определить степень надежности намечаемого плана, оценить затраты на его выполнение и прибыль от реализации.

Достоинство метода аналитического моделирования в том, что данный метод требует минимума затрат, позволяет проводить исследование оперативно, является достаточно гибким: не составляет большого труда изменить состав факторов, содержание модели и применить любые способы оценки факторов.

Основная трудность, связанная с применением метода аналитического моделирования, состоит в том, что рассчитанная на основе используемой модели прогнозная оценка плана может иметь **неприемлемо большую погрешность**, иногда превышающую саму оценку. Это объясняется тем, что значения факторов модели могут быть оценены лишь приближенно (на основе мнений экспертов), а допущенные при этом погрешности в оценках факторов многократно увеличиваются в процессе расчетов искомого плана по формулам используемой оптимизационной модели (см. пример в Приложении 2).

Применение в задачах планирования информационной технологии оптимизационного типа, реализованной в системе **«Decision»**[*] в модуле **«Equilibrium»**[**], позволяет:

• рассчитать близкие к оптимальным объемы продаж и цены реализуемых товаров, не погружаясь в сложный математический аппарат[***];

• снизить погрешность оптимизационных расчетов плана на один–два (минимум в 10 раз) по сравнению с погрешностью расчетов методом аналитического моделирования.

Оптимизационные расчеты выполняются в диалоге с модулем **«Equilibrium»**, предназначенным для решения класса задач, для которых характерны следующие основные признаки:

• результат выполнения принимаемого решения зависит от внешних случайных факторов, которые делают ситуацию не вполне определенной;

• последствия возможного несовпадения принимаемого решения и результата его выполнения двояки: с одной стороны, решение может оказаться излишне оптимистическим и быть не полностью выполненным, и тогда возникнут потери от неэффективного использования вложенных денежных средств; с другой стороны, реше-

---

[*] Инструментальная система **«Decision»** является единственным отечественным коммерческим программным пакетом, обеспечивающим решение управленческих, экономических и исследовательских задач оптимизационного типа. В диалоге с модулем **«Equilibrium»** системы **«Decision»** можно эффективно решать экономические задачи, связанные с исследованием емкости рынка, ожидаемого объема продаж, поиска оптимальной цены, прогноза прибыли и затрат, анализа тактики и стратегии конкурентной борьбы. При этом погрешность оптимизационных расчетов на один-два порядка меньше погрешности, свойственной аналитическим методам.

[**] Для работы с модулем **«Equilibrium»** нет необходимости осваивать достаточно сложные математические методы решения оптимизационных задач, а интерфейс модуля совпадает с общеизвестной электронной таблицей Excel, поэтому овладение навыками работы с системой не вызывает затруднений.

[***] В модуле «**Equilibrium**» реализован эволюционно-симулятивный метод (ЭСМ), сочетающий в себе методы теории статистической оптимизации, имитационного моделирования и экспертных оценок.

ние может оказаться излишне пессимистическим, и тогда возникнут потери в виде упущенных возможностей;

• необходимо на стадии принятия решения (до начала планового периода) комплексно учесть ожидаемые противоречивые последствия.

Более кратко задачи, относящиеся к данному классу, можно сформулировать так:

необходимо принять решение, которое минимизирует риски, связанные с вероятным завышением либо занижением этого решения

2.2. Риски завышения и занижения плана. Оценка рисков на основе метода статистических испытаний

Выполнение намеченного плана продаж зависит от многих случайных факторов, информация о которых по объективным причинам всегда **не полна, не точна или даже ошибочна**, поэтому фактический объем продаж, как правило, будет отличаться от составленного плана. При этом план может оказаться невыполненным (при неудачном стечении обстоятельств), либо перевыполненным (при особенно удачном стечении обстоятельств). В обоих случаях для фирмы возникают негативные последствия.

1. Если план оказался завышенным (то есть не удалось продать все, что было заготовлено), то непроданный товар будет пролеживать на складе или вовсе пропадет. В этом случае фирма понесет потери, связанные с затратами на хранение нереализованного товара, снижение качества во время хранения, замораживанием денежных средств, вложенных в непроданный товар. Эти потери составляют **издержки завышения плана** относительно спроса.

2. Если план был занижен (то есть не было заготовлено такого количества товара, которое можно было бы продать), то возникают издержки в виде упущенного дохода, упущенной прибыли, упущенной доли фирмы на рынке. Такого рода потери образуют **издержки занижения плана** относительно спроса.

Размер издержек завышения и занижения плана есть количественное выражение тех негативных последствий, которые возникают из-за несовпадения запланированного объема продаж с фактическим[*].

В оптимальном плане продаж размеры возможных издержек завышения и занижения плана должны быть минимизированы. Однако в момент составления плана не известна не только величина издержек, но и то, какого рода издержки (завышения или занижения плана) будут иметь место по окончании планового периода, то есть когда станет известен фактический объем продаж. Следовательно, при разработке плана имеется возможность учесть и минимизировать лишь **риск возникновения издержек**, а не конкретную величину самих издержек.

Таким образом, задача оценки и минимизации величины рисков завышения и занижения плана является необходимой составляющей задачи нахождения оптимального плана продаж.

В модуле **«Equilibrium»** реализован ряд моделей, позволяющих для любого плана продаж достаточно точно оценить величину возникающих рисков.

[*] Размер издержек завышения и занижения плана зависит от экономических условий, в которых действует фирма (от технологии и организации производства, способов хранения и транспортировки товара, конъюнктуры рынка, налогового законодательства и др.).

В моделях используются следующие обозначения:

PL – плановый объем продаж, называемый далее **планом**;

Fa – фактический объем продаж, называемый далее **фактом**;

R_1 – размер издержек завышения плана, возникающих при $PL > Fa$;

R_2 – размер издержек занижения плана, возникающих при $PL < Fa$.

Величина PL устанавливается до начала планового периода и остается неизменной на всем его протяжении. В то же время фактический объем продаж Fa, а также размеры издержек R_1, R_2 в момент составления плана неизвестны. Они зависят от случайных факторов и, следовательно, являются случайными величинами, значения которых можно оценить лишь вероятностно.

Таким образом, риск завышения (занижения) плана представляет собой ожидаемые размеры издержек, возникающих из-за несовпадения плана PL с фактом Fa.

Ожидаемые размеры R_1 и R_2 ожно оценить с помощью показателей математического ожидания $M[R_1]$ и $M[R_2]$ соответственно*.

В модуле **«Equilibrium»** эти показатели рассчитываются на основе **метода статистических испытаний**. Статистическое испытание – это розыгрыш фактической ситуации с помощью датчика случайных чисел и имитационной модели. Важно подчеркнуть, что результаты розыгрышей всегда различны, то есть для одних и тех же исходных данных в разных испытаниях генерируются разные результаты.

В модуле **«Equilibrium»** в каждом отдельном статистическом испытании, называемом реализацией, автоматически генерируются:

1) случайные значения факторов, влияющих на спрос;

2) зависящее от факторов, влияющих на спрос случайное значение Fa;

3) либо случайное значение R_1, либо случайное значение R_2 (при $PL > Fa$ генерируется значение R_1, при $PL < Fa$ – значение R_2).

По завершении серии всех статистических испытаний образуются два непересекающихся множества издержек $\{R_1\}$ и $\{R_2\}$, для которых рассчитываются их средние арифметические значения $\bar{R}_1$ и $\bar{R}_2$ соответственно.

Если испытаний достаточно много, то средние арифметические $\bar{R}_1$ и $\bar{R}_2$ представляют собой, как правило, достаточно хорошие приближения для показателей $M[R_1]$ и $M[R_2]$. В стандартном варианте модуль **«Equilibrium»** генерирует 7000 испытаний, поэтому можно считать справедливыми равенства

$$M[R_1] \text{ и } = \bar{R}_1;$$
$$M[R_2] = \bar{R}_2. \qquad (2.1)$$

Исходя из вышесказанного, риски вероятных издержек оцениваются следующим образом:

• **риск завышения плана** – это усредненная величина, рассчитанная по множеству $\{R_1\}$ всех тех и только тех случайных издержек R_1, для которых $PL > Fa$;

• **риск занижения плана** – это усредненная величина, рассчитанная по множеству $\{R_2\}$ всех тех и только тех случайных издержек R_2, для которых $PL < Fa$.

---

* Согласно определению, математическое ожидание $M[X]$ случайной величины X есть среднее значение для генеральной совокупности значений X (то есть для множества всевозможных значений X). При оценке результатов экономических экспериментов величина $M[X]$ рассчитывается обычно как среднее арифметическое конечного числа наблюдаемых значений $x_1, x_2, ..., x_n$ величины X. С содержательной точки зрения математическое ожидание $M[X]$ – это точка на числовой оси, вокруг которой концентрируется большинство случайных значений x_1, x_2, x_3 ... величины X.

2.3. Оптимизационная модель планирования продаж на основе минимаксного критерия выбора решений

При составлении оптимизационной модели планирования необходимо учитывать следующие теоретические положения.

1. По окончанию планового периода, то есть после того, как принятый план продаж *PL* будет реализован и станет известным фактический объем продаж *Fa*, в результате сравнения *PL* с *Fa* фактически реализованным окажется только один из ожидаемых рисков: либо риск завышения плана $\bar{R}_1$, либо риск занижения плана $\bar{R}_2$, но не оба риска вместе. В таких случаях теория принятия решений в условиях неопределенности рекомендует применять минимаксную стратегию поведения фирмы на рынке:

действовать так, чтобы при наихудшем стечении обстоятельств понести наименьшие потери, то есть следует минимизировать больший из рисков

2. Как следует из графика, представленного на рис. 2.1, чем больше план продаж *PL*, тем больше риск, что этот план окажется завышенным, и тем меньше риск, что план окажется заниженным. Иными словами, с увеличением плана *PL* риск завышения не может уменьшиться, а риск занижения не может увеличиться.

Рис. 2.1. Конъюнктурные риски и минимаксная стратегия фирмы на рынке

Следовательно, в самом общем виде критерий оптимальности плана выражается следующем образом:

оптимальным окажется план, при котором риск завышения и риск занижения плана уравновешиваются

Любое отклонение от такого плана увеличивает либо один, либо другой риск. Условие

$$\boxed{\text{риск завышения плана = риск занижения плана}} \quad (2.2)$$

представляет собой наиболее общей вид **модели оптимальности плана продаж**. В формализованном виде это условие выражается уравнением

$$|\bar{R}_1 - \bar{R}_2| = 0, \quad (2.3)$$

которое означает, что при оптимальном плане продаж $PL_{опт}$ абсолютная разность рисков завышения и занижения плана обращается в ноль.

Таким образом, стратегия поиска $PL_{опт}$ состоит в том, что для каждого из возможных планов продаж PL генерируются риск завышения $\bar{R}_1$ и риск занижения $\bar{R}_2$, определяется их абсолютная разность $|\bar{R}_1 - \bar{R}_2|$ и в качестве $PL_{опт}$ выбирается план PL, доставляющий минимум разности рисков (табл. 2.1).

Таблица 2.1

Пример расчета оптимального плана продаж*

План	Риск завышения	Риск занижения	|Риск завышения – Риск занижения|
19,70	543,93	174 265,86	17 3721,94
29,11	1 631,79	155 832,50	154 200,72
38,56	2 719,63	137 400,50	134 680,88
47,92	3 807,48	119 060,65	115 253,16
57,37	4 899,52	101 258,73	96 359,21
66,73	6 105,90	84 306,91	78 201,07
76,14	7 690,12	67 803,53	60 113,41
85,54	96 47,38	51 484,38	41 837,00
94,95	11 738,39	35 934,38	24 195,98
104,36	13 986,66	22 507,38	8 520,72
113,76	16 850,68	12 428,59	4 422,09
123,19	21 090,44	6 008,29	15 082,15
132,57	27 128,95	2 557,18	24 517,77
141,98	34 929,14	929,56	33 999,58
151,39	43 959,57	319,40	43 640,16
160,79	53 752,53	94,54	53 658,00
170,20	63 891,37	19,41	63 871,97
179,60	74 165,38	0,00	74 165,37

* Цветом выделены значения, соответствующие минимуму абсолютной разности рисков.

3. При расчете оптимального плана $PL_{опт}$ на основе уравнений (2.1) — (2.3) значения случайной величины Fa, а также случайных издержек R_1 и R_2, зависящих от разности $|PL - Fa|$, генерируются с помощью датчика случайных чисел и имитационных моделей (далее — IM).

В модуле **«Equilibrium»** для этой — цели используются три имитационные модели — IM_0, IM_1, IM_2. С помощью IM_0 генерируются случайные значения фактических продаж Fa_i. Модели IM_1 и IM_2 предназначены для расчета случайных издержек R_1 и R_2 соответственно: IM_1 используется в случае $PL > Fa_i$, IM_2 — в случае $PL < Fa_i$. При этом сгенерированные в результате всей серии статистических испытаний случайные значения R_1 и R_2 образуют два непересекающихся множества

$$\{R_1\} \cap \{R_2\} = \varnothing.$$

Риск завышения плана $\overline{R}_1$ рассчитывается путем усреднения случайных издержек по множеству $\{R_1\}$, риск занижения плана $\overline{R}_2$ — усреднением по множеству $\{R_2\}$.

4. Каков бы ни был план PL, если он реалистичен, он должен быть сопоставим с генерируемыми случайными фактами Fa_i, то есть должно выполняться условие

$$\mathrm{PL} \in \{Fa_i\},$$

где $\{Fa_i\}$ — множество всех сгенерированных случайных значений Fa. Иными словами, полученный в результате расчетов оптимальный план продаж $PL_{опт}$ является **случайной величиной**.

Исходя из вышесказанного (см. пп. 1—4), модель поведения фирмы на рынке (модель экономического планирования) можно представить в общем виде следующей системой соотношений:

$$\overline{f} = f_1, f_2, \ldots, f_n; \tag{2.4}$$

$$\overline{p} = p_1, p_2, \ldots, p_m; \tag{2.5}$$

$$F\alpha = IM_0(\overline{f}, \overline{p}); \tag{2.6}$$

$$R_1 = IM_1(PL, F\alpha, \overline{f}, \overline{p}); \tag{2.7}$$

$$R_2 = IM_2(PL, F\alpha, \overline{f}, \overline{p}); \tag{2.8}$$

$$S(PL, F\alpha) = \begin{cases} R_1, \text{если } PL > F\alpha \\ R_2, \text{если } PL < F\alpha \end{cases}; \tag{2.9}$$

$$\min_{PL}\{\max_i\{M[S(PL, F\alpha_i)]\}\}. \tag{2.10}$$

Равенства (2.4), (2.5) вводят обозначения выбранного для модели набора факторов (случайных величин) и показателей (условно-постоянных величин) соответственно.

Для каждого фактора $f_k \in \overline{f}$ необходимо указать минимальное и максимальное значения, в пределах которых изменяется его величина при генерации случайных фактов $F\alpha_i{:}f_k \in [f_k^{\min};\ f_k^{\max}]$. В отличие от факторов каждому показателю $p_k \in \overline{p}$ приписывается лишь одно определенное значение, которое остается неизменным (фиксированным) на протяжении расчета $PL_{опт}$ применительно к заданному набору входных данных $\overline{f}$, $\overline{p}$.

Из равенства (2.6) следует, что фактический объем продаж Fa зависит от случайных факторов $\bar{f}$ и фиксированных показателей $\bar{p}$ и что прогнозные (случайные) значения Fa_i получают с применением имитационной модели IM_0.

Равенства (2.7) и (2.8) показывают, что величина издержек завышения плана R_1 и издержек занижения плана R_2 зависят от намечаемого плана продаж PL, фактического объема продаж Fa, случайных факторов $\bar{f}$, фиксированных значений исходных показателей $\bar{p}$ и что R_1 и R_2 рассчитываются с применением имитационных моделей IM_1 и IM_2 соответственно.

Условие (2.9) есть формальное выражение некоторого (принятого в модели) правила расчета величины издержек R_1, R_2 в зависимости от значений PL и Fa.

Условие (2.10) представляет собой **критерий оптимальности плана**, выражающий минимаксную стратегию поведения фирмы на рынке. Оно равносильно условию (2.2) равенства рисков завышения и занижения плана.

В рассмотренной математической формулировке (2.4) — (2.10) оптимизационной задачи планирования содержание имитационных моделей IM_0, IM_1, IM_2 не раскрывается. Допускаются различные способы конкретизации этих моделей.

Кроме того необходимо, чтобы соблюдались следующие условия моделирования:

• PL и Fa должны принадлежать одному и тому же вполне упорядоченному множеству W:

$$PL \in W;\ Fa \in W;$$

• для любых PL и Fa справедливо либо $PL > Fa$, либо $PL < Fa$;

• существует процедура сопоставления PL с Fa;

• имитационная модель IM_0 дает возможность получать реализации $Fa \in W$;

• имитационные модели IM_1, IM_2, позволяют рассчитывать случайные издержки R_1 и R_2 соответственно;

• R_1 и R_2 должны принадлежать некоторому множеству $\{R\}$, которое допускает операции сложения и сопоставления элементов из и $\{R\}$.

2.4. Надежность выполнения оптимального плана $PL_{опт}$

Надежность P плана $PL_{опт}$ — это вероятность того, что план окажется выполненным или перевыполненным на заданном секторе рынка:

$$P = P(F\alpha \geq PL).$$

Модуль **«Equilibrium»** при определении оптимального плана $PL_{опт}$ рассчитывает также надежность выполнения $PL_{опт}$. Для этой цели используется формула

$$P = P(F\alpha \geq PL) = 1 - \frac{PL - F\alpha^{\min}}{F\alpha^{\max} - F\alpha^{\min}}, \tag{2.11}$$

где $F\alpha^{\min}$ и $F\alpha^{\max}$ Fa — соответственно минимальное и максимальное из случайных значений Fa, сгенерированных при расчетах плана $PL_{опт}$.

Соотношение между количественными и качественными характеристиками надежности задается определенной оценочной шкалой, зависящей от решаемой задачи и субъективного мнения пользователя. Одна из возможных оценочных шкал приведена в табл. 2.2.

Таблица 2.2

Оценочная шкала показателя надежности P

Количественная оценка *P*	**Качественная характеристика надежности**
0,1—0,3	Низкая
0,3—0,5	Удовлетворительная
0,5—0,75	Достаточная хорошая
0,75—1,00	Высокая

2.5. Индикатор прироста рисков при отклонении от плана $PL_{опт}$

Оптимальный план $PL_{опт}$. находится (согласно теории равновесных процессов) при выполнении условия (2.2):

риск завышения плана = риск занижения плана

Если рассчитано оптимальное значение плана $PL_{опт}$, то можно оценить, в какой мере отклонение от найденного оптимума (в сторону его завышения или занижения) изменяет величину рисков издержек $\overline{R}_1$ и $\overline{R}_2$.

Для этой цели используется относительный показатель

$$Z = \frac{\text{риск завышения } PL_{опт}}{\text{риск занижения } PL_{опт}}, \tag{2.12}$$

который служит **индикатором прироста рисков** при отклонении от $PL_{опт}$.

Если $Z = \frac{\text{риск завышения } PL_{опт}}{\text{риск занижения } PL_{опт}} > 1,$ **то *Z* показывает, насколько превышение найденного оптимума $PL_{опт}$ на некоторую величину Δ более рискованно, чем занижение $PL_{опт}$ на ту же величину, то есть насколько план ($PL_{опт}$ + Δ) рискованнее плана ($PL_{опт}$. − Δ).**

Таким образом, индикатор *Z* указывает, как изменяется риск издержек (увеличивается или уменьшается) при отклонении от оптимального плана в сторону его завышения или занижения.

2.6. Точность и достоверность результатов оптимизационных расчетов

Погрешность Δf_k оценке отдельного фактора Δf_k равна разности между предельными экспертными оценками значений данного фактора:

$$\Delta f_k = \left| f_k^{\max} - f_k^{\min} \right|.$$

Погрешность $\Delta F\alpha$ в оценке ожидаемого объема продаж $F\alpha$ выражается через погрешности в оценках отдельных факторов Δf_k и вычисляется по формуле

$$\Delta F\alpha = F\alpha^{\max} - F\alpha^{\min}, \tag{2.13}$$

где величина $\Delta F\alpha$ составляет исходную погрешность расчетов, свойственную методу аналитического моделирования, причем $\Delta F\alpha$, как правило, оказывается неприемлемо большой (Приложение 2).

Методу статистических испытаний, реализованному в модуле **«Equilibrium»**, также свойственна погрешность в оценке $F\alpha$, возникающая за счет применения датчика случайных чисел. Погрешность метода проявляется в том, что для одних и тех же исходных данных при повторных испытаниях генерируются разные значения $F\alpha$. При этом колебания значений $F\alpha$ находятся в некоторых границах, составляющих фактическую погрешность расчетов, называемую также точностью расчетов.

Величина фактической погрешности (интервала колеблемости случайных значений $F\alpha$) зависит как от погрешности в оценках факторов и особенностей применяемой модели, так и от числа статистических испытаний. Выполнив расчет плана $PL_{опт}$, модуль **«Equilibrium»** всегда указывает его интервальную оценку (интервал неопределенности $PL_{опт}$).

В процессе оптимизационных расчетов используются еще два вида погрешности — приемлемая и допустимая.

Приемлемой погрешностью результатов расчетов считается 5%-ная погрешность. В модуле **«Equilibrium»** она всегда может быть достигнута при достаточном количестве испытаний.

Фактическая погрешность расчетов $F\alpha$ может превышать приемлемую 5%-ную погрешность, что указывает на недостаточное количество испытаний, выполненных в процессе расчетов.

Допустимая погрешность — это выраженное в процентах допустимое отклонение плана $PL_{опт}$ от теоретического оптимума. Если, например, допустимая погрешность установлена в размере 10%, то это означает, что отклонения искомого плана PL в пределах интервала

$$[PL - 0{,}1PL;\ PL + 0{,}1PL]$$

считаются допустимыми.

Качество полученного решения, кроме погрешности результатов, характеризуется также достоверностью результатов. Достоверность — это вероятность того, что найденный план $PL_{опт}$ будет отличаться от теоретического оптимума в пределах допустимой погрешности. Обычно достоверность (доверительная вероятность) принимается равной 0,95.

При выполнении расчетов пользователь в режиме диалога с модулем **«Equilibrium»** задает требуемую для его задачи величину допустимой погрешности и достоверности результатов.

В процессе оптимизационных расчетов плана продаж, выполняемых модулем **«Equilibrium»**, неприемлемо большая исходная погрешность $\Delta F\alpha_{исх}$. многократно снижается до уровня заданной допустимой погрешности $\Delta F\alpha_{доп}$. В результате рассчитанный модулем оптимальный план имеет на один–два порядка меньшую погрешность, чем исходная

$$10 \le \frac{\Delta F\alpha_{исх}}{\Delta F\alpha_{доп}} \le 100. \tag{2.14}$$

Многократное снижение погрешности расчетов в модуле **«Equilibrium»** достигается за счет применения высокоэффективного математического аппарата моделирования равновесия рыночных процессов, реализованного в библиотеке равновесных моделей этого модуля.

2.7. Задача исследования емкости товарного рынка или рынка услуг

2.7.1. Содержательная постановка задачи

На фирму поступило предложение о поставке партии из 50 детских велосипедов для детей до 7 лет по цене 1 100 руб. за штуку с учетом доставки. Требуется оценить, выгодно ли это предложение, принимая во внимание, что плановый период составляет один месяц и розничная цена детских велосипедов у конкурентов составляет от 2 300 до 2 500 руб.

Данная задача относится к классу задач планирования, поскольку для принятия решения необходимо ответить на вопросы: сколько детских велосипедов и по какой цене можно продать при сложившейся конъюнктуре и том секторе рынка, на котором работает фирма.

Следовательно, задача заключается в оценке ожидаемого спроса на товар, предлагаемый поставщиком. Оценив спрос на детские велосипеды, нетрудно заключить, выгодна ли предлагаемая коммерческая сделка.

Чтобы разработать математическую модель зависимости плана продаж велосипедов от факторов спроса, необходимо выделить главное и отбросить второстепенное. В частности, на объем продаж велосипедов влияет их цвет, дизайн, имидж производителя и др. Однако если товар у конкурентов обладает примерно теми же основными качествами, то наиболее существенными факторами, влияющими на спрос, становятся количество покупателей, их доходы, структура потребления и конкуренция.

Учитывая сказанное, в модель планирования целесообразно включить следующий набор факторов:

- f_1 — количество семей, проживающих в микрорайоне и посещающих магазин;
- f_2 — доля семей, имеющих детей в возрасте до 7 лет (так как велосипеды рассчитаны на этот возраст);
- f_3 — доля семей, имеющих средний и высокий доход (так как велосипеды достаточно дорогие и ориентированы на эту категорию покупателей);
- f_4 — средний доход семьи (рассматриваемой категории);
- f_5 — доля дохода, которую семья готова тратить на игрушки;
- f_6 — склонность к покупке (вероятность, что потенциальный покупатель пожелает стать реальным);
- f_7 — доля рынка, занятая конкурентами.

Для расчетов планов продаж факторам $f_1 - f_7$ необходимо присвоить числовые значения, учитывая при этом сопоставимость исходных данных и результатов. Так, доход, доля затрат на игрушки и склонность к покупке должны быть соотнесены с плановым периодом. Поскольку разрабатывается план продаж на месяц, то доход должен быть ежемесячным, затраты на игрушки (в долях) — ежемесячными и склонность к покупке также ежемесячной.

Наиболее простой, доступный, оперативный и дешевый способ получения исходных данных — применение **метода экспертных оценок**.

В табл. 2.3. даны экспертные оценки предельных значений факторов, то есть минимального и максимального значения. Комментарии, поясняющие эти экспертные оценки, приведены в Приложении 2.

Таблица 2.3

Экспертные оценки факторов спроса на детские велосипеды

Факторы и исходные показатели	Размерность	Значение
Количество семей, min	Семья	13 000,00
Количество семей, max	Семья	15 000,00
С детьми до 7 лет, min	%	20,00
С детьми до 7 лет, max	%	25,00
Со средним и высоким доходом, min	%	25,00
Со средним и высоким доходом, max	%	35,00
Доход, min	Руб./мес.	35 000,00
Доход, max	Руб./мес.	50 000,00
Доля затрат на игрушки, min	%	13,00
Доля затрат на игрушки, max	%	15,00
Склонность к покупке, min	Доля ед.	0,70
Склонность к покупке, max	Доля ед.	0,90
Доля конкурентов, min	%	93,00
Доля конкурентов, max	%	100,00
Цена конкурентов, min	Руб.	2 300,00
Цена конкурентов, max	Руб.	2 500,00

Данные табл. 2.3 достаточно приближенные и, возможно, содержат грубые ошибки. Чтобы выявить ошибки и уточнить исходные данные, необходимо провести пробные расчеты плана, установив пробную цену.

При назначении пробной цены следует исходить из цены конкурентов, сравнительного качества товаров и мотивов поведения потребителей.

Средняя розничная цена велосипеда у конкурентов — 2 400 руб., поэтому целесообразно назначить пробную цену несколько ниже этой цены, например, 2 350 руб.

Закупочная и розничная цены велосипедов фиксированы на протяжении планового периода (месяц) и включаются в модель как условно-постоянные показатели p_1 и p_2 (табл. 2.4).

Таблица 2.4

Оценки исходных показателей

Факторы и исходные показатели	Размерность	Значение
Закупочная цена велосипеда (p_1)	Руб.	1 100,00
Намеченная пробная розничная цена велосипеда (p_2)	Руб.	2 350,00

2.7.2. Формализация задачи в виде модели экономического планирования

Математическая модель экономического планирования представлена соотношениями (2.4) — (2.10). Необходимо ее конкретизировать применительно к рассматриваемой задаче с учетом выбранных факторов и условно-постоянных показателей, значения которых представлены в таблицах 2.3 и 2.4 соответственно.

Для расчета значений $F\alpha$ в имитационной модели (2.6) модуль **«Equilibrium»** используют формулу, рекомендуемую для оценки ожидаемого объема продаж:

$$F\alpha = \frac{Sd_1 d_2}{p_2}, \tag{2.15}$$

где S — суммарный доход семей со средним и высоким доходом, имеющих детей до 7 лет;

d_1 — доля дохода, которая предположительно будет затрачена на покупку детских велосипедов;

d_2 — доля рынков велосипедов, которую предположительно можно занять;

p_2 — розничная цена велосипеда.

Величины S, d_1 и d_2 рассчитываются по формулам (2.16а) — (2.16в):

$$S = f_1 \frac{f_2}{100} \cdot \frac{f_3}{100} f_4; \tag{2.16а}$$

$$d_1 = \frac{f_5}{100} f_6, \tag{2.16б}$$

$$d_2 = \frac{100 - f_7}{100}. \tag{2.16в}$$

Факторы $f_1 - f_7$ являются случайными величинами, следовательно, значения $F\alpha$ также случайны.

При разработке плана продаж велосипедов неизбежно возникает риск, который носит двоякий характер. Какой бы план продаж PL ни был принят, он может оказаться меньше спроса $F\alpha$ и фирма понесет денежные потери, связанные с закупкой велосипедов, оказавшихся непроданными. Стоимость невостребованных велосипедов — это издержки R_1 завышения плана относительно спроса ($PL > Fa$). Но этот же самый план продаж может оказаться и меньше спроса на велосипеды ($PL < Fa$), в этом случае возникает упущенная выгода (недополученная прибыль, упущенная доля фирмы на рынке). Размер упущенной выгоды — это издержки R_2 занижения плана относительно спроса.

Величина R_1 рассчитывается по формуле

$$R_1 = p_1(PL - F\alpha), \tag{2.17}$$

где p_1 — закупочная цена велосипеда и $PL > Fa$.

Величина R_2 рассчитывается по формуле

$$R_2 = (p_2 - p_1)(F\alpha - PL), \tag{2.18}$$

где p_2 — розничная цена велосипеда и $PL < Fa$.

Формулы (2.17) и (2.18) являются конкретизациями соотношений (2.7) и (2.8) соответственно.

В фактической ситуации (по окончании планового периода продажи велосипедов) возникают либо издержки R_1, либо издержки R_2. Чтобы учесть возможность появления любой из этих альтернатив в расчетах $PL_{опт}$ используется функция $S(PL, F\alpha)$, которая любому принятому плану PL и любым реализациям спроса $F\alpha$ ставит в соответствие величину издержек:

$$S(PL,\ F\alpha) = p_1(PL - F\alpha), \text{ если } PL > F\alpha; \tag{2.19}$$

$$S(PL,\ F\alpha) = (p_2 - p_1)(F\alpha - PL), \text{ если } PL < F\alpha. \tag{2.20}$$

Формула (2.19) является конкретизацией функции разветвления (2.9).

Так как размеры издержек R_1, R_2 зависят от случайных значений $F\alpha$, то они являются случайными величинами и, следовательно, оцениваются лишь вероятностно, с помощью показателей математического ожидания $M[R_1]$ и $M[R_2]$ соответственно.

В модуле **«Equilibrium»** эти показатели рассчитываются **методом статистических испытаний** на основе формул (2.1). При этом в каждом отдельном испытании автоматически генерируются случайные значения $f_1 - f_7$, R_1 или R_1.

Таким образом, ***риск завышения плана продаж*** велосипедов оценивается величиной $\bar{R}_1$, где R_1 рассчитывается по формуле (2.17) и усреднение производится по множеству сгенерированных значений $\{F\alpha_i\}$, для которых $PL > F\alpha_i$ (i = 1,2,3, ...).

Риск занижения плана продаж велосипедов оценивается величиной $\bar{R}_2$, где R_2 рассчитывается по формуле (2.18) и усреднение производится по множеству сгенерированных значений $\{F\alpha_i\}$, для которых $PL > F\alpha_i$ (i = 1,2,3, ...).

Исходя из условия оптимальности плана (2.2) и (2.3) может быть получена формула (2.21), позволяющая рассчитать оптимальный объем продаж, который удовлетворяет минимаксному критерию оптимальности (2.10)

$$PL = F\alpha^{\min} + (F\alpha^{\max} - F\alpha^{\min})\frac{1}{1 + \sqrt{\dfrac{p_1}{p_2 - p_1}}}. \tag{2.21}$$

Именно эта формула используется в модуле **«Equilibrium»** для расчета $PL_{опт}$.

Как следует из неравенства (2.14), расчет $PL_{опт}$ по формуле (2.21) выполняется с погрешностью, которая на один—два порядка меньше, чем исходная погрешность.

Надежность выполнения плана $PL_{опт}$ рассчитывается по формуле (2.11). Качественные характеристики надежности приведенные в оценочной табл. 2.1.

На основе найденного плана продаж $PL_{опт}$ рассчитываются ожидаемые для него прибыль и затраты:

$$\textit{прибыль} = (p_2 - p_1)PL_{опт}; \tag{2.22}$$

$$\textit{затраты} = p_1 PL_{опт}. \tag{2.23}$$

При намерении увеличить (или уменьшить) план $PL_{опт}$ на N шт. велосипедов возникает дополнительный риск (прирост риска):

1) при ($PL_{опт}$ + N) возрастает угроза потерь дополнительных денежных средств, вложенных в N велосипедов, которые могут оказаться непроданными;

2) при ($PL_{опт}$ – N) возрастает риск недополучения прибыли, которую обеспечивает реализация N велосипедов в случаи их поступления в продажу.

Для анализа прироста рисков при отклонении от плана $PL_{опт}$ используется индикатор Z, вычисляемый по формуле (2.12):

$$Z = \frac{\text{Риск завышение } PL_{опт}}{\text{Риск занижение } PL_{опт}},$$

Если $Z > 1$, то увеличение плана $PL_{опт}$ на несколько велосипедов создает больший прирост риска, чем уменьшение $PL_{опт}$ на то же количество велосипедов;

Если $Z < 1$, то уменьшение плана $PL_{опт}$ на несколько велосипедов создает больший прирост риска, чем увеличения $PL_{опт}$ на то же количество велосипедов;

Если $Z \approx 1$, то как увеличение, так и уменьшение $PL_{опт}$ создает примерно одинаковый дополнительный риск.

2.8. Инструментальные средства модуля «Equilibrium» системы «Decision» для задачи планирования

1. Загрузка системы **«Decision»** и модуля **«Equilibrium»** выполняются с помощью диалоговой процедуры:

Пуск => Все программы => Decision => Equilibrium

После загрузки модуля **«Equilibrium»** появляется панель главного меню модуля (рис. 2.2 и рис. 2.3), содержащее список меню, посредством которых инициируется выполнение различных команд:

- «*Библиотека равновесных моделей*»;
- «*Расчет*»;
- «*Сервис*»;
- «*База данных*» (в данной практической работе не используется);
- «?» — вызов справки для получения информации о моделях, процедурах, тематическом поиске и других возможностях модуля **«Equilibrium»** системы **«Decision»**.

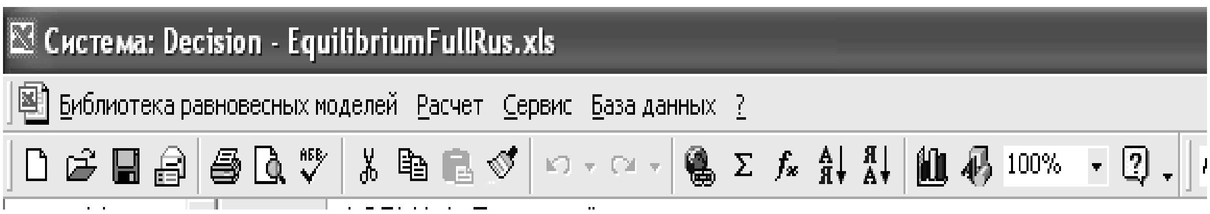


Рис. 2.2. Панель главного меню модуля **«Equilibrium»** в Microsoft Excel 2003 и более ранних

Рис. 2.3. Панель главного меню модуля **«Equilibrium»** в Microsoft Excel 2007 и более поздних

При запуске модуля **«Equilibrium»** возможно появление предупреждения системы безопасности (рис. 2.4). Необходимо нажать «Включить содержимое». Панель главного меню появится на вкладке «Надстройки».

Рис. 2.4. Предупреждение системы безопасности Microsoft Excel 2007 и более поздних

2. Меню **«Библиотека равновесных моделей»** содержит команды выбора 12 моделей, предназначенных для решения задач с применением теории равновесных процессов (Приложение 3). Кроме того, в этом меню предусмотрены пять команд *«Моя модель 1» — «Моя модель 5»* для включения в нее авторских моделей, разрабатываемых с помощью средств системы **«Decision»**.

В настоящем практикуме используется учебная версия *«Тест»* модели **«Товарные рынки и рынки услуг»**, предназначенной для исследования емкости товарного рынка и рынка услуг, оценки ожидаемого объема продаж, поиска оптимальной цены, прогноза прибыли и затрат, анализа тактики и стратегии конкурентной борьбы.

Вместе с выбором модели **«Тест»** автоматически решается вопрос о составе факторов и показателей. Обращение к модели **«Тест»** и демонстрационному примеру с экспертными оценками производится с помощью диалоговой процедуры:

1. Библиотека равновесных моделей => Тест =>

2. Пример заполнения данных → **ОК** → **ОК**

Выполнение в диалоговой процедуре команды *«Пример заполнения данных»* позволяет автоматически ввести исходные данные демонстрационного примера (табл. 2.3 и табл. 2.4).

В результате выполнения процедуры автоматически открывается лист «Варианты», где в разделе «ФОРМА 1: Прямой расчет» размещена таблица Excel со входными данными демонстрационного примера (рис. 2.5).

	A	B	C	D	G	H	I	J
1	**ФОРМА 1: Прямой расчет**				**ТЕСТ**	**0,00**		
2		ИСХОДНЫЕ ДАННЫЕ			РЕЗУЛЬТАТ ОПТИМИЗАЦИОННОГО РАСЧЕТА			Количество
3		ФАКТОРЫ И	Размер-		ПЛАН	Размерность	Значение	Факторов
4	№	ИСХОДНЫЕ ПОКАЗАТЕЛИ	ность	*Значение*	Оптимум	Шт	78,21	8,00
5	1	Количество семей, min	Семей	13 000,00	Завышение/Занижение	Доли ед.	1,06	Показателе
6	2	Количество семей, max	Семей	15 000,00	Надежность по завышению	Доли ед.	0,52	3,00
7	3	С детьми до 7 лет, min	%	20,00	Надежность по занижению	Доли ед.	0,40	Расчетны
8	4	С детьми до 7 лет, max	%	25,00	НОРМАТИВ			показателе
9	5	С средним и высоким доходом, min	%	50,00	Оптимум	Шт	78,21	2,00
10	6	С средним и высоким доходом, max	%	70,00	Завышение/Занижение	Доли ед.	1,06	Знаков при
11	7	Доход, min	руб/мес	35 000,00	Надежность по завышению	Доли ед.	0,52	расчете
12	8	Доход, max	руб/мес	50 000,00	Надежность по занижению	Доли ед.	0,40	плана и
13	9	Доля затрат на игрушки, min	%	13,00				норматива
14	10	Доля затрат на игрушки, max	%	15,00	Расчетные показатели			2,00
15	11	Склонность к покупке, min	доли ед.	0,70	Название	Размерн.		
16	12	Склонность к покупке, max	доли ед.	0,90	Прибыль	руб.	195 537,09	ИНТЕРВАЛ
17	13	Доля конкурентов, min	%	93,00	Затраты	руб.	172 072,64	НЕОПРЕ-
18	14	Доля конкурентов, max	%	100,00	<>	<>	<>	ДЕЛЕН-
19	15	Цена конкурентов, min	Руб	4 600,00	<>	<>	<>	НОСТИ:
20	16	Цена конкурентов, max	Руб	5 000,00	<>	<>	<>	
21	17	По договорам	Шт	15,00	<>	<>	<>	ПЛАН
22	18	Цена	руб.	4 700,00	<>	<>	<>	от:
23	19	Себестоимость	руб.	2 200,00	<>	<>	<>	72,47
24	20	<>	<>	<>	<>	<>	<>	до:
25	21	<>	<>	<>	<>	<>	<>	92,58
26	22	<>	<>	<>	<>	<>	<>	

Рис. 2.5. Структура листа «Варианты» (ФОРМА 1: Прямой расчет)

В случае использования входных данных, отличных от демонстрационного примера, их значения вводятся с клавиатуры в столбец D (вместо команды *«Пример заполнения данных»*).

3. При обращении к меню **«Расчет»** главного меню модуля открывается диалоговое окно, позволяющее выбрать команду выполнения варианта оптимизационного расчета из следующего списка:

- *«Прямой/Обратный»;*
- *«Исследование»;*
- *«Зависимости».*

3.1. Команда *«Прямой/Обратный»* реализует прямой и обратный варианты расчета плана продаж. Прямой расчет плана — это поиск оптимального плана, обратный — поиск плана (возможно, не оптимального) при заданных характеристиках модели расчета. Результаты расчетов размещаются на листе «Варианты» в ФОРМАХ 1—4 в соответствии с выбранным вариантом расчета (рис. 2.6—2.8).

При выборе команды *«Прямой расчет»* система осуществляет оптимизационный расчет плана $PL_{\text{опт}}$, а также рассчитывает соответствующие плану показатели: прибыль, затраты, надежность выполнения плана, прирост рисков при отклонении от $PL_{\text{опт}}$. Затем система выдает в специальных окнах сообщения о результатах расчета, комментарии к ним, открывает лист «Варианты» и помещает результаты в ФОРМУ 1 (см. рис. 2.5), а также строит графики рисков на листе «Риски» (рис. 2.9).

При обратном расчете диалоговом окне доступны команды:

- *«Обратный расчет по заданному Плану/Нормативу»;*
- *«Обратный расчет по заданной надежности»;*

• *«Обратный расчет по заданному соотношению рисков завышения и занижения плана»* (в практической работе не используется).

По команде *«Обратный расчет по заданному Плану/Нормативу»* для заданного пользователем плана *PL* определяется надежность его выполнения и рассчитываются показатели прибыли и затрат. Результаты записываются в ФОРМУ 2 листа «Варианты» (см. рис. 2.6).

По команде *«Обратный расчет по заданной надежности»* для введенного пользователем уровня надежности рассчитывается план *PL*, показатели прибыли и затрат. Результаты помещаются в Форму 3 листа «Варианты» (см. рис. 2.7).

При выборе команды *«Обратный расчет по заданному соотношению рисков завышения и занижения плана»* для заданного пользователем значения прироста рисков *Z* система рассчитывает план *PL*, показатели надежности его выполнения, прибыли и затрат. Результаты помещаются в Форму 4 листа «Варианты» (см. рис. 2.8).

M	N	O	P	Q	R	S
ФОРМА 2: Обратный расчет по заданному значению ПЛАНА (НОРМАТИВА)						**0,00**
	ИСХОДНЫЕ ДАННЫЕ			РЕЗУЛЬТАТ РАСЧЕТА		
	ФАКТОРЫ И	Размер-		ПЛАН	Размерность	Значение
№	ИСХОДНЫЕ ПОКАЗАТЕЛИ	ность	*Значение*	Завышение/Занижение	Доли ед.	0,00
1	Количество семей, min	Семей	13 000,00	Надежность по завышению	Доли ед.	1,00
2	Количество семей, max	Семей	15 000,00	Надежность по занижению	Доли ед.	1,00
3	С детьми до 7 лет, min	%	20,00	НОРМАТИВ		
4	С детьми до 7 лет, max	%	25,00	Завышение/Занижение	Доли ед.	0,97
5	С средним и высоким доходом,	%	25,00	Надежность по завышению	Доли ед.	1,00
6	С средним и высоким доходом,	%	35,00	Надежность по занижению	Доли ед.	1,00
7	Доход, min	руб/мес	35 000,00			
8	Доход, max	руб/мес	50 000,00	Расчетные показатели		
9	Доля затрат на игрушки, min	%	13,00	Название	Размерность	Значение
10	Доля затрат на игрушки, max	%	15,00	Прибыль	руб.	0,00
11	Склонность к покупке, min	доли ед.	0,70	Затраты	руб.	0,00
12	Склонность к покупке, max	доли ед.	0,90	<>	<>	<>
13	Доля конкурентов, min	%	93,00	<>	<>	<>
14	Доля конкурентов, max	%	100,00	<>	<>	<>
15	Цена конкурентов, min	Руб	2 300,00	<>	<>	<>
16	Цена конкурентов, max	Руб	2 500,00	<>	<>	<>
17	По договорам	Шт	15,00	<>	<>	<>
18	Цена	руб.	2 830,00	<>	<>	<>
19	Себестоимость	руб.	1 100,00	<>	<>	<>
20	<>	<>	<>	<>	<>	<>

Рис. 2.6. Структура листа «Варианты» (ФОРМА 2: Обратный расчет по заданному значению ПЛАНА/НОРМАТИВА)

U	V	W	X	Y	Z	AA	AB
	ФОРМА 3: Обратный расчет по заданной НАДЕЖНОСТИ:					**...**	
		ИСХОДНЫЕ ДАННЫЕ			РЕЗУЛЬТАТ РАСЧЕТА		
		ФАКТОРЫ И	Размер-		ПЛАН (НОРМАТИВ)	Размерность	Значение
	№	ИСХОДНЫЕ ПОКАЗАТЕЛИ	ность	*Значение*	по завышению (1)	Шт	
	1	Количество семей, min	Семей		по занижению (2)	Шт	
	2	Количество семей, max	Семей		Зав/Зан (для 1)	Доли ед.	
	3	С детьми до 7 лет, min	%		Зав/Зан (для 2)	Доли ед.	
	4	С детьми до 7 лет, max	%		Прирост совокупного риска	Абс. ед.	
	5	С средним и высоким доход	%		Прирост совокупного риска	%	
	6	С средним и высоким доход	%				
	7	Доход, min	руб/мес		Расчетные показатели		
	8	Доход, max	руб/мес		Название	Размерность	Значение
	9	Доля затрат на игрушки, min	%		Прибыль	руб.	
	10	Доля затрат на игрушки, max	%		Затраты	руб.	
	11	Склонность к покупке, min	доли ед.		<>	<>	<>
	12	Склонность к покупке, max	доли ед.		<>	<>	<>
	13	Доля конкурентов, min	%		<>	<>	<>
	14	Доля конкурентов, max	%		<>	<>	<>
	15	Цена конкурентов, min	Руб		<>	<>	<>
	16	Цена конкурентов, max	Руб		<>	<>	<>
	17	По договорам	Шт		<>	<>	<>
	18	Цена	руб.		<>	<>	<>
	19	Себестоимость	руб.		<>	<>	<>
	20	<>	<>	<>	<>	<>	<>

Рис. 2.7. Структура листа «Варианты» (ФОРМА 3: Обратный расчет по заданной НАДЕЖНОСТИ)

	AE	AF	AG	AH	AI	AJ	AK
1	**ФОРМА 4: Обратный расчет по заданному знач. пок-я Завышение/Занижение:**						
2		ИСХОДНЫЕ ДАННЫЕ			РЕЗУЛЬТАТ РАСЧЕТА		
3		ФАКТОРЫ И	Размер-		ПОКАЗАТЕЛЬ	Размерность	Значение
4	№	ИСХОДНЫЕ ПОКАЗАТЕЛИ	ность	*Значение*	ПЛАН	Шт	
5	1	Количество семей, min	Семей		Надежность по завышению	Доли ед.	
6	2	Количество семей, max	Семей		Надежность по занижению	Доли ед.	
7	3	С детьми до 7 лет, min	%				
8	4	С детьми до 7 лет, max	%		НОРМАТИВ	Шт	
9	5	С средним и высоким доходо	%		Надежность по завышению	Доли ед.	
10	6	С средним и высоким доходо	%		Надежность по занижению	Доли ед.	
11	7	Доход, min	руб/мес				
12	8	Доход, max	руб/мес		Расчетные показатели		
13	9	Доля затрат на игрушки, min	%		Название	Размерность	Значение
14	10	Доля затрат на игрушки, max	%		Прибыль	руб.	
15	11	Склонность к покупке, min	доли ед.		Затраты	руб.	
16	12	Склонность к покупке, max	доли ед.		<>	<>	<>
17	13	Доля конкурентов, min	%		<>	<>	<>
18	14	Доля конкурентов, max	%		<>	<>	<>
19	15	Цена конкурентов, min	Руб		<>	<>	<>
20	16	Цена конкурентов, max	Руб		<>	<>	<>
21	17	По договорам	Шт		<>	<>	<>
22	18	Цена	руб.		<>	<>	<>
23	19	Себестоимость	руб.		<>	<>	<>
24	20	<>	<>	<>	<>	<>	<>

Рис. 2.8. Структура листа «Варианты» (ФОРМА 4: Обратный расчет по заданному значению показателя «Завышение/Занижение»)

3.2. Обращение к команде *«Исследование»* меню **«Расчет»** позволяет установить степень влияния факторов и исходных показателей на план продаж, чтобы попытаться снизить погрешность именно тех факторов, которые влияют на план продаж в наибольшей мере.

3.3. Команда *«Зависимости»* позволяет проводить оперативные (прикидочные) расчеты путем построения графиков зависимостей и представлять эти зависимости в процентном выражении.

При построении зависимости в качестве изменяемой величины (аргумента) могут выступать:

- минимальное (максимальное) значение любого из факторов $f_1 - f_7$;
- исходные показатели p_1 или p_2;
- план продаж PL;
- показатели прибыли, затрат и надежности P выполнения плана;
- индикатор прироста рисков Z.

В качестве зависимой величины (функции) могут фигурировать:

- план продаж PL;
- показатели прибыли, затрат и надежности P выполнения плана**;**
- индикатор прироста рисков Z.

Для выбора зависимости используется диалоговое окно «**Вариант»** содержащее команды:

- *«План/норматив»*;
- *«Надежность»*;
- *«Соотношение рисков завышения и занижения плана»*;
- *«Фактор или показатель»*.

Аргумент для зависимости указывается в диалоговом окне «Изменяемая величина», доступ к окну осуществляется путем выполнения процедуры

Зависимости => Вариант => Фактор или показатель

Выбор зависимой величины производится в диалоговом окне **«Выбор показателя»**. Степень аппроксимирующей кривой (линейной, квадратической, кубической) задается в диалоговом окне **«Построить тренд»**.

При построение графиков система предоставляет возможность выразить искомую зависимость как в процентной, так и в натуральной форме. Для этой цели служит диалоговое окно **«Вопрос»**, в котором система запрашивает пользователя, следует ли строить график зависимости в процентном выражении, и при выборе ответа **«Да»** строит таковой.

4. При обращении к меню **«Сервис»** главного меню открывается диалоговое окно содержащее следующие команды:

- *«Тренд»*;
- *«Значения факторов, совместимые с оптимумом»*;
- *«Задание фактора в виде: МО и СКО, гистограммы, массива»*;
- *«Игра с конкурентом»*;
- *«Открыть лист»*.

4.1. Команда *«Тренд»* предназначена для выбора степени аппроксимирующей функции (линейной, квадратической, кубической).

4.2. Команда *«Значения факторов, совместимые с оптимумом»* позволяет сузить интервал неопределенности (погрешность) рассчитанного показателя $PL_{опт}$ путем поэтапного снижения интервалов неопределенности факторов.

4.3. Команда *«Задание фактора в виде: МО и СКО, гистограммы, массива»* предназначена для задания факторов $f_1 - f_7$ в том случае, если кроме экспертных оценок факторов известна некоторая дополнительная информация об их значениях:

• *МО* и *СКО* — если имеется основание предположить, что закон распределения вероятностей фактора близок к нормальному (известны математическое ожидание и среднеквадратическое отклонение значений фактора);

• *Гистограмма* — при наличии возможности уточнить информацию о факторе;

• *Массив* — если имеется массив реализаций значения фактора;

• *Тренд* — если необходимо прогнозировать фактор на основе имеющегося динамического ряда.

В практической работе используется только способ задания гистограммы для фактора.

4.4. Команда *«Игра с конкурентом»* предназначена для анализа ситуации на рынке, складывающейся в результате оперативного изменения цен конкурентами и ответной реакции фирмы в виде изменения цен (в данной практической работе не используется).

4.5. Команда *«Открыть лист»* — это вспомогательная команда для работы с диалоговыми процедурами. Используются следующие листы (табл. 2.5):

Таблица 2.5

Листы модуля «Equilibrium»

Лист	Назначение
«Графики»	Построение и изображение графиков зависимостей и рисков
«Факторы»	Задание входных и промежуточных значений факторов
«Имитаторы»	Задание факторов в виде МО и СКО, гистограммы и массива
«Игра с конкурентом»	Выполнение итераций при ценовой игре с конкурентами
«Таблицы»	Представление результатов расчетов в табличной форме

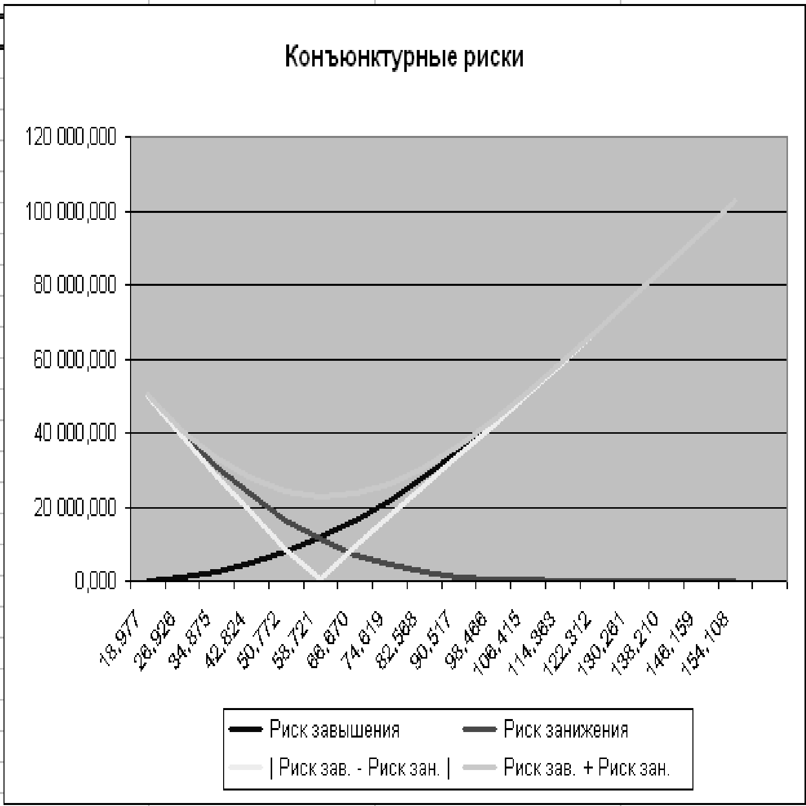
Кроме перечисленных, в процессе выполнения диалоговых процедур используются листы «Варианты» (Формы 1—4) и «Риски». Структура этих листов представлена на рис. 2.5—2.9, структура листа «Имитаторы» — на рис. 2.10.

5. В случае необходимости прервать работу с модулем и не потерять при этом данные и результаты, выполняется следующая диалоговая процедура:

1. Активизировать в правом верхнем углу окна модуля команду $\otimes$.
2. В появившемся диалоговом окне, содержащем вопрос «Вы хотите сохранить изменения в файле EquilibriumFullRus.xls?», выбрать вариант **«Да»**.

6. Выход из системы производится по следующей процедуре:

1. Активизировать в правом верхнем углу окна модуля команду $\otimes$.
2. В появившемся диалоговом окне, содержащем вопрос «Вы хотите сохранить изменения в файле EquilibriumFullRus.xls?» выбрать вариант **«Нет»**.



ПЛАН (НОРМАТИВ)	Р И С К И			
	Риск завышения	Риск занижения	| Риск зав. - Риск зан. |	Риск зав. + Риск зан.
18,977	99,525	50 112,297	50 012,773	50 211,820
26,926	875,213	39 834,629	38 959,418	40 709,840
34,875	2 423,649	30 793,594	28 369,945	33 217,242
42,824	4 716,649	22 979,588	18 262,938	27 696,238
50,772	7 802,623	16 405,365	8 602,742	24 207,988
58,721	11 664,354	11 147,099	517,255	22 811,453
66,670	16 324,203	7 148,440	9 175,763	23 472,645
74,619	21 814,537	4 277,392	17 537,145	26 091,930
82,568	28 103,391	2 378,383	25 725,008	30 481,773
90,517	35 095,086	1 215,398	33 879,688	36 310,484
98,466	42 694,176	584,560	42 109,617	43 278,734
106,415	50 742,426	255,875	50 486,551	50 998,301
114,363	59 112,648	103,224	59 009,426	59 215,871
122,312	67 664,227	33,260	67 630,969	67 697,484
130,261	76 318,563	12,546	76 306,016	76 331,109
138,210	85 022,930	5,699	85 017,234	85 028,625
146,159	93 755,172	1,965	93 753,211	93 757,133
154,108	102 495,023	0,000	102 495,023	102 495,023

Варианты | Риски

Рис. 2.9. Структура листа «Риски»

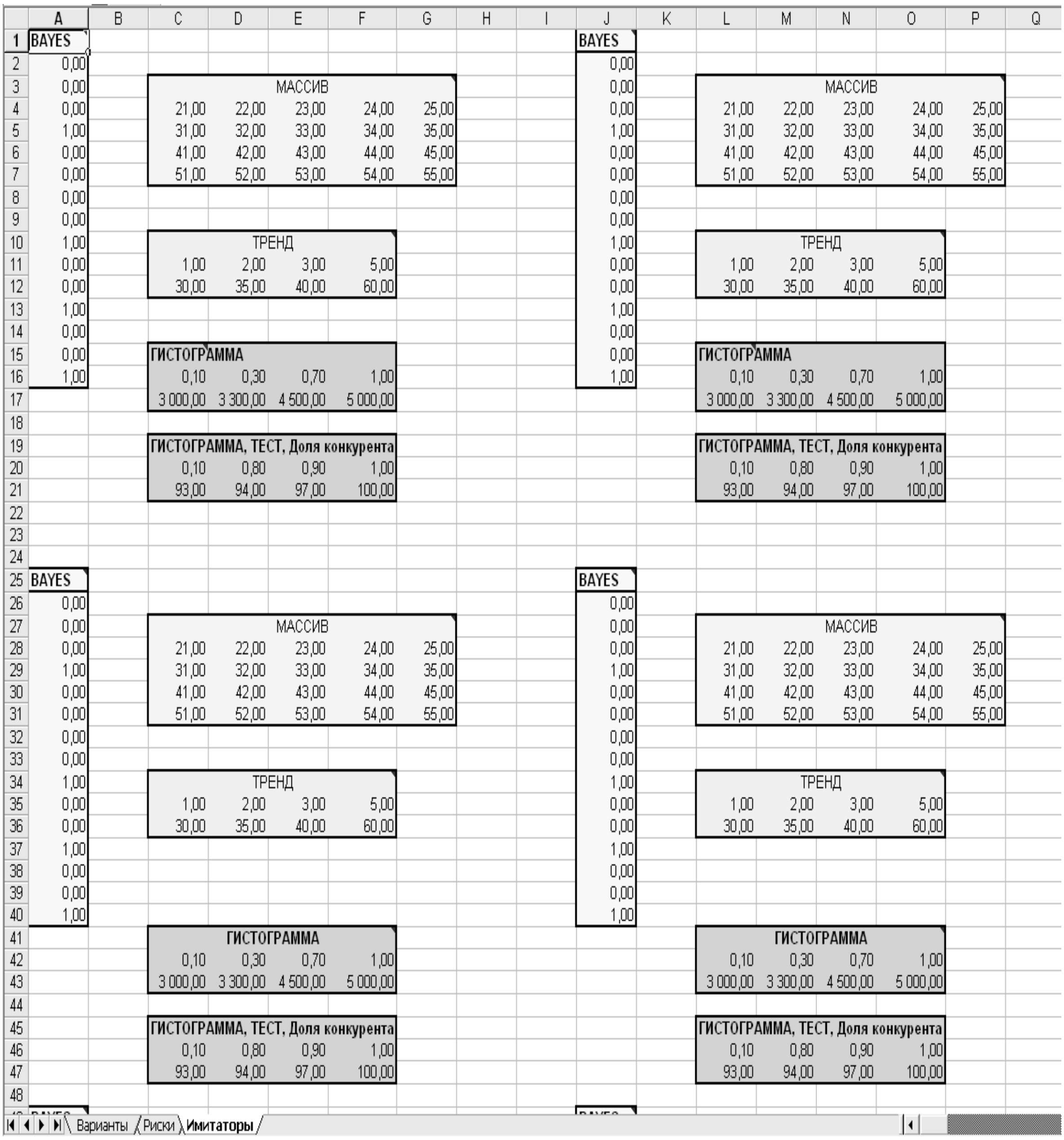


Рис. 2.10. Форма листа «Имитаторы»

2.9. Технология разработки плана продаж в среде модуля «Equilibrium» системы «Decision»

2.9.1. Состав и содержание информационной технологии разработки плана продаж с использованием модели «Тест»

Информационная технология планирования включает в качестве составляющих компонент информационные процессы I—VI, реализуемые в виде технологических диалоговых процедур.

I. Выбор модели для задачи планирования и предварительный анализ рыночной ситуации

Для рассматриваемой задачи применяется модель **«Тест»**, которая является упрощенной (учебной) версией модели **«Товарные рынки и рынки услуг»**.

Предварительный анализ рыночной ситуации заключается в том, чтобы дать ответы на вопросы, касающиеся аспектов изучения товарного рынка, покупателей и конкурентов. Для рассматриваемой задачи характеристика рыночной ситуации дана в Приложении 3.

II. Экспертиза факторов и пробный расчет оптимального плана продаж

Содержание процесса. Исходя из проведенного анализа ситуации, необходимо дать экспертные оценки предельных значений факторов и показателей, фигурирующих в выбранной модели, а также выполнить пробные оптимизационные расчеты плана продаж, в частности, оценить емкость рынка.

Реализация процесса. Для рассматриваемой задачи экспертные оценки даны в табл. 2.2. Их обоснование приведено в Приложении 4.

Ввод исходных данных для нахождения оптимального плана продаж осуществляется на этапе 1, пробные оптимизационные расчеты плана — на этапе 2:

Этап 1	Загрузка модели **«Тест»**, ввод исходных данных и оценка погрешности факторов
Этап 2	Пробный оптимизационный расчет плана

III. Уточнение предварительных оценок факторов и показателей, оптимизационные расчеты на их основе

Содержание процесса. Необходимо сузить интервалы предельных оценок факторов или заменить их гистограммами, массивами или прогнозами факторов с применением трендов; выполнить оптимизационные расчеты для определения ценовой политики фирмы.

Реализация процесса: диалоговые процедуры этапов 3—5:

Этап 3	Исследование оптимизационной модели расчета плана
Этап 4	Анализ влияния на план продаж уточнения данных о факторе **«Доля конкурентов»**
Этап 5	Вероятностная оценка реализации партии товара

IV. Определение ценовой политики фирмы на основе анализа зависимости прибыли и объема продаж от цены

Содержание процесса. Необходимо найти диапазон цен, максимизирующих прибыль, и установить, каковы объемы продаж, соответствующие этому диапазону цен.

Реализация процесса: диалоговые процедуры этапов 6—8:

Этап 6	Анализ зависимости прибыли от цены товара
Этап 7	Анализ зависимости объемов продаж от цены товара
Этап 8	Анализ зависимости надежности выполнения плана продаж от объема продаж

V. Исследование ожидаемых сценариев поведения покупателей и конкурентов в зависимости от рыночной конъюнктуры

Реализация процесса: диалоговые процедуры этапа 9:

Этап 9	Анализ зависимости объемов продаж от нижней границы цены товара у конкурентов

VI. Контрольный оптимизационный расчет

Содержание процесса. Коррекция исходных данных, расчет оптимального плана продаж, цены и сопутствующих характеристик плана.

Реализация процесса: диалоговые процедуры этапа 10:

Этап 10	Контроль принятого решения об оптимальном плане продаж

Таким образом, информационная технология разработки плана продаж включает десяти этапов, каждый из которых реализуется путем выполнения определенных технологических диалоговых процедур.

2.9.2. Этапы разработки плана продаж

Этап 1. Загрузка модели «Тест», ввод исходных данных и оценка погрешности факторов

Цель. Загрузка модели **«Тест»**, ввод в модель начальных значений факторов и показателей, оценка интервалов неопределенности (исходной погрешности) факторов.

Комментарий. Модель **«Тест»** имеет фиксированный состав входных данных — 8 факторов и два показателя (рис. 2.11). Для каждого из факторов должны быть введены предельные значения «min» и «max», устанавливающие интервал неопределенности фактора. В границах этих интервалов будут генерироваться в дальнейшем случайные значения факторов, необходимые для расчета плана $PL_{\text{опт}}$. Величина интервала неопределенности фактора — это погрешность в оценке данного фактора.

В отличие от факторов для каждого показателя указывается лишь одно значение. В показателе «Себестоимость» фиксируется закупочная цена велосипеда, в показателе «Цена» — розничная цена велосипеда.

Введенные значения факторов и показателей рассматриваются как пробные исходные данные для предварительного (пробного) расчета плана $PL_{\text{опт}}$.

Для задачи разработки плана продаж детских велосипедов модель **«Тест»** содержит демонстрационный пример экспертных оценок значений факторов. В диалоговом окне **«Тест»** имеется команда *«Пример заполнения данных»*, которая автоматически загружает в модель данные этого примера.

Диалоговая процедура 1.1 (выполняется при расчетах по данным демонстрационного примера):

1. Библиотека равновесных моделей => Тест =>
2. Пример заполнения данных → **ОК** → **ОК**

Диалоговая процедура 1.2 (выполняется при расчетах по индивидуальным входным данным):

1. Библиотека равновесных моделей => Тест =>
2. В столбец D листа «Варианты» ввести исходные данные → **ОК** → **ОК**

Результат

1. Таблица Excel, содержащая в массиве [B5:D23] перечень факторов и показателей модели **«Тест»**, а также их начальные (пробные) значения. Экранная форма введенных данных представлена на рис. 2.11.

2. Значения исходной погрешности факторов, рассчитанные по входным данным, представлены в табл. 2.6.

	A	B	C	D
1	**ФОРМА 1: Прямой расчет**			
2		**ИСХОДНЫЕ ДАННЫЕ**		
3		**ФАКТОРЫ И**	**Размер-**	
4	**№**	**ИСХОДНЫЕ ПОКАЗАТЕЛИ**	**ность**	*Значение*
5	1	Количество семей, min	Семей	13 000,00
6	2	Количество семей, max	Семей	15 000,00
7	3	С детьми до 7 лет, min	%	20,00
8	4	С детьми до 7 лет, max	%	25,00
9	5	С средним и высоким доходом, min	%	25,00
10	6	С средним и высоким доходом, max	%	35,00
11	7	Доход, min	руб/мес	35 000,00
12	8	Доход, max	руб/мес	50 000,00
13	9	Доля затрат на игрушки, min	%	13,00
14	10	Доля затрат на игрушки, max	%	15,00
15	11	Склонность к покупке, min	доли ед.	0,70
16	12	Склонность к покупке, max	доли ед.	0,90
17	13	Доля конкурентов, min	%	93,00
18	14	Доля конкурентов, max	%	100,00
19	15	Цена конкурентов, min	Руб	2 300,00
20	16	Цена конкурентов, max	Руб	2 500,00
21	17	По договорам	Шт	15,00
22	18	Цена	руб.	2 350,00
23	19	Себестоимость	руб.	1 100,00

Рис. 2.11. Экранная форма входных данных модели «**Тест**»

Таблица 2.6

Исходная погрешность факторов

Фактор	Размерность	Величина погрешности в оценке фактора
Количество семей	Семья	2000,00
С детьми до 7 лет	%	5,00
Со средним и высоким доходом	%	10,00
Доход	Руб./мес.	15000,00
Доля затрат на игрушки	%	2,00
Склонность к покупке	Доля ед.	0,20
Доля конкурентов	%	7,00
Цена конкурентов	Руб.	200,00

Этап 2. Пробный оптимизационный расчет плана

Цель. Произвести пробный расчет плана $PL_{опт}$ на основе исходных значений факторов и показателей, оценить зависимые от плана показатели надежности выполнения плана, прибыли и затрат, а также точность выполненных расчетов.

Комментарий. Результаты пробного расчета плана рассматриваются *как предварительные* (прикидочные) и будут использоваться для последующего уточнения плана на основе коррекции исходных данных.

Поиск плана $PL_{опт}$ производится в варианте прямого расчета.

Диалоговая процедура 2.1:

1. Расчет => Прямой/Обратный => Прямой расчет → ОК
2. При появлении окон-сообщений и окон с комментариями → **ОК**

Результат

Оптимальный по минимаксному критерию показатель $PL_{опт}$, его интервальная оценка, качественные характеристики расчетной модели плана, показатели прибыли и затрат.

Экранная форма результата оптимизационного расчета представлена на рис. 2.12, табличная форма — в табл. 2.7.

G	H	I	J
ТЕСТ	0,00		
РЕЗУЛЬТАТ ОПТИМИЗАЦИОННОГО РАСЧЕТА			Количество:
ПЛАН	Размерность	Значение	Факторов
Оптимум	Шт	78,21	8,00
Завышение/Занижение	Доли ед.	1,06	Показателей
Надежность по завышению	Доли ед.	0,52	3,00
Надежность по занижению	Доли ед.	0,40	Расчетнных
НОРМАТИВ			показателей
Оптимум	Шт	78,21	2,00
Завышение/Занижение	Доли ед.	1,06	Знаков при
Надежность по завышению	Доли ед.	0,52	расчете
Надежность по занижению	Доли ед.	0,40	плана и
			норматива
Расчетные показатели			2,00
Название	Размерн.		
Прибыль	руб.	97 768,55	ИНТЕРВАЛ
Затраты	руб.	86 036,32	НЕОПРЕ-
<>	<>	<>	ДЕЛЕН-
<>	<>	<>	НОСТИ:
<>	<>	<>	
<>	<>	<>	ПЛАН
<>	<>	<>	от:
<>	<>	<>	72,47
<>	<>	<>	до:
<>	<>	<>	92,58
<>	<>	<>	

Рис. 2.12. Результаты пробного расчета плана

Таблица 2.7

Результаты пробного расчета плана

Параметр	Обозначение параметра в модуле «Equilibrium» и его адрес на листе «Варианты»	Размерность	Значение
Рассчитанный оптимальный план продаж ($PL_{опт}$)	Оптимум [G4:I4]	Штуки	78,21
Надежность Выполнения ($PL_{опт}$)	Надежность позанижению [G7:I7]	Доля ед.	0,40
Интервальная оценка ($PL_{опт}$)	Интервал неопределенности плана [J15:J25]	Штуки	[72,47;92,58]
Прирост рисков при отклонении плана от $PL_{опт}$	Завышение/Занижение [G5:I5]	Доля ед.	1,06
Ожидаемая прибыль при выполнении $PL_{опт}$	Прибыль [G16:I16]	Руб.	97 768
Ожидаемые затраты на закупку партии велосипедов	Затраты [G17:I17]	Руб.	86 036

Выводы

1. Рассчитанный оптимальный план продажи велосипедов $PL_{опт}$ составляет 78 шт. (с учетом округления). Случайные колебания значений плана лежат в интервале от 72 шт. до 93 шт.

2. Затраты на закупку партии в 78 шт. велосипедов составляют 86 036 руб., ожидаемая прибыль — 97 768 руб.

3. Надежность плана (вероятность 100%-ного выполнения) составляет P = 0,40. Согласно шкале надежности, ее можно считать *удовлетворительной*.

4. Индикатор прироста рисков $Z = 1{,}06 \approx 1$ указывает на то, что как увеличение, так и уменьшение плана в окрестности $PL_{опт}$ = 78 шт. создает примерно одинаковый дополнительный риск.

5. Погрешность в оценке $PL_{опт}$ составляет 21 велосипед (93 шт. — 72 шт.) или в процентном выражении ≈ 27% ($\frac{21}{78}100$). Так как погрешность расчетов (точность результата) значительно превышает приемлемую 5%-ную погрешность, интервал неопределенности плана желательно сузить путем сужения интервалов неопределенности факторов.

6. Поскольку $PL_{опт}$ = 78 шт. больше предложенной партии в 50 шт. велосипедов, *закупку партии велосипедов следует считать выгодной.* Даже нижняя граница интервала неопределенности (72 велосипеда) подтверждает выгодность закупки.

Этап 3. Исследование оптимизационной модели расчета плана

Цель. Получить информацию о значениях параметров модели, определяющих точность и достоверность оценки $PL_{опт}$, а также установить меру влияния на план продаж каждого из факторов и исходных показателей.

Комментарий. В составе модуля **«Equilibrium»** имеется команда *«Исследование»*, информирующая пользователя о тех значениях параметров, которые определяют качество оптимизационной модели, а также о границах допустимого изменения исходных данных. В этих границах генерируются случайные значения факторов при расчете степени их влияния на $PL_{опт}$.

Прежде всего система информирует в трех диалоговых окнах о значениях качественных параметров модели, которые предполагается использовать при расчета $PL_{опт}$:

• **«Допустимая погрешность»** (допустимое отклонение плана от $PL_{опт}$, выраженное в процентах) — 10%;

• **«Стандартное отклонение»** (исследуемое изменение значений факторов и показателей, выраженное в процентах) — 75%;

• **«Достоверность»** (желаемая вероятность попадания в окрестность оптимума с допустимой погрешностью) — 0,95;

Согласие пользователя с предложенными значениями этих трех параметров подтверждается выбором варианта **ОК** и означает, что план $PL_{опт}$ должен быть найден с вероятностью 0,95 и иметь погрешность на более чем

$$[PL - 0{,}1PL;\ PL + 0{,}1PL].$$

Ниже приводится пример экранной формы запросов системы (рис. 2.13).

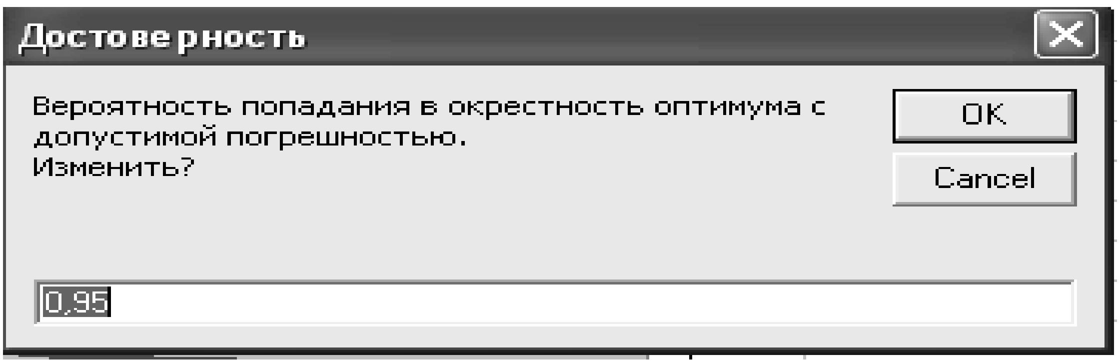


Рис. 2.13. Экранная форма запросов системы

Далее система для каждого из восьми факторов и двух показателей указывает в соответствующих диалоговых окнах минимальное и максимальное значения параметров. Эти значения устанавливают, в каких пределах допустимы пробные изменения факторов и показателей при выявлении степени их влияния на $PL_{опт}$ (для демонстрационного примера модели **«Тест»** граничные значения параметров приведены в табл. 2.8).

Таблица 2.8

Граничные значения параметров демонстрационного примера в диалоговой процедуре исследования

Параметр	Размерность	Знание	
		min	max
Количество семей	Семья	3 250	26 250
С детьми до 7 лет	%	5	43,75
Со средним и высоким доходом	%	6,25	61,25
Доход	Руб./мес.	8 750	87 500
Доля затрат на игрушки	%	3,25	26,25
Склонность к покупке	Доля ед.	0,175	1
Доля конкурентов	%	23,25	100
Цена конкурентов	Руб.	575	4 375
Цена	Руб.	—	4 112,5
Себестоимость	Руб.	—	1 925

При появлении последовательности диалоговых окон с запросами
о допустимых границах изменения параметров → **ОК**

Как правило, пользователь соглашается с предложенными системой граничными значениями параметров. Коррекция вносится лишь в тех случаях, когда предлагаемый системой вариант не имеет содержательного смысла (например, пробная закупочная цена оказывается выше розничной цены). В практической работе рекомендуется соглашаться с предложенными вариантами изменения исходных данных, для чего следует выбрать вариант **ОК** в диалоговом окне **«Допустимое изменение исходных данных»**.

Согласовав с пользователем допустимые изменения значений каждого фактора и показателя, система в автоматическом режиме рассчитывает коэффициенты эластичности плана $PL_{опт}$ относительно каждого фактора и показателя. Коэффициент эластичности показывает, на сколько процентов изменится $PL_{опт}$ при изменении фактора (показателя) на 1%.

Степень влияния фактора (показателя) на $PL_{опт}$ для наглядности выделяется цветом. Соответствие между цветовой гаммой и степенью влияния показано на рис. 2.14.

Цвет	Влияние
← желтый	Пренебрежимо малое
← коричневый	Слабое
← зеленый	Умеренное
← синий	Сильное
← красный	Определяющее

Рис. 2.14. Соответствие цветовой гаммы и степени влияния параметров

Диалоговая процедура 3.1:

1. Расчет => Исследование =>
2. Диалоговое окно **«Допустимая погрешность» → ОК =>**
3. Диалоговое окно **«Стандартная отклонение» → ОК =>**
4. Диалоговое окно **«Отклонение» → ОК =>**
5. Окно-сообщение **«Внимание» → ОК=>**
6. Диалоговые окна **«Допустимость изменений исходных данных» → ОК =>**
7. Окна-сообщения **→ ОК**

Результат

Экранная форма результата представлена на рис. 2.15, табличная форма — в табл. 2.9.

	A	B	C	D	E	G	H	I	J
1	ФОРМА 1: Прямой расчет					ТЕСТ	0,00		
2		ИСХОДНЫЕ ДАННЫЕ			ИССЛЕДОВАНИЕ	РЕЗУЛЬТАТ ОПТИМИЗАЦИОННОГО РАСЧЕТА			Количество:
3		ФАКТОРЫ И	азмер-		Степень влияния, %/%	ПЛАН	Размерность	Значение	Факторов
4	№	ИСХОДНЫЕ ПОКАЗАТЕЛИ	ность	*Значение*	*План*	Оптимум	Шт	79,46	8,00
5	1	Количество семей, min	Сем	13 000,00	0,38	Завышение/Занижение	Доли ед.	1,15	Показателей
6	2	Количество семей, max	Сем	15 000,00	0,58	Надежность по завышению	Доли ед.	0,53	3,00
7	3	С детьми до 7 лет, min	%	20,00	0,34	Надежность по занижению	Доли ед.	0,36	Расчетных
8	4	С детьми до 7 лет, max	%	25,00	0,53	НОРМАТИВ			показателей
9	5	С средним и высоким доходом, min	%	25,00	0,36	Оптимум	Шт	79,46	2,00
10	6	С средним и высоким доходом, max	%	35,00	0,65	Завышение/Занижение	Доли ед.	1,15	Знаков при
11	7	Доход, min	руб/	35 000,00	0,36	Надежность по завышению	Доли ед.	0,53	расчете
12	8	Доход, max	руб/	50 000,00	0,66	Надежность по занижению	Доли ед.	0,36	плана и
13	9	Доля затрат на игрушки, min	%	13,00	0,38				норматива
14	10	Доля затрат на игрушки, max	%	15,00	0,53	Расчетные показатели			2,00
15	11	Склонность к покупке, min	доли	0,70	0,33	Название	Размерн.		
16	12	Склонность к покупке, max	доли	0,90	0,14	Прибыль	руб.	99 320,17	ИНТЕРВАЛ
17	13	Доля конкурентов, min	%	94,00	12,02	Затраты	руб.	87 401,75	НЕОПРЕ-
18	14	Доля конкурентов, max	%	99,00	0,14	<>	<>	<>	ДЕЛЕН-
19	15	Цена конкурентов, min	Руб	2 300,00	0,36	<>	<>	<>	НОСТИ:
20	16	Цена конкурентов, max	Руб	2 500,00	0,54	<>	<>	<>	
21	17	По договорам	Шт	15,00	0,10	<>	<>	<>	ПЛАН
22	18	Цена	руб.	2 350,00	0,74	<>	<>	<>	от:
23	19	Себестоимость	руб.	1 100,00	0,24	<>	<>	<>	
24	20	<>	<	<>	<>	<>	<>	<>	до:

Рис. 2.15. Результаты процедуры «Исследование»

В экранной форме в столбце E* указаны **коэффициенты эластичности**, характеризующие степень влияния факторов и показателей на $PL_{\text{опт}}$**.**

* Если значения из столбца E на экране не видны, столбец необходимо расширить.

В табл. 2.9. последние три строки отвечают трем максимальным коэффициентам эластичности (в порядке убывания).

Таблица 2.9

Результаты диалоговой процедуры

Параметр	Обозначение параметра в модуле «Equilibrium»	Размерность	Значение
Рассчитанный оптимальный план продаж ($PL_{\text{опт}}$)	Оптимум	Штуки	79,46
Надежность выполнение ($PL_{\text{опт}}$)	Надежность по занижению	Доли ед.	0,36
Прирост рисков при отклонении плана от $PL_{\text{опт}}$	Завышение/ Занижение	Доли ед.	1,15
Ожидаемая прибыль при выполнении $PL_{\text{опт}}$	Прибыль	Руб.	99 320,17
Ожидаемые затраты на закупку партии велосипедов	Затраты	Руб.	87 401,75
Наиболее значимые параметры			
Доля конкурентов, минимальное значение	Доля конкурентов, min	%	12,02
Розничная цена велосипеда	Цена	Руб.	0,74
Максимальный доход семей	Доход, max	Руб.	0,66

Выводы

1. План $PL_{\text{опт}}$ = 79 шт. найден с достоверностью 0,95 и имеет интервал неопределенности не более чем

$$[PL - 0{,}1PL; PL + 0{,}1PL] = [79{,}46 - 0{,}1 \cdot 79{,}46;$$
$$79{,}46 + 0{,}1 \cdot 79{,}46] = [79{,}46;\ 87{,}3] \text{ шт.},$$

то есть погрешность расчета $PL_{\text{опт}}$ не превышает 16 шт. велосипедов или в процентном выражении ≈ 20% (16/79 100). Следовательно, погрешность расчетов плана по сравнению с предыдущем этапом (27%) уменьшилась.

2. План продаж $PL_{\text{опт}}$ = 79 шт. незначительно изменился по сравнению с предыдущим расчетом $PL_{\text{опт}}$ = 78 шт., хотя оба плана рассчитаны для одних и тех же исходных данных. Колебания значений $PL_{\text{опт}}$, наблюдаемые при повторных расчетах плана, объясняются тем, что план рассчитывается с использованием случайных значений факторов, генерируемых датчиком случайных чисел (колебания плана наблюдаются в пределах его погрешности).

3. Доминирующее влияние на $PL_{\text{опт}}$ оказывает фактор **«Доля конкурентов, min»** — 12%. Следовательно, при изменении минимальной рыночной доли конкурентов на 1% план $PL_{\text{опт}}$ = 79 шт. изменяется на 12%, что составляет 9 шт. велосипедов (79 0,12).

4. Влияние всех других факторов меньше 1%, в их числе наибольшее влияние на план оказывают показатель «Розничная цена велосипеда» и фактор «Максимальный доход семей».

Этап 4. Анализ влияния на план продаж уточнения данных о факторе «Доля конкурентов»

Цель. Провести оптимизационный расчет плана продаж при использовании уточненных оценок фактора **«Доля конкурентов»**. Проанализировать, в какой мере уточнение исходной экспертной оценки данного фактора влияет на результаты расчета.

Комментарий. Если требуется получить более полную информацию о каком-либо факторе, чтобы уточнить его влияние на объем продаж, используется способ задания данного фактора в виде гистограммы. В системе **«Decision»** гистограммой называется набор значений фактора с указанием вероятностей этих значений.

Согласно результатам исследования, проведенного на предыдущем этапе, наибольшую степень влияния на объем продаж оказывает **доля конкурентов**, в частности, ее минимальное значение. Поэтому целесообразно получить дополнительную информацию именно об этом факторе с тем, чтобы проанализировать, повлияет ли уточнение экспертных оценок на план продаж, и каково это влияние.

Минимальное и максимальное значения фактора известны — 93% и 100% соответственно, но можно уточнить, что конкуренты скорее всего займут не менее 94% и не более 97% (другие значения на интервале от 93 до 100% также возможны). Гистограмма для фактора «Доля конкурентов» представлена в табл. 2.10.

Таблица 2.10

Гистограмма для фактора «Доля конкурентов»

Вероятности (накопительным итогом), ед.	Доля конкурентов, %
0,10	93,00
0,80	94,00
0,90	97,00
1,00	100,00

Гистограмма вводится на лист «Имитаторы» в виде массива 2 × N (то есть 2 строки по N элементов в каждой), где в верхней строке указываются вероятности значений фактора накопительным итогом от 0 до 1, а в нижней строке — значения фактора. Для демонстрационного примера модели **«Тест»** гистограмма табл. 2.10 зафиксирована на листе «Имитаторы» в ячейках [C20: F21] (см. рис. 2.10).

Диалоговая процедура 4.1 (выполняется, если необходимая гистограмма на листе «Имитаторы» отсутствует):

> **1. Сервис => Открыть лист => Имитаторы → ОК**
> **2.** В открывшемся листе **«Имитаторы»** ввести в массив [C80:F81] необходимую гистограмму → **ОК**

Для того чтобы произвести расчет $PL_{\text{опт}}$ на основе уточненных оценок фактора, они должны поступить из листа «Имитаторы» в таблицу с исходными данными (в **Форму 1** листа «Варианты»).

Диалоговая процедура 4.2:

1. Сервис => Задание факторов в виде: МО и СКО, гистограммы, массива → ОК
2. В открывшемся диалоговом окне **«Факторы»** выбрать пункт «*Доля конкурентов*» → **ОК**
3. В открывшемся окне **«Значения факторов»** выбрать пункт «*Гистограмма*» → **ОК.**
4. В открывшемся листе «Имитаторы» в поле диалогового окна **«Гистограмма»** ввести адрес [C20:F21] (или адрес [C80:F81], если предварительно выполнялась процедура 4.1.) → **ОК**

Результат

Экранная форма результата представлена на рис. 2.16. Надпись **«Гистограмма»** в верхней части листа указывает на применение способа задания исходных данных в виде гистограммы.

D4 fx Гистограмма

	A	B	C	D
1	**ФОРМА 1: Прямой расчет**			
2		**ИСХОДНЫЕ ДАННЫЕ**		
3		**ФАКТОРЫ И**	**Размер-**	***-30.00***
4	**№**	**ИСХОДНЫЕ ПОКАЗАТЕЛИ**	**ность**	***стограм***
5	**1**	**Количество семей, min**	**Семей**	**13`000.00**
6	**2**	**Количество семей, max**	**Семей**	**15`000.00**
7	**3**	**С детьми до 7 лет, min**	**%**	**20.00**
8	**4**	**С детьми до 7 лет, max**	**%**	**25.00**
9	**5**	**С средним и высоким доходом, min**	**%**	**25.00**
10	**6**	**С средним и высоким доходом, max**	**%**	**35.00**
11	**7**	**Доход, min**	**руб/мес**	**35`000.00**
12	**8**	**Доход, max**	**руб/мес**	**50`000.00**
13	**9**	**Доля затрат на игрушки, min**	**%**	**13.00**
14	**10**	**Доля затрат на игрушки, max**	**%**	**15.00**
15	**11**	**Склонность к покупке, min**	**доли ед.**	**0.70**
16	**12**	**Склонность к покупке, max**	**доли ед.**	**0.90**
17	**13**	**Доля конкурентов, min**	**%**	**93.00**
18	**14**	**Доля конкурентов, max**	**%**	**100.00**
19	**15**	**Цена конкурентов, min**	**Руб**	**2`300.00**
20	**16**	**Цена конкурентов, max**	**Руб**	**2`500.00**
21	**17**	**По договорам**	**Шт**	**15.00**
22	**18**	**Цена**	**руб.**	**2`350.00**
23	**19**	**Себестоимость**	**руб.**	**1`100.00**

Рис. 2.16. Задание фактора «Доля конкурентов» в виде гистограммы

Для оптимизационного расчета плана продаж по уточненным оценкам фактора «Доля конкурентов» (или другого фактора) используется еще одна диалоговая процедура.

Диалоговая процедура 4.3:

1. Расчет => Прямой/Обратный => Прямой расчет → ОК
2. При появлении окон-сообщений и окон с комментариями → **ОК**

Результат

Экранная форма результата представлена на рис. 2.17, табличная форма — в табл. 2.11.

В экранной форме вместо граничных значений фактора **«Доля конкурентов»** появляется информация, указывающая на применение гистограммы: в ячейке по адресу D17 появляется число – 30, а по адресу D18 – надпись **«Гистограмма»**.

1	**ФОРМА 1: Прямой расчет**				ТЕСТ		0,00	
2		ИСХОДНЫЕ ДАННЫЕ			РЕЗУЛЬТАТ ОПТИМИЗАЦИОННОГО РАСЧЕТА			Количество:
3		ФАКТОРЫ И	Размер-		ПЛАН	Размерность	Значение	Факторов
4	№	ИСХОДНЫЕ ПОКАЗАТЕЛИ	ность	*Значение*	Оптимум	Шт	109,43	8,00
5	1	Количество семей, min	Семей	13 000,00	Завышение/Занижение	Доли ед.	0,52	Показателей
6	2	Количество семей, max	Семей	15 000,00	Надежность по завышению	Доли ед.	0,69	3,00
7	3	С детьми до 7 лет, min	%	20,00	Надежность по занижению	Доли ед.	0,54	Расчетных
8	4	С детьми до 7 лет, max	%	25,00	НОРМАТИВ			показателей
9	5	С средним и высоким доходом, min	%	25,00	Оптимум	Шт	124,34	2,00
10	6	С средним и высоким доходом, max	%	35,00	Завышение/Занижение	Доли ед.	1,16	Знаков при
11	7	Доход, min	руб/мес	35 000,00	Надежность по завышению	Доли ед.	0,55	расчете
12	8	Доход, max	руб/мес	50 000,00	Надежность по занижению	Доли ед.	0,35	плана и
13	9	Доля затрат на игрушки, min	%	13,00				норматива
14	10	Доля затрат на игрушки, max	%	15,00	Расчетные показатели			2,00
15	11	Склонность к покупке, min	доли ед.	0,70	Название	Размерн.		
16	12	Склонность к покупке, max	доли ед.	0,90	Прибыль	руб.	136 782,29	**ИНТЕРВАЛ**
17	13	Доля конкурентов, min	%	-30,00	Затраты	руб.	120 368,42	**НЕОПРЕ-**
18	14	Доля конкурентов, max	%	стограмма	<>	<>	<>	**ДЕЛЕН-**
19	15	Цена конкурентов, min	Руб	2 300,00	<>	<>	<>	**НОСТИ:**
20	16	Цена конкурентов, max	Руб	2 500,00	<>	<>	<>	
21	17	По договорам	Шт	15,00	<>	<>	<>	**ПЛАН**
22	18	Цена	руб.	2 350,00	<>	<>	<>	**от:**
23	19	Себестоимость	руб.	1 100,00	<>	<>	<>	**104,46**
24	20	<>	<>	<>	<>	<>	<>	**до:**
25	21	<>	<>	<>	<>	<>	<>	**121,85**

Рис. 2.17. Результат оптимизационного расчета плана после замены граничных значений фактора «Доля конкурентов» на гистограмму

Таблица 2.11

Результаты расчета плана с применением гистограммы для задания фактора «Доля конкурентов»

Параметр	Обозначение параметра в модуле «Equilibrium» и его адрес на листе «Варианты»	Размерность	Значение
Рассчитанный оптимальный план продаж ($PL_{опт}$)	Оптимум [G4:I4]	Штуки	109,43
Надежность выполнения ($PL_{опт}$)	Надежность по занижению [G7:I7]	Доля ед.	0,54
Интервальная оценка ($PL_{опт}$)	Интервал неопределенности плана [J15:J25]	Штуки	[104,46; 121,85]
Прирост рисков при отклонении плана от $PL_{опт}$	Завышение/Занижение [G5:I5]	Доля ед.	0,52
Ожидаемая прибыль при выполнении $PL_{опт}$	Прибыль [G16:I16]	Руб.	136 782,29
Ожидаемые затраты на закупку партии велосипедов	Затраты [G17:I17]	Руб.	120 368,42

Вывод

1. Сопоставление результатов оптимизационных расчетов из табл. 2.7 и табл. 2.11 показывает, что использование более полных данных о доле конкурентов существенно повлияло на результат: план продаж $PL_{опт}$ увеличился с 78 шт. велосипедов до 109 шт. При этом прибыль возрастает с 97 768 руб. до 136 782 руб.

2. Надежность выполнения плана возросла с P = 0,40 до P = 0,54, изменив свою качественную характеристику с удовлетворительной надежности до достаточно хорошей.

3. Индикатор прироста рисков Z уменьшился с 1,06 до 0,52. Это указывает на то, что уменьшение плана в окрестности $PL_{опт}$ = 109 шт. создает значительно больший дополнительный риск, чем его увеличение.

4. Погрешность в оценке $PL_{опт}$ составляет 17 шт. велосипедов или в процентном выражении ≈ 16% (17/109 · 100), что меньше погрешности планов, рассчитанных на этапе 2 (27%) и на этапе 3 (19%).

5. Уточнение оценок наиболее значимого фактора «Доля конкурентов» еще более убеждает в том, что скорее всего удастся продать предложенную партию 50 шт. велосипедов по намеченной цене 2 350 руб.

6. Проведенный на этапах 3 и 4 анализ подтверждает вывод о выгодности закупки предложенной партии велосипедов, сделанный на этапе 2 (п. 6). Можно также ставить вопрос об увеличении объема закупаемой партии.

Этап 5. Вероятностная оценка реализации партии товара

Цель. Оценить надежность реализации (вероятность 100%-ной продажи) предложенной партии велосипедов (50 шт.) по намеченной цене в 2 350 руб. Определить, какое максимальное количество велосипедов может быть продано с приемлемой (достаточно хорошей) степенью надежности ($P \geq 0,50$).

Комментарий. Вероятность 100%-ной реализации 50 шт. велосипедов можно оценить, используя окно **«Обратный расчет по заданному плану»**, если в качестве плана принять размер предложенной партии (50 шт.).

Обратный расчет по заданной надежности позволяет установить, какое максимальное количество велосипедов можно закупить, чтобы была гарантирована их 100%-ная продажа с надежностью $P \geq 0,50$ (в практической работе выбрана надежность P = 0,52).

Диалоговая процедура 5.1:

1. Расчет => Прямой/Обратный => Обратный по заданному плану/Нормативу → ОК
2. В поле открывшегося окна **«Обратный расчет»** ввести 50 → **ОК**

Результат

Экранная форма результата представлена на рис. 2.18, табличная форма — в табл. 2.12.

M	N	O	P	Q	R	S
ФОРМА 2: Обратный расчет по заданному значению ПЛАНА (НОРМАТИВА)						**50,00**
	ИСХОДНЫЕ ДАННЫЕ			РЕЗУЛЬТАТ РАСЧЕТА		
	ФАКТОРЫ И	Размер-		ПЛАН	Размерность	Значение
№	ИСХОДНЫЕ ПОКАЗАТЕЛИ	ность	*Значение*	Завышение/Занижение	Доли ед.	0,10
1	Количество семей, min	Семей	13 000,00	Надежность по завышению	Доли ед.	0,90
2	Количество семей, max	Семей	15 000,00	Надежность по занижению	Доли ед.	0,88
3	С детьми до 7 лет, min	%	20,00	НОРМАТИВ		
4	С детьми до 7 лет, max	%	25,00	Завышение/Занижение	Доли ед.	0,97
5	С средним и высоким доходом,	%	25,00	Надежность по завышению	Доли ед.	0,90
6	С средним и высоким доходом,	%	35,00	Надежность по занижению	Доли ед.	0,88
7	Доход, min	руб/мес	35 000,00			
8	Доход, max	руб/мес	50 000,00	Расчетные показатели		
9	Доля затрат на игрушки, min	%	13,00	Название	Размерность	Значение
10	Доля затрат на игрушки, max	%	15,00	Прибыль	руб.	62 500,00
11	Склонность к покупке, min	доли ед.	0,70	Затраты	руб.	55 000,00
12	Склонность к покупке, max	доли ед.	0,90	<>	<>	<>
13	Доля конкурентов, min	%	-30,00	<>	<>	<>
14	Доля конкурентов, max	%	Гистограмма	<>	<>	<>
15	Цена конкурентов, min	Руб	2 300,00	<>	<>	<>
16	Цена конкурентов, max	Руб	2 500,00	<>	<>	<>
17	По договорам	Шт	15,00	<>	<>	<>
18	Цена	руб.	2 350,00	<>	<>	<>
19	Себестоимость	руб.	1 100,00	<>	<>	<>

Рис. 2.18. Выходная экранная форма процедуры расчета надежности реализации партии велосипедов при размере партии N = 50 шт.

Таблица 2.12

Результаты расчета надежности реализации партии велосипедов при размере партии N = 50 шт.

Параметр	Обозначение параметра в модуле «Equilibrium» и его адрес на листе «Варианты»	Размерность	Значение
Надежность 100%-ной реализации 50 шт. велосипедов по намеченной цене 2 350 руб.	Надежность по занижению [Q6:S6]	Доля ед.	0,88
Ожидаемая прибыль от продажи 50 шт. велосипедов	Прибыль [Q14:S14]	Руб.	62 500,01
Затраты на закупку 50 шт. велосипедов	Затраты [Q15:S15]	Руб.	55 000,01
Прирост рисков при отклонении от размера партии в 50 шт.	Завышение/Занижение [Q4:S4]	Доля ед.	0,1

Выводы

1. Продажа всей партии в 50 шт. велосипедов по намеченной цене 2350 руб. гарантируется с вероятностью 0,88. В соответствии с оценочной шкалой надежности степень надежности можно считать высокой.

2. Значение индикатора прироста рисков Z = 0,1 указывает на то, что уменьшение размера партии в окрестности N = 50 шт. создает существенно больший (крайне высокий) дополнительный риск, чем увеличение размера партии.

Диалоговая процедура 5.2:

1. Расчет =>Прямой/Обратный => Обратный по заданной надежности → ОК **2.** В поле открывшегося окна **«Обратный расчет»** ввести 0,52 → **ОК**

Результат

Экранная форма результата представлена на рис. 2.19, табличная форма — в табл. 2.13.

V	W	X	Y	Z	AA	AB
ФОРМА 3: Обратный расчет по заданной НАДЕЖНОСТИ:					**0,52**	
	ИСХОДНЫЕ ДАННЫЕ			РЕЗУЛЬТАТ РАСЧЕТА		
	ФАКТОРЫ И	Размер-		ПЛАН (НОРМАТИВ)	Размерность	Значение
№	ИСХОДНЫЕ ПОКАЗАТЕЛИ	ность	*Значение*	по завышению (1)	Шт	124,68
1	Количество семей, min	Семей	13 000,00	по занижению (2)	Шт	109,68
2	Количество семей, max	Семей	15 000,00	Зав/Зан (для 1)	Доли ед.	1,34
3	С детьми до 7 лет, min	%	20,00	Зав/Зан (для 2)	Доли ед.	0,65
4	С детьми до 7 лет, max	%	25,00	Прирост совокупного риска при	Абс. ед.	15 754,00
5	С средним и высоким доход	%	25,00	изменении норматива: 73,12057=>146,2411	%	-74,00
6	С средним и высоким доход	%	35,00			
7	Доход, min	руб/мес	35 000,00	Расчетные показатели		
8	Доход, max	руб/мес	50 000,00	Название	Размерность	Значение
9	Доля затрат на игрушки, min	%	13,00	Прибыль	руб.	155 850,00
10	Доля затрат на игрушки, max	%	15,00	Затраты	руб.	137 148,00
11	Склонность к покупке, min	доли ед.	0,70	<>	<>	<>
12	Склонность к покупке, max	доли ед.	0,90	<>	<>	<>
13	Доля конкурентов, min	%	-30,00	<>	<>	<>
14	Доля конкурентов, max	%	стограмма	<>	<>	<>
15	Цена конкурентов, min	Руб	2 300,00	<>	<>	<>
16	Цена конкурентов, max	Руб	2 500,00	<>	<>	<>
17	По договорам	Шт	15,00	<>	<>	<>
18	Цена	руб.	2 350,00	<>	<>	<>
19	Себестоимость	руб.	1 100,00	<>	<>	<>
20	<>	<>	<>	<>	<>	<>

Рис. 2.19. Выходная экранная форма расчета максимального размера партии с гарантированной надежностью продажи P = 0,52

Таблица 2.13

Результаты расчета максимального размера партии с гарантированной надежностью продажи *P* = 0,52

Параметр	Обозначение параметра в модуле «Equilibrium» и его адрес на листе «Варианты»	Размерность	Значение
Максимальный размер партии с гарантированной надежностью продажи P = 0,52	По занижению (2) [Z5:AB5]	Штуки	109,68
Ожидаемая прибыль от продажи максимальной партии велосипедов	Прибыль [Z13:AB13]	Руб.	155 850,00
Затраты на закупку максимальной партии велосипедов	Затраты [Z14:AB14]	Руб.	137 148,00
Прирост рисков при изменении максимального размера партии велосипедов	Завышение/ Занижение (для 2) [Z7:AB7]	Доля ед.	0,65

Выводы

1. При заданной надежности P = 0,52 по намеченной цене 2 350 руб. может быть продано со 100%-ной гарантией 110 шт. велосипедов, что более чем в два раза превышает размер предложенной партии.

2. Значение индикатора Z = 0,65 указывает на то, что уменьшение максимального размера закупочной партии (110 шт.) создает больший дополнительный риск, чем увеличение этого размера.

Этап 6. Анализ зависимости прибыли от цены товара

Цель. Построить график функции, выражающий зависимость прибыли от цены, и на его основе выявить диапазон цен, максимизирующих прибыль.

Комментарий. Поскольку цель намечаемой сделки — получение максимально возможной прибыли, необходимо найти цену или диапазон цен, максимизирующих прибыль фирмы при условии продажи всей партии в 50 шт. велосипедов с достаточно хорошей надежностью ($P \geq 0{,}52$).

При построении графика зависимости целесообразно варьировать цену в широком диапазоне, используя различные наценки на закупочную цену 1 100 руб., например, от 20 до 220%, что соответствует диапазону цен от 1 320 до 3 520 руб.

При построении графиков система предоставляет возможность выразить искомую зависимость как в процентной, так и в натуральной форме. Для этой цели служит диалоговое окно **«Вопрос»**, в котором система запрашивает пользователя, следует ли строить график зависимости в процентном выражении, и при выборе ответа **«Да»** строит таковой.

Для целей проводимого исследования нет необходимости выражать зависимость между анализируемыми параметрами в процентах, поэтому на данном и последующих этапах на запрос системы в диалоговом окне **«Вопрос»** выбирается ответ **«Нет»**.

Диалоговая процедура 6.1:

1. Расчет => **Зависимости** =>
2. В окне **«Вариант»** выбрать пункт *Фактор или показатель* → **ОК.**
3. Для заполнения поля открывшегося окна **«Построение зависимости»** в столбце **В** отметить показатель «Цена», установив курсор на ячейку **В 22 → ОК.**
4. В поле окна **«Минимальное значение»** ввести показатель «Цену» 1 320 → **ОК.**
5. В поле окна **«Максимальное значение»** ввести «Цену» 3 520 → **ОК.**
6. В поле окна **«Количество точек»** ввести **10 → ОК.**
6. В окне **«Выбор показателя»** выбрать пункт *«Прибыль, руб.»* → **ОК.**
7. В окне **«Построить тренд»** выбрать пункт *«Кубический»* → **ОК.**
8. В открывшемся окне **«Вопрос»** активизировать **«Нет»** → **ОК.**

Примечание. При расчете по индивидуальному варианту данных на шагах 4 и 5 вводятся значения $(p_1 + 0,2p_2)$ и $(p_2 + 2,2p_2)$ соответственно,
где p_2 — закупочная цена велосипеда.
При p_2 = 2200 руб. вводятся соответственно данные 2 640 руб. и 7 040 руб.

Результат

Экранная форма графика зависимости прибыли от цены представлена на рис. 2.20.

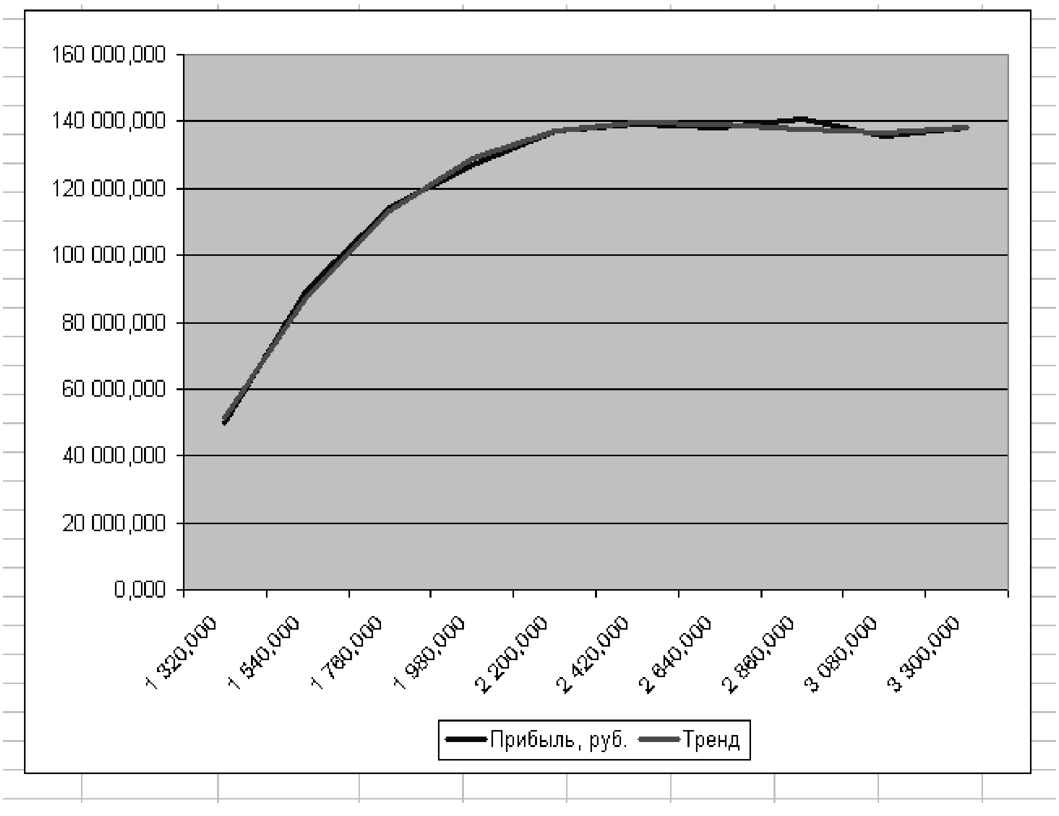


Рис. 2.20. График зависимости прибыли от цены

Выводы

1. График прибыли вначале быстро возрастает, а затем становится почти горизонтальным, то есть начиная с некоторого значения (≈ 140 000 руб.) прибыль не увеличивается и стабилизируется в некотором диапазоне цен.

2. Интервал цен, соответствующий диапазону стабилизации максимальной прибыли, составляет от ≈ 2 480 руб. до ≈ 3 000 руб.

3. При первоначально намеченной цене в 2 350 руб. прибыль составляет ≈ 138 тыс. руб., то есть не достигает максимума.

4. Наилучшим способом достижения максимальной прибыли является установление цены в 2 480 руб. Ее дальнейшее увеличение не приводит к увеличению прибыли, однако может снизить объем продаж. Следовательно, необходимо выявить зависимость объемов продаж (значит и планов $PL_{\text{опт}}$) от розничной цены велосипеда.

Этап 7. Анализ зависимости объемов продаж от цены товара

Цель. Построить график функции, выражающий зависимость объемов продаж от цены, и на его основе выявить объемы продаж, соответствующие ценовому диапазону, максимизирующему прибыль, а также установить цену, соответствующую размеру предлагаемой партии.

Комментарий. Установив на предыдущем этапе диапазон цен, максимизирующих прибыль (от ≈ 2 480 руб. до ≈ 3 000 руб.), по графику зависимости объема продаж от цены можно определить, каковы объемы продаж, соответствующие максимальной прибыли, и каков объем продаж, соответствующий скорректированной цене 2 480 руб.

Варьирование объема продаж равнозначно варьированию оптимальных планов, так как $PL_{\text{опт}}$ выбирается из множества возможных объемов продаж. Поэтому в диалоговой процедуре 7 при выборе зависимой величины в окне «**Выбор показателя**» выбирается пункт «**Оптимум (П)**» («П» обозначает план).

Диалоговая процедура 7.1:

1. Сервис => Тренд =>
2. В открывшемся окне **«Выбор показателя»** выбрать пункт *«Оптимум (П)»* **=> ОК**
3. В открывшемся окне «Построить тренд» выбрать пункт *«Кубический»* **=> ОК**
4. В открывшемся окне **«Вопрос»** активизировать **«Нет» => ОК**

Результат

Экранная форма графика зависимости объемов продаж от цены представлена на рис. 2.21.

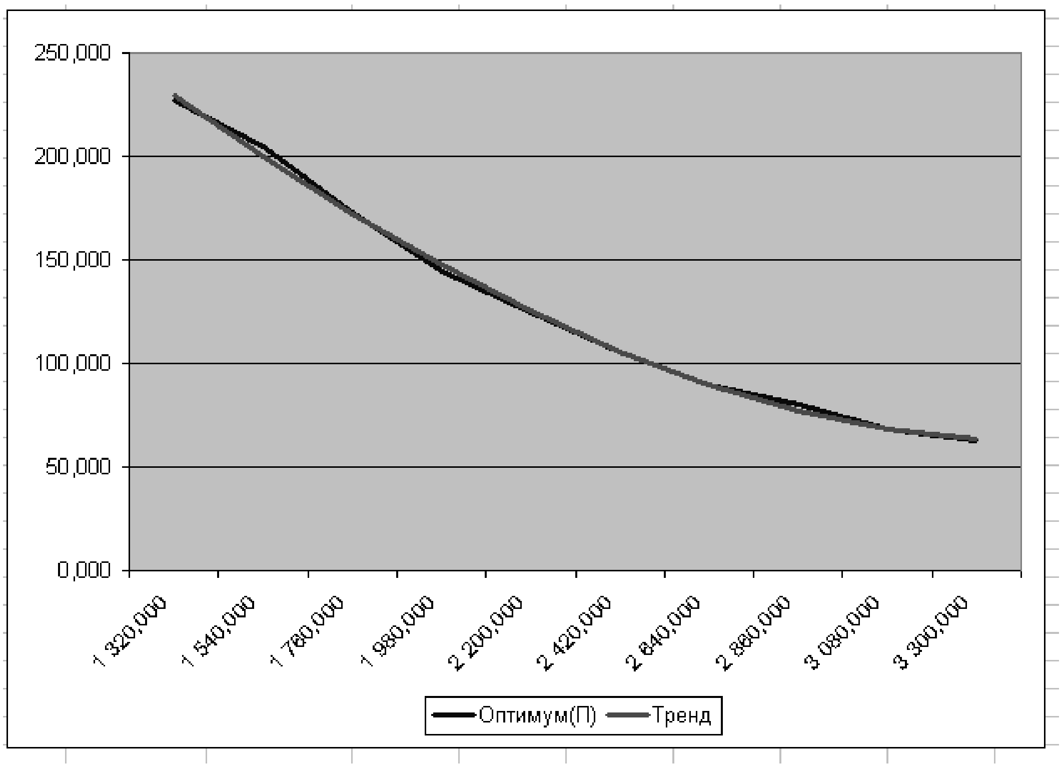


Рис. 2.21. График зависимости объемов продаж от цены

Выводы

1. Диапазон объемов продаж, соответствующих максимальной прибыли, составляет от ≈75 до ≈100 шт., что соответственно в 1,5—2 раза больше размера предложенной партии 50 шт.

2. По первоначально намеченной цене в 2 350 руб. можно продать ≈120 шт. велосипедов, по скорректированной цене в 2 480 руб. — 100 шт. Однако, согласно графику на рис. 2.20, в первом случае прибыль не достигает максимума, а во втором — будет максимальной.

3. Соотношение между прибылью, ценой и объемом продаж, полученная на этапах 6 и 7, представлена в табл. 2.14.

Таблица 2.14

Результаты расчетов зависимости между прибылью, ценой и объемом продаж

Прибыль, тыс.руб.	Цена, руб.	Объем продаж, шт. (план *PL*)
140	3000	75
140	2480	100
138	2350	120

4. Размер предложенной партии велосипедов (50 шт.) в два раза меньше возможного объема продаж по скорректированной цене 2 480 руб., поэтому 50 шт. велосипедов можно продать и по существенно более высокой цене. В соответствии с графиком на рис. 2.21 цену можно увеличить до 3 000 руб., однако при этом значительно возрастает риск, что часть партии может оказаться непроданной.

5. Увеличивая цену велосипеда, необходимо учитывать погрешность расчетов $PL_{опт}$ (то есть неопределенность плана). При проведенных на этапе 4 расчетах погрешность $PL_{опт}$ составила 17 шт. Следовательно, розничную цену велосипеда для партии 50 шт. необходимо определять, исходя из объема продаж в 67 шт. (с запасом в 17 шт.). Согласно графику на рис. 2.20, эта цена ≈ 3 100 руб. При этом прибыль составляет ≈ 136 000 руб.

6. Для выбора наиболее выгодной цены на основе соотношения «прибыль — цена — объем продаж» (см. табл. 2.14.) необходимо выявить зависимость надежности выполнения плана продаж от объема продаж.

Этап 8. Анализ зависимости надежности выполнения плана продаж от объема продаж

Цель. Построить график функции, отражающей влияние объемов продаж на надежность выполнения планов продаж с учетом установленного соотношения *«цена — объем продаж»*.

Комментарий. Вероятность 100%-ной реализации N шт. велосипедов по заданной цене можно определить, используя окно **«Обратный расчет по заданному плану»**, установив в ячейке D22 значение цены (поочередно 2 350, 2 480, 3 000, 3 100 руб.) и указывая в окне **«Обратный расчет»** соответствующий размер партии (120, 100, 75, 67 шт.).

Для цены 2 350 руб. надежность установлена на этапе 2 и составляет P = 0,54. Для остальных трех значений цены надежность устанавливается по следующей типовой процедуре, которую надо выполнить три раза при соответствующих значениях «Цена» и «Объем продаж N».

Типовая диалоговая процедура 8.1 (выполняется при различных значениях цены и объема продаж N):

1. В ячейку D22 листа «Варианты» ввести значение «*Цена*» → **ОК.** **2. Расчет =>Прямой/Обратный =>** **3. Обратный по заданному плану/Нормативу → ОК** **4.** В поле открывшегося окна **«Обратный расчет»** ввести **N → ОК**

Результат

Три экранные формы, аналогичные экранной форме на рис. 2.17 этапа 5.

После каждого выполнения типовой процедуры 8.1 рассчитанное значение надежности (столбец S Формы 2 значение «*Надежность по занижению*» для плана) заносится в табл. 2.15.

Таблица 2.15

Итоговая таблица аналитических расчетов

Прибыль, тыс.руб.	**Цена, руб.**	**Объем продаж, (план PL)**	**Надежность выполнения плана**
140	3 000	75	0,36
140	2 480	100	0,51
138	2 350	120	0,54
136	≈ 3100	67 (50)	0,41

Вывод

1. Учитывая, что надежность выполнения плана при цене велосипедов в 2 350 руб. и 2 480 руб. практически одинакова, а объем партии всего 50 шт., целесообразно выбрать цену максимизирующую прибыль, то есть 2 480 руб.

Этап 9. Анализ зависимости объемов продаж от нижней границы цены товара у конкурентов

Цель. Получить график зависимости объемов продаж от нижней границы цены велосипеда у конкурентов и на его основе оценить ожидаемые объемы продаж велосипедов в предположении, что конкуренты могут существенно снизить цену (до 1 320 руб.).

Комментарий. Оптимизационные расчеты плана проводились при условии, что нижняя граница цены велосипеда у конкурентов составляет 2 300 руб. (то есть при закупочной цене 1 100 руб. розничная наценка в 1 200 руб. составляет 110%). Необходимо выяснить, как изменится ситуация на рынке и шансы продать предложенную партию в 50 шт. велосипедов по скорректированной цене в 2 480 руб., если некоторые конкуренты существенно снизят цену. Диапазон снижения цен может составить от 2 300 руб. (минимальная цена у конкурентов в исходной ситуации со 110%-ной наценкой) до 1 320 руб. (что соответствует 20%-ной розничной наценке).

Диалоговая процедура 9.1:

1. В ячейку D22 листа «Варианты» ввести значение *«Цена»* (2 480) → **ОК.**
2. Расчет => Зависимости =>
3. В открывшемся окне «Варианты» выбрать пункт
«Фактор или показатель» → **ОК.**
4. Для заполнения поля открывшегося окна **«Построение зависимости»**
в столбце **В** отметить показатель ***«Цена конкурентов,min.»***, установив
курсор на ячейку **В19** → **ОК.**
5. В поле окна **«Минимальное значение»** ввести показатель «Цена» 1 320 → **ОК.**
6. В поле окна **«Максимальное значение»** ввести показатель «Цена» 2 300 → **ОК.**
7. В поле окна **«Количество точек»** ввести **10** → **ОК**.
8. В окне **«Выбор показателя»** выбрать пункт «Оптимум (П)» → **ОК.**
9. В окне **«Построить тренд»** выбрать пункт *«Квадратичный»* → **ОК**.
10. В открывшемся окне **«Вопрос»** активизировать **«Нет»** → **ОК**.

Примечание. На шаге 1 в ячейку D22 вводится розничная цена, выбранная на предыдущем этапе.

Результат

Экранная форма графика зависимости объемов продаж от нижней границы цены товара у конкурентов приведена на рис. 2.22.

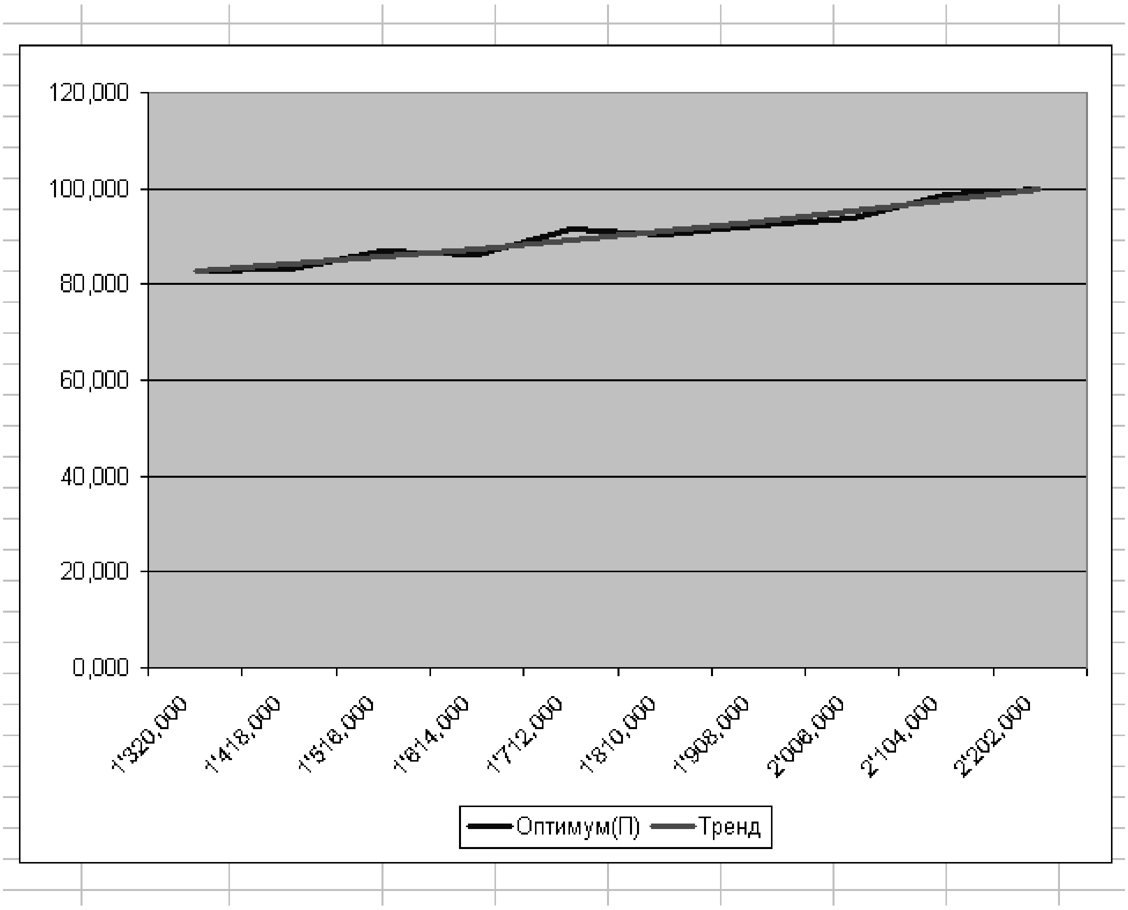


Рис. 2.22. График зависимости объема продаж от нижней границы цены велосипеда у конкурентов

Выводы

1. При снижении нижней границы цены велосипеда у конкурентов объемы продаж варьируется в пределах от ≈ 80 и до ≈ 100 шт. (при заданной доле конкурентов 93—100% и цене велосипеда 2 480 руб.).

2. При снижении цены у конкурентов до максимально низкой в 1 320 руб. объем продаж составит не менее 80 шт. Следовательно, есть основание полагать, что предложенная партия в 50 шт. велосипедов будет реализованной и в этой ситуации.

Этап 10. Контрольный расчет оптимального плана продаж для принятого решения о розничной цене товара

Цель. Рассчитать план $PL_{опт}$ для установленной розничной цены велосипеда и уточнить его значение путем сужения интервала неопределенности плана.

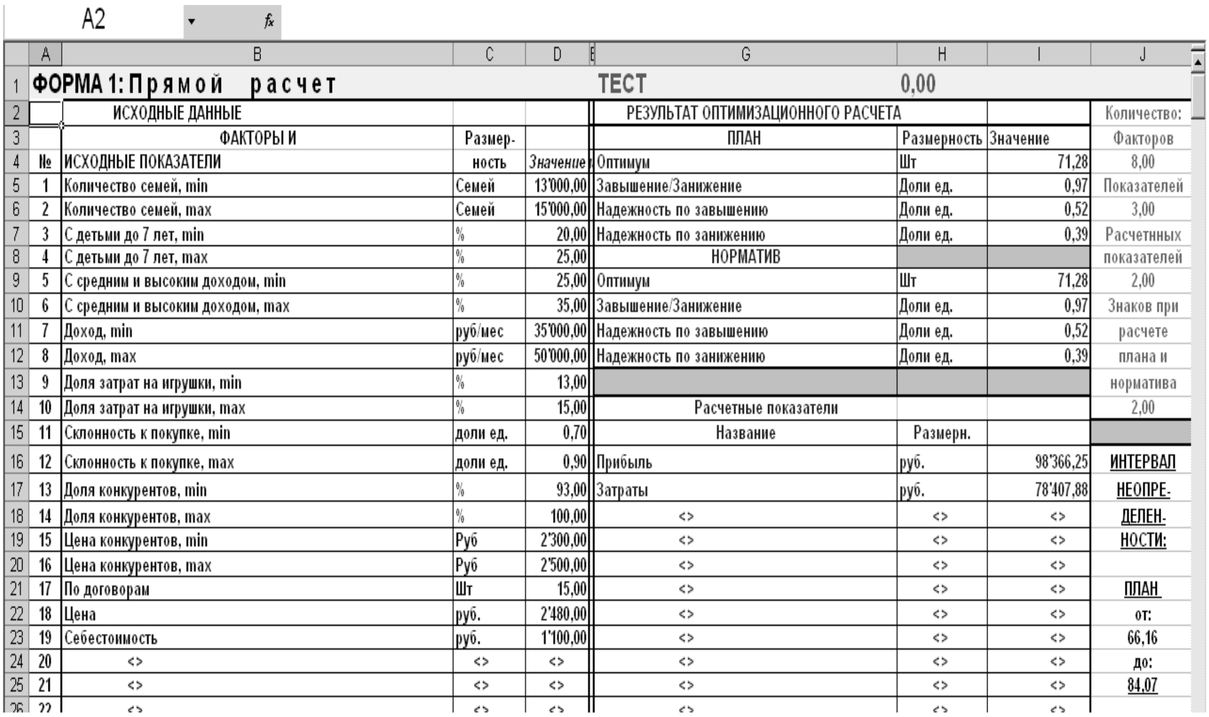
Комментарий. Для принятого решения о розничной цене велосипеда (2 480 руб.) необходимо произвести итоговый расчет оптимального плана $PL_{опт}$ и попытаться снизить его погрешность путем сужения интервалов неопределенности факторов. Для этой цели служит команда «*Значения факторов, совместные с оптимумом*», которая выполняется только после оптимизационного расчета. При обращении к данной команде система выполняет расчеты, в которых поэтапно сокращаются интервалы допустимых значений для факторов и вместе с тем последовательно уточняется значение $PL_{опт}$.

Диалоговая процедура 10.1:

1. Библиотека равновесных моделей => Тест =>
2. Пример заполнения данных (или индивидуальный вариант данных) → **ОК** → **ОК.**
3. В ячейку D22 листа **«Варианты»** ввести значение *«Цена»* (2 480) → **ОК.**
4. Расчет => Прямой/Обратный => Прямой расчет → ОК
5. При появлении окон-сообщений и окон с комментариями → **ОК**
Примечание. На шаге 3 в ячейку D22 вводится розничная цена, выбранная на предыдущем этапе из нескольких вариантов цен.

Результат

Экранная форма результата расчета представлена на рис. 2.23, табличная форма — табл. 2.16.



A2

	A	B	C	D	G	H	I	J
1	**ФОРМА 1: Прямой расчет**				**ТЕСТ**	0,00		
2		ИСХОДНЫЕ ДАННЫЕ			РЕЗУЛЬТАТ ОПТИМИЗАЦИОННОГО РАСЧЕТА			Количество:
3		ФАКТОРЫ И	Размер-		ПЛАН	Размерность	Значение	Факторов
4	№	ИСХОДНЫЕ ПОКАЗАТЕЛИ	ность	*Значение*	Оптимум	Шт	71,28	8,00
5	1	Количество семей, min	Семей	13'000,00	Завышение/Занижение	Доли ед.	0,97	Показателей
6	2	Количество семей, max	Семей	15'000,00	Надежность по завышению	Доли ед.	0,52	3,00
7	3	С детьми до 7 лет, min	%	20,00	Надежность по занижению	Доли ед.	0,39	Расчетнных
8	4	С детьми до 7 лет, max	%	25,00	НОРМАТИВ			показателей
9	5	С средним и высоким доходом, min	%	25,00	Оптимум	Шт	71,28	2,00
10	6	С средним и высоким доходом, max	%	35,00	Завышение/Занижение	Доли ед.	0,97	Знаков при
11	7	Доход, min	руб/мес	35'000,00	Надежность по завышению	Доли ед.	0,52	расчете
12	8	Доход, max	руб/мес	50'000,00	Надежность по занижению	Доли ед.	0,39	плана и
13	9	Доля затрат на игрушки, min	%	13,00				норматива
14	10	Доля затрат на игрушки, max	%	15,00	Расчетные показатели			2,00
15	11	Склонность к покупке, min	доли ед.	0,70	Название	Размерн.		
16	12	Склонность к покупке, max	доли ед.	0,90	Прибыль	руб.	98'366,25	ИНТЕРВАЛ
17	13	Доля конкурентов, min	%	93,00	Затраты	руб.	78'407,88	НЕОПРЕ-
18	14	Доля конкурентов, max	%	100,00	<>	<>	<>	ДЕЛЕН-
19	15	Цена конкурентов, min	Руб	2'300,00	<>	<>	<>	НОСТИ:
20	16	Цена конкурентов, max	Руб	2'500,00	<>	<>	<>	
21	17	По договорам	Шт	15,00	<>	<>	<>	ПЛАН
22	18	Цена	руб.	2'480,00	<>	<>	<>	от:
23	19	Себестоимость	руб.	1'100,00	<>	<>	<>	66,16
24	20	<>	<>	<>	<>	<>	<>	до:
25	21	<>	<>	<>	<>	<>	<>	84,07
26	22	<>	<>	<>	<>	<>	<>	

Рис. 2.23. Результат оптимизационного расчета плана для установленной розничной цены велосипеда (2 480 руб.)

Таблица 2.16

Результаты контрольного расчета плана

Параметр	Обозначение параметра в модуле «Equilibrium» и его адрес на листе «Варианты»	Размерность	Значение
Рассчитанный оптимальный план продаж $PL_{опт}$	Оптимум [G4:I4]	Штуки	71,28
Интервальная оценка $PL_{опт}$	Интервал неопределенности плана [J15:J25]	Штуки	[66,16;84,07]

Выводы

1. Оптимальный план продажи велосипедов $PL_{опт}$ по цене 2 480 руб. составляет 71 шт. Случайные колебания значений плана лежат в интервале от 66 шт. до 84 шт. Следовательно, даже нижняя граница интервала неопределенности подтверждает выгодность закупки предложенной партии велосипедов.

2. Погрешность в оценке $PL_{опт}$ составляет ≈ 25% (18/71 · 100). Так как она значительно превышает приемлемую 5%-ную погрешность, интервал неопределенности плана $PL_{опт}$ желательно сузить путем сужения интервалов неопределенности факторов.

Диалоговая процедура 10.2:

1. Сервис => Значения факторов, совместные с оптимумом → ОК **2.** В открывшемся окне «**Значение факторов»** выбрать команду «*Оптимизационный расчет*» → **ОК**. **3.** В открывшемся окне **«Вопрос»** активизировать **«Да»** на каждом из этапов сужения интервалов факторов до появления окна с сообщением «Результаты последней итерации и дальнейшее сокращение интервалов теряет содержательны смысл» → **ОК**. **4.** В открывшемся окне **«Вопрос»** с предложением системы восстановить исходные данные активизировать **«Нет»**

Результат

Сужение интервалов неопределенности факторов в процессе итерационной процедуры и соответствующие параметры расчета $PL_{опт}$ представлены в экранной форме на рис. 2.24. Цветом выделен результирующий столбец. Значения результирующей погрешности факторов представлены в табл. 2.17, приведены также исходные погрешности (см. табл. 2.5).

	A	B	C	D	E	F
1	Количество семей, min	Семей	13'000,00	13'000,00	13'294,42	13'294,4
2	Количество семей, max	Семей	15'000,00	13731,56	13'660,20	13'412,7
3	С детьми до 7 лет, min	%	20,00	22,76	23,81	23,8
4	С детьми до 7 лет, max	%	25,00	25,00	24,93	24,3
5	С средним и высоким доходом, min	%	25,00	25,27	26,23	27,0
6	С средним и высоким доходом, max	%	35,00	30,27	28,73	28,2
7	Доход, min	руб/мес	35'000,00	41'554,14	46791,28	47'579,5
8	Доход, max	руб/мес	50'000,00	49'054,14	49'054,14	48710,9
9	Доля затрат на игрушки, min	%	13,00	13,77	13,77	13,7
10	Доля затрат на игрушки, max	%	15,00	14,77	14,12	13,9
11	Склонность к покупке, min	доли ед.	0,70	0,70	0,73	0,7
12	Склонность к покупке, max	доли ед.	0,90	0,79	0,77	0,7
13	Доля конкурентов, min	%	93,00	94,73	96,13	96,9
14	Доля конкурентов, max	%	100,00	98,23	97,88	97,8
15	Цена конкурентов, min	Руб	2'300,00	2'300,00	2'312,48	2'327,0
16	Цена конкурентов, max	Руб	2'500,00	2'370,17	2'347,56	2'344,5
17	По договорам	Шт	15,00			
18	Цена	руб.	2'480,00			
19	Себестоимость	руб.	1'100,00			
20	ПЛАН	Размерность				
21	Оптимум	Шт	71,00	66,80	59,22	52,1
22	Завышение/Занижение	Доли ед.	0,91	1,02	0,93	1,1
23	Надежность по завышению	Доли ед.	0,53	0,59	0,73	0,9
24	Надежность по занижению	Доли ед.	0,41	0,32	0,23	0,0
25	НОРМАТИВ					
26	Оптимум	Шт	76,59	66,80	59,89	52,1
27	Завышение/Занижение	Доли ед.	1,13	1,02	1,10	1,1
28	Надежность по завышению	Доли ед.	0,49	0,59	0,71	0,9
29	Надежность по занижению	Доли ед.	0,36	0,32	0,21	0,0
30						
31	Расчетные показатели					
32	Название	Размерн.				
33	Прибыль	руб.	97'973,71	92'189,44	81728,18	72'026,7
34	Затраты	руб.	78'094,99	73'484,34	65'145,65	57'412,6

Рис. 2.24. Результаты итерационной процедуры снижения интервала неопределенности плана $PL_{опт}$

Выводы

1. В процессе выполнения итерационной процедуры происходит достаточно сильное сужение интервалов неопределенности факторов (табл. 2.17).

Таблица 2.17

Сравнительные значения исходной и результирующей погрешности факторов

Фактор	Размерность	Величина погрешности в оценке фактора	
		Исходная	Результирующая
Количество семей	Семья	2 000,00	365,78
С детьми до 7 лет	%	5,00	1,12
Со средним и высоким доходом	%	10,00	0,50
Доход	Руб./мес.	15 000,00	2 262,86
Доля затрат на игрушки	%	2,00	0,35
Склонность к покупке	Доля ед.	0,20	0,04
Доля конкурентов	%	7,00	1,75
Цена конкурентов	Руб.	200,00	35,08

2. Дальнейшее снижение погрешности теряет содержательный смысл, так как минимальные значения факторов могут превысить максимальные.

3. Сужение интервалов неопределенности факторов приводит к значительному снижению погрешности рассчитываемого плана, в результате чего значения $PL_{опт}$ последовательно снижаются с 71 шт. до 59 шт. велосипедов.

4. Уточненная оценка $PL_{опт}$ = 59 шт. превышает размер предложенной партии, что подтверждает выгодность закупки партии по цене 2 480 руб.

Итоговый вывод по результатам проведенного аналитического исследования:

Имеются все основания для приобретения предложенной партии велосипедов в 50 шт. с тем, чтобы продавать их по цене 2 480 руб. Кроме того, можно вести переговоры о закупке большей партии (до 100 шт.)

2.10. Организация выполнения и отчетность по практической работе

1. Практическая работа выполняется по вариантам, номер которого назначается преподавателем. Исходные данные вариантов для проведения оптимизационных расчетов плана продаж, представлены в табл. 2.18.

2. Практическая работа состоит из трех этапов — подготовительного, расчетного и заключительного (отчетного).

На подготовительном этапе выполняются следующие действия.

1) Создается персональная папка студента и в нее копируются два файла — результатный и отчетный. Стандартизованные форматы этих файлов размещены в директории, указанной в инструкции к работе.

В результатный файл в процессе выполнения работы будут заноситься полученные экранные формы, а также заполняться результирующие таблицы.

Отчетный файл используется на заключительном этапе при составлении отчета.

2) Осуществляется запуск системы **«Decision»** и модуля **«Equilibrium»** с моделью **«Тест»** и данными демонстрационного примера.

3) Студент подготавливается к работе в технологической среде модуля **«Equilibrium»** путем ознакомления с инструментальными средствами модуля в соответствии с разделом 2.8. «Инструментальные средства модуля **«Equilibrium»** системы **«Decision»** для задачи планирования».

На расчетном этапе производятся оптимизационные расчеты плана согласно технологии, изложенной в разделе 2.9. «Технология разработки плана продаж в среде модуля **«Equilibrium»** системы **«Decision»**. На данном этапе реализуются два вычислительных процесса:

• выполнение расчетов по исходным данным демонстрационного примера модели **«Тест»** на основе диалоговых процедур технологических этапов 1—10;

• выполнение расчетов по индивидуальным исходным данным согласно диалоговым процедурам этапов 1—10 с занесение результатов выполнения каждого этапа в результатный файл.

По окончанию расчетного этапа результатный и отчетный файлы копируются студентом из его персональной папки на съемный flesh-носитель для использования при подготовке отчета.

На заключительном этапе анализируются полученные результаты и составляется отчет.

1) Формат отчета стандартизован: для каждого из 10 этапов технологии разработки плана продаж в отчете должны быть представлены ***цель, комментарий, результаты и выводы*** (диалоговые процедуры в отчет не включаются).

2) Экранные формы, заполненные таблицы и выводы, соответствующие десяти этапам технологии оптимизационных расчетов и представленные в отчетном файле для демонстрационного примера, служат образцами, которыми должен руководствоваться студент при составлении отчета для своего варианта исходных данных (см. табл. 2.18).

Таблица 2.18

Исходные данные вариантов выполнения оптимизационных расчетов плана

Факторы и исходные показатели	**Размерность**	**Варианты**									
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Количество семей, min	Тыс. семей	13	13,1	13,2	13,3	13,4	13	13,6	13,7	13,8	14
Количество семей, мах	Тыс. семей	15	15,1	15,2	14,5	15,3	15,4	14,7	15,4	15,6	15,8
С детьми до 7 лет, min	%	20	20	22	23	24	24	20	24	12,4	12,5
С детьми до 7 лет, мах	%	25	26	25	25	25	26	27	28	28	20
Со средним и высоким доходом, min	%	50	52	54	56	58	50	50	50	66	68
Со средним и высоким доходом, max	%	70	70	70	70	70	60	62	64	70	70
Доход, min	Тыс. руб./мес.	35	36	37	38	39	40	41	42	43	44
Доход, max	Тыс. руб./мес.	50	50	50	50	50	55	56	58	59	58
Доля затрат на игрушки, min	%	13	13	13	13	13	13	13	13	13	13

Окончание табл. 2.18

Факторы и исходные показатели	Размерность	Варианты									
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Доля затрат на игрушки, max	%	15	15	15	15	15	15	15	15	15	15
Склонность к покупке, min	Доля ед.	0,7	0,7	0,7	0,7	0,7	0,7	0,7	0,7	0,7	0,7
Склонность к покупке, max	Доля ед.	0,9	0,9	0,9	0,9	0,9	0,9	0,9	0,9	0,9	0,9
Доля конкурентов, min	%	93	93	93	93	93	93	93	93	93	93
Доля конкурентов, max	%	100	100	100	100	100	100	100	100	100	100
Цена конкурентов, min	Руб.	4600	4600	4600	4600	4600	4600	4600	4600	4600	4600
Цена конкурентов, max	Руб.	5000	5000	5000	5000	5000	5000	5000	5000	5000	5000
По договорам	Штуки	15	15	15	15	15	15	15	15	15	15
Цена	Руб.	4700	4700	4700	4700	4700	4700	4700	4700	4700	4700
Себестоимость	Руб.	2200	2200	2200	2200	2200	2200	2200	2200	2200	2200
Количество семей, min	Тыс. семей	14,1	14,2	14,3	14,5	14,6	14,8	14,7	14,8	14,9	15
Количество семей, max	Тыс. семей	17	17,2	17,5	16,5	15,9	16	16,1	16,2	16,3	16,5
С детьми до 7 лет, min	%	13,0	13,1	13,2	13,3	13,4	13,2	13,5	13,4	13,2	13,2
С детьми до 7 лет, max	%	25	26	27	28	19	17	20	19	17	18
Со средним и высоким доходом, min	%	50	50	50	50	50	50	50	50	50	50
Со средним и высоким доходом, max	%	70	70	70	70	70	70	70	70	70	70
Доход, min	Тыс. руб./мес.	50	40	50	42	44	50	42	44	50	45
Доход, max	Тыс. руб./мес	60	55	65	58	56	66	58	56	66	59
Доля затрат на игрушки, min	%	13	13	13	13	13	13	13	13	13	13
Доля затрат на игрушки, max	%	15	15	15	15	15	15	15	15	15	15
Склонность к покупке, min	Доля ед.	0,7	0,7	0,7	0,7	0,7	0,7	0,7	0,7	0,7	0,7
Склонность к покупке, max	Доля ед.	0,9	0,9	0,9	0,9	0,9	0,9	0,9	0,9	0,9	0,9
Доля конкурентов, min	%	93	93	93	93	93	93	93	93	93	93
Доля конкурентов, max	%	100	100	100	100	100	100	100	100	100	100
Цена конкурентов, min	Руб.	4600	4600	4600	4600	4600	4600	4600	4600	4600	4600
Цена конкурентов, max	Руб.	5000	5000	5000	5000	5000	5000	5000	5000	5000	5000
По договорам	Штуки	15	15	15	15	15	15	15	15	15	15
Цена	Руб.	4700	4700	4700	4700	4700	4700	4700	4700	4700	4700
Себестоимость	Руб.	2200	2200	2200	2200	2200	2200	2200	2200	2200	2200

Литература к главе 2

1. **Лихтенштейн В.Е., Росс Г.В.** Информационные технологии в бизнесе. Практикум. — М.: Финансы и статистика, 2008.
2. **Лихтенштейн В.Е., Росс Г.В.** Информационные технологии в бизнесе. Применение системы Decision в микро- и макроэкономике. — М.: Финансы и статистика, 2008.
3. **Лихтенштейн В.Е., Росс Г.В.** Информационные технологии в бизнесе. Применение системы Decision в решении прикладных экономических задач. — М.: Финансы и статистика, 2009.
4. **Лихтенштейн В.Е., Росс Г.В.** Математическое доказательство необходимости перемен в экономике. [Электронный ресурс]. — URL: http://www.decision-online.ru/files/Mathematical_proof_of_the_need_changes_in_the_economy_ru.pdf (Дата обращения: 31-07-2016).
5. **Лихтенштейн В.Е., Росс Г.В.** Новые подходы в экономике. [Электронный ресурс] М.: Финансы и статистика, 2013. — URL: http://www.decision-online.ru/files/new_approaches_in_economics.pdf. (Дата обращения: 31.07.2016).
6. **Росс Г.В., Лихтенштейн В.Е., Бич М.Г.** Информационные технологии диагностики региональных экономических систем // Информатизация и связь. — 2014. — № 3. — С. 69—73.

Приложения

Приложение 1

Значения статистик Дарбина-Уотсона при 5%-ном уровне значимости*
(*n* – число наблюдений, *k* – число объясняющих переменных без учета постоянного члена; *d* – статистика Дарбина-Уотсона: d_L и $d_{U.}$)

n	k = 1		k = 2		k = 3		k = 4		k = 5		k = 6	
	d_1	d_2	d_1	d_2	d_1	d_2	d_1	d_2	d_1	d_2	d_1	d_2
6	0,610	1,400	—	—	—	—	—	—	—	—	—	—
7	0,700	1,356	0,467	1,896	—	—	—	—	—	—	—	—
8	0,763	1,332	0,559	1,777	0,368	2,287	—	—	—	—	—	—
9	0,824	1,320	0,629	1,699	0,455	2,128	0,296	2,588	—	—	—	—
10	0,879	1,320	0,697	1,641	0,525	2,016	0,376	2,414	0,243	2,822	—	—
11	0,927	1,324	0,658	1,604	0,595	1,928	0,444	2,283	0,320	2,645	0,203	3,005
12	0,971	1,331	0,812	1,579	0,658	1,864	0,512	2,177	0,379	2,506	0,268	2,832
13	1,010	1,340	0,861	1,562	0,715	1,816	0,574	2,094	0,445	2,390	0,328	2,692
14	1,045	1,350	0,905	1,551	0,767	1,779	0,632	2,030	0,505	2,296	0,389	2,572
15	1,077	1,361	0,946	1,543	0,814	1,750	0,685	1,977	0,562	2,220	0,447	2,472
16	1,106	1,371	0,982	1,539	0,857	1,728	0,734	1,935	0,615	2,157	0,502	2,388
17	1,133	1,381	1,015	1,536	0,897	1,710	0,779	1,900	0,664	2,104	0,554	2,318
18	1,158	1,391	1,046	1,535	0,933	1,696	0,820	1,872	0,710	2,060	0,603	2,257
19	1,180	1,401	1,074	1,536	0,967	1,685	0,859	1,849	0,752	2,023	0,649	2,206
20	1,201	1,411	1,100	1,537	0,998	1,676	0,984	1,828	0,792	1,991	0,692	2,162
21	1,222	1,420	1,125	1,538	1,026	1,669	0,927	1,812	0,829	1,964	0,732	2,124
22	1,239	1,429	1,147	1,541	1,053	1,664	0,958	1,797	0,863	1,940	0,769	2,090
23	1,257	1,437	1,168	1,543	1,078	1,660	0,986	1,785	0,895	1,920	0,804	2,061
24	1,273	1,446	1,188	1,546	1,101	1,656	1,013	1,775	0,925	1,902	0,837	2,035
25	1,288	1,454	1,206	1,550	1,123	1,654	1,038	1,767	0,953	1,886	0,868	2,012
26	1,302	1,461	1,224	1,553	1,143	1,652	1,062	1,759	0,979	1,873	0,897	1,992
27	1,316	1,469	1,240	1,556	1,162	1,651	1,084	1,753	1,004	1,861	0,925	1,974
28	1,328	1,476	1,255	1,560	1,181	1,650	1,104	1,747	1,028	1,850	0,951	1,958
29	1,341	1,483	1,270	1,563	1,198	1,650	1,124	1,743	1,050	1,841	0,975	1,944
30	1,352	1,489	1,284	1,567	1,214	1,650	1,143	1,739	1,071	1,833	0,998	1,931
31	1,363	1,496	1,297	1,570	1,229	1,650	1,160	1,735	1,090	1,825	1,020	1,920
32	1,373	1,502	1,309	1,574	1,244	1,650	1,177	1,732	1,109	1,819	1,041	1,909

Окончание табл.

n	**k = 1**		**k = 2**		**k = 3**		**k = 4**		**k = 5**		**k = 6**	
	d_1	**d_2**	**d_1**	**d_2**	**d_1**	**d_2**	**d_1**	**d_2**	**d_1**	**d_2**	**d_1**	**d_2**
33	1,383	1,508	1,321	1,577	1,258	1,651	1,193	1,730	1,127	1,813	1,061	1,900
34	1,393	1,514	1,333	1,580	1,271	1,652	1,028	1,728	1,144	1,808	1,080	1,891
35	1,402	1,519	1,343	1,584	1,283	1,653	1,222	1,726	1,160	1,803	1,097	1,884
36	1,411	1,525	1,354	1,587	1,295	1,654	1,236	1,724	1,175	1,799	1,114	1,877
37	1,419	1,530	1,364	1,590	1,307	1,655	1,249	1,723	1,190	1,795	1,131	1,870
38	1,427	1,535	1,373	1,594	1,318	1,656	1,261	1,722	1,204	1,792	1,146	1,864
39	1,435	1,540	1,382	1,597	1,328	1,658	1,273	1,722	1,218	1,789	1,161	1,859
40	1,442	1,544	1,391	1,600	1,338	1,659	1,285	1,721	1,230	1,786	1,175	1,854
45	1,475	1,566	1,430	1,615	1,383	1,666	1,336	1,720	1,287	1,776	1,238	1,835
50	1,503	1,585	1,462	1,628	1,421	1,674	1,378	1,721	1,335	1,771	1,291	1,822
55	1,528	1,601	1,490	1,641	1,452	1,681	1,414	1,724	1,374	1,768	1,334	1,814
60	1,549	1,616	1,514	1,652	1,480	1,689	1,444	1,727	1,408	1,767	1,372	1,808
65	1,567	1,629	1,536	1,662	1,503	1,696	1,471	1,731	1,438	1,767	1,404	1,805
70	1,583	1,641	1,554	1,672	1,525	1,703	1,494	1,735	1,464	1,768	1,433	1,802
75	1,598	1,652	1,571	1,680	1,543	1,709	1,515	1,739	1,487	1,770	1,458	1,801
80	1,611	1,662	1,586	1,688	1,560	1,715	1,534	1,743	1,507	1,772	1,480	1,801
85	1,624	1,671	1,600	1,696	1,575	1,721	1,550	1,747	1,525	1,774	1,500	1,801
90	1,635	1,679	1,612	1,703	1,589	1,726	1,566	1,751	1,542	1,776	1,518	1,801
95	1,645	1,687	1,623	1,709	1,602	1,732	1,579	1,755	1,557	1,778	1,535	1,802
100	1,654	1,694	1,634	1,715	1,613	1,736	1,592	1,758	1,571	1,780	1,550	1,803
150	1,720	1,746	1,706	1,760	1,693	1,774	1,679	1,788	1,665	1,802	1,651	1,817

* Источник: **Суслов В.И; Ибрагимов Н.М, Талышева Л.П.; Цыплаков А.А.** Эконометрия: Учебник. — Новосибирск: Из-во СО РАН, 2005. «—» обозначена невозможность получения оценки статистики *dw* из-за отсутствия автокорреляции остатков с помощью данного критерия при соответствующих комбинациях значений объема выборки *n* и числа объясняющих переменных *k* в модели.

Приложение 2

Пример расчета погрешности оценки ожидаемого объема продаж

Среди основных факторов, влияющих на объем продаж товара или услуги, как правило, выделяются следующие:

- количество потенциальных покупателей товара или услуги (физических и юридических лиц) на рассматриваемом секторе рынка (f_1);
- средний годовой доход потенциальных покупателей (f_2);
- доля дохода, которую предполагаемые клиенты готовы расходовать на покупку этого товара или услуги (f_3);
- доля затрат на товар или услугу, которую потенциальные клиенты готовы будут потратить на товар или услугу из всего комплекса товаров и услуг (f_4);
- доля рынка, занятая конкурентом (f_5).

При намечаемой средней цене товара (услуги) C ожидаемый объем продаж выражается формулой

$$F\alpha = f_1 f_2 f_3 f_4 (1 - f_5) / C. \tag{1}$$

Основным способом оценки факторов $f_1 - f_5$ являются экспертные оценки. В условиях недостаточности исходной информации, как правило, не представляется возможным оценить конкретные значения факторов, а только их предельные (минимальное и максимальное) значения.

Интервал неопределенности в оценке i-го фактора (погрешность фактора) выражается разностью между минимальным и максимальным значениями фактора (для i = 1, ... ,5)

$$\Delta f_i = f_i^{\max} - f_i^{\min} \tag{2}$$

Учитывая, что в оценке ожидаемого объема продаж $F\alpha$ погрешности факторов могут сочетаться любым произвольным образом (как наименее, так и наиболее удачным), погрешность $\Delta F\alpha$ может быть вычислена по формуле

$$\Delta F\alpha = F\alpha^{\max} - F\alpha^{\min}), \tag{3}$$

где $F\alpha^{\max} = f_1^{\max} f_2^{\max} f_3^{\max} f_4^{\max} (1 - f_5^{\min}) / C,$

$F\alpha^{\min} = f_1^{\min} f_2^{\min} f_3^{\min} f_4^{\min} (1 - f_5^{\max}) / C.$

Пусть экспертами даны оценки факторов спроса, представленные ниже:

Цена, руб. (*C*)	**Факторы спроса**									
	Количество клиентов, шт.		**Средний годовой доход, руб**		**Доля дохода на покупку товара или услугу, доля ед.**		**Доля затрат на предлагаемую услугу (в том числе), доля ед.**		**Доля рынка, занятая конкурентами, доля ед.**	
	min	*max*	*min*	*max*	*min*	*max*	*min*	*max*	*min*	*max*
30	10000	30000	5000	20000	0,1	0,35	0,3	0,75	0,50	1,00

Расчет по формуле (1) дает следующие значения $F\alpha^{\max}$ и $F\alpha^{\min}$):

$$F\alpha^{\max} = f_1^{\max} f_2^{\max} f_3^{\max} f_4^{\max} (1 - f_5^{\min}) / C = 30\ 000 \cdot 20\ 000 \cdot 0{,}35 \cdot 0{,}75\ (1 - 0{,}5) / 30 - 2\ 625\ 000,$$

$$F\alpha^{\min} = f_1^{\min} f_2^{\min} f_3^{\min} f_4^{\min} (1 - f_5^{\max}) / C = 10\ 000 \cdot 5\ 000 \cdot 0{,}1 \cdot 0{,}3\ (1 - 1) / 30 = 0.$$

Подстановка этих значений в (3) приводит к следующему значению $\Delta F\alpha$

$$\Delta F\alpha = F\alpha^{\max} - F\alpha^{\min} = 2\ 625\ 000 - 0 = 2\ 625\ 000.$$

Следовательно, при имеющихся предельных значениях факторов **ожидаемый план продаж колеблется от 0 до 2 625 тыс. шт.** Поскольку погрешность в абсолютном выражении (2 625 тыс. шт.) превосходит оцениваемую величину (1 млн шт.), **то полученный прогноз не имеет никакого значения и не дает возможности принять какое-либо обоснованное решение относительно плана продаж.**

Приложение 3

Результаты экспертизы факторов спроса на детские велосипеды

Фактор	Минимальное значение (min)	Максимальное значение (max)	Размерность	Комментарий
f_1 f_2 f_3 f_4	13000 20 25 35000	15000 25 35 50000	Семья % % Руб./мес.	Данные о количестве семей (f_1), доле, семей с детьми до 7 л. (f_2) доли семей со средним и высоким доходом (f_3), а также размере этого дохода (f_4) в той или иной форме публикуются в периодической печати и различных сводках статистических органов. Исходя из анализа этой информации, можно экспертно оценить предельные значения факторов
f_5 f_6	13 0,7	15 0,9	% Доли ед.	Для оценки доли дохода, которую семья готова тратить на игрушки (f_5), можно выполнить собственные ограниченные обследования (опросить друзей и знакомых) или воспользоваться специальными обследованиями, которые публикуют различные социологические службы. Аналогичным образом могут быть получены исходные данные для оценки склонности к покупке (f_6)
f_7	93	100	%	Поскольку на рынке нет дефицита детских велосипедов, то очевидно, что конкуренты занимают все 100% покупательского спроса. Оборот магазина составляет не более 7% суммарного оборота других магазинов аналогичного профиля в микрорайоне. Можно предположить, что при удачном стечении обстоятельств удастся занять до 7% спроса. От 93 до 100% рынка при всех обстоятельствах останется у конкурентов

Приложение 4

Анализ рыночной ситуации при разработке плана продаж детских велосипедов

Направление исследований		Комментарии
Товар	Новизна	Товар имеет новый дизайн и сделан из современных материалов
	Соответствие требованиям	Соответствует необходимым требованиям к товарам данной категории
	Соответствие запросам покупателей	Соответствует
	Необходимость модификации	Нет
Рынок	Географическое положение	Город *N* (районный центр)
	Емкость	500 шт./месяц в сумме у всех поставщиков (мнение специалистов магазина)
	Товарная и фирменная структуры	Имеются два аналога отечественного производства
	Конъюнктура на 6-18 месяцев	—
	Тенденции развития на 5, 10 и 15 лет	—
	Покупатели (3-4 характеристики)	Семьи с детьми до 7 лет
Покупатели	Способы использования товара	Приобретение для собственного ребенка, приобретение для подарка
	Мотивы приобретения товара	Физическое воспитание ребенка
	Факторы, задающие предпочтения	Качество товара, марка, цена
	Деление покупателей на категории	Семьи со средним доходом
	способ покупки (по категориям)	Покупка за наличные
	Неудовлетворенность нашим товаром	Товар примерно аналогичен товару у конкурентов
	Неудовлетворенность товаром конкурентов	—
	Влияние НТР на потребности покупателей	—
Конкуренты	Основные конкуренты (3-4 фирмы)	Имеются еще пять магазинов, торгующих детскими велосипедами
	Наиболее динамичные фирмы (2—3 фирмы)	Магазин «Центральный» ведет активную рекламную кампанию
	Торговые марки	—
	Преимущества товаров конкурентов	—
	Формы и методы сбыта	Продажа за наличные
	Ценовая политика	Получение максимальной прибыли

Приложение 5

Назначение моделей, реализованных в модуле «Equilibrium»

Модель	Комментарий
Товарные рынки и рынки услуг (Good)	Исследование емкости товарного рынка или рынка услуг, ожидаемого объема продаж, поиск оптимальной цены, прогноз прибыли и затрат, анализ тактики и стратегии конкурентной борьбы
Страхование рисков (Insurance)	Исследование кредитного рынка и процентной политики банка. Имеются два варианта модели: Анализ рынка банковских услуг и Ставка процента по депозиту
Банковские услуги (Bank)	Исследование рынка страховых услуг и оптимизация условий страхования
Ценные бумаги (Paper)	Прогнозирование эмитентом платежеспособного спроса на ценную бумагу; расчет целесообразного объема удерживаемых ценных бумаг
Спрос на гостинечные услуги (Hotel)	Анализ рынка непромышленного строительства (на примере анализа потребности в гостиничной площади определенной категории)
Запасы (Zapas)	Нормирование производственных запасов (по каждому маркосорторазмеру); расчет оптимальной нормы запаса фирмы, работающей по системе КАН-БАН
Равновесие на денежном рынке (Finance)	Исследование макроэкономического равновесия на денежном рынке (финансовый сектор экономики)
Макроэкономическое равновесие (Macro)	Анализ и прогнозирование макроэкономического равновесия и динамики в реальном секторе экономики
Количество АЗС в регионе (AZS)	Оптимизация дилерской или трейдерской сети компании, в частности сети АЗС
Байесовского подход (Bayes)	Решение задачи Байеса
Стохастическое программирование (Rand)	Решение задач статистической оптимизации
Расчет центра масс и решение интегральных уравнений (Centr)	Решение инженерно-технических и математических задач (расчет центра масс сложной динамической системы; расчет давления в сосудах с упругими стенками и случайным поступлением и истечением жидкости или газа); решение интегральных уравнений

Оглавление

Contents

Учебное издание

Орлова Ирина Владленовна,
Рытиков Сергей Александрович,
Щепетова Светлана Евгеньевна,
Росс Геннадий Викторович,
Бич Михаил Геннадиевич

ОСНОВЫ МАТЕМАТИЧЕСКОГО МОДЕЛИРОВАНИЯ СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ ПРОЦЕССОВ

Практикум. Часть 2

Учебное пособие

Редактирование Е.Б. Егоровой
Оформление обложки Т.А. Антоновой

Подписано в печать 28.11.2016. Формат 60×90 1/8.
Бумана офсетная. Гарнитура Petersburg
Усл.печ.л. 16,50. Уч.-изд.л. 7,74
Тираж 100 экз. Заказ № 1232

Финансовый университет
Ленинградский просп., 49, 125993 (ГСП-3), Москва
Отпечатано в Издательстве Финансового университета

Для заметок