Министерство образования и науки Российской Федерации
Ярославский государственный университет им. П. Г. Демидова

О. В. Зеткина

Эконометрика: основы математического моделирования социально-экономических процессов

(Эконометрические модели
в анализе социально-экономических процессов)

Учебное пособие

Рекомендовано
Научно-методическим советом университета
для студентов, обучающихся по направлениям
Экономика, Государственное и муниципальное управление

Ярославль
ЯрГУ
2013

УДК 330.43(075.8)
ББК У.в611я73
З-58

Рекомендовано
Редакционно-издательским советом университета
в качестве учебного издания. План 2013 года

Рецензенты:
Кузнецова С. А., кандидат экономических наук, доцент
кафедры экономики и математики
Тутаевского филиала РГАТУ имени П. А. Соловьева;
кафедра математики и естественно-научных дисциплин ЯфМЭСИ

Зеткина, Оксана Валерьевна.

З-58 Эконометрика: основы математического моделирования социально-экономических процессов (Эконометрические модели в анализе социально-экономических процессов): учебное пособие / О. В. Зеткина; Яросл. гос. ун-т им. П. Г. Демидова. – Ярославль : ЯрГУ, 2013. – 124 с.

ISBN 978-5-8397-0969-0

Современные экономические теории и исследования, опирающиеся в значительной степени на использование математических моделей и методов анализа, требуют от экономистов достаточно свободного владения математическим аппаратом изучения статистических данных, что значительно расширяет необходимые знания, навыки, компетенции выпускников экономических вузов. Данное пособие имеет широкое практическое применение и позволяет приобрести необходимые навыки в области обработки эмпирических данных.

Предназначено для студентов, обучающихся по направлениям 080100.62 Экономика (дисциплина «Эконометрика», цикл Б3), 081100.62 Государственное и муниципальное управление (дисциплина «Основы математического моделирования социально-экономических процессов», цикл Б2), очной, очно-заочной и заочной форм обучения.

УДК 330.43(075.8)
ББК У.в611я73

ISBN 978-5-8397-0969-0

Оглавление

Введение

Постоянно усложняющиеся экономические процессы потребовали широкого использования моделирования и количественного анализа эмпирических данных. Выделилось и сформировалось одно из направлений экономических исследований – эконометрика.

Она объединяет и использует достижения ряда наук для выявления количественных взаимосвязей в экономической жизни общества. Данная наука помогает уточнить наши представления об окружающей действительности и развить теоретические модели, т. е. с ее помощью происходит процесс познания мира. Эконометрика способствует обоснованному выбору оптимальных инструментов управления окружающей действительностью, достижению экономических и политических целей.

Как показывает опыт, именно эконометрика помогает студентам экономического вуза понять и уверенно использовать на практике категории теории вероятностей и математической статистики, ранее казавшиеся им оторванными от жизни. Современные экономические теории и исследования, опирающиеся в значительной степени на использование математических моделей и методов анализа, требуют от экономистов достаточно свободного владения математическим аппаратом изучения статистических данных. Эконометрика стала одним из базовых курсов в системе экономического образования. Настоящее пособие ориентировано на студентов экономического профиля. Оно также может быть полезно аспирантам и преподавателям экономических дисциплин, занимающихся статистическими методами анализа экономических процессов. Предполагается, что студенты, изучающие эконометрику, владеют материалом базовых курсов по микро- и макроэкономике, социально-экономической статистике, линейной алгебре, теории вероятностей и математической статистике.

В главе 1 даны определения эконометрики как отдельной научной дисциплины, рассмотрены ее задачи, предмет и методы исследования. Подробно проанализирована история создания эконометрики и этапы ее развития. Представлены основные этапы эконометрического моделирования в анализе социально-экономических процессов. В рамках каждого этапа указано, на что следует обратить особое внимание.

Глава 2 посвящена изучению классической модели регрессии и является основной в данной схеме изложения материала. Рассматриваются предпосылки классической линейной регрессионной модели, выполнимость которых обеспечивает получение качественных оценок параметров линейных уравнений регрессии на базе метода наименьших квадратов (МНК).

Опыт показывает, что многим начинающим изучение вводного курса эконометрики необходимо восстановить знания основных положений теории вероятностей и математической статистики, без которых невозможно понимание излагаемого материала. Именно на ликвидацию пробелов в этой области направлена глава 2.1 данного пособия. При этом особое внимание уделяется экономическим приложениям рассматриваемых понятий. В главе 2.2 изучаются условия существования классической модели регрессии. Анализируются прогнозные качества парной линейной регрессии. Приводится схема определения точности оценок коэффициентов. В главе 2.3 классифицируются виды нелинейных моделей и рассматриваются правила их линеаризации. Анализируется экономический смысл коэффициентов регрессии на основе примеров и заданий. Отмечается важность и критерии выбора адекватной формы эконометрической модели.

В главе 3 рассмотрены основные индикаторы качества модели: коэффициент детерминации и средняя ошибка аппроксимации. Описывается алгоритм оценки общего качества уравнения регрессии с помощью коэффициента детерминации. Он традиционно более популярен на практике, так как оценивает качество подгонки регрессии. Хотя повышение значения этого коэффициента не самоцель эконометрического исследования, многие исследователи по-прежнему придают ему повышенное значение. В пособии уделено внимание описанию последствий, к которым могут привести некоторые нарушения спецификации и условий классической модели. Значительное внимание в главе 3.2 уделено решению проблемы спецификации модели. Анализируются важность и критерии выбора адекватной формы эконометрической модели. Описываются виды и последствия ошибок спецификации – неправильного выбора регрессионной модели.

В главе 4 изучаются элементы теории принятия статистических решений. Рассмотрены правила проверки гипотез – мощного инструмента эконометрики. С учетом обычно низкого качества данных проверка гипотез часто остается наиболее убедительным способом получения важных для практического использования результатов. В главе представлена теория статистической проверки гипотез, приведены основы t- и F-тестов и описаны примеры их использования. Приводится схема оценки значимости коэффициентов регрессии. Рассмотрена проблема получения качественных статистических оценок параметров исследуемых величин, что является одной из фундаментальных предпосылок получения эконометрических моделей, максимально соответствующих реальности.

В конце каждой главы пособия приводятся вопросы и задания, позволяющие более глубоко разобраться и осмыслить пройденный материал, а также кратко описываются ключевые понятия и проблемы.

Построение эконометрических моделей при большом объеме исходных данных обусловливает существенный объем вычислений, требует использования вычислительной техники и соответствующего программного обеспечения. Универсальным инструментом для решения задач эконометрики является табличный процессор Excel. При его использовании многие студенты сталкиваются с существенными трудностями реализации расчетных соотношений. Это вызвано тем, что изучению практических основ эконометрики в учебной литературе уделяется крайне мало внимания, что затрудняет использование современных алгоритмов решения эконометрических задач на практике. Поэтому основной целью данного пособия является теоретическое изложение основных разделов эконометрики и представление практических разработок по решению основных задач регрессионного анализа.

В Приложение включены таблицы с исходными статистическими данными по личному располагаемому доходу и расходам на основные товары (услуги) за ряд лет, предназначенные для проведения лабораторных работ.

1. Эконометрика как наука

1.1. Определение эконометрики как научной дисциплины

Становление эконометрики как научной дисциплины представляет значительный интерес с точки зрения как определения объектов исследования, так и формирования набора методов. Сам термин «эконометрика» сформировался из двух частей: «эконо-» – от «экономика» и «-метрика» – от «измерение». Поэтому статистический анализ экономических данных называется эконометрикой, что буквально означает *«наука об экономических измерениях»*.

Термин «эконометрика» («эконометрия») был впервые введен в 1910 г. австро-венгерским бухгалтером П. Цьемпой. Он считал, что если к данным бухучета применить методы алгебры и геометрии, то будет получено новое, более глубокое представление о результатах хозяйственной деятельности. Это употребление термина, как и сама концепция, не прижилось, но название «эконометрика» оказалось весьма удачно для определения нового направления в экономической науке, которое выделилось в 1930 г.

Существуют различные определения эконометрики. Их можно дать на основе задач, опубликованных Эконометрическим обществом в первом выпуске своего журнала[1]:

- унифицирование теоретико-количественных и эмпирико-количественных подходов к решению экономических проблем;
- развитие методов конструктивного точного анализа этих проблем, подобных существующим в естественных науках.

В этом же легендарном номере журнала, давшем рождение новой научной дисциплине – эконометрике, один из ее создателей Р. Фриш подчеркивает, что ни один из следующих подходов, взятых в отдельности: статистика, экономическая теория,

---

[1] Frisch O. R. The Responsibility of the Econometrician // H Econometrica. 1946. Vol. 14, № 1. P. 1–4 (Arrow K. J. The Work of Ragan Frisch, Econometrician // Econometrica. Vol. 28, № 2. P. 175–192).

математика, – не представляет собой эконометрику. «Эконометрика – это не то же самое, что экономическая статистика. Она не идентична и тому, что мы называем экономической теорией, хотя значительная часть этой теории носит количественный характер. Эконометрика не является синонимом приложений математики к экономике. Как показывает опыт, каждая из трех отправных точек – статистика, экономическая теория и математика – необходимое, но не достаточное условие для понимания количественных соотношений в современной экономической жизни. Это – единство всех трех составляющих. И это единство образует эконометрику»[2]. Эконометрика объединяет и использует достижения всех перечисленных выше наук для выявления количественных взаимосвязей в экономической жизни общества.

Наиболее общим является определение: эконометрика – это наука, которая дает количественное выражение взаимосвязей экономических явлений и процессов. Эконометрическое исследование основывается на экономической теории и на фактах, относящихся к событиям, имевшим место в реальном экономическом мире. Э. Маленво придерживался широкого понимания, интерпретируя эконометрику как любое приложение математики или статистических методов к изучению экономических явлений[3].

Таким образом, эконометрика – это наука, связанная с эмпирическим выводом экономических законов. Мы используем данные или наблюдения для того, чтобы получить количественные зависимости для экономических соотношений, т. е. формулировать экономические модели. Но это – только малая часть задач, решаемых эконометрикой. Она также позволяет, основываясь на эмпирических данных, оценивать неизвестные величины (параметры) в этих моделях, делать прогнозы и давать рекомендации по экономической политике.

До 1930-х гг. сложились все предпосылки для выделения эконометрики в отдельную науку. Стало ясно, что для более глубокого понимания экономических процессов стоит использовать

---

[2] Frisch O. R. Op. cit.

[3] Маленво Э. Статистические методы эконометрии: пер. с фр. 1976. Вып. 2. 328 с.

в той или иной мере статистику и математику. Возникла необходимость появления новой науки со своим *объектом*, *предметом* и *методом исследования*, которая объединяет все исследования в этом направлении.

Объектом эконометрики выступают экономические процессы и явления, происходящие в экономической системе общества.

Предметом эконометрики является количественная оценка взаимосвязи между случайными событиями, признаками, факторами, переменными экономических объектов, проверка теоретических моделей реальных экономический процессов для получения прогнозов, имитации и анализа деятельности экономических. С этой целью исследуются массовые экономические явления и процессы. Эконометрика через математические и статистические методы анализирует экономические закономерности, доказанные экономической теорией.

Целью эконометрики является оценка точечных и интервальных прогнозов деятельности генеральной совокупности объектов экономической системы на основании расчетов по данным выборочной совокупности при реализации случайного процесса[4].

К основным *задачам* эконометрики можно отнести следующие:

- Построение эконометрических моделей, т. е. представление экономических моделей в математической форме, удобной для проведения эмпирического анализа. Данную проблему принято называть проблемой спецификации. Отметим, что зачастую она может быть решена несколькими способами. Подробнее об этом в гл. 3.2.

- Оценка параметров построенной модели, делающих выбранную модель наиболее адекватной реальным данным. Это так называемый этап параметризации/

- Проверка качества найденных параметров модели и самой модели в целом. Иногда этот этап анализа называют этапом верификации.

- Использование построенных моделей для объяснения поведения исследуемых экономических показателей, прогнозирования и предсказания, а также для осмысленного проведения экономической политики.

[4] Валентинов В. А. Эконометрика: учебник. М.: Дашков и К, 2006. С. 24–25.

Статистический подход к эконометрическим измерениям является доминирующим, что и отражено в *основных методах*.

Основные эконометрические методы:

1) сводка и группировка информации;

2) вариационный и дисперсионный анализ;

3) регрессионный и корреляционный анализ;

4) статистические уравнения зависимости;

5) статистические индексы и др.

Специальные эконометрические методы классифицируют по следующим признакам[5]:

По методам определения коэффициентов регрессии

1) метод наименьших квадратов;

2) двухшаговый метод наименьших квадратов;

3) обобщенный метод наименьших квадратов;

4) метод максимального правдоподобия;

5) метод наименьших отклонений и наименьших расстояний.

По методам построения деятельности экономического объекта

1) методы построения модели сверху вниз;

2) методы построения модели снизу вверх.

По методам выборки данных из генеральной совокупности

1) анализ всей генеральной совокупности;

2) выборочный метод анализа генеральной совокупности

По методам анализа выборочных данных: пространственных, временных, пространственно-временных

1) метод корреляционного анализа;

2) метод регрессионного анализа;

3) методы с участием лаговых переменных (объясняемых, зависимых переменных и остатков);

4) методы с участием фиктивных, инструментальных переменных;

5) методы линеаризации нелинейных регрессий;

6) методы устранения мультиколлинеарности факторов, гетероскедастичности и автокорреляции остатков;

7) методы выявления периодических составляющих временных рядов;

---

[5] Там же. С. 25–27.

8) методы сглаживания временных рядов;

9) методы адаптивного прогнозирования;

10) методы воспроизведения стационарных случайных процессов с помощью белого шума;

11) методы анализа системы одновременных уравнений;

12) методы связи между качественными переменными.

Принципы эконометрики

правильной постановки проблемы;

системной направленности эконометрических расчетов;

учета рыночной неопределенности;

улучшения имеющихся альтернатив и поиска новых.

Перечисленные выше возможности эконометрика реализует с помощью следующих инструментов:

количественной оценки экономических взаимосвязей;

статистического тестирования гипотез;

прогнозирования.

Рассмотрим примеры реализаций этих возможностей[6].

Количественные оценки экономических взаимосвязей.

К этим оценкам относят расчеты, позволяющие ответить на следующие и им подобные вопросы.

На сколько увеличивается объем продаж конкретного товара или услуги в результате роста расходов на рекламу на 1000 руб. в день?

Как изменится индекс потребительских цен, если денежный агрегат Ml возрастает на 1 млрд руб. при прочих равных условиях?

В каком интервале лежат значения эластичностей спроса на товар или услугу по цене и по доходу потребителя?

Чему равен мультипликатор государственных расходов?

Статистическое тестирование гипотез

С математической точки зрения эту сторону эконометрики можно охарактеризовать как попытку суждения о математических ожиданиях (теоретических значениях) случайных величин по ограниченной выборке наблюдений. Этот формальный аппарат позволяет, например, ответить на следующие вопросы:

[6] Аистов А. В., Максимов А. Г. Эконометрика шаг за шагом: учеб. пособие для вузов. М.: ГУВШЭ, 2006. С. 10–13.

Оказывает ли влияние рекламная компания определенной продукции на доходы от продажи этой продукции?

Зависит ли индекс потребительских цен от величины денежного агрегата Ml?

Следует ли повышать цену на продукцию для увеличения выручки фирмы?

Эффективны ли фискальная и монетарная политики правительства и Центробанка?

Влияет ли смена процессора новой модели компьютера на его рыночную цену?

Как зависит объем продаж от расположения товара на торговых рядах в супермаркете (высоты полки, на которой расположен товар; удаленности от начала торгового ряда и т. п.)?

Зависит ли объем продаж товара от упаковки?

Прогнозирование.

Исследователи иногда различают термины «прогноз» и «предсказание». При прогнозировании строится точечная или интервальная оценка экономического показателя при условии, что переменные, оказывающие влияние на искомую величину, принимают значения в пределах верхней и нижней границ имеющейся выборки наблюдений. Предсказание – это попытка построения оценки значения показателя за пределами интервала наблюдавшихся выборочных значений его аргументов. Ниже приведены некоторые примеры прогнозирования.

Предсказание выручки от продажи продукции фирмы в следующем месяце.

Предсказание доходности проекта.

Предсказание собираемости налогов после вступления в силу нового налогового законодательства.

Вычисление рыночной цены автомобиля с заданным набором потребительских и эксплуатационных качеств.

Прогноз вероятностей выбора определенной категории товара из предложенных альтернатив покупателем, имеющим определенные социальные и экономические характеристики.

В основе прогнозирования лежат три взаимодополняющих источника информации о будущем:

- оценка перспектив развития будущего состояния прогнозируемого явления на основе опыта, чаще всего при помощи аналогии с достаточно хорошо известными сходными явлениями и процессами;

- условное продолжение в будущее (экстраполяция) тенденций, закономерности развития которых в прошлом и настоящем обладают высокой степенью инертности;

- модель будущего состояния того или иного явления или процесса, построенная сообразно ожидаемым или желательным изменениям ряда условий, перспективы развития которых достаточно хорошо известны.

В соответствии с этим существуют три дополняющих друг друга способа разработки прогнозов:

- *анкетирование* (интервью, опрос) – метод изучения мнений населения, специалистов (экспертов) с целью упорядочить, объективизировать субъективные оценки прогнозного характера. Особенно большое значение имеют экспертные оценки; опросы населения в практике прогнозирования применяются сравнительно редко;

- *экстраполирование* и *интерполирование* – построение динамических рядов развития показателей прогнозируемого явления на протяжении периодов основания прогноза в прошлом и упреждения прогноза в будущем (ретроспекции и проспекции прогнозных разработок);

- *моделирование* – построение поисковых и нормативных моделей с учетом вероятного или желательного изменения прогнозируемого явления на период упреждения прогноза по имеющимся прямым или косвенным данным о масштабах и направлениях изменений.

Наиболее эффективная прогнозная модель – система уравнений. Однако имеют значение все возможные виды моделей в широком смысле этого термина: сценарии, имитации, графы, матрицы, подборки показателей, графические изображения и т. д.

Приведенное разделение способов прогнозирования условно, потому что на практике эти способы взаимно перекрещиваются и дополняют друг друга. Прогнозная оценка обязательно включает в себя элементы экстраполяции и моделирования. Процесс экстраполяции невозможен без элементов оценки и моделирования. Моделирование подразумевает предварительную оценку и экстраполирование.

Различают следующие основные понятия инструментария прогнозирования[7]:

- *Прием прогнозирования* – конкретная форма теоретического или практического подхода к разработке прогноза, одна или несколько математических или логических операций, направленных на получение конкретного результата в процессе разработки прогноза.

- *Процедура* – ряд приемов, обеспечивающих выполнение определенной совокупности операций.

- *Метод* – сложный прием, упорядоченная совокупность простых приемов, направленных на разработку прогноза в целом. *Методика* – упорядоченная совокупность приемов, процедур, операций, правил исследования на основе одного или чаще определенного сочетания нескольких методов. *Методология прогнозирования* – область знания о методах, способах, системах прогнозирования. *Способ прогнозирования* – получение и обработка информации о будущем на основе однородных методов разработки прогноза.

- *Система прогнозирования* (прогнозирующая система) – упорядоченная совокупность методик, технических средств, предназначенная для прогнозирования сложных явлений или процессов.

Опыт показывает, что ни один из названных способов и тем более методов, взятый сам по себе, не может обеспечить значительную степень достоверности точности горизонта прогноза. Но в определенных сочетаниях они оказываются в высокой степени эффективными.

Общая логическая последовательность важнейших операций разработки прогноза сводится к следующим основным этапам:

- *Предпрогнозная ориентация* (программа исследования). Уточнение задания на прогноз: характер, масштабы, объект, периоды основания и упреждения и т. д. Формулирование целей

---

[7] Арженовский С. В., Молчанов И. Н. Статистические методы прогнозирования: учеб. пособие. Ростов-н/Д., 2001. С. 8–10.

и задач, предмета, проблемы и рабочих гипотез, определение методов, структуры и организации исследования.

- *Построение исходной (базовой) модели* прогнозируемого объекта методами системного анализа. Для уточнения модели возможен опрос населения и экспертов.

- *Сбор данных* прогнозного фона методами, о которых говорилось выше.

- *Построение динамических рядов показателей* – основы, стержня будущих прогнозных моделей методами экстраполяции; возможно обобщение этого материала в виде прогнозных предмодельных сценариев.

- *Построение серии гипотетических (предварительных) поисковых моделей* прогнозируемого объекта методами поискового анализа профильных и фоновых показателей с конкретизацией минимального, максимального и наиболее вероятного значений.

- *Построение серии гипотетических нормативных моделей* прогнозируемого объекта методами нормативного анализа с конкретизацией значений абсолютного (т. е. не ограниченного рамками прогнозного фона) и относительного (т. е. привязанного к этим рамкам) оптимума по заранее определенным критериям сообразно заданным нормам, идеалам, целям.

- *Оценка достоверности и точности,* а также обоснованности прогноза (его верификация) посредством уточнения гипотетических моделей обычно методами опроса экспертов.

- *Выработка рекомендаций для вариантных решений в сфере управления* на основе сопоставления поисковых и нормативных моделей. Для уточнения рекомендаций возможно проведение опросов населения и экспертов. Иногда при этом строятся серии поствероятностных прогнозных моделей – сценариев с учетом возможных последствий реализации выработанных рекомендаций для их дальнейшего уточнения.

- *Экспертное обсуждение (экспертиза) прогноза* и научных рекомендаций, их доработка с учетом выявленных несовершенств и сдача заказчику.

- *Вновь предпрогнозная ориентация* на основе сопоставления материалов уже разработанного прогноза с новыми данными прогнозного фона и новый цикл исследования, ибо прогнозирование должно быть таким же непрерывным, как управление, повышению эффективности которого оно призвано служить.

Все вычисления, прогнозы и выводы имеют *вероятностный характер*, этому в эконометрике уделяется особое внимание.

Для понимания прикладного применения эконометрики для анализа социально-экономических процессов рассмотрим историю ее возникновения и развития.

29 декабря 1930 г. по инициативе И. Фишера, Р. Фриша, Я. Тинбергена, И. Шумпетера, О. Андерсона и других ученых было создано Эконометрическое общество. В 1933 г. Р. Фриш учредил журнал «Эконометрика», который и в настоящий момент имеет большое значение для развития эконометрики. В 1941 г. появляется первый учебник по новой научной дисциплине, написанный Я. Тинбергеном.

Первые попытки количественных исследований в экономике относятся к XVII в. Они были связаны с представителями нового направления в экономической теории – политической арифметики. В. Пэтти, Ч. Давенант, Г. Кинг использовали конкретные экономические данные в своих исследованиях, в первую очередь при вычислении национального дохода. Это направление побуждало к поиску экономических законов по аналогии с физическими, астрономическими и другими естественно-научными законами. При этом существование неопределенности в экономике еще не осознавалась.

Важным этапом возникновения эконометрики стало развитие статистической теории в трудах Ф. Гальтона, К. Пирсона, Ф. Еджворта. Эти ученые опередили первые применения парной корреляции. Дж. Е. Юл определял связь между уровнем бедности и формами помощи бедным. Г. Хукер, в свою очередь, измерял связь между уровнем женитьбы и благосостоянием, в котором использовалось несколько индикаторов благосостояния. Также он исследовал часовые ряды экономических переменных.

С 1830-х гг. наиболее развитые страны начали чувствовать непонятные с точки зрения экономической науки потрясения – упадок деловой активности, возникновение массовой безработицы. Быстрое промышленное развитие и урбанизация обнаружили огромный пласт нерешенных социальных проблем. Уже в конце XIX в. неоклассическая теория стала восприниматься как слишком удаленная от действительности. Теория могла стать убедительной в том случае, если бы она смогла объяснить изменения, которые происходят в экономике. Для ее практического приложения были нужны количественные выражения базовых экономических сроков.

В 1911 г. американский экономист Г. Мур в первом труде по эконометрике «Законы заработной платы: эссе по статистической экономике» показал, что с помощью сложных математических конструкций, основанных на фактических данных, можно разработать основу для социальной политики. В это же время итальянский экономист Р. Бенини впервые использовал множественную регрессию при оценке функции спроса.

Значительный вклад в становление эконометрики внесли исследования цикличности экономики. Первым цикличность экономики проанализировал К. Жугляр, установив 7–11-летние циклы инвестиций. Сразу после него С. Китчин обнаружил 3–5-летнюю периодичность обновления оборотных средств. С. Кузнец проследил 15–20-летние циклы в строительстве.

Выдающийся советский экономист Николай Дмитриевич Кондратьев установил «длинные волны» длительностью 45–60 лет. Он проанализировал некоторые макроэкономические показатели стран Западной Европы и США с 1790 по 1920 гг. Построив и сгладив графики, устранив краткосрочные колебания, он обнаружил, что значения этих показателей синхронно движутся в долгосрочном периоде. Кроме того, во время подъема длинной волны возрастало количество войн и восстаний и происходило вовлечение новых стран и регионов в мировую торговлю и в мировое разделение труда. На основании этих наблюдений ученый сделал долгосрочный прогноз до 2010 г., предсказав, в частности, Великую депрессию 1930-х гг.[8]

---

[8] ЭСКО: электронный журнал энергосервисной компании «Экологические системы». 2009. № 2, февраль. URL:http://esco.co.ua/journal/2009_2/art067.htm

Важным этапом формирования эконометрики стала разработка экономических барометров. Они основаны на идее, что существуют показатели, которые изменяются раньше других и потому могут служить сигналами изменений последних. Первым и самым известным стал Гарвардский барометр, который был создан в 1903 г. под руководством У. Персонса и В. Митчелла. Он состоял из кривых, характеризующих фондовый, товарный и денежный рынки. Каждая из этих кривых представляла среднюю арифметическую из входящих в нее нескольких показателей. Эти ряды предварительно обрабатывались путем исключения трендов, сезонных колебаний и приведения колебаний отдельных кривых к сравнительному масштабу. Успех использования Гарвардского барометра вызывал появление многих аналогичных барометров в других странах. Однако приблизительно в 1925 г. он потерял свою чувствительность. Его крах объясняется появлением мощного регулирующего фактора в экономике США. В этих условиях основным методом макроэкономического анализа становится метод построения межотраслевого баланса В. В. Леонтьева. В это же время начали строиться экономические модели, которые используют методы гармонического анализа. Эти методы были перенесены в экономику из астрономии, метеорологии и физики.

До 1970-х гг. эконометрика воспринималась как эмпирическая оценка моделей, созданных в рамках экономической теории. По мнению эконометристов того времени, статистические данные должны были защитить теорию от догматизма. При этом подавляющее большинство экономических моделей, построенных в этот период, были кейнсианскими. Но начиная с 1970-х гг. формальные методы стали использоваться при выборе причинности теоретических концепций. При этом эконометрикой стали активно пользоваться и монетаристы. Важным событием для развития эконометрики стало появление компьютеров. Благодаря им, мощное развитие получил статистический анализ числовых рядов. Исследователи Г. Бокс и Г. Дженкинс создали ARIMA-модель в 1970 г., а К. Симс и некоторые другие ученые – VAR-модели в начале 1980-х гг. Стимулировало эконометрические исследование и бурное развитие финансовых рынков и производных инструментов. Это привело лауреата Нобелевской премии по экономике за 1981 г. Дж. Тобина к разработке моделей с использованием цензируемых данных.

Большое влияние на современную эконометрику оказал норвежский экономист Трюгве Хаавельмо. Он показал, как можно использовать методы математической статистики для того, чтобы получать обоснованные выводы о сложных экономических взаимосвязях, исходя из случайной выборки эмпирических наблюдений. Эти методы можно использовать и для оценки соотношений, полученных на основе экономических теорий, и для проверки этих теорий. В 1989 г. ему присудили Нобелевскую премию по экономике «за прояснение вероятностных основ эконометрики и анализ одновременных экономических структур». Т. Хаавельмо рассматривал экономические ряды как реализацию случайных процессов. Главные проблемы, которые возникают при работе с такими данными, – это нестационарность и сильная волатильность. Если переменные нестационарные, то существует риск установить связь там, где ее нет. Вариантом решения этой проблемы является переход от уровня ряда к разницам рядов. Недостатком данного метода является сложность экономической интерпретации полученных результатов. Для решения этой проблемы английский ученый Клайв Гренджер в 1981 г. вввел концепцию коинтеграции – свойства нескольких *нестационарных* (интегрированных) временных рядов, заключающегося в существовании некоторой их *стационарной* линейной комбинации. Им была предложена модель коррекции отклонений, для которой он разработал методы оценивания параметров, обобщения и тестирования. Коинтеграция применяется в случае, если краткосрочная динамика отображает значительные дестабилизировавшие факторы, а долгосрочная стремится к экономическому равновесию. Модели, созданные К. Гренджером, в 1990 г. были обобщены С. Йохансеном для многомерного случая. В 2003 г. Гренджер совместно с американским экономистом Р. Инглом получили Нобелевскую премию. Р. Ингл, в свою очередь, известен как исследователь моделей с переменной во времени волатильностью (ARCH-модели), получивших широкое распространение на финансовых рынках. В современной экономике анализ таких моделей крайне актуален: каждый день решения миллионов инвесторов о покупке или продаже тех или иных финансовых инструментов влияют на движения рынка. Рыночные цены изменяются во времени. Волатильность показывает изменение доходности финансового инструмента и отражает уровень колебаний доходности, т. е. меру риска. Высокий уровень волатильности означает, что доходность изменяется в широком диапазоне. Низкий уровень волатильности актива говорит о том, что его доходность изменяется несущественно. Активы с более низким уровнем волатильности являются менее рискованными, чем активы с высоким уровнем волатильности[9].

Мирового признания в становлении эконометрики удостоены многие ученые, в число которых вошли представители ряда стран, включая Россию:

- в 1973 г. – Василий Васильевич Леонтьев, американский экономист российского происхождения, – за разработку метода прогнозного экономического анализа «затраты – выпуск»;

- в 1975 г. – Леонид Витальевич Канторович, советский экономист и математик, – за введение в экономическую науку моделей линейного программирования и разработку подходов к оптимизации использования ресурсов[10].

Сегодня в мире выпускается ряд научных журналов, освещающих современные проблемы эконометрики, в числе которых:

Journal of Econometrics (Швеция),

Econometric Reviews (США), Econometrica (США),

Sankhya Indian Journal of Statistics, Ser.D. Quantitative Economics (Индия),

Publications Econometriques (Франция),

Прикладная эконометрика (Россия).

Главной заслугой эконометрики стала разработка специальных эконометрических методов в конкретной области науки и практики. В настоящее время эконометрические методы исследования используются не только в экономике, но и в социологии, психологии, медицине, физике, химии и других науках.

---

[9] URL:http://www.egartech.ru/fields/derivatives/riskfactors

[10] URL:http://www.cpe.une.edu/projects/rlms

1.2. Этапы эконометрического моделирования социально-экономических процессов

Можно выделить следующие виды социально-экономических объектов прогнозирования[11]:

1) с полным обеспечением количественной информацией – для них имеется в наличии ретроспективная количественная информация в необходимом объеме;

2) с неполным обеспечением количественной информацией – для них имеющаяся в наличии ретроспективная информация допускает использование статистических методов, однако не обеспечивает на заданном времени упреждения заданную точность прогноза;

3) с наличием качественной ретроспективной информации – относительно их прошлого развития имеется только качественная информация и полностью отсутствует, либо очень ограничена количественная;

4) с полным отсутствием ретроспективной информации – это, как правило, несуществующие проектируемые объекты.

Статистические методы могут с уверенностью применяться для первого вида, с некоторым уменьшением точности прогноза – для второго вида. Для двух последних видов более эффективно применение экспертных методов.

Исследование и прогнозирование социально-экономических процессов включает следующие составляющие[12]:

- Построение эконометрических моделей, т. е. представление экономических моделей в математической форме, удобной для проведения эмпирического анализа. Данная проблема называется проблемой спецификации.

- Оценка параметров построенной модели, делающих выбранную модель наиболее адекватной реальным данным. Это этап параметризации.

- Проверка качества найденных параметров модели и самой модели в целом – этап верификации.

[11] Арженовский С. В., Молчанов И. Н. Указ. соч. С. 10–11.

[12] Доугерти К. Введение в эконометрику: пер. с англ. М.: ИНФРА-М, 1999. С. 30–35.

- Использование построенных моделей для объяснения поведения исследуемых экономических показателей, прогнозирования и предсказания, а также для осмысленного проведения экономической политики.

Порядок проведения эмпирического (эконометрического) исследования можно представить в виде блок-схемы (рис. 1.1)[13].

Рис. 1.1. Алгоритм эконометрического моделирования

Согласно приведенному алгоритму, эконометрическое исследование включает следующие 6 этапов[14]:

1. Постановочный. Формулировка проблемы (качественный анализ связей экономических переменных – выделение зависимых (y_i) и независимых переменных (x_{ik})).

---

[13] Аистов А. В., Максимов А. Г. Указ. соч. С. 13.

[14] Составлено автором на основе: Кремер Н. Ш., Путко Б. А. Эконометрика: учебник для вузов / под ред. проф. Н. Ш. Кремера. М.: ЮНИТИ-ДАНА, 2002. С. 21–22.

2. Априорный (доопытный). Анализ сущности изучаемого объекта, определение главных и второстепенных факторов, влияющих на проблему. Анализ проблемы с выявлением, какая переменная отражает количественно значение этой проблемы; определить, какие факторы влияют на зависимую переменную – результат.

3. Параметризация (спецификация). Осуществляется выбор общего вида модели, т. е. задается форма связи между зависимой переменной y_i и независимыми переменными – факторами x_{ik}. Определение такой математической функции, которая должна воспроизводить определенное количество закономерностей зависимой переменной. В итоге подбирается вид функции, адекватно описывающей влияние факторов на зависимую переменную. Практика решения проблемы спецификации рассмотрена в гл. 3.2.

4. Информационный. Получение данных, анализ их качества. Проводится визуальных анализ графиков для всех переменных и причинно-следственных связей для определения тенденций и вида зависимостей между переменными. Расчет матрицы парных коэффициентов корреляции и проверка наличия линейной связи между переменными.

5. Идентификация модели. Производится статистический анализ модели и оценка ее параметров. Известно несколько методов оценки параметров, наиболее популярным является метод наименьших квадратов, подробно рассмотренный в гл. 2.2.

6. Верификация. Осуществляется интерпретация результатов. Проводится проверка истинности (адекватности) модели. С этой целью анализируется, насколько удачно решены проблемы спецификации, идентификации и идентифицируемости модели, какова точность расчетов по этой модели и, в конечном счете, насколько соответствует построенная модель исследуемому реальному экономическому объекту или процессу.

Последовательность выполнения исследований проиллюстрируем следующим примером[15]. Необходимо проанализиро-

---

[15] Бородич С. А. Эконометрика: учеб. пособие. 3-е изд. Минск: Новое знание, 2006. С. 9–10.

вать зависимость спроса Q на некоторое (нормальное) благо от цены P этого блага. Согласно экономической теории с ростом цены объем (величина) спроса уменьшается. На основе этого утверждения на этапе спецификации формулируются несколько математических зависимостей. Например,

$$Q = \alpha + \beta P \quad , \beta < 0; \tag{1.1}$$

$$Q = \alpha P^{\beta} \quad , \beta < 0; \tag{1.2}$$

$$\ln Q = \alpha + \beta \ln P \quad , \beta < 0 \tag{1.3}$$

Выбор формы зависимости (математической модели) является важнейшей отправной точкой для дальнейшего анализа. Исследователь опирается на базовые положения экономической теории, знания о характере зависимостей на предыдущих этапах исследований, некоторые субъективные предположения. Любая из моделей является всего лишь упрощением реальности и всегда содержит определенную погрешность. Поэтому из всех предлагаемых моделей статистическими методами отбирается та, которая в наибольшей степени соответствует реальным эмпирическим данным и характеру зависимости.

Далее оцениваются параметры α, β выбранной зависимости (этап параметризации). Эта оценка осуществляется на основе имеющихся статистических данных. Поэтому вопрос точности статистической информации является одним из ключевых для построения работоспособной модели. Обычно для получения количественных оценок используются методы регрессионного анализа.

После этого проверяется качество найденных оценок, а также соответствие всей модели эмпирическим данным и теоретическим предпосылкам (этап верификации). Данный анализ в основном осуществляется по схеме проверки статистических гипотез. На этом этапе не только совершенствуется форма модели, но и уточняется состав ее объясняющих переменных (возможно, объем спроса на товар определяется не только его ценой, но и ценой на товары-заменители, располагаемым доходом и другими факторами). Если модель удовлетворяет всем необходимым требованиям качества, то она может быть использована либо для прогнозирования, либо для объяснения внутренних механизмов исследуемых процессов. Такая модель позволяет с определенной надежностью предсказывать среднее значение исследуемого экономического показателя (в нашем примере это Q) на основе прогнозируемых или фиксированных значений других показателей (P), предвидеть вероятности отклонений конкретных значений изучаемой величины от предсказываемых по модели. Она поможет определить, на какие факторы, в каком направлении и объеме следует воздействовать, чтобы значение исследуемого показателя лежало в определенных границах. Отметим, что, вскрывая механизмы и взаимосвязи изучаемых процессов, эконометрические модели не решают вопрос о причине этих взаимосвязей.

Постановка проблемы включает в себя следующие действия:

1. Формулируются предмет и цели исследования.

2. В рассматриваемой экономической системе выделяются структурные или функциональные элементы, соответствующие данной цели, выявляются наиболее важные качественные характеристики этих элементов.

3. Качественно описываются взаимосвязи между элементами модели.

Для спецификации модели вводятся символические обозначения для учитываемых характеристик экономического объекта и формализуются, насколько возможно, взаимосвязи между ними. Тем самым формулируется *математическая* модель.

При исследовании социально-экономических процессов широко используется построение моделей. Модель – это специально подобранный объект, который имеет с реальным объектом некоторые общие свойства, интересующие исследователя. Различают модели натуральные и знаковые. Натуральная модель – это реальный (физический, биологический, химический и др.) объект, характеристики которого изменяются по тем же законам, по которым изменяются показатели экономической системы. Знаковая модель состоит из графических объектов (схемы, графики, символы, формулы и т. д.), связанных определенными правилами и преобразованиями. Математическая (знаковая) модель составляется на языке математики с использованием математических законов и правил.

Цели и задачи построения модели:

- исследование и изучение на моделях экономических процессов и законов;
- предсказание последствий принимаемых решений;
- автоматизация расчетов в проектировании, прогнозировании, планировании, управлении, подготовке решений.

В качестве моделируемых целей и критериев субъектов экономики, например, для экономистов или менеджеров могут выступать:

- максимизация прибыли, рентабельность;
- снижение затрат;
- минимизация налогов;
- обеспечение устойчивости в нестабильной среде и др.

Разработка модели решения проблемы включает следующие этапы:

- определение объекта моделирования;
- изучение внешней среды объекта;
- характеристику системы управления объектом;
- детализацию описания подсистем и элементов модели.

В общем случае формально-логическая модель системы разрабатывается для получения некоторой новой информации о системе-оригинале с целью решения исходной проблемы (см. рис. 1.2). При решении экономических задач для этой цели строится некоторая экономико-математическая модель, анализ которой предполагает установление характерных свойств отдельных элементов этой модели. Такими элементами могут служить переменные, ограничения, целевая функция модели и множество допустимых наборов значений переменных. *Экономико-математическая* модель представляет собой описание экономического по сущности явления или процесса математическим языком.

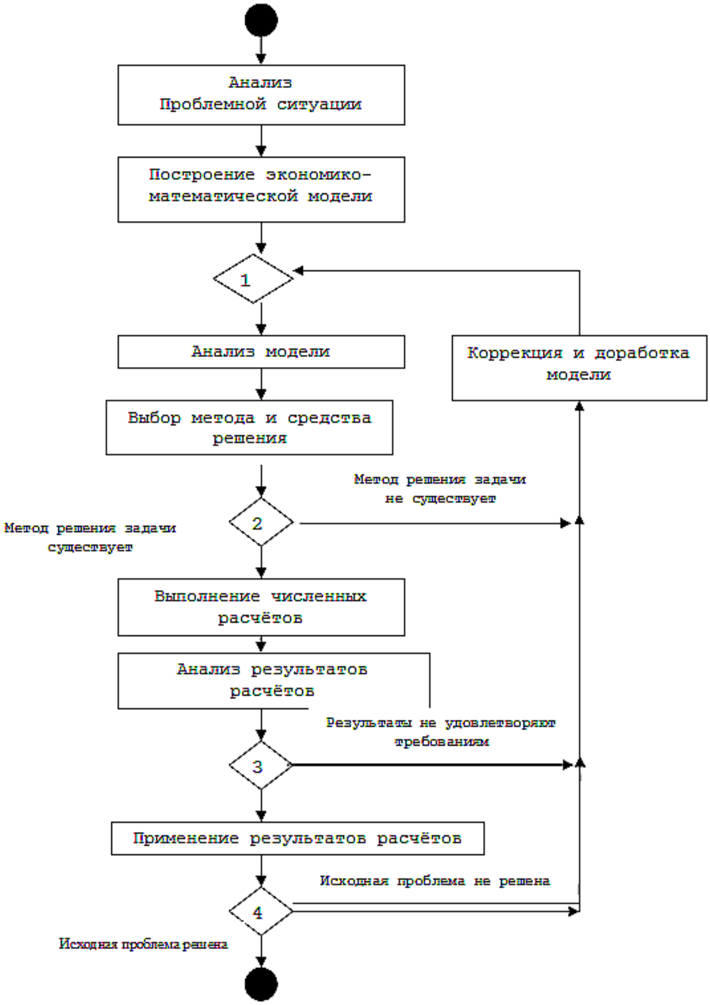


Рис. 1.2. Общая схема построения модели
для решения экономической задачи[16]:

1 – Логическая проверка возможности построения модели; 2 – Проверка существования решения задачи; 3 – Соответствуют ли результаты расчетов требованиям; 4 – Проверка возможности применения результатов расчетов для решения исходной проблемы

[16] Леоненков А. В. Решение задач оптимизации в среде MS Excel. СПб.: БХВ-Петербург, 2005. С. 32.

Эконометрической экономико-математическая модель становится в случае оценки параметров модели с помощью подходящего для стохастической связи метода (см. рис. 1.3).

Рис. 1.3. Последовательность эконометрических исследований[17]

Предлагаемая схема наглядно демонстрирует суть и последовательность эконометрических исследований, где отражается циклический характер современных экономических исследований: от экономической теории к моделированию, от моделирования к совершенствованию теории и более глубокому пониманию сути происходящих процессов; от понимания сути к осуществлению продуманной и целенаправленной экономической политики.

При оценке параметров модели определяются значения неизвестных одним из допустимых в каждом конкретном случае методов. Затем проверяется качество найденных оценок, т. е. вычисляется вероятность того, насколько верно определены параметры, их значимость и другие статистические характеристики, а также соответствие модели эмпирическим данным и теорети-

[17] Бородич С. А. Указ. соч. С. 10–11.

ческим предпосылкам. Данный анализ в основном осуществляется по схеме проверки статистических гипотез. На этом этапе совершенствуется не только форма модели, но и уточняется состав объясняющих ее переменных. Например, спрос на товар определяется не только его ценой, но и другими факторами, например располагаемым доходом. Именно такая зависимость исследуется студентами в рамках проведения лабораторных работ по дисциплинам «Эконометрика» и «Основы математического моделирования социально-экономических процессов».

Интерпретация результатов заключается в том, что проводятся расчеты по математической модели и анализ полученного решения. Если модель удовлетворяет требованиям качества, то она может быть использована для прогнозирования либо для анализа внутреннего механизма исследуемых процессов. Оцененная эконометрическая модель может использоваться как для структурного анализа, включая обратное влияние на экономическую теорию, так и для прогнозирования и связанной с ним выработки экономической политики.

Основным инструментом исследования в эконометрике является эконометрическая модель, имеющая основополагающую логическую структуру. Первоначально модель может быть сформулирована как в явном виде, так и в неформализованной форме – в форме наших представлений об окружающей реальности. Источниками формирования модели могут выступать опыт, интуиция, экономическая теория и другие факторы.

В экономике, как и в естественных науках, модель формализуется математическим языком в форме уравнений, связывающих экономические показатели: как отдельных уравнений, так и системы уравнений. В них могут входить как детерминированные величины и константы, так и случайные (в математическом смысле этого слова) переменные.

Эконометрика не тестирует тождества, являющиеся, по существу, определениями экономических показателей.

В соответствии с методами сбора информации источники данных для эконометрического исследования можно разделить на три типа[18]:

[18] Аистов А. В., Максимов А. Г. Указ. соч. С. 19.

1) эксперимент,

2) опрос,

3) наблюдение.

1. *Эксперимент* должен обладать свойством повторяемости, т. е. воспроизводимостью условий его проведения с целью наблюдения результатов желаемое число раз. Реализация этого свойства достаточно сложна и в большинстве случаев невозможна в экономике.

В качестве примера повторяющегося эксперимента можно привести испытания скольжения лыж, покрытых новым парафином, разработанным фабрикой спортивного снаряжения. Например, проводятся измерения коэффициента трения лыж о снег в разных погодных условиях, оценивается влияние различных улучшающих скольжение добавок на скорость движения лыж, на коэффициент сцепления парафина с поверхностью лыж и т. д.

Другой пример эксперимента – фирма-производитель поставляет на рынок товары с различными потребительскими свойствами мелкими экспериментальными партиями с целью выявления предпочтений потребителей для завоевания в дальнейшем конкурентного преимущества.

2. Опросы – более распространенная форма сбора информации для эмпирического исследования, которая широко используется в маркетинге, социологии, психологии. Опросы могут проводиться в форме единовременных акций, например, с целью выявления предпочтений потребителей или их отношения к тому или иному событию, а также в форме мониторинга. В последнем случае это регулярно, например, ежегодно проводимые опросы, по возможности одной и той же группы респондентов. Пример мониторинга – The Russia Longitudinal Monitoring Survey *(RLMS*). Это совместный проект Государственного учреждения института питания Российской академии медицинских наук, Университета штата Северная Каролина и Института социологии Российской академии наук[19].

3. Наблюдение – наиболее доступная форма сбора информации – предоставляет исследователю неэкспериментальные данные. К ним относятся наблюдения за ВВП страны, объемами продаж фирмы, котировками акций, биржевыми индексами и т. п.

[19] URL:http://www.cpe.une.edu/projects/rlms

С точки зрения методов оценивания моделей для эконометрики более важен не источник данных или способ их получения, а наличие определенной упорядоченности данных (наблюдений). По этому признаку различают:

1) временной срез,

2) временные ряды,

3) панельные данные.

1. Временной срез (cross section data) – наблюдения, выполненные в определенный достаточно короткий период времени, если измеряются потоки, или мгновенно – если измеряются запасы. Период времени измерений выбирается настолько малым, что в рамках используемой модели все наблюдения можно считать выполненными одновременно. Например, пытаясь оценить эффективность бренда, мы измеряем объем продаж товара за неделю в разных торговых точках и одновременно собираем информацию о ценах и неценовых детерминантах спроса и предложения.

2. Временной ряд – это совокупность численных значений отдельного экономического показателя, входящего в соответствующие модели. Он возникает, если аппарат эконометрики используется для выявления закономерностей изменения экономических показателей во времени, оценивания порядка предшествования событий и т. п., исследователь обязан проводить измерения входящих в модель величин в разные моменты времени. В таких случаях важен порядок следования наблюдений.

Например, управляющий кафе пытается проанализировать, выявить закономерности и построить прогноз объема продаж мороженого. С этой целью могут быть проведены измерения объемов продаж по месяцам за несколько лет для определения сезонных изменений. Для выявления неравномерности продаж по дням недели необходимы ежедневные наблюдения и т. д.

3. Панельные данные представляют собой регулярно повторяющиеся наблюдения за одними и теми же объектами (индивидуумами, фирмами, домохозяйствами и т. п.). Другими словами, панельные данные можно представить как упорядоченную во времени совокупность временных срезов. При создании математических моделей, использующих панельные данные, одного индекса нумерации значений переменных уже недостаточно. Необходимо вводить дополнительную размерность – переменные распределены как в пространстве, так и во времени. В некоторых социологических или маркетинговых исследованиях в результате проведения опроса (например, при выявлении предпочтений потребителей) отсутствует упорядочение наблюдений (временных срезов) во времени, но присутствует их упорядоченность по другой переменной, например по номеру вопроса. Такой набор данных также можно интерпретировать как панельные данные. Это позволяет использовать методы оценивания моделей, аналогичные используемым при анализе панельных данных. Иногда при обработке результатов опросов, проводимых дня выявления предпочтений потребителей, приходится нумеровать наблюдения не только по пространству (номер индивидуума) и времени (номер вопроса), но и по категории продукции, о которой высказывают свое мнение респонденты. Так возникает еще одна размерность данных (еще один индекс в модели).

В приведенных примерах порядок следования наблюдений по любой координате, отличающейся от времени, скорее всего, был не важен. Существует особая категория моделей, где порядок следования наблюдений по «невременной» координате принципиален и не может быть изменен без изменения смысла модели. Например, при проведении социологических опросов может наблюдаться коррелированность ответов на вопросы анкет у тех респондентов, которые находятся в схожих климатических, экономических или политических условиях. Методы оценивания моделей, учитывающих подобные эффекты, разрабатываются в рамках пространственной эконометрики.

В соответствии со схемой, приведенной на рис. 1.1, следующий шаг выполнения эмпирического исследования заключается в *оценке искомых параметров модели.* Под этим в первую очередь подразумевается оценка интересующих исследователя параметров, входящих в модель.

Существует большое количество методов оценивания параметров моделей. В общем случае выбор метода зависит от природы решаемой проблемы, количества и качества данных.

Следующий шаг проведения исследования – *тестирование гипотез* (см. рис. 1.1) – позволяет по полученным до этого точечным оценкам (выборочным значениям случайных величин) с определенной долей вероятности сделать вывод о генеральной совокупности, т. е. о тех закономерностях, на изучение которых направлено данное исследование. Эти суждения очень часто противоречат нашим априорным представлениям, на основе которых и была сформулирована первоначальная модель. Даже если это противоречие не обнаружено, необходимо выполнить тесты, позволяющие проверить соответствие основных предпосылок, заложенных в модели, тем наблюдениям, на основе которых выполнены оценки (и наоборот – соответствие наблюдений предпосылкам). Необходимо также проверить, допускают ли наблюдения использование выбранных нами методов оценивания параметров модели и тестирования гипотез.

В ходе эмпирического исследования необходимо добиться полного взаимного соответствия (не противоречия) наблюдений, предпосылок модели, методов оценивания параметров и тестирования гипотез. Для достижения этого соответствия допустим, а порой и необходим возврат на любой из предшествующих этапов исследования (см. рис. 1.1). Возможно, придется дополнить или уточнить наблюдения, изменить спецификацию модели или метод оценивания параметров, провести новые тесты и т. д.

Таким образом, первоначальная теоретическая модель часто является лишь основой (скелетом) будущего исследования.

Финальный этап эмпирического исследования – *интерпретация результатов*. Выводы могут не противоречить экономической теории или могут опровергнуть ее. В последнем случае следует уточнить теоретические предпосылки.

Практическое использование результатов исследования может быть самым разнообразным – начиная с построения прогнозов и оценок до выбора инструментов экономической политики фирмы и государства.

Необходимо подчеркнуть, что эмпирическое исследование положительно отвечает на следующие вопросы:

- Имеет ли модель экономический смысл?

- Учитывает ли модель все взаимодействия, которые порождают использованную реализацию наблюдений?

- Надежен ли источник данных и сами данные?

- Удовлетворяет ли цели исследования выбранный метод оценивания искомых параметров? В частности, обеспечивает ли он необходимую точность количественных оценок[20]?

- Удовлетворяют ли полученные результаты экономической теории и здравому смыслу?

Кроме того, необходимо сравнить результаты исследования с результатами, полученными из других источников, другими методами, на основе других моделей, другими авторами.

На практике при построении прогнозов социально-экономических явлений исследователь чаще всего имеет дело с исходными данными поперечного или продольного срезов. В первом случае он применяет регрессионные модели, во втором – модели временных рядов. Если же имеется недостаток количественной информации, то наиболее распространенными по применению являются экспертные методы.

Контрольные вопросы и задания

1. Дайте определение эконометрики на основе сложившихся подходов.

2. Назовите основные этапы становления эконометрического моделирования.

3. Проанализируйте алгоритм эконометрического исследования. Какие этапы на ваш взгляд являются наиболее существенными?

4. Перечислите источники данных.

5. Назовите основные типы данных для эконометрического моделирования.

[20] Иногда смещенная оценка с малой дисперсией может оказаться лучше несмещенной, но имеющей большую дисперсию.

2. Модели регрессии

2.1. Элементы теории вероятности

Случайной величиной называется величина, которая в результате наблюдения принимает то или иное значение, заранее неизвестное и зависящее от случайных обстоятельств[21]. Объем ВНП, количество реализованной продукции, прибыль фирмы, размер чистого экспорта за год и т. д. являются случайными величинами.

Различают *дискретные* и *непрерывные* случайные величины. Дискретной называют такую случайную величину (СВ), которая принимает отдельные, изолированные значения с определенными вероятностями (такая СВ имеет счетное количество значений). Непрерывной называют СВ, которая может принимать любое значение из некоторого конечного или бесконечного промежутка (т. е. число возможных значений непрерывной СВ бесконечно). Например, можно считать, что число покупателей в магазине, побывавших там в течение дня; число автомобилей, ремонтируемых еженедельно в данной мастерской; число находящихся в аэропорту самолетов являются дискретными СВ. Однако большинство СВ, рассматриваемых в экономике, имеют настолько большое число возможных значений, что их удобнее представлять в виде непрерывных СВ. Например, курсы валют, доход, объемы ВНП, ВВП и т. п. обычно рассматриваются как непрерывные СВ.

Для описания дискретной СВ необходимо установить соответствие между всеми возможными значениями СВ и их вероятностями. Такое соответствие называется *законом распределения* дискретной СВ. Его можно задать таблично, аналитически (в виде формулы) либо графически. При табличном задании закона распределения дискретной случайной величины X первая строка таблицы содержит ее возможные значения ($x_1, x_2, \dots, x_n$), а вторая – их вероятности ($p_1, p_2, \dots, p_n$). Как правило, $x_1 < x_2 < \dots < x_n$. Обязательно выполняется равенство:

$$p_1 + p_2 + \dots + p_к = 1. \qquad (2.1)$$

[21] Вентцель Е. С.Теория вероятностей: учебник для вузов. 9-е изд. М.: Академия, 2003. С. 25.

Пример 1.1[22]. На станции технического обслуживания анализируются затраты времени на ремонт автомобилей. На основании данных, полученных по 100 автомобилям, выяснилось, что для 25 из них требуется 1 ч для проведения профилактических работ. Мелкий ремонт требуется для 40 автомобилей, что занимает 2 ч. Для 20 автомобилей требуется ремонт с заменой отдельных узлов, что занимает в среднем 5 ч. 10 автомобилей могут быть отремонтированы за 10 ч. Для 5 автомобилей необходимое время ремонта составляет 20 ч. Построить закон распределения случайной величины X – времени обслуживания случайно выбранного автомобиля.

Решение задачи представим в виде таблицы.

Таблица 2.1

Распределение случайной величины

x_i	1	2	5	10	20
p_i	0,25	0,40	0,20	0,10	0,05

Аналитически случайная величина задается либо функцией распределения, либо плотностью вероятностей. *Функцией распределения* случайной величины X называют функцию $F(x)$, определяющую вероятность того, что случайная величина X принимает значение меньше, чем x, т. е.

$F(x) = P(X<x)$. (2.2)

Иногда эту функцию называют функцией накопленной вероятности или кумулятивной функцией распределения, что отражает ее суть.

Из определения вытекают *свойства функции распределения*:

1. $0 \le F(x) \le 1$;

2. $F(x)$ – неубывающая функция, т. е. $(x_1 < x_2) \Rightarrow F(x_1) < F(x_2)$;

3. $\lim\limits_{x \to -\infty} F(X) = 0$; $\lim\limits_{x \to +\infty} F(x) = 1$;

4. $P(a \le X \le b) = F(b) - F(a)$;

5. $P(X \ge x) = 1 - F(x)$;

[22] Бородич С. А. Указ. соч. С. 16.

6. Если возможные значения случайной величины *X* принадлежат отрезку [*а, в*], то

$$F(x) = \begin{cases} 0, если\, x < a; \\ 1, если\, x < b \end{cases} \qquad (2.3)$$

График функции распределения дает наглядное представление о вероятности изменения значений случайной величины. Для примера 2.1 функция распределения *F(x)* и ее график имеют вид[23]:

$$F(x) = \begin{cases} 0 \quad при\, x \le 1; \\ 0{,}25\; при\, 1 < x \le 2; \\ 0{,}65\; при\, 2 < x \le 5; \\ 0{,}85\; при\, 5 < x \le 10; \\ 0{,}95\; при\, 10 < x \le 20; \\ 1 \quad при\, x > 20 \end{cases}$$

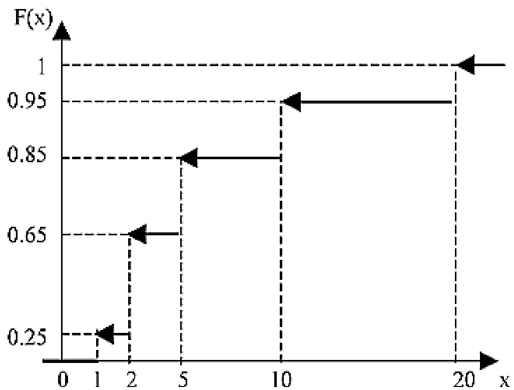


Рис. 2.1. График функции распределения случайной величины

Для непрерывной СВ нельзя определить вероятность того, что она примет некоторое конкретное значение (точечную вероятность). Так как в любом интервале содержится бесконечное число значений, то вероятность выпадения одного из них асимптотически равна нулю.

[23] Бородич С. А. Указ. соч. С. 17.

В результате непрерывную случайную величину нельзя задать таблично. Однако функция распределения может быть использована для описания непрерывной СВ. При этом она является непрерывной неубывающей функцией, изменяющейся от 0 до 1 (рис. 2.2).

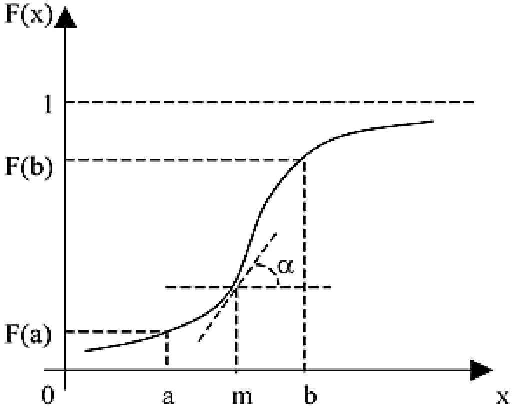


Рис. 2.2. График функции распределения

Плотностью вероятности (*плотностью распределения вероятностей*) непрерывной случайной величины X называют функцию

$$f(x) = \lim_{\Delta x \to 0} \frac{P(x \leq X < x + \Delta x)}{\Delta x} \qquad (2.4)$$

Плотность вероятности равна производной от функции распределения (поэтому иногда ее называют дифференциальной функцией распределения).

Свойства плотности вероятности:

1. $f(x) \geq 0$;

2. $P(a \leq X \leq b) = \int_a^b f(x)dx$;

Если $f(x)$ – плотность вероятности непрерывной случайной величины, то функция распределения $F(x) = \int_{-\infty}^{x} f(x)dx$;

$\int_{-\infty}^{+\infty} f(x)dx = 1$ – условие нормировки.

Для непрерывной случайной величины справедливы равенства:

$$P(a \le X \le b) = P(a < X < b) = P(a \le X < b) = P(a < X \le b) \qquad (2.5)$$

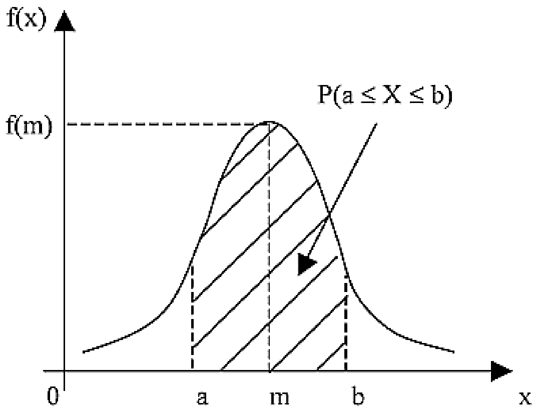


Рис. 2.3. Плотность вероятности
непрерывной случайной величины

Площадь под графиком кривой плотности вероятности *f(x)* равна единице. Площадь заштрихованной фигуры S равна:

$$S = \int_a^b f(x)dx = P(a \le X \le b); \qquad (2.6)$$

Вероятность попадания в «хвосты» распределения случайной величины равна

$$1 - P(\, a \le X \le в). \qquad (2.7)$$

Зная плотность вероятности *f(x)* непрерывной случайной величины *X*, можно определить вероятность ее попадания в заданный интервал, что имеет большое прикладное значение.

Во многих случаях удобнее пользоваться *числовыми характеристиками* случайной величины. Условно их подразделяют на:

- характеристики положения (математическое ожидание, мода, медиана, начальные моменты различных порядков);
- характеристики рассеивания (дисперсия, среднее квадратическое отклонение, центральные моменты различных порядков).

Наибольшее прикладное значение в моделировании получили математическое ожидание, дисперсия, среднее квадратическое отклонение.

Математическое ожидание характеризует среднее ожидаемое значение случайной величины и приближенно равно ее среднему значению. Например, при оценивании покупательной способности населения достаточно знание среднего дохода. При анализе выгодности двух видов деятельности можно ограничиться сравнением их средних прибыльностей. Знание того, что выпускники данного университета зарабатывают в среднем больше выпускников другого университета, может послужить основанием для принятия решения о поступлении в высшее учебное заведение и т. д.

Однако в ряде случаев недостаточно знание среднего значения. Существуют отличные друг от друга случайные величины, имеющие одинаковые математические ожидания. Например, средний уровень жизни в Швеции и США приблизительно одинаков, однако разброс в доходах в этих странах существенно отличается. Акции двух компаний могут приносить в среднем одинаковые дивиденды, однако вложение денег в одну из них может быть гораздо более рискованной операцией, чем в другую. Следовательно, требуется числовая характеристика, оценивающая разброс возможных значений случайной величины относительно ее среднего значения (математического ожидания). Такой характеристикой является дисперсия. *Дисперсией $D(X)$* случайной величины X называется математическое ожидание квадрата отклонения случайной величины от ее математического ожидания:

$$D(X)=M(X-M(X))^2=M(X^2)-M^2(X) \qquad (2.8)$$

Дисперсия имеет размерность, равную квадрату размерности случайной величины. С правилами расчета математического ожидания $M(X)$ и дисперсии $D(X)$ случайной величины студенты знакомы из курса теории вероятностей и математической статистики.

Для того чтобы представить разброс значений случайной величины в тех же единицах, что и саму случайную величину, вводится другая числовая характеристика – среднее квадратическое отклонение. Средним квадратическим отклонением *σ(X)* случайной величины *X* называют квадратный корень из дисперсии *D(X)*:

$$\sigma(x) = \sqrt{D(X)}$$

Для оценки разброса значений случайной величины в процентах относительно ее среднего значения вводится коэффициент вариации *V(X),* рассчитываемый по формуле:

$$V(X) = \frac{\sigma(X)}{|M(X)|} \cdot 100\% . \tag{2.9}$$

Меры разброса, такие как дисперсия, среднее квадратическое отклонение, коэффициент вариации, кроме оценивания рассеивания значений случайной величины, обычно применяются при изучении риска различных действий со случайным исходом, в частности при анализе риска инвестирования в ту или иную отрасль, при оценивании различных активов в портфеле и портфеля активов в целом в финансовом анализе и т. д.

Большинство случайных величин подчиняется определенному закону распределения, на основании знания которого можно предвидеть вероятности попадания исследуемой случайной величины в определенные интервалы. Такое предсказание весьма желательно при анализе экономических показателей, т. к. в этом случае появляется возможность осуществлять продуманную политику с учетом возможности возникновения той или иной ситуации. В эконометрическом анализе используются несколько законов распределения: нормальное, распределения Стьюдента, Фишера. Для удобства использования данных законов были разработаны таблицы так называемых критических точек, которые позволяют быстро и эффективно оценивать соответствующие вероятности.

Нормальное распределение (распределение Гаусса) является предельным случаем почти всех реальных распределений вероятности. Поэтому оно используется в очень большом числе реальных приложений теории вероятностей. Говорят, что случайная величина *x* имеет нормальное распределение, если ее плотность вероятности имеет вид:

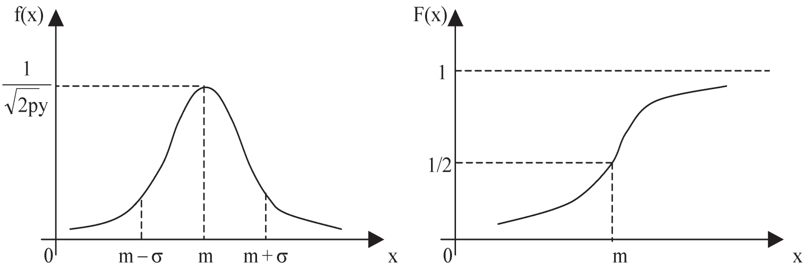
$$f(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{\frac{-(x-m)^2}{2\sigma^2}} \tag{2.10}$$

Это равносильно формуле:

$$F(x) = \frac{1}{\sqrt{2\pi\sigma}} \int_{-\infty}^{x} e^{-\frac{(t-m)^2}{2\sigma^2}} dt \tag{2.11}$$

Как видно из формул (2.10), (2.11), нормальное распределение зависит от параметров *m* и *σ* и полностью определяется ими. При этом *m = M(X)*, *σ= σ(x),* т. е. *D(x)*=σ^2, π=3,14159... , *e* = 2,71828... .

Если случайная величина *X* имеет нормальное распределение с параметрами *M(x)= m* и *σ(x) = σ*, то символически это можно записать так: *X ~ N(m, σ)* или $X \sim N(m,\sigma^2)$.



а *б*

Рис. 2.4. График функции вероятности (а)
и функции распределения (б) нормальной случайной величины

Важным частным случаем нормального распределения является ситуация, когда *m=0* и *σ=1.* В этом случае говорят о *стандартизированном (стандартном) нормальном распределении*. Стандартизированную нормальную случайную величину обозначают *U* (*U ~ N* (*0;1*)), учитывая при этом, что

$$f(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} \quad ; \quad F(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} e^{-\frac{t^2}{2}} dt \tag{2.12}$$

Для практических расчетов специально разработаны таблицы функций $f(u)$, $F(u)$ стандартизированного нормального распределения, однако чаще используется так называемая таблица значений функции Лапласа $\Phi(u)$.

Как видно из рис. 2.4, нормально распределенная случайная величина X ведет себя достаточно предсказуемо. График ее плотности вероятности симметричен относительно прямой $X = m$. Площадь фигуры под графиком плотности вероятности должна оставаться равной единице при любых значениях m и σ. Следовательно, чем меньше значение σ, тем более крутым является график. Кроме того, справедливы следующие соотношения:

$$P(\left|x - M(x)\right| < \sigma = 0,68\,;$$

$$P(\left|x - M(x)\right| < 2\sigma = 0,95\,;$$

$$P(\left|x - M(x)\right| < 3\sigma = 0,9973\,. \qquad (2.13)$$

В моделировании социально-экономических процессов значительное прикладное значение имеют следующие свойства:

- значения нормально распределенной случайной величины $X(X \sim N(m, \sigma))$ на 99,73 % сосредоточены в области $[m - 3\ \sigma, m + 3\ \sigma]$. Данный факт получил название правила трех сигма[24];

- линейная комбинация произвольного количества нормальных случайных величин имеет нормальное распределение.

2.2. Основные предпосылки классической модели

Основные идеи экономики – взаимосвязь между экономическими переменными. Спрос на товар на рынке рассматривается как функция его цены. Затраты на изготовление какого-либо продукта – функция от объема производства. Потребительские расходы – функция дохода и т. п. Это – примеры взаимосвязей между двумя переменными, одна из которых, например спрос на товар, производственные затраты, потребительские расходы, является объясняемой переменной (результирующим показателем), другая – объясняющей переменной (фактор – аргумент).

Как правило, в каждое такое соотношение приходится вводить несколько объясняющих переменных и остаточную случай-

[24] Вентцель Е. С. Указ. соч. С. 120.

ную составляющую, отражающую влияние на результирующий признак всех неучтенных факторов. Например, спрос на товар можно рассматривать как функцию его цены, потребительского дохода и цен на конкурирующие и дополнительные товары. Производственные затраты зависят от объема производства, его динамики, цен на основные производственные ресурсы. Потребительские расходы – функция дохода, ликвидных активов и предыдущего уровня потребления.

Участвующая в каждом из этих соотношений случайная составляющая, отражающая влияние на результирующий показатель всех неучтенных факторов, обусловливает *стохастический* характер зависимости. Даже фиксировав значения объясняющих переменных (например, цена на сам товар и на конкурирующие с ним или дополняющие товары, а также потребительский доход), мы не можем ожидать однозначно, каким будет спрос на данный товар. Иначе говоря, переходя в своих наблюдениях спроса от одного временного (или пространственного) промежутка к другому, мы увидим случайное варьирование спроса около некоторого определенного уровня даже при фиксировании всех объясняющих переменных.

Начальным пунктом эконометрического анализа зависимостей обычно является оценка *линейной* связи переменных, что объясняется простотой исследования. Поэтому проверка наличия такой зависимости, оценивание ее индикаторов и параметров является одним из важнейших направлений приложения математической статистики.

Наиболее простым для изучения является случай взаимосвязи двух переменных x и y – *парная модель* вида (2.14)

$$y_i = f(x_i) + u_i. \qquad (2.14)$$

Если это реальные статистические данные, то мы никогда не получим простую линию – линейную, квадратичную, экспоненциальную и т. д. Всегда будут присутствовать отклонения зависимой переменной, вызванные ошибками измерения, влиянием неучтенных величин или случайных факторов. Связь переменных, на которую накладываются воздействия случайных факторов, называется статистической связью. Наличие такой связи заключается в том, что изменение одной переменной приводит к изменению математического ожидания другой переменной[25].

Выделяют два типа взаимосвязей между переменными x и y:

переменные равноправны, т. е. может быть неизвестно, какая из двух переменных является независимой, а какая – зависимой;

две исследуемые переменные неравноправны, но одна из них рассматривается как объясняющая (или независимая), а другая как объясняемая (или зависящая от первой).

В первом случае говорят о статистической взаимосвязи *корреляционного* типа. При этом возникают проблемы оценки связи между переменными. Например, связь показателей безработицы и инфляции в данной стране за определенный период времени. Может стоять вопрос, связаны ли между собой эти показатели, и при положительном ответе на него встает задача нахождения формы связи. Вопрос о наличии связи между экономическими переменными сводится к определению конкретной формулы (спецификации) такой связи, устойчивой к изменению числа наблюдений. Для этого используются специальные статистические методы и, соответственно, показатели, значения которых определенным образом (и с определенной вероятностью) свидетельствуют о наличии или отсутствии линейной связи между переменными.

Во втором случае, когда изменение одной из переменных служит причиной для изменения другой, должно быть оценено уравнение регрессии. Для точного уравнения регрессии необходимо знать закон распределения результативного показателя y.

На практике приходится ограничиваться поиском подходящих аппроксимаций для неизвестной истинной функции регрессии $f(x)$, так как исследователь не знает точно условный закон распределения вероятностей анализируемого результирующего показателя y при заданных значениях аргумента x.

Наиболее распространенную формализацию стохастической зависимости между результирующим показателем y и объ-

[25] Шикин Е. В., Чхартишвили А. Г. Математические методы и модели в управлении: учеб. пособие. 3-е изд. М.: Дело, 2004. (Серия «Классический университетский учебник»). С. 285–286.

ясняющими переменными $x_1, x_2, \ldots , x_k$ представляет *аддитивная линейная форма*:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k + u \, , \tag{2.15}$$

где $\beta_0, \beta_1, \ldots, \beta_k$ – некоторые параметры, обычно неизвестные до проведения соответствующего статистического анализа, характеризующие разницу между модельным и наблюдаемым значениями анализируемой результирующей переменной; u – случайная величина, отражающая влияние неучтенных в модели факторов.

Для удобства интерпретации часто вместо β_0 используют обозначение α, называя, как и традиционно в математической модели, начальной составляющей или свободным членом.

$$y = \alpha + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k + u \, . \tag{2.16}$$

После выявления отдельных соотношений их группируют в модель. Математическая модель – упрощенное формализованное представление реальности (объекта, явления, процесса). При ее построении следует совмещать как можно большую лаконичность параметризации тех сторон моделируемой реальности, которые интересуют исследователя. Количество связей в модели зависит от условий, в которых она конструируется, от подробности объяснения, к которой исследователь стремится.

Например, рассмотрим модель спроса и предложения. Она должна объяснять соотношения между ценой и объемом выпуска, характерные для данного рынка. Модель содержит 3 уравнения:

1. Уравнение спроса;
2. Уравнение предложения;
3. Уравнение реакции рынка.

В них, помимо переменных объема выпуска и цены, входят потребительский доход (в уравнение спроса), цена (в уравнения спроса и предложения). Достигаемое при помощи модели объяснение обусловливается некоторыми «внешними» переменными. Поэтому данная модель неполная (условная).

Все экономические модели имеют общие особенности:

1. Они основаны на предположении, что поведение экономических переменных определяется с помощью совместных

и одновременных операций с некоторым числом экономических соотношений;

2. Принимается, что модель улавливает главные характеристики изучаемого объекта;

3. Полагается, что на основе достигнутого с ее помощью понимания реальной системы удастся предсказать будущее движение экономических показателей.

Наиболее популярным на практике методом исследования эконометрической модели является *регрессионный анализ*. Термин «регрессия» – движение назад, возвращение в прежнее состояние – был введен Фрэнсисом Галтоном в конце XIX в. при анализе зависимости между ростом родителей и ростом детей; в любом случае средний рост детей – и у низких, и у высоких родителей – стремится (возвращается) к среднему росту людей в данном регионе. В настоящее время под *регрессией* понимается зависимость между объясняющими переменными и условным математическим ожиданием (средним значением) зависимой переменной, которая строится с целью предсказания (прогнозирования) этого среднего значения при фиксированных значениях первых[26].

Функция $f(x_1, x_2, \ldots, x_k)$, описывающая зависимость среднего значения результативного признака y от заданных значений аргументов $x_1, x_2, \ldots, x_k$, называется *функцией* или *уравнением регрессии.* Уравнение регрессии – это формула статистической связи между переменными. Она называется *парной регрессией*, в случае двух переменных, если зависимость от нескольких переменных, – *множественной регрессией*. Например. Дж. Кейнсом была предложена линейная формула зависимости частного потребления C от располагаемого личного дохода:

$Y_d : C = C_0 + b\ Y_d$,

где $C_0 > 0$ – величина автономного потребления, $0 < b < 1$ – предельная склонность к потреблению.

Рассмотрим модель вида:

$$y = f(x_1, x_2, \ldots, x_k) + u \tag{2.17}$$

$$\hat{y} = f(x_1, x_2, \ldots, x_k) \tag{2.18}$$

[26] Бородич С. А. Указ. соч. С. 101–102.

Тогда для выборки из *n* наблюдений функция результативной переменной *y* примет общий вид и с учетом подстановок:

$$y_i = \hat{y}_i + u_i \ , \tag{2.19}$$

где $i = \overline{1,n}$.

С учетом выполненных подстановок получим:

$$\hat{y}_i = b_0 + b_1 x_{1_i} + b_2 x_{2_i} + \ldots + b_k x_{k_i} \tag{2.20}$$

или с учетом замены $a = b_0$:

$$\hat{y}_i = \alpha + b_1 x_{1_i} + b_2 x_{2_i} + \ldots + b_k x_{k_i} \tag{2.21}$$

Наиболее простым для изучения является случай взаимосвязи двух переменных *x* и *y*. Если это реальные статистические данные, то мы никогда не получим простую линию – линейную, квадратичную, экспоненциальную и т. д. Всегда будут присутствовать отклонения зависимой переменной, вызванные ошибками измерения, влиянием неучтенных величин или случайных факторов. Связь переменных, на которую накладываются воздействия случайных факторов, называется *статистической связью*. Наличие такой связи заключается в том, что изменение одной переменной приводит к изменению математического ожидания другой переменной.

Выделяют два типа взаимосвязей между переменными *x* и *y*:

переменные равноправны, т. е. может быть неизвестно, какая из двух переменных является независимой, а какая – зависимой;

две исследуемые переменные неравноправны, но одна из них рассматривается как объясняющая (или независимая), а другая как объясняемая (или зависящая от первой).

Выбор формулы связи переменных называется *спецификацией уравнения регрессии*. Решение проблемы спецификации подробно изучено в гл. 3.

Рассмотрим линейный регрессионный анализ, для которого функция $f(x_i)$ линейна относительно оцениваемых параметров:

$$M\left(y_i\right) = a + bx_i \, . \tag{2.23}$$

Предположим, что для оценки параметров линейной функции регрессии взята выборка, содержащая n пар значений переменных x_i,y_i , где $i = \overline{1,n}$.

В этом случае *линейная парная модель* имеет вид:

$$y_i = \alpha + \beta x_i + u_i \, . \tag{2.24}$$

где α, β – параметры, u_i – стохастическая случайная величина.

Линейная парная регрессионная модель имеет вид:

$$\hat{y}_i = a + bx_i \, . \qquad (2.25),$$

где a и $в$ – коэффициенты модели, подлежащие определению.

Необходимые условия для построения качественной модели регрессии получили название условий (предпосылок) Гаусса – Маркова и являются основными предпосылками регрессионного анализа. Пользуясь методом наименьших квадратов (OLS), необходимо помнить о том, что он даст BLUE – *оценки коэффициентов регрессии* (Best Linear Unbiased Estimators – наилучшие линейные несмещенные оценки), в смысле наименьшей их дисперсии, только при выполнении условий Гаусса – Маркова.

1. В модели (9) составляющая (возмущение) u_i – случайная величина, а объясняющая переменная x_i – величина неслучайная.

2. Математическое ожидание возмущения u_i равно нулю:

$$M(u_i) = 0 \tag{2.26}$$

Или математическое ожидание зависимой переменной y_i равно функции линейной регрессии:

$$M\left(y_i\right) = a + bx_i$$

Дисперсия возмущения u_i (или зависимой переменной y_i) постоянна для любого i.

$$D(u_i) = \sigma^2 \tag{2.27}$$

Или $D(u_i) = \sigma^2$ – условие гомоскедастичности или равноизменчивости возмущения (зависимой переменной).

3. Возмущения u_i и u_j (или переменные y_i и y_j не коррелированны):

$$M(u_i\, u_j) = 0 \; (i \neq j) \tag{2.28}$$

4. Возмущение u_i (или зависимая переменная y_i) нормально распределенная случайная величина.

При выполнении вышеперечисленных условий модель (2.25) называется *классической нормальной линейной регрессионной моделью*.

Для получения уравнения регрессии достаточно предпосылок 1–4. Они носят название *условий Гаусса – Маркова*. Требование предпосылки 5 необходимо для оценки точности уравнения регрессии и его параметров.

Теорема Гаусса – Маркова: если регрессионная модель вида (2.25) удовлетворяет условиям 1–4, то оценки a, b являются:

1) несмещенными, т. е.

$$M(a) = \alpha\ ,\ M(b) = \beta. \tag{2.29}$$

Это вытекает из того, что $M(u_i) = 0$, и говорит об отсутствии систематической ошибки в определении положения линии регрессии;

2) состоятельными, так как дисперсия оценок параметров при росте числа наблюдений стремится к нулю:

$$D(a) \to 0\ ,\ D(b) \to 0 \text{ при } n \to \infty \tag{2.30}$$

Другими словами, при увеличении объема выборки надежность оценок увеличивается;

3) эффективными, т. е. они имеют наименьшую дисперсию по сравнению с любыми другими оценками данных параметров, линейных относительно величин y_i .

Рассмотрим каждое свойство, изучив понятия несмещенности, состоятельности и эффективности случайной величины более подробно на основе случайной величины x.

Существует два варианта сведений о случайной величине:

- точная информация о рассматриваемой случайной переменной, в частности об ее распределении вероятностей (в случае дискретной переменной) или о функции плотности распределения (в случае непрерывной переменной). С помощью этой информации можно рассчитать теоретическое математическое ожидание, дисперсию и любые другие характеристики, дающие ее описание;

- оценки случайной величины, получаемые в случаях, когда мы не знаем точного вероятностного распределения или плотности распределения вероятностей. Это означает, что неизвестны также теоретическое математическое ожидание и дисперсия.

Алгоритм процедуры оценивания заключается в следующем:

- строится выборка из n наблюдений,
- с помощью соответствующей формулы рассчитывается оценка определенной характеристики.

Существует различие между способом или формулой оценивания и рассчитанным по ней для данной выборки числом, являющимся значением оценки. *Способ оценивания* – это общее правило, или формула, в то время как *значение оценки* – это конкретное число, которое меняется от выборки к выборке.

В табл. 2.2 приведены формулы оценивания для двух важнейших характеристик генеральной совокупности. *Выборочное среднее* $\overline{x}$ обычно дает оценку для математического ожидания, а формула s^2 – оценку дисперсии генеральной совокупности.

Таблица 2.2

Характеристики и формулы оценивания

Характеристики генеральной совокупности	*Формулы оценивания*
Среднее, μ	$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$
Дисперсия, σ^2	$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2$

Отметим, что это обычные формулы оценки математического ожидания и дисперсии генеральной совокупности, однако не единственные. Возможно, на практике привыкли использовать выборочное среднее $\overline{x}$ в качестве оценки для μ, что даже не рассматривают альтернативы. Конечно, не все формулы оценки, которые можно представить, одинаково хороши. Причина, по которой в действительности используется среднее $\overline{x}$, в том, что эта оценка в наилучшей степени соответствует двум очень важным критериям – несмещенности и эффективности.

Получаемая *оценка* представляет частный случай случайной переменной. Причина здесь в том, что сочетание значений x_i в выборке случайно, поскольку x – случайная переменная и, следовательно, случайной величиной является и функция набора ее значений. Возьмем, например, $\overline{x}$ – оценку математического ожидания:

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \ . \qquad (2.31)$$

Выше мы показали, что величина x_i может быть разложена на две составляющие – постоянную часть μ и чисто случайную составляющую u_i:

$$x_i = \mu + u_i \, . \qquad (2.32)$$

Следовательно,

$$\overline{x} = \overline{\mu} + \overline{u} \ , \qquad (2.33)$$

где $\overline{u}$ – выборочное среднее величин u_i.

Отметим, что $\overline{x}$, как и x, имеет как фиксированную, так и чисто случайную составляющие. Ее фиксированная составляющая – μ, т. е. математическое ожидание x, а ее случайная составляющая – μ, т. е. среднее значение чисто случайной составляющей в выборке.

Функции плотности вероятности для x и $\overline{x}$ показаны на одинаковых графиках (рис. 2.5). Величина x имеет нормальное распределение. Можно видеть, что распределения как x, так и $\overline{x}$ симметричны относительно μ – теоретического среднего. Разница между ними в том, что распределение $\overline{x}$ уже и выше. Величина $\overline{x}$, вероятно, должна быть ближе к μ, чем значение единичного наблюдения x, поскольку ее случайная составляющая $\overline{u}$ есть среднее от чисто случайных составляющих u_1, u_2, ..., u_n в выборке. Теоретическая дисперсия величины $\overline{u}$ составляет лишь часть теоретической дисперсии u.

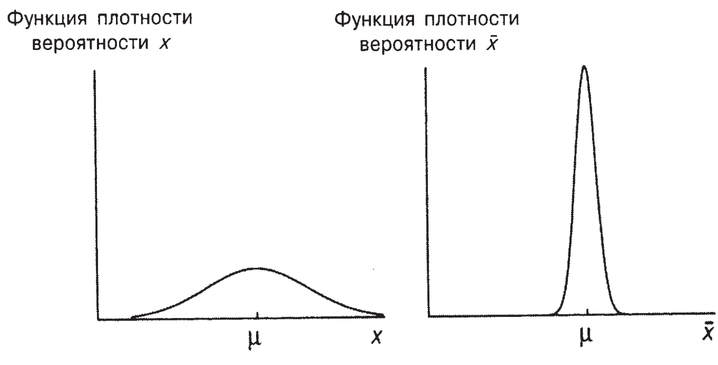


Рис. 2.5. Сравнение дисперсий случайных величин[27]

[27] Доугерти К. Указ. соч.

Величина s^2 – оценка теоретической дисперсии x – также является случайной переменной. Выполнив соответствующие преобразования, получим:

$$x_i - \bar{x} = u_i - \bar{u}. \tag{2.34}$$

Следовательно,

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 = \frac{1}{n-1}\sum_{i=1}^{n}(u_i - \bar{u})^2. \tag{2.35}$$

Таким образом, s^2 зависит от случайной составляющей наблюдений x в выборке; поскольку эти составляющие меняются от выборки к выборке, также от выборки к выборке меняется и величина оценки s^2.

Свойство *несмещенности* является одним из самых важных. Поскольку оценки являются случайными переменными, их значения лишь по случайному совпадению могут в точности равняться характеристикам генеральной совокупности. Обычно будет присутствовать определенная ошибка, которая может быть большой или малой, положительной или отрицательной, в зависимости от чисто случайных составляющих величин x в выборке.

Хотя это и неизбежно, на интуитивном уровне желательно, тем не менее, чтобы оценка в среднем за достаточно длительный период была аккуратной. Выражаясь формально, мы хотели бы, чтобы математическое ожидание оценки равнялось бы соответствующей характеристике генеральной совокупности. Если это так, то оценка называется *несмещенной*. Если это не так, то оценка называется *смещенной*, и разница между ее математическим ожиданием и соответствующей теоретической характеристикой генеральной совокупности называется *смещением*.

Выборочное среднее является несмещенной оценкой теоретического среднего, т. е. выполняется равенство:

$$M(x) = \mu. \tag{2.36}$$

Число несмещенных оценок бесконечно, однако на практике используют выборочное среднее.

Величина s^2, определяемая в соответствии с табл. 2.2, является оценкой теоретической дисперсии σ^2. Математическое ожидание s^2 равно σ^2, и эта величина является несмещенной оценкой теоретической дисперсии, если наблюдения в выборке независимы друг от друга.

Важным качеством оценок является их надежность. Оценка с максимально возможной вероятностью должна иметь близкое значение к теоретической характеристике, что означает желание получить функцию плотности вероятности, как можно более «сжатую» вокруг истинного значения. Один из способов выразить это требование – сказать, что мы хотели бы получить сколь возможно малую дисперсию.

Предположим, что мы имеем две оценки теоретического среднего, рассчитанные на основе одной и той же информации, что обе они являются несмещенными и что их функции плотности вероятности показаны на рис. 2.6. Поскольку функция плотности вероятности для оценки *B* более «сжата», чем для оценки *A*, с ее помощью мы скорее получим более точное значение. Формально говоря, эта оценка более эффективна.

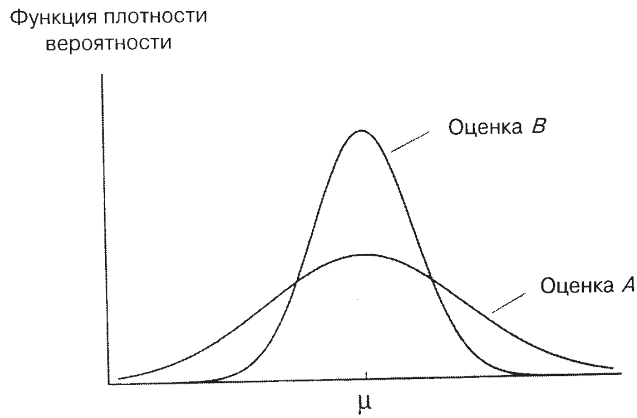


Рис. 2.6. Оценки теоретического среднего[28]

Важно заметить, что мы использовали здесь слово «скорее». Даже хотя оценка *B* более эффективна, это не означает, что она всегда дает более точное значение. При определенном стечении обстоятельств значение оценки *A* может быть ближе к истине.

[28] Доугерти К. Указ. соч.

Однако вероятность того, что оценка A окажется более точной, чем B, составляет менее 50 %.

Это напоминает вопрос о том, пользоваться ли ремнями безопасности при управлении автомобилем. Множество обзоров в разных странах показало, что значительно менее вероятно погибнуть или получить увечья в дорожном происшествии, если воспользоваться ремнями безопасности. В то же время не раз отмечались странные случаи, когда не сделавший этого индивид чудесным образом уцелел, но погиб бы, будучи пристегнут ремнями. Упомянутые обзоры не отрицают этого. В них лишь делается вывод, что преимущество на стороне тех, кто пользуется ремнями безопасности. Подобным же преимуществом обладает и эффективная оценка. (Неприятный комментарий: в тех странах, где пользование ремнями безопасности сделано обязательным, сократилось предложение для трансплантации почек людей, ставших жертвами аварий.)

Выборочное среднее имеет наименьшую дисперсию среди оценок рассматриваемого типа. Это означает, что оно имеет наиболее «сжатое» вероятностное распределение вокруг истинного среднего и, следовательно (в вероятностном смысле), наиболее точно. Строго говоря, выборочное среднее – это наиболее эффективная оценка среди всех несмещенных оценок.

Эффективность оценок можно сравнивать лишь тогда, когда они используют одну и ту же информацию, например один и тот же набор наблюдений нескольких случайных переменных. Если одна из оценок использует в 10 раз больше информации, чем другая, то она вполне может иметь меньшую дисперсию, но было бы неправильно считать ее более эффективной. Во-вторых, мы ограничиваем понятие эффективности сравнением распределений несмещенных оценок.

Итак, для оценки желательны несмещенность и наименьшая возможная дисперсия. Эти критерии совершенно различны, и иногда они могут противоречить друг другу. Может случиться так, что имеются две оценки теоретической характеристики, одна из которых является несмещенной (A на рис. 2.7), другая же смещена, но имеет меньшую дисперсию (B).

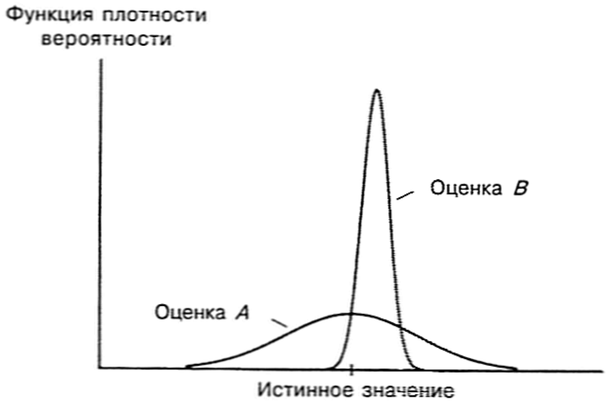


Рис. 2.7. Сравнение двух оценок[29]

Оценка *A* хороша своей несмещенностью, но преимуществом оценки *B* является то, что ее значения практически всегда близки к истинному значению. Какую из них выбрать, зависит от обстоятельств. Если возможные ошибки исследователя не беспокоят при условии, что за длительный период они «погасят» друг друга, то, выбирают *A*. С другой стороны, если для исследователя приемлемы малые ошибки, но неприемлемы большие, то выбирают *B*.

Формально говоря, выбор определяется функцией потерь, стоимостью сделанной ошибки как функцией ее размера. Обычно выбирают оценку, дающую наименьшее ожидание потерь, и делается это путем взвешивания функции потерь по функции плотности вероятности. Если вы не любите риск, то можете также пожелать учесть дисперсию потерь.

Точность оценки x зависит от числа наблюдений n: при увеличении n оценка $\bar{x}$ становится более точной. В единичном эксперименте большая по размеру выборка необязательно даст более точную оценку, чем меньшая выборка, – всегда может присутствовать элемент везения, – но общая тенденция должна быть именно такой. Поскольку дисперсия $\bar{x}$ выражается формулой σ^2/n,

[29] Доугерти К. Указ. соч.

она тем меньше, чем больше размер выборки, и, значит, тем сильнее «сжата» функция плотности вероятности для $\bar{x}$ (рис. 2.7). Предполагается, что x нормально распределена со средним 25 и стандартным отклонением 50. Если размер выборки равен 25, то стандартное отклонение величины $\bar{x}$, равное $\sigma/\sqrt{n}$, составит: $50/\sqrt{25}=10$. Если размер выборки равен 100, то это стандартное отклонение равно 5. На рис. 2.8. показаны соответствующие функции плотности вероятности. Вторая (n=100) выше первой в окрестности μ, что говорит о более высокой вероятности получения с ее помощью аккуратной оценки. За пределами этой окрестности вторая функция всюду ниже первой.

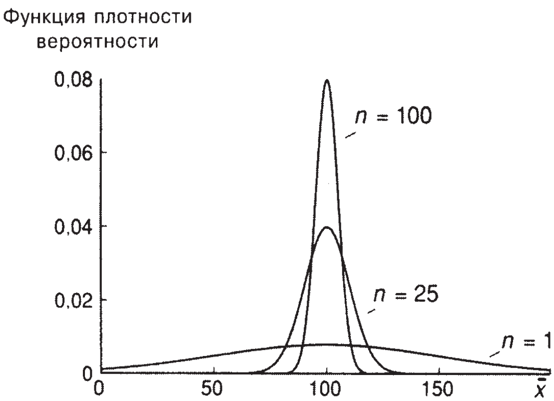


Рис. 2.8. Функции плотности вероятностей
в зависимости от числа наблюдений[30]

Чем больше размер выборки, тем уже и выше будет график функции плотности вероятности для $\bar{x}$. Если $n \to \infty$, то график функции плотности вероятности будет неотличим от вертикальной прямой, соответствующей $\bar{x}=\mu$. Для такой выборки случайная составляющая x становится действительно очень малой, и поэтому $\bar{x}$ обязательно будет очень близкой к μ. Это вытекает из того факта, что стандартное отклонение $\bar{x}$, равное $\sigma/\sqrt{n}$, становится очень

[30] Доугерти К. Указ. соч.

малым при больших n. В пределе, при стремлении n к бесконечности, $\sigma / \sqrt{n}$ стремится к нулю и $\bar{x}$ стремится в точности к μ.

Если предел оценки по вероятности равен истинному значению характеристики генеральной совокупности, то эта оценка называется *состоятельной*. Иначе говоря, состоятельной называется такая оценка, которая дает точное значение для большой выборки независимо от входящих в нее конкретных наблюдений.

В большинстве конкретных случаев несмещенная оценка является и состоятельной. Для этого можно построить контрпримеры, но они, как правило, будут носить искусственный характер.

Иногда бывает, что оценка, смещенная на малых выборках, является состоятельной (иногда состоятельной может быть даже оценка, не имеющая на малых выборках конечного математического ожидания). На рис. 2.9 показано, как при различных размерах выборки может выглядеть распределение вероятностей. Тот факт, что при увеличении размера выборки распределение становится симметричным вокруг истинного значения, указывает на асимптотическую несмещенность. То, что в конечном счете оно превращается в единственную точку истинного значения, говорит о состоятельности оценки.

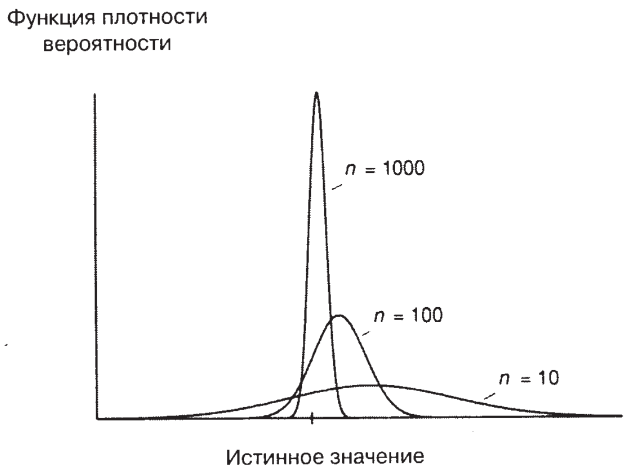


Рис. 2.9. Распределение вероятностей
при различных размерах выборки

Оценки, подобные изображенным на рис. 2.9, весьма важны в регрессионном анализе. Иногда невозможно найти оценку, не смещенную на малых выборках. Если при этом вы можете найти хотя бы состоятельную оценку, это может быть лучше, чем не иметь никакой оценки, особенно если вы можете предположить направление смещения на малых выборках.

Однако надо иметь в виду, что состоятельная оценка может на малых выборках работать хуже, чем несостоятельная (например, иметь большую среднеквадратичную ошибку), и поэтому требуется осторожность. Подобно тому, как вы можно предпочесть смещенную оценку несмещенной, если ее дисперсия меньше, возможно предпочесть состоятельную, но смещенную оценку несмещенной или несостоятельную оценку им обеим (также в случае меньшей дисперсии).

На практике вышеописанные свойства оценок применяются к принятию управленческих решений. Например, ставится задача определения самой эффективной из оценок, представленных на графике: m_2 или m_3. Здесь $f(x)$ – функция плотности вероятности (плотность распределения) случайной величины x, m_1 – среднее значение случайной величины x.

Выбираем оценку с наименьшей дисперсией – m_2 . Визуальное определение сводится к поиску той оценки, для которой плотность вероятности более «сжата» около своего среднего (рис. 2.10).

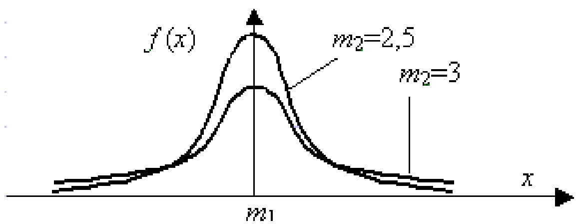


Рис. 2.10. Дисперсии оценок случайной величины x

Таким образом, если регрессионная модель вида (2.25) удовлетворяет условиям 1–4, то оценки a, b являются наилучшими среди всех линейных несмещенных оценок (BLUE).

Однако результаты применения классических статистических методов могут искажаться в случае возникновения:

асимметричности связей;

мультиколлинеарности связей;

эффекта гетероскедастичности;

автокорреляции;

ложной корреляции;

наличия лагов.

Перечисленные проблемы нарушают классическую регрессионную модель. Они диагностируются и устраняются с помощью специальных методов и критериев. Например, проблема автокорреляции решается с помощью критерия Дарбина – Уотсона.

2.3. Применение метода наименьших квадратов

Одним из простейших и распространенных методов построения регрессии является *метод наименьших квадратов (МНК, OLS – Ordinary Least Squares)*. Выполняя практическое исследование, как правило, используют компьютерные программы, которые выдают необходимые оценки коэффициентов.

Метод наименьших квадратов был предложен для воспроизведения движения траектории движения планет более 200 лет назад немецким ученым, «королём математиков», механиком, физиком и астрономом Карлом Фридрихом Гауссом. До начала XIX в. учёные не имели определённых правил для решения системы уравнений, в которой число неизвестных меньше, чем число уравнений; до этого времени употреблялись частные приёмы, зависевшие от вида уравнений и от остроумия вычислителей, и потому разные вычислители, исходя из тех же данных наблюдений, приходили к различным выводам. Гауссу (1795) принадлежит первое применение метода, а Лежандр (1805) независимо открыл и опубликовал его под современным названием. Лаплас связал метод с теорией вероятностей, а американский математик Эдрейн (1808) рассмотрел его теоретико-вероятностные приложения[31].

[31] Вилейтнер Г. История математики от Декарта до середины XIX столетия. М.: Наука, 1966. 507 с.

В дальнейшем этот метод стал использоваться для биологических объектов, экономических и социальных систем. МНК продолжает активно использоваться в эконометрических исследованиях благодаря тому, что в его основе лежит предположение о нормальном законе распределения анализируемых случайных величин.

Построение уравнения регрессии сводится к оценке ее параметров. Метод наименьших квадратов позволяет получить такие оценки параметров, при которых сумма квадратов отклонений фактических значений y_i результативного признака y от теоретических $\hat{y}_i$ минимальна, т. е.

$$\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \longrightarrow \min. \qquad (2.37)$$

В уравнении взаимосвязи двух переменных (парная регрессионная модель), представленной в виде приведенной выше формулы (2.14):

$$y_i = f(x_i) + u_i,$$

функция $f(x_i)$ может быть линейной или нелинейной.

В линейном случае $\hat{y}_i = a + bx_i$ задача сводится к решению следующей системы линейных уравнений:

$$\begin{cases} na + b\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i \\ a\sum_{i=1}^{n} x_i + b\sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} y_i x_i \end{cases} \qquad (2.38)$$

С теоретическим изложением метода наименьших квадратов студенты знакомы из курса математических дисциплин.

Для нахождения *а* и *в* можно воспользоваться формулами, которые получаются решением системы: $\begin{cases} a + b\bar{x} = \bar{y} \\ a\bar{x} + b\overline{x^2} = \overline{xy} \end{cases}$

$$a = \bar{y} - b\,\bar{x}, \quad b = \frac{\overline{\mathrm{y} \cdot \mathrm{x}} - \bar{\mathrm{y}} \cdot \bar{\mathrm{x}}}{\sigma_{\mathrm{x}}^2}, \qquad (2.39),$$

где $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$, $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$, $\overline{yx} = \frac{1}{n}\sum_{i=1}^{n} y_i x_i$, $i = \overline{1,n}$. (2.40)

Доверие к методу наименьших квадратов можно проверить с помощью эксперимента по методу Монте-Карло, который связан с попыткой учета в модели влияния на результат фактора неопределенности. Данный метод позволяет моделировать ситуации, не определенные в данный момент времени, и много раз «проигрывать» их на компьютере. В лекционной части дисциплины «Эконометрика» рассматриваются теоретические основы моделирования методом Монте-Карло. На лабораторных занятиях студентами проводится простейший эксперимент по данному методу, суть которого заключается в следующем[32].

Будем использовать формулы из предыдущих разделов с сохранением нумерации. За основу берется наиболее распространенная простейшая парная стохастическая зависимость между результирующим показателем y_i и единственной объясняющей переменной $x_{i\,в}$ в приведенной выше формуле (2.24)

Линейная парная регрессионная *модель согласно приведенной выше формуле* (2.25) имеет вид*:*

$$\hat{y}_i = a + bx_i,$$

где *a* и *в* – коэффициенты модели, подлежащие определению.

Выбирают метод нахождения коэффициентов таким образом, чтобы расхождения между коэффициентами *а*, *в* и параметрами *α*, *β* соответственно были минимальными. Наиболее подходящим с точки зрения расчета и интерпретации результатов является метод *регрессионного анализа.* Проблема состоит в том, что никогда не знают истинных значений *α* и *β*. Поэтому нельзя сказать, хорошие или плохие оценки дает выбранный метод. Эксперимент по методу Монте-Карло – это искусственный контролируемый эксперимент, дающий возможность такой проверки.

Простейший возможный эксперимент по методу Монте-Карло состоит из трех частей. Во-первых:

1) выбираются истинные значения *α, β*;

---

[32] Доугерти К. Введение в эконометрику : пер. с англ. М.: ИНФРА-М, 1999. С. 30–32.

2) в каждом наблюдении i выбирается детерминированное значение x_i;

3) используется некоторый процесс генерации случайных чисел (или берется последовательность из таблицы случайных чисел) для получения значений случайного фактора u_i в каждом из наблюдений.

Во-вторых, в каждом наблюдении формируется значение y_i с использованием соотношения (9) и значений α, β, x_i и u_i.

В-третьих, применяется регрессионный анализ для оценивания параметров *а, в* с использованием только полученных указанным образом значений y_i и данных для x_i.

Сравнив значения отклонений соответственно α от *а*, β от *в*, отмечают, являются ли *а* и *в* хорошими оценками α, β. Полученный вывод позволит определить пригодность метода построения регрессии.

На первых двух шагах проводится подготовка к применению регрессионного метода. Мы полностью контролируем модель, которую создаем, и знаем истинные значения параметров α, β, потому что сами их определили. На третьем этапе мы определяем, может ли поставленная нами задача решаться с помощью метода регрессии, т. е. могут ли быть получены хорошие оценки для α, β при использовании данных об y_i и x_i.

Заметим, что проблема возникает вследствие включения случайного фактора в процесс получения y_i. Если бы этот фактор отсутствовал, то точки, соответствующие значениям каждого наблюдения, лежали бы точно на прямой, и точные значения α, β, можно было бы определить по значениям y_i и x_i. Заметим, несмотря на то что в одних случаях оценки принимают заниженные значения, а в других –завышенные, в целом значения *а* и *в* группируются вокруг истинных значений α и β. При этом хороших оценок получено больше, чем плохих. Например, фиксируя значения b при очень большом числе повторений эксперимента, можно построить таблицу частот и получить аппроксимацию функции плотности вероятности.

Выше говорилось, что расхождения между коэффициентами регрессии и истинными значениями параметров вызваны случайным членом u_i. Отсюда следует, что чем больше элемент случайности, тем, вообще говоря, менее точными являются оценки.

Важной составляющей моделирования методом Монте-Карло является построение случайной величины с заданными характеристиками. Для моделирования случайных чисел в MS Excel можно использовать формулу =СЛЧИС(), которая позволяет получить число, с одинаковой вероятностью принимающее значение в диапазоне [0,1].

Уделим внимание практической применимости моделирования по методу Монте-Карло[33]. Моделирование можно использовать для оценки «реальных возможностей», например возможности развития, принятия обязательств или отсрочки проекта. Многие компании применяют моделирование методом Монте-Карло как важное средство принятия решений:

- компании General Motors, Procter and Gamble и Eli Lilly моделируют оценки средней доходности и риск, связанный с выпуском новых товаров;
- в General Motors прогнозируют чистую прибыль корпорации, структурные затраты и затраты на приобретение; определяют подверженность корпорации различным видам рисков (например, изменению процентных ставок и колебаниям валютного курса);
- Eli Lilly прогнозируют определение оптимальной производственной мощности, требуемой для производства каждого лекарства;
- компании с Wall Street строят модели для оценки сложных финансовых показателей и суммы под риском (СПР) их инвестиционных портфелей;
- Procter and Gamble применяют моделирование для примерной оценки и оптимального хеджирования рисков, связанных с изменением курса иностранной валюты;
- специалисты по финансовому планированию используют моделирование методом Монте-Карло для определения оптимальной инвестиционной стратегии для пенсионных вкладов.

Если между экономическими явлениями существуют нелинейные соотношения, то они выражаются с помощью соответ-

[33] Уэйн Л. Винстон. Microsoft Excel: анализ данных и построение бизнес моделей: пер. с англ. М.: Русская Редакция, 2005. С. 77.

ствующих нелинейных функций. В эконометрическом моделировании нелинейные модели подразделяют на 2 класса:

1. Нелинейные относительно включенных в анализ объясняющих переменных, но линейные по оцениваемым параметрам, например:

- полиномы различных степеней

$$y_i = \alpha + \beta_1 x_i + \beta_2 x_i^2 + \ldots + \beta_k x_i^k + u_i\,; \qquad (2.41)$$

- равносторонняя гипербола

$$y_i = \alpha + \frac{\beta}{x_i} + u_i\,; \qquad (2.42)$$

- полулогарифмическая функция

$$y_i = \alpha + \beta \ln x_i + u_i\,. \qquad (2.43)$$

Нелинейные модели путем соответствующих преобразований приводятся к линейным, т. е. производится линеаризация переменных.

Регрессии нелинейные по включенным переменным приводятся к линейному виду простой заменой переменных, а дальнейшая оценка параметров производится с помощью метода наименьших квадратов.

Так, парабола второй степени:

$$y_i = \alpha + \beta_1 x_i + \beta_2 x_i^2 + u_i \qquad (2.44)$$

приводится к линейному виду с помощью замены:

$x_1 = x$:, $x_2 = x^2$.

Полученное двухфакторное уравнение

$$y_i = \alpha + \beta_1 x_{1i} + \beta_2 x_{2i} + u_i \qquad (2.45)$$

решается МНК через систему нормальных уравнений:

$$\begin{cases} an + b_1 \sum_{i=1}^{n} x_{1i} + b_2 \sum_{i=1}^{n} x_{2i} = \sum_{i=1}^{n} y_i \\ a \sum_{i=1}^{n} x_{1i} + b_1 \sum_{i=1}^{n} x_{1i}^2 + b_2 \sum_{i=1}^{n} x_{1i} x_{2i} = \sum_{i=1}^{n} x_{1i} y_i \\ a \sum_{i=1}^{n} x_{2i} + b_1 \sum_{i=1}^{n} x_{1i} x_{2i} + b_2 \sum_{i=1}^{n} x_{2i}^2 = \sum_{i=1}^{n} x_{2i} y_i \end{cases} \qquad (2.46)$$

Равносторонняя гипербола приводится к линейному уравнению заменой $z = 1/x$. Система линейных уравнений для МНК будет иметь вид:

$$\begin{cases} na + b\sum_{i=1}^{n} z_i = \sum_{i=1}^{n} y_i \\ a\sum_{i=1}^{n} z_i + b\sum_{i=1}^{n} {z_i}^2 = \sum_{i=1}^{n} y_i z_i \end{cases} \qquad (2.47)$$

2. Нелинейные по оцениваемым параметрам, которые подразделяются на два типа: внутренне линейные и внутренне нелинейные. К внутренне линейным моделям относятся:

- степенная

$$y_i = \alpha {x_i}^{\beta} \cdot u_i\,; \qquad (2.48)$$

- показательная

$$y_i = \alpha \beta^{x_i} \cdot u_i\,. \qquad (2.49)$$

Нелинейные модели внутренне линейные приводятся к линейному виду с помощью соответствующих преобразований, например логарифмированием. Отметим, что в данных моделях связь между результирующей переменной y и факторной x не аддитивная, а мультипликативная.

Для степенной модели (2.48) преобразование будет иметь вид:

$$\ln y_i = \ln(\alpha\, x_i^{\beta}\, u_i)$$

$$\ln y_i = \ln \alpha + \beta \ln x_i + \ln u_i\,. \qquad (2.50)$$

Для показательной модели преобразование будет выглядеть:

$$\ln y_i = \ln(\alpha\, \beta^{x_i}\, u_i)$$

$$\ln y_i = \ln \alpha + x_i \ln \beta + \ln u_i\,. \qquad (2.51)$$

К нелинейным моделям внутренне нелинейным относятся, например:

- $y_i = \alpha + \beta x_i^c + u_i\,,$ (2.52)

где c – константа;

- $y_i = \alpha\left(1 - \dfrac{1}{1 - x_i^b}\right) + u_i\,.$ (2.53)

Нелинейные модели внутренне нелинейные к линейному виду не приводятся.

После построения уравнения регрессии дается его интерпретация. Существуют два этапа интерпретации уравнения регрессии.

1. Первый состоит в словесном истолковании уравнения так, чтобы оно было понятно человеку, не являющемуся специалистом в области эконометрики и статистики.

2. На втором этапе необходимо решить, следует ли ограничиться первым этапом или провести более детальное исследование зависимости.

Первый этап будет проиллюстрирован моделью регрессии для функции спроса: регрессией между расходами потребителя на питание *y* и располагаемым личным доходом *x* по данным, приведенным в табл. 1 (см. Приложение) для США за период с 1959 по 1983[34].

Предположим, что истинная модель представлена в аддитивной линейной форме вида:

$$y_i = \alpha + \beta x_i + u_i$$

и оценена регрессия: $\hat{y}_i = 55,009 + 0,093\, x_i$

Коэффициент при *x*, называемый *коэффициентом регрессии*, показывает, что при увеличении *x* на единицу, *y* возрастает на 0,093 единиц. Как *x*, так и *y* измеряются в миллиардах долларов в постоянных ценах, т. о., коэффициент наклона показывает, что если доход увеличивается на 1 млрд долл., то расходы на питание возрастают на 93 млн долл. Другими словами, из каждого дополнительного доллара дохода 9,3 цента будут израсходованы на питание. Данный вывод называется *интерпретацией уравнения регрессии*. Свободная составляющая *a* показывает прогнозируемый уровень *y*, если *x = 0*. Когда *x = 0*, то буквальная интерпретация может привести к неверным результатам; даже если линия регрессии достаточно точно описывает значения наблюдаемой выборки, нет гарантии, что так же будет при экстраполяции влево или вправо. В данном случае константа выполняет единственную функцию: она позволяет определить положение линии регрессии на графике.

[34] Данные взяты из: Доугерти К. Указ. соч.

Во многих практических случаях моделирование экономических зависимостей линейными уравнениями в силу их многообразия нередко не приводит к построению качественных моделей. Например, при рассмотрении зависимости спроса y_i на некоторый товар от цены этого товара x_i в ряде случаев можно ограничиться линейным уравнением регрессии. Если же требуется проанализировать эластичность спроса по цене, то приведенное уравнение не позволит этого сделать и в этом случае целесообразнее рассмотреть логарифмическую модель. При анализе издержек y_i от объема выпуска x_i наиболее обоснованной является полиномиальная (точнее, кубическая) модель. При рассмотрении производственной функции линейная модель также не является адекватной. Обычно используют степенные модели, например, широко известную функцию Кобба – Дугласа:

$$У = AK^{\alpha} L^{\beta}, \tag{2.54}$$

где *A, α, β* – параметры модели; *K, L* – затраты на труд и капитал; *У* – объем выпуска.

Достаточно широко применяются в современном эконометрическом анализе и многие другие модели, в частности обратная и экспоненциальная модели.

Модель равносторонней гиперболы применяется в тех случаях, когда неограниченное увеличение объясняющей переменной x ассимптотически приближает зависимую переменную y к некоторому пределу, к α. Важным приложением модели является кривая Филипса, отражающая зависимость между уровнем безработицы x в процентах и процентным изменением заработной платы y. При этом точка пересечения кривой с осью OX определяет естественный уровень безработицы.

Широкое практическое применение нашли полиномиальные функции. Например, кубическая функция

$$y = \alpha + \beta x + \gamma x^2 + \eta x^3 + u \tag{2.55}$$

в микроэкономике моделирует зависимость общих издержек TC от объема выпуска Q.

Квадратичная функция

$$y = \alpha + \beta x + \gamma x^2 + u \tag{2.56}$$

может отражать зависимость между объемом выпуска и средними *AC* либо предельными *MC* издержками; или между расходами на рекламу *C* и прибылью *П*.

Контрольные вопросы и задания

1. Как определяется число степеней свободы случайной величины СВ?

2. Как связаны между собой СВ, имеющие стандартизированное нормальное распределение, распределения Стьюдента и Фишера?

3. Справедливо или ложно утверждение, что при увеличении числа степеней свободы распределения Стьюдента и Фишера стремятся к стандартизированному нормальному распределению?

4. Приведите примеры описания экономической ситуации с помощью линейных и нелинейных моделей.

5. На основе статистических данных постройте модель регрессии линейную, степенную, полиномиальную 2-й степени и показательную. Дайте интерпретацию коэффициента регрессии в каждом случае.

6. Распределите предложенные функции на 2 класса нелинейных моделей:

- нелинейные относительно включенных в анализ объясняющих переменных, но линейные по оцениваемым параметрам;
- нелинейные по оцениваемым параметрам.

7. Выполните эксперимент по методу Монте-Карло. Какова роль эксперимента по методу Монте-Карло в моделировании?

8. Сформулируйте условия Гаусса-Маркова.

3. Оценка качества построенной регрессии

3.1. Коэффициент детерминации

В наиболее общем виде при изучении взаимосвязей исследователя интересует количественная оценка их наличия и направления, а также характеристика силы и формы влияния одних факторов на другие. Для решения этого вопроса применяется две группы методов, одна из которых включает в себя методы ***корреляционного анализа***, а другого – ***регрессионнщго анализа***. В то же время ряд исследователей объединяет эти методы в *корреляционно-регрессионный* анализ, что объясняется наличием целого ряда схожих вычислительных процедур, взаимодополнения при интерпретации результатов и др.

Задачи собственно корреляционного анализа сводятся к измерению *тесноты связи* между изменяющимися признаками, определению неизвестных причинных связей и оценке факторов, оказывающих наибольшее влияние на результативный признак. Задачи регрессионного анализа лежат в сфере установления *формы зависимости*, определения *функции регрессии*, оценки неизвестных значений зависимой переменной.

Решение указанных выше задач опирается на соответствующие приемы, алгоритмы, показатели, применение которых дает основание говорить о статистическом изучении взаимосвязей. Вычислительные процедуры представляют самостоятельный интерес, но знание принципов изучения взаимосвязей, возможностей и ограничений тех или иных методов интерпретации результатов являются обязательным условием исследования.

Построение, проверку и улучшение экономических моделей необходимо проводить с использованием статистического анализа переменных на основе реальных статистических данных. Вся сфера экономических исследований может быть в определенном смысле охарактеризована как *изучение взаимосвязей экономических переменных*. При этом инструментарием их базового анализа являются методы статистики и эконометрики.

Методы оценки тесноты связи подразделяются на корреляционные (параметрические) и непараметрические. Параметрические методы основаны на использовании, как правило, оценок нормального распределения и применяются в случаях, когда изучаемая совокупность состоит из величин, которые подчиняются закону нормального распределения. Непараметрические методы не накладывают ограничений на закон распределения изучаемых величин.

Простейшим приемом выявления связи между изучаемыми признаками x и y является построение корреляционной таблицы. Ее наглядным изображением служит корреляционное поле, представляющее собой график, где на оси абсцисс откладываются значения x_i, по оси ординат y_i. По расположению точек, их концентрации в определенном направлении можно судить о наличии связи между изучаемыми признаками x и y.

Последовательность точек x_i ($i = 1, ..., n$) и среднего значения y_i, обозначаемого $\overline{y}$, позволяет построить график, который иллюстрирует зависимость среднего значения результативного признака y от факторного x, – эмпирическую линию регрессии.

По существу, корреляционная таблица, корреляционное поле, эмпирическая линия регрессии предварительно уже характеризуют взаимосвязь, когда выбраны факторный и результативный признаки и требуется сформулировать предположение о форме и направленности связи.

На практике для количественной оценки тесноты связи для линейной регрессии используется линейный коэффициент парной корреляции rxy (-1J

На практике для количественной оценки тесноты связи для линейной регрессии используется линейный коэффициент парной корреляции r_{xy} ($-1 \le r_{xy} \le 1$), который может определяться следующим образом:

$$r_{xy} = b\frac{\sigma_x}{\sigma_y} = \frac{\operatorname{cov}(x,y)}{\sigma_x\ \sigma_y} = \frac{\overline{yx} - \overline{y}\ \overline{x}}{\sigma_x \sigma_y} \tag{3.1}$$

$$b = \frac{\operatorname{cov}(x,y)}{{\sigma_x}^2} = \frac{\overline{yx} - \overline{y}\ \overline{x}}{\overline{x^2} - \overline{x}^2} \tag{3.2}$$

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i\ ;\ \overline{y} = \frac{1}{n}\sum_{i=1}^{n} y_i\ ;\ \overline{yx} = \frac{1}{n}\sum_{i=1}^{n} y_i\ x_i \tag{3.3}$$

где b – коэффициент линейной регрессии $\hat{y}_i = a + bx_i$; σ_x, σ_y – среднее квадратическое отклонение соответствующей случайной величины; σ_x^2 – дисперсия признака x.

Ковариацией (или *корреляционным моментом*) случайных величин x и y называется математическое ожидание произведения отклонений этих величин от своих математических ожиданий, т. е.:

$$\operatorname{cov}(x, y) = M\left[\left(x - M(x)\right)\left(y - M(y)\right)\right] \qquad (3.4)$$

где $cov\ (x, y)$ – ковариация признаков x и y; $M(x)$, $M(y)$ – математическое ожидание случайных величин x и y соответственно.

Для оценки тесноты связи *нелинейной регрессии* строится индекс корреляции ρ_{xy} $(0 \le \rho_{xy} \le 1)$:

$$\rho_{xy} = \sqrt{1 - \frac{\sigma_{ост}{}^2}{\sigma_y{}^2}} = \sqrt{1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}} \qquad (3.5)$$

где $\hat{y}_i = a + bx_i$

Коэффициент (индекс) корреляции является безразмерной величиной, так как его значение не зависит от выбора единиц измерения обеих переменных.

Близкая к нулю величина коэффициента корреляции свидетельствует об отсутствии линейной связи переменных, но не об отсутствии связи между ними вообще. Например, если показатель корреляции величин уровней инфляции и безработицы для периода 1970–1980-х гг. для экономики некоторой страны практически равен нулю, не следует говорить сразу о независимости этих показателей в данный период. Следует попытаться построить более сложную модель их связи, учитывающую, возможно, как *нелинейность* самой зависимости, так и наличие в ней *запаздываний во времени* (*лагов*), а также *инерционность динамики* соответствующих величин.

Равенство нулю коэффициента корреляции для генеральной совокупности еще не означает, что он будет в точности нулевым для выборки. Наоборот, он обязательно будет отклоняться от истинного значения, но чем больше такое отклонение, тем менее оно вероятно при данном объеме выборки. При каждом конкретном значении коэффициента корреляции величин x и y для генеральной совокупности выборочный коэффициент корреляции является *случайной величиной*. Следовательно, случайной величиной является также любая его функция и требуется указать такую функцию, которая имела бы одно из известных распределений, удобное для табличного анализа. Для выборочного коэффициента корреляции r_{xy} такой функцией является *t-статистика*, рассчитываемая по формуле

$$t = r_{xy}\frac{\sqrt{n-2}}{\sqrt{1-r_{xy}^{\ 2}}} \tag{3.6}$$

и имеющая распределение Стьюдента с (*n-2*) степенями свободы. Число степеней свободы меньше числа наблюдений на 2, поскольку в формулу коэффициента корреляции входят средние значения x и y, для расчета которых используются две линейные формулы их зависимости от наблюдений случайных величин.

Для коэффициента корреляции будет проверяться *нулевая гипотеза* H_0 , т. е. гипотеза о *равенстве его нулю* в генеральной совокупности (более подробно см. следующую тему).

Оценку качества построенной модели даст *коэффициент* $R^2 = r_{xy}^{\ 2}$ ($R^2 = \rho_{xy}^{\ 2}$ индекс), а также *средняя ошибка аппроксимации*:

$$\overline{A} = \frac{1}{n}\sum_{i=1}^{n} A_i = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| 100\% \tag{3.7}$$

Традиционно считается, что допустимый предел значений не более 8–10 %. В этом случае модель оценивается как *достаточно точная*, в противном случае говорят о *плохом качестве* построенной модели.

Одной из наиболее эффективных оценок адекватности регрессионной модели, *мерой качества уравнения регрессии* или, как говорят, *мерой качества подгонки* регрессионной модели к наблюдаемым значениям, характеристикой прогностической силы анализируемой регрессионной модели является *коэффициент детерминации* R^2, принимающий значения в интервале $0 \le R^2 \le 1$.

$$R^2 = \frac{\sigma^2{}_{y\,объясн}}{\sigma^2{}_{y.общ}} \tag{3.8}$$

Например, если $R^2=0{,}982$, уравнением регрессии объясняется 98,2 % результативного признака, а на долю прочих факторов приходится лишь 1,8 % ее дисперсии (так называемая остаточная дисперсия). Чем ближе значение R^2 к единице, тем большую долю изменения результативного фактора y можно объяснить за счет вариации включенного в модель фактора x, меньше роль прочих факторов, и, следовательно, линейная модель хорошо аппроксимирует исходные данные (наблюдения «теснее примыкают» к линии регрессии), и модель можно использовать для прогноза значений результативного признака.

Заметим, что коэффициент детерминации R^2 имеет смысл рассматривать только при наличии свободного члена в уравнении регрессии, так как лишь в этом случае верны равенства:

$$Q = Q_R + Q_e, \tag{3.9}$$

$$R^2 = \frac{Q_R}{Q} = 1 - \frac{Q_e}{Q} \tag{3.10}$$

где Q – общая сумма квадратов отклонений; Q_R – сумма квадратов отклонений, обусловленная регрессией; Q_e – остаточная сумма квадратов отклонений.

В случае парной линейной модели коэффициент детерминации равен квадрату коэффициента корреляции. Тогда выполняется равенство:

$$R^2 = r^2{}_{xy}\,. \tag{3.11}$$

Коэффициент детерминации традиционно на практике оценивает качество подгонки регрессии. Хотя повышение значения этого коэффициента не самоцель эконометрического исследования, многие исследователи по-прежнему уделяют ему повышенное внимание.

Рассмотрим практику использования коэффициента детерминации R^2 на примере исследования промышленности и сельского хозяйства Ярославской области. В качестве расчетного инструмента использовалась производственная функция Кобба

– Дугласа, на основе которой строится трехфакторная модель регрессии[35]:

$$Y = AK^{\alpha}L^{\beta}I^{\gamma} \tag{3.12}$$

$$\alpha > 0, \beta > 0, \gamma > 0,$$

где Y – объем национального дохода или продукции; K – объем основных производственных фондов; L – численность занятых (работающих); I – научно-технический прогресс (НТП); A – параметр, приводящий к единому фактору национальный доход (продукцию) и используемые факторы, зависящий от выбранных единиц измерения; α, β, γ – коэффициенты эластичности фондов, труда и НТП.

Расчет выполнен по ограниченному временному периоду, за который удалось получить сопоставимые по методологии данные по более широкому кругу показателей, в том числе и тех, которыми можно количественно оценить фактор НТП. По промышленности Ярославской области они приведены в табл. 3.1.

В качестве фактора *НТП* берется количество использованных передовых технологий, которое, по нашему мнению, более точно отражает влияние данного фактора на экономический рост и, вместе с тем, не подвержено искажающему влиянию изменения стоимостных измерителей, чем, например, инвестиции в новые виды продукции, технологии и т. д.

В качестве фактора *труд* предлагается рассмотреть два варианта. *В первом варианте* за фактор труд примем численность человек, занятых в науке и научном обслуживании: $L=L2$.

С учетом ограничений на факторы полагаем, что

$$\alpha + \beta + \gamma = 1$$

$$\gamma = 1-(\alpha + \beta), \tag{3.13}$$

то есть фактор НТП является комбинацией факторов труд и капитал:

$$\begin{gathered} \lg Y = \lg A + \alpha \lg K + \beta \lg L + (1 - \alpha - \beta) \lg I \\ \lg Y - \lg I = \lg A + \alpha(\lg K - \lg I) + \beta(\lg L - \lg I) \end{gathered} \tag{3.14}$$

[35] Россия и ВТО: региональный уровень оценки последствий интеграции / под ред. Д. С. Лебедева. Ярославль, 2013. С. 90–97.

Таблица 3.1

Данные для расчета производственной функции по промышленности Ярославской области (в сопоставимой методологии) за 2000–2009 гг.

Годы	*Y – стоимость промышленной продукции, млрд рублей*	*K – стоимость основного капитала в промышленности, млрд рублей*	*L1 – численность занятых в промышленности, человек*	*L2 – численность занятых в науке и научном обслуживании, человек*	*I – количество используемых передовых технологий, шт.*
2000	4553,4	166,562	196873	9259	719
2001	4840,3	174,573	197986	9145	1031
2002	4874,1	166,702	194230	8545	1106
2003	4996,0	187,540	184251	7876	1407
2004	5370,7	171,411	175980	7593	1256
2005	5666,1	185,809	163616	6608	1401
2006	5983,4	173,361	160705	6660	1647
2007	6390,3	176,828	158208	7190	2690
2008	6326,4	171,878	154752	6739	2856
2009	5276,2	180,986	131034	6358	3283

Выполним соответствующие преобразования, представив расчеты в табл. 3.2, 3.4.

Рассчитаем статистические характеристики: значения оценок параметров, их стандартные ошибки, значение коэффициента детерминации R^2 (табл. 3.3, 3.5).

Таблица 3.2

Значения логарифмов факторов

lg Y	*lg K*	*lg L*	*lg I*	*lgY-lgI*	*lgK-lgI*	*lgL-lgI*
3,658336	2,221575927	3,966564	2,856729	0,801607	-0,63515	1,109835
3,684872	2,241937269	3,961184	3,013259	0,671614	-0,77132	0,947925
3,687894	2,22194081	3,931712	3,043755	0,644139	-0,82181	0,887957
3,698622	2,273093912	3,896306	3,148294	0,550328	-0,8752	0,748012
3,730031	2,234038689	3,880413	3,09899	0,631041	-0,86495	0,781424
3,753284	2,269066746	3,82007	3,146438	0,606846	-0,87737	0,673632
3,776948	2,238951403	3,823474	3,216694	0,560254	-0,97774	0,606781
3,805521	2,247551035	3,856729	3,429752	0,375769	-1,1822	0,426977
3,801157	2,235220292	3,828595	3,455758	0,345398	-1,22054	0,372837
3,722321	2,257644982	3,803321	3,516271	0,20605	-1,25863	0,28705

Таблица 3.3

Данные статистического анализа в Excel

статистики	*β*	*α*	*lgA*	*γ*	*A*
коэффициенты	0,033077	0,796898	1,272523	0,170025	18,72935
стандартные ошибки	0,308187	0,393197	0,581882		
R^2	0,954688	0,043045	#Н/Д		
F	73,74287	7	#Н/Д		
суммы квадратов	0,273268	0,01297	#Н/Д		

Модель с ограничениями на *I*: γ = 1-(α + β) имеет вид

$$Y = 18,73K^{0,797}L^{0,033}I^{0,17} \qquad (3.15)$$

$$R^2 = 0,95$$

Ошибка модели составляет примерно 4 % и является незначительной.

Полученная модель показывает, что каждый фактор вносит свой вклад в экономический рост. Фактор капитал обеспечивает почти

80 % относительного экономического роста. Фактор НТП обеспечил только 17 % относительного экономического роста в промышленности области. Фактор труд не является эффективным, так как составляет 3,3 % относительного экономического роста.

Во втором варианте за фактор труд примем численность человек, занятых в промышленности $L=L1$.

Таблица 3.4

Значения логарифмов факторов

lg Y	*lg K*	*lg L*	*lg I*	*lgY-lgI*	*lgL-lgI*	*lgK-lgI*
3,658336	2,221575927	5,294186	2,856729	0,801607	2,437457	-0,63515
3,684872	2,241937269	5,296634	3,013259	0,671614	2,283376	-0,77132
3,687894	2,22194081	5,288316	3,043755	0,644139	2,244561	-0,82181
3,698622	2,273093912	5,26541	3,148294	0,550328	2,117116	-0,8752
3,730031	2,234038689	5,245463	3,09899	0,631041	2,146474	-0,86495
3,753284	2,269066746	5,213826	3,146438	0,606846	2,067388	-0,87737
3,776948	2,238951403	5,206029	3,216694	0,560254	1,989336	-0,97774
3,805521	2,247551035	5,199228	3,429752	0,375769	1,769476	-1,1822
3,801157	2,235220292	5,189636	3,455758	0,345398	1,733878	-1,22054
3,722321	2,257644982	5,117384	3,516271	0,20605	1,601113	-1,25863

Таблица 3.5

Данные статистического анализа в Excel

Статистики	α	β	*lgA*	γ	*A*
коэффициенты	0,27435	0,443698	-0,10518	0,281952	0,784902
стандартные ошибки	0,483409	0,376977	1,224774		
R^2	0,962112	0,039361	#Н/Д		
F	88,87722	7	#Н/Д		
суммы квадратов	0,275393	0,010845	#Н/Д		

Трехфакторная модель с ограничением на параметры имеет вид:

$$Y = 0,7849 K^{0,274} L^{0,444} I^{0,282} \qquad (3.16)$$

$$R^2 = 0,96$$

Ошибка модели составляет столько же, сколько и в предыдущей модели, – 4 % и является незначительной.

Согласно модели, почти половину относительного экономического роста (44,4 %) обеспечивает увеличение трудозатрат. В рассмотренные годы фактор НТП обеспечивал лишь около трети (28,2 %) относительного экономического роста в промышленности области. Вносимый капитал также не является эффективным, так как составляет 27,4 % относительного экономического роста. И величина стандартной ошибки почти в 2 раза превышает непосредственное значение α.

В обоих вариантах получено высокое значение коэффициента детерминации R^2, что указывает на хорошее качество моделей, так как в первом варианте 95 %, во втором варианте 96,2 % изменения объема промышленной продукции объясняется суммарной вариацией факторов.

По аналогии с исследованием промышленности рассмотрим модель для сельского хозяйства Ярославской области. Рассмотрим взаимозамещаемость факторов с учетом ограничений на параметры: $\alpha + \beta + \gamma = 1$ (табл. 3.6–3.8).

Получена модель следующего вида:

$$Y = 72,7 \cdot K^{-0,05} L^{0,84} I^{0,21}, \qquad (3.17)$$

$R^2 = 0,99$.

Отметим высокое качество модели, так как на основе R^2 99 % результата объясняется влиянием суммарного изменения факторов.

Ошибка модели составляет около 2 % и является незначительной.

Таблица 3.6

Исходные данные для расчета производственной функции по Ярославской области, отрасль сельское хозяйство, за период 2000–2010 гг.

Годы	*Y – стоимость произведенной продукции (в сопоставимой методологии)*[36]*, млн руб.*	*K – стоимость основного капитала на конец года (в сопоставимых ценах)*[37]*, млн руб.*	*L – среднегодовая численность занятых, тыс. чел.*	*I – число используемых передовых технологий, шт.*
2000	6855	7417	72,3	719
2001	7061	8425	72,7	1031
2002	6637	9571	73,1	1106
2003	6730	9914	62,5	1407
2004	6266	9281	55,4	1256
2005	6510	10742	59,8	1401
2006	6653	11080	58,2	1647
2007	6726	13363	57,5	2690
2008	6726	15863	53,1	2856
2009	6982	21655	54,2	3283
2010	6472	26532	53,7	3267

Таблица 3.7

Значения логарифмов факторов

lg Y	*lg K*	*lg L*	*lg I*	*lgY-lgI*	*lgK-lgI*	*lgL-lgI*
3,836007	3,870228279	1,859138	2,856729	0,979279	1,013499	-0,99759
3,848866	3,92556991	1,861534	3,013259	0,835608	0,912311	-1,15172
3,821972	3,980957316	1,863917	3,043755	0,778217	0,937202	-1,17984
3,828015	3,996248915	1,79588	3,148294	0,679721	0,847955	-1,35241

[36] Стоимость продукции приведена к сопоставимому виду путем пересчета по индексу физического объема.

[37] Пересчет в сопоставимые цены сделан по индексу цен производителей в строительстве (строительно-монтажные работы).

3,79699	3,967594773	1,74351	3,09899	0,698001	0,868605	-1,35548
3,813581	4,031085148	1,776701	3,146438	0,667143	0,884647	-1,36974
3,823018	4,04453976	1,764923	3,216694	0,606324	0,827846	-1,45177
3,827757	4,125903968	1,759668	3,429752	0,398005	0,696152	-1,67008
3,827757	4,200385324	1,725095	3,455758	0,371999	0,744627	-1,73066
3,84398	4,335558188	1,733999	3,516271	0,327709	0,819287	-1,78227
3,811039	4,423769989	1,729974	3,514149	0,296889	0,909621	-1,78417

Таблица 3.8

Данные статистического анализа

статистики	β	*α*	lgA	γ	A
коэффициенты	0,840994	-0,05572	1,861402241	0,214723	72,67788
стандартные ошибки	0,033934	0,104214	0,129396946		
R^2	0,993539	0,0203	#Н/Д		
F	615,1371	8	#Н/Д		
суммы квадратов	0,507003	0,003297	#Н/Д		

Коэффициенты эластичности показывают, что за 2000–2010 гг. увеличение объема сельскохозяйственной продукции в области на 84,1 % шло за счет увеличения занятости и на 21,5 % за счет НТП. В то же время объем сельскохозяйственной продукции не проявлял статистически достоверной зависимости от основных производственных фондов.

Сравнивая модели по анализу промышленности и сельского хозяйства Ярославской области, отметим следующие факты:

1. Параметр при факторе *труд* для модели по сельскому хозяйству в 2 раза выше, чем для модели по промышленности. Можно предположить, что для сельского хозяйства крайне актуальной является проблема неэффективного труда в связи с низкой его производительностью;

2. Внедрение передовых технологий и других факторов НТП является малоэффективным как для промышленности, так и для

сельского хозяйства. В обоих случаях рост продукции происходит примерно на 20–28 %.

Практическая направленность построенных моделей и прогнозирование на их основе сводятся к оптимизации структуры государственных расходов в развитие промышленности и сельского хозяйства. Достаточно неоднозначно нами оценивается, насколько эти вложения соответствуют современному состоянию экономической ситуации.

3.2. Спецификация модели

Выбор общего вида модели и выявление входящих в нее связей осуществляется через подбор функции $f(x)$. В частности, рассматривается возможность применения линейной модели как наиболее простой и надежной. На основе построенных моделей регрессии рассчитываются коэффициенты детерминации, и если его значение неудовлетворительно, то рассматривают построение нелинейных моделей. В случае улучшения значения коэффициента детерминации проводят дальнейшую проверку качества модели на основе статистических критериев значимости, проверка на мультиколлинеарность, автокорреляцию и гетероскедастичность – свойства, нарушающие условия классической модели регрессии, рассмотренные в гл. 2.1.

Рассмотрим, как решается проблема спецификации на примере моделирования экономической устойчивости предприятия.

В качестве примера моделирования социально-экономической системы определим зависимость изменения положения предприятия в «Рейтинге-200» журнала «Эксперт» от следующих параметров:

1) производительности труда;

2) рентабельности реализации;

3) темпа прироста объема реализации, определенного как отношение объема реализации этого года к предыдущему за вычетом единицы;

4) прибыли после налогообложения в расчете на одного человека.

Показатели прибыли и рентабельности реализации являются основополагающими при рассмотрении как социально-экономического, так и производственно-технического уровней при моделировании экономических систем.

Нами вводится система частных показателей (факторных переменных), из которых формируется интегральный показатель экономической устойчивости (результирующая переменная модели).

Для решения поставленной задачи необходимо привести параметры к сопоставимому виду, что позволяет сделать введение нами индексов эффективности. Для этого как частные (локальные), так и обобщенные (интегральные) показатели ранжировались, и им присваивалось значение индекса эффективности, исходя из традиционного правила: чем выше значение показателя, тем более высокий индекс эффективности предприятия по данному показателю.

На основе приведенных выше показателей предлагается построить систему относительных индексов эффективности, включающую следующие индексы (рис. 3.1):

- эффективности производительности труда;
- эффективности рентабельности реализации;
- эффективности темпа прироста объема реализации;
- эффективности прибыли после налогообложения в расчете на одного человека;
- интегральных показателей, характеризующих деятельность компании по перечисленным выше четырем индексам эффективности.

Выбор частных показателей устойчивости

Определение формы связи частных показателей

Расчет весовых коэффициентов

Интегральный показатель экономической устойчивости предприятия

Определение допустимых вариаций частных показателей

Подбор границ устойчивости интегрального показателя

Рис. 3.1. Составляющие моделирования интегрального показателя экономической устойчивости предприятия

Интегральные показатели строились с использованием четырех индексов эффективности, присвоенных предприятиям, исходя из следующих формул:

$$I_1 = \sqrt{Мпроизв^2 + Мрент^2 + Мприрост^2 + Мчп^2} \quad (3.18)$$

$$I_{11} = \sqrt{(Мпроизв - 1)^2 + (Мрент - 1)^2 + (Мприрост - 1)^2 + (Мчп - 1)^2} \quad (3.19)$$

$$I_2 = \max\{Мпроизв; Мрент; Мприрост; Мчп\} \quad (3.20)$$

$$I_3 = \min\{Мпроизв; Мрент; Мприрост; Мчп\} \quad (3.21)$$

Данные индексы рассчитываются как евклидово расстояние и имеют следующее экономическое содержание:

I_1 – отражает отклонение предприятия от идеального, лидирующего по всем выбранным исследователем показателям;

I_{11} – характеризует отклонение от предприятия с единичными значениями индексов эффективности и является модификацией показателя *I1*;

I_2 – показывает самое «слабое место» у предприятия, тот показатель, по которому предприятие наиболее отстает;

I_3 – представляет самое «сильное место» у предприятия, тот показатель, по которому предприятие выглядит наиболее успешным;

I_4 – отражает среднее положение предприятия по всем предложенным нами индексам эффективности;

$M_{произв}$ – индекс эффективности производительности;

$M_{рент}$ – индекс эффективности рентабельности реализации;

$M_{прирост}$ – индекс эффективности темпа прироста реализации;

$M_{чп}$ – индекс эффективности чистой прибыли в расчете на человека.

Параметры приводились к сопоставимому виду через переход от системы абсолютных и относительных показателей к относительной системе индексов эффективности.

Так как нами исследовалась тенденция изменения индексов эффективности локальных и интегральных показателей, строились тренды и анализировались тенденции изменения показателей в динамике. При этом мы исходили из того, что наибольшую точность аппроксимации при построении тренда, описывающего тенденции развития предприятия, дает исследование нелинейных функций.

При построении интегральных показателей не учитывалось положение предприятия по объему реализации, так как априорно предполагалось, что объем реализации не является главным и единственным индикатором успешности деятельности предприятия. Проведя исследование на основе значительного статистического материала с использованием компьютерных технологий, в частности табличного процессора Excel[38], на основе корреляционно-регрессионного анализа нами выявлена достаточно низкая зависимость изменения индекса эффективности предприятия по введенным относительным частным и интегральным показателям от индекса эффективности реализации, следовательно, и от объема реализации.

Далее рассматривалась зависимость изменения положения предприятия в «Рейтинге» от следующих параметров:

- индекса эффективности локальных относительных показателей: производительности труда, рентабельности, темпа прироста реализации, прибыли после налогообложения в расчете на одного человека;
- индекса эффективности интегральных показателей, построенных по формулам (3.18–3.21).

В результате проведенного исследования были выявлены локальные показатели, оказывающие наиболее существенное влияние на построенный интегральный показатель I_{11}.

Для анализа по всем перечисленным выше направлениям с целью обоснования общей тенденции использовались несколько видов функций. Выбор того или иного вида функции делался с помощью значения коэффициента детерминации R^2. Гипотеза о нелинейности трендов локальных и интегральных показателей подтвердилась: значение R^2 для нелинейных функций выше, чем для линейной. Расчеты по всем годам сравнительных характеристик точности аппроксимации линейных и нелинейных трендов приведены ниже в таблицах. Согласно вычисленным значениям коэффициента детерминации R^2 полиномиальный тренд 6-й степени точнее представляет тенденцию интегральных показателей,

[38] Тюрин Ю. Н., Макаров А. А. Анализ данных на компьютере / под ред. В. Э. Фигурнова. 3-е изд, перераб. и доп. М.: Инфра-М, 2003. С. 44–46.

чем линейный. Остальные рассмотренные нами нелинейные функции также значительно превосходят по точности аппроксимации линейную. Исключение составляет экспоненциальная функция, которая даже уступает в точности описания линейной. Как показывает опыт большинства исследователей, среди нелинейной полиномиальной регрессии чаще всего используется парабола второй степени, в отдельных случаях – полином третьего порядка.

Ограничения в использовании полиномов более высоких степеней связаны с требованием однородности исследуемой совокупности: чем выше порядок полинома, тем больше изгибов имеет кривая и соответственно менее однородна совокупность по результативному признаку. Поэтому исследователь должен учитывать данный факт при выборе функции для аппроксимации. Например, парабола 2-й степени целесообразна к применению, если для определенного интервала значений фактора меняется характер связи рассматриваемых признаков: прямая связь меняется на обратную или наоборот. Если же исходные данные не обнаруживают изменения направленности связи, то параметры параболы 2-го порядка, а тем более полиномов более высоких степеней становятся трудно интерпретируемыми, форма связи часто заменяется другими нелинейными моделями.

Мы исходим из того, что наиболее простая функция поддается более точной экономической интерпретации. С учетом вышесказанного построенные функции будут проранжированы в порядке убывания предпочтений следующим образом:

- линейная;
- полиномиальная;
- степенная;
- экспоненциальная;
- логарифмическая.

Сравнительная характеристика точности аппроксимации трендов линейных и нелинейных функций для интегральных показателей за 1999–2003 гг. позволяет сделать следующие выводы:

1. Если исследователь выбирает надежность экономической интерпретации результатов, несмотря на усложнение функции, в данном случае наилучшим трендом является полином 6-й степени;

2. Если для нас более существенным является использование «простой» функции, несмотря на некоторую возможную неточность аппроксимации, то следует воспользоваться линейной функцией;

3. Точность аппроксимации трендов для интегральных показателей нелинейных функций в некоторых видах совпадает или близко по значениям. Тогда при выборе функции для построения тренда мы будем рассматривать следующие варианты:

- точность аппроксимации и сложность построения функции и экономической интерпретацией результатов,
- более простая функция и наглядная экономическая интерпретация результатов (если брать полиномы, то чем больше точек «перегибов», тем менее точна интерпретация), однако в ущерб точности аппроксимации тренда.

Значительный интерес для нас представлял вопрос выявления тенденции зависимости относительных локальных и интегральных показателей от объема реализации, то есть наличие связи между индексом эффективности предприятия по введенной нами системе показателей и индексом объема реализации. В случае обнаружения данной зависимости ставилась задача определения вида функции для построения тренда. В результате исследования:

1) выявлено, что практически отсутствует связь между изменением положения предприятия в «Рейтинге-200» и, следовательно, абсолютного показателя объема реализации и изменением введенных нами интегральных показателей.

Зависимость изменения интегральных показателей от объема реализации на основе показателя точности аппроксимации линейного и полиномиального трендов представлена в таблице. Даже выбранная нами как наиболее надежная функция аппроксимации – полином 6-й степени – показала, что под действием изменения места компании в «Рейтинге-200», т. е. под действием изменения объема реализации, можно объяснить лишь незначительные доли изменения показателей:

I_{11} в 2000 г. – 6,6 %, в 2001 г. –11,2 %, в 2002 г. – 3,7 %, в 2003 г. – 8,2 %;

I_2 в 2000 г. – 7,6 %, в 2001 г. – 12 %, в 2002 г. – 2,9 %, в 2003 г. – 6,2 %;

I_3 в 2000 г. – 3,9 %, в 2001 г. – 4,4 %, в 2002 г. – 2,9 %, в 2003 г. на 6,8 %;

I_4 в 2000 г. – 6 %, в 2001 г. – 11 %, в 2002 г. – 4,6 %, в 2003 г. – 8,2 %;

2) обнаружено совпадение значений коэффициента детерминации R^2 для некоторых нелинейных функций.

Так, в 2000 г. – для I_2 полиномы 3 и 4-й степеней имеют показатель R^2 = 0,0741-0,0742 и для логарифмической и степенной 0,0551-0,0505. Для I_3 у полиномов 2 и 3-й степеней показатель R^2 = 0,0222. Для I_4 полиномы 5 и 6-й степеней имеют показатель R^2 = 0,0602-0,0604, для логарифмической и степенной 0,0492-0,0505. Следует отметить, что показатель точности аппроксимации для I_4 полиномов 4, 5 и 6-й степеней также имеет близкие значение 0,0598; 0,0602; 0,0604 соответственно (табл. 3.9).

В 2001 г. – для I_{11} полиномы 4 и 5-й степеней имеют показатель R^2 = 0,1105-0,1109. Для I_2 полиномы 4 и 5-й степеней имеют показатель R^2 = 0,1177-0,1178. Прослеживается аналогия с 2000 г., так как показатель точности аппроксимации для I_4 полиномов 4, 5 и 6-й степеней также имеет близкие значение 0,1075; 0,1085;0,1092 соответственно (табл. 3.10).

В 2002 г. – для показателей I_2 и I_3 полиномы 6-й степени по точности совпадают R^2 =0,0287, полиномы 3 и 4-й степеней имеют показатель R^2 = 0,1178, следовательно, для анализа не существенно предпочтение построения полинома четвертой степени. Отметим очень низкую информативность построения линейного тренда по всем 4 интегральным показателям: для I_{11} R^2 = 0,0028; для I_2 R^2 = 0,0005; для I_3 R^2 = 0,0007; для I_4 R^2 =0,0041. Полином 6-й степени, несмотря на его лучшую информативность по сравнению с другими функциями, также дает низкие показатели коэффициента R^2: для I_{11} R^2 = 0,0374; для I_2 R^2 = 0,0287; для I_3 R^2 = 0,0287; для I_4 R^2 = 0,0455 (табл. 3.11).

Таблица 3.9

Значения R^2 для интегральных показателей в 2000 г.

Функция / *Показатели*	*Линейная*	*Полиномиальная степени n*					*Экспоненциальная*	*Логарифмическая*	*Степенная*
		2	*3*	*4*	*5*	*6*			
I_{11}	0,0367	0,0610	0,0618	0,0642	0,0649	0,0657	0,0359	0,0515	0,0540
I_2	0,0345	0,0699	0,0741	0,0742	0,0752	0,0759	0,0344	0,0551	0,0550
I_3	0,0181	0,0222	0,0222	0,0376	0,0386	0,0388	0,0134	0,0247	0,0189
I_4	0,0360	0,0564	0,0569	0,0598	0,0602	0,0604	0,0344	0,0492	0,0505

Таблица 3.10

Значения R^2 для интегральных показателей в 2001 г.

Функция / *Показатели*	*Линейная*	*Полиномиальная степени n*					*Экспоненциальная*	*Логарифмическая*	*Степенная*
		2	*3*	*4*	*5*	*6*			
I_{11}	0,0426	0,0915	0,0933	0,1105	0,1109	0,1116	0,0334	0,0831	0,0869
I_2	0,0388	0,1007	0,1045	0,1177	0,1178	0,1197	0,0294	0,0853	0,0810
I_3	0,0130	0,0343	0,0363	0,0395	0,0423	0,0439	0,0080	0,0239	0,0221
I_4	0,0430	0,0893	0,0900	0,1075	0,1085	0,1092	0,0337	0,0807	0,0868

Таблица 3.11

Значения R^2 для интегральных показателей в 2002 г.

Функция / *Показатели*	*Линейная*	*Полиномиальная степени n*					*Экспоненциальная*	*Логарифмическая*	*Степенная*
		2	*3*	*4*	*5*	*6*			
I_{11}	0,0028	0,0174	0,0198	0,0232	0,0279	0,0374	0,0034	0,0113	0,0087
I_2	0,0005	0,0029	0,0039	0,0117	0,0124	0,0287	0,0014	0,0011	0,0016
I_3	0,0007	0,0124	0,0178	0,0178	0,0279	0,0287	7E-0,5	0,0076	0,0034
I_4	0,0041	0,0241	0,0303	0,0319	0,0383	0,0455	0,0053	0,0171	0,0155

В 2003 г. – для показателя I_{11} полиномы 3 и 4-й степеней имеют одинаковый показатель R^2 = 0,0614. Для показателя I_2 полиномы 4 и 5-й степеней по точности совпадают R^2 =0,0485, следовательно, для анализа можно остановиться на полиноме 4-й степени. По сравнению с предыдущими годами повысилась информативность построения линейного тренда по всем 4 интегральным показателям: для I_{11} R^2 = 0,0462; для I_2 R^2 = 0,0330; для I_3 R^2 = 0,0341; для I_4 R^2 = 0,0451. Полином 6-й степени также повысил информативность по сравнению с другими функциями: показатели коэффициента для I_{11} R^2 = 0,0812; для I_2 R^2 = 0,0623; для I_3 R^2 = 0,0684; для I_4 R^2 = 0,0816 (табл. 3.12).

На основе проведенного исследования считаем возможным ограничиться двумя трендами для моделирования деятельности предприятия:

- линейным – как наиболее простым по построению и экономической интерпретации;

- полиномом 6-й степени – как наиболее точным из возможных для построения в табличном процессоре Excel.

Обобщив данные по отдельным годам исследования, построим табл. 3.13.

Используя значения интегральных показателей из таблицы выше можно сделать следующие существенные выводы о динамике показателей в период с 1999 по 2003 гг. По всему исследуемому периоду наблюдается достаточно низкая зависимость изменения интегральных показателей предприятия от его позиции в рейтинге крупнейших по реализации предприятий. Несмотря на то что применение полиномиального тренда дает более высокое значение R^2, чем использование линейного тренда, все же его значение невелико, а в 2002 г. практически не наблюдается связи между изменением объема реализации предприятия и вариацией интегральных показателей.

На следующем этапе исследования ставилась задача рассмотрения зависимости между изменением каждого из выбранных локальных и интегральных и вариации остальных участвующих в исследовании показателей (табл. 3.14–3.17).

Таблица 3.12

Значения R2 для интегральных показателей в 2003 г.

Функция	*Линей-ная*	*Полиномиальная степени n*					*Экспонен-циальная*	*Логарифми-ческая*	*Степен-ная*
Показатели		*2*	*3*	*4*	*5*	*6*			
I_{11}	0,0462	0,0569	0,0614	0,0614	0,0633	0,0812	0,0317	0,0619	0,0448
I_2	0,0330	0,0468	0,0471	0,0485	0,0485	0,0623	0,0263	0,0455	0,0338
I_3	0,0341	0,0358	0,0426	0,0515	0,0606	0,0684	0,0164	0,0373	0,0182
I_4	0,0451	0,0544	0,0608	0,0611	0,0639	0,0816	0,0275	0,0608	0,0428

Таблица 3.13

Зависимость интегральных показателей от объема реализации на основе показателя R2 в 2000–2003 гг.

Функция	*Годы*							
	2000		*2001*		*2002*		*2003*	
Показатели	*Линей-ный*	*Полином*	*Линей-ный*	*Полином*	*Линей-ный*	*Полином*	*Линей-ный*	*Полином*
I_{11}	0,0367	0,0657	0,0426	0,1116	0,0028	0,0374	0,0462	0,0812
I_2	0,0345	0,0759	0,0388	0,1197	0,0005	0,0287	0,0330	0,0623
I_3	0,0181	0,0388	0,0130	0,0439	0,0007	0,0287	0,0341	0,0684
I_4	0,0360	0,0604	0,0430	0,1092	0,0041	0,0455	0,0451	0,0816

Таблица 3.14

Зависимость локальных и интегральных показателей от индекса эффективности производительности в 2000 – 2003 гг.

Показатель R^2	Годы							
	2000		2001		2002		2003	
	Линейный	Полином	Линейный	Полином	Линейный	Полином	Линейный	Полином
I_{11}	0,5191	0,5217	0,4856	0,4964	0,5082	0,5101	0,5375	0,5466
I_2	0,0155	0,0518	0,3523	0,3759	0,2874	0,3166	0,3981	0,4194
I_3	0,0063	0,0296	0,3227	0,397	0,3683	0,4003	0,3407	0,3751
I_4	0,016	0,0383	0,5213	0,5314	0,5537	0,5586	0,5492	0,5641
Мпроизв	1	1	1	1	1	1	1	1
Мрентаб	0,0751	0,0867	0,0837	0,1246	0,1135	0,1415	0,1331	0,1715
Мприрост	0,2166	0,2265	0,0294	0,0582	0,0116	0,0374	0,0596	0,1227
МЧП/чел	0,3627	0,3738	0,429	0,4317	0,4537	0,4684	0,4113	0,4298
Мсп1	0,0728	0,0938	0,0257	0,0569	0,0175	0,0422	0,0369	0,0981

Таблица 3.15

Зависимость локальных и интегральных показателей от индекса эффективности рентабельности в 2000–2003 гг.

Показатель R^2	Годы							
	2000		2001		2002		2003	
	Линейный	Полином	Линейный	Полином	Линейный	Полином	Линейный	Полином
I_{11}	0,6339	0,639	0,5928	0,6012	0,5597	0,58	0,5713	0,6029
I_2	0,0371	0,0654	0,3783	0,3878	0,2493	0,2796	0,2977	0,3213
I_3	0,012	0,0287	0,3319	0,3922	0,3635	0,3824	0,461	0,4953
I_4	0,036	0,0669	0,6128	0,622	0,6078	0,6142	0,6127	0,6368
Мпроив	0,0751	0,0862	0,0837	0,1191	0,1135	0,1153	0,1331	0,1753
Мрентаб	1	1	1	1	1	1	1	1
Мприрост	0,1564	0,1644	0,0359	0,0759	0,0006	0,0215	0,0441	0,0932
МЧП/чел	0,7616	0,7766	0,6645	0,6983	0,7364	0,7572	0,643	0,698
Мсп1	0,0088	0,0718	0,0537	0,0710	0,0026	0,0178	0,0360	0,0534

Таблица 3.16

Зависимость локальных и интегральных показателей от индекса эффективности темпа прироста в 2000 – 2003 гг.

Показатель R^2	Годы							
	2000		2001		2002		2003	
	Линейный	Полином	Линейный	Полином	Линейный	Полином	Линейный	Полином
I_{11}	0,5072	0,5217	0,2604	0,3035	0,1952	0,2167	0,3559	0,3636
I_2	0,0059	0,0148	0,2741	0,3318	0,2601	0,3052	0,352	0,364
I_3	0,0073	0,033	0,3237	0,3615	0,2233	0,3297	0,3723	0,3991
I_4	0,0061	0,0141	0,2742	0,3028	0,1727	0,1959	0,3522	0,3626
Мпроив	0,2166	0,222	0,0294	0,0377	0,0116	0,0483	0,0596	0,0655
Мрентаб	0,1564	0,1804	0,0359	0,093	0,0006	0,0183	0,0441	0,0784
Мприрост	1	1	1	1	1	1	1	1
МЧП/чел	0,2382	0,25	0,0301	0,0750	0,0026	0,0283	0,1210	0,1321
Мсп1	0,0035	0,0376	0,0002	0,0360	0,0048	0,0249	0,0059	0,0288

Таблица 3.17

Зависимость локальных и интегральных показателей от индекса эффективности чистой прибыли на человека в 2000–2003 гг.

Показатель R^2	Годы							
	2000		2001		2002		2003	
	Линейный	Полином	Линейный	Полином	Линейный	Полином	Линейный	Полином
I_{11}	0,87750	0,8910	0,8282	0,8385	0,7902	0,7953	0,8362	0,8411
I_2	0,0431	0,0731	0,6009	0,6215	0,4136	0,4326	0,5546	0,5911
I_3	0,0188	0,0295	0,3181	0,3673	0,4082	0,4316	0,4522	0,4764
I_4	0,0449	0,0577	0,8137	0,8331	0,0233	0,8289	0,8449	0,8499
Мпроизв	0,3627	0,511	0,4290	0,5085	0,4537	0,5469	0,4113	0,4871
Мрентаб	0,7616	0,7699	0,6645	0,6950	0,7364	0,7459	0,6430	0,6722
Мприрост	0,2382	0,272	0,0301	0,0623	0,0026	0,0298	0,1210	0,1981
МЧП/чел	1	1	1	1	1	1	1	1
Мсп1	0,0336	0,0732	0,0408	0,0867	0,0045	0,0249	0,0338	0,0414

Кроме приведенного выше анализа изучалась и обратная зависимость, т. е. насколько вариация интегральных показателей ведет к изменению составляющих локальных параметров. Результаты исследования в динамике представлены в табл. 3.18, 3.19.

По итогам данных таблиц можно сделать следующие выводы:

1. Наблюдается значительное влияние как вариации интегрального показателя I_{11} под действием изменения рентабельности и чистой прибыли в расчете на одного человека, так и обратная тенденция: изменение данных локальных показателей приводит к значительному изменению показателя I_{11};

2. Промежуточное положение занимает взаимное влияние интегрального показателя I_{11} и локального производительности труда;

3. Минимальное изменение I_{11} наблюдается под действием вариации темпа прироста реализации;

4. При рассмотрении попарного воздействия локальных показателей друг на друга для линейной зависимости наблюдается одинаковая сила влияния.

Таким образом, при управлении деятельностью предприятия на основе исследования экономической устойчивости с помощью построенной модели следует учитывать различную степень влияния факторов на итоговый показатель. Построенная модель требует корректировки через введение весовых коэффициентов как для каждого года в отдельности, так и для общей модели управления экономической устойчивостью промышленного предприятия на основе интегрального показателя.

На основе анализа данных следует отметить, что в течение исследуемого временного промежутка подтвердилась ожидаемая зависимость между изменением среднего места компании – показатель I_4 и итоговым показателем I_{11}. Однако следует отметить, что характеризующий самые «сильные» стороны компании показатель I_3 гораздо в меньшей степени зависит от изменения I_{11}, чем самые «слабые» места – I_2. Тем не менее более половины изменения I_3 происходят за счет вариации I_{11}, поэтому при рассмотрении влияния локальных показателей на интегральные ограничимся одним интегральным показателем I_{11}.

Таблица 3.18

Зависимость локальных показателей от интегрального I_{11} за 2000–2003 гг.

Показатель R^2	Годы									
	1999		2000		2001		2002		2003	
	Линейный	Полином	Линейный	Полином	Линейный	Полином	Линейный	Полином	Линейный	Полином
Мпроизв	0,3689	0,411	0,4959	0,5348	0,4803	0,496	0,5094	0,5157	0,5177	0,5484
Мрентаб	0,6486	0,6578	0,643	0,6508	0,5843	0,5989	0,572	0,584	0,575	0,6003
Мприрост	0,3627	0,4062	0,4976	0,5248	0,2255	0,3102	0,1648	0,2639	0,333	0,3741
МЧП/чел	0,7746	0,7839	0,885	0,8883	0,8491	0,8505	0,8131	0,8209	0,8526	0,8597
Мсп1	0,0164	0,0334	0,0335	0,058	0,0428	0,0511	0,0025	0,0051	0,0441	0,0736

Таблица 3.19

Зависимость интегральных показателей от изменения I_{11} за 2000–2003 гг.

Показатель R^2	Годы							
	2000		2001		2002		2003	
	Линейный	Полином	Линейный	Полином	Линейный	Полином	Линейный	Полином
I_{11}	0,9876	0,999	0,9728	0,9988	0,9734	0,9985	0,9778	0,998
I_2	0,8444	0,9106	0,7096	0,8853	0,687	0,8232	0,7558	0,8631
I_3	0,5887	0,637	0,4396	0,5478	0,5175	0,5801	0,5745	0,649
I_4	0,9746	0,9837	0,9483	0,9736	0,9523	0,9655	0,962	0,9762

В табл. 3.20, 3.21 представлены показатели влияния изменения результата от фактора по линейной и нелинейной функции. Рассмотрев в динамике изменение интегрального показателя I_{11} в зависимости от локальных показателей, можно сделать следующие выводы по каждому году:

1. В 2000 г наиболее сильное изменение интегрального показателя I_{11} наблюдается под действием изменения чистой прибыли на человека – 87,8 и 89,1 % для линейной и нелинейной функций соответственно; рентабельности продаж – 63,4 и 63,9 % для линейной и нелинейной функций соответственно. Меньшее влияние оказывает изменение производительности – 51,9 и 52,2 % для линейной и нелинейной функций соответственно; изменение темпа прироста реализации– 50,7 и 52,2 %;

2. В 2001 г сохраняется тенденция 2000 г: наиболее сильное изменение интегрального показателя I_{11} наблюдается под действием вариации чистой прибыли на человека и рентабельности. Меньшее влияние оказывает изменение производительности; изменение темпа прироста;

3. В 2002 г сохраняется тенденция 2000 г: наиболее сильное изменение интегрального показателя I_{11} наблюдается под действием вариации чистой прибыли на человека – 79 и 79,5% для линейной и нелинейной функций соответственно; рентабельности – 57 и 58 % для линейной и нелинейной функций соответственно. Промежуточное по влиянию положение занимает изменение производительности – 52 % для линейной и нелинейной функций; наименьшее влияние на вариацию I_{11} оказало изменение темпа прироста – 19,5 и 21,7 % для линейной и нелинейной функций соответственно. Причем данное влияние по сравнению с 2000 г. уменьшилось примерно в 2,5 раза по обеим функциям. Обращает на себя внимание тот факт, что в отличие от остальных лет (как до 2002 г., так и после) нелинейная функция дает на целый порядок оценку надежнее, чем линейная;

4. В 2003 г продолжается отмеченная тенденция: наиболее сильное изменение интегрального показателя I_{11} наблюдается под действием вариации чистой прибыли на человека 83,6 и 84,1 % – для линейной и нелинейной функций соответственно;

рентабельности – 57 и 60,3 % для линейной и нелинейной функций соответственно. Промежуточное положение по влиянию оказало изменение производительности – 54 и 55 % для линейной и нелинейной функций соответственно; наименьшее влияние на вариацию I_{11} оказало изменение темпа прироста реализации – 35,6 и 36,4 % для линейной и нелинейной функций соответственно; причем данное влияние по сравнению с 2002 г. увеличилось примерно в 2 раза по обеим функциям.

В результате анализа данных таблиц представляется возможным проранжировать на основе R^2 локальные показатели по степени влияния на интегральный I_{11} в порядке убывания степени влияния следующим образом:

- чистая прибыль в расчете на одного человека;
- рентабельность реализации продукции;
- производительность труда;
- темп прироста реализации продукции.

Таблица 3.20

Изменение линейного тренда интегрального показателя $\mathbf{I_{11}}$ в зависимости от изменения локальных показателей

Показатель	*Годы*			
	2000	*2001*	*2002*	*2003*
Чистая прибыль на чел.	88,5	84,9	81,3	85,3
Рентабельность	64,3	58,4	57,2	57,5
Производительность	50	48	50,9	51,8
Темп прироста	49,8	22,6	16,5	33,3

Таблица 3.21

Изменение полиномиального тренда интегрального показателя $\mathbf{I_{11}}$ в зависимости от изменения локальных показателей

Показатель	*Годы*			
	2000	*2001*	*2002*	*2003*
Чистая прибыль на чел.	88,8	85,1	82,1	86,0
Рентабельность	65,1	59,9	58,4	60,0
Производительность	53,3	49,6	51,6	54,8
Темп прироста	52,5	31,0	26,4	37,4

Из вышесказанного следует, что для целей управления деятельностью предприятия на основе исследования экономической устойчивости через интегральный показатель может быть предложена следующая модель:

$$I_{11} = \sqrt{a_1(Мпроизв-1)^2 + a_2(Мрент-1)^2 + a_3(Мприрост-1)^2 + a_4(Мчп-1)^2}\ , \quad (3.22)$$

$$\sum_{i=1}^{n} a_i = 1,$$

где n – число наблюдений.

При анализе социально-экономического процесса на основе модели (3.19) получим для каждого исследуемого года периода 1999–2003 гг. модели с определенными коэффициентами при выбранных факторах (3.23):

1999 г.

$$I_{11} = \sqrt{0{,}226(Мпроизв-1)^2 + 0{,}094(Мрент-1)^2 + 0{,}235(Мприрост-1)^2 + 0{,}445(Мчп-1)^2}$$

2000г.

$$I_{11} = \sqrt{0{,}237(Мпроизв-1)^2 + 0{,}177(Мрент-1)^2 + 0{,}209(Мприрост-1)^2 + 0{,}378(Мчп-1)^2}$$

2001г.

$$I_{11} = \sqrt{0{,}175(Мпроизв-1)^2 + 0{,}154(Мрент-1)^2 + 0{,}260(Мприрост-1)^2 + 0{,}41(Мчп-1)^2}$$

2002 г.

$$I_{11} = \sqrt{0{,}180(Мпроизв-1)^2 + 0{,}144(Мрент-1)^2 + 0{,}282(Мприрост-1)^2 + 0{,}394(Мчп-1)^2}$$

2003 г.

$$I_{11} = \sqrt{0{,}234(Мпроизв-1)^2 + 0{,}196(Мрент-1)^2 + 0{,}256(Мприрост-1)^2 + 0{,}313(Мчп-1)^2}$$

В качестве индикатора степени устойчивости определен показатель I_{11}. Если менеджер может рассчитать приемлемые для предприятия границы изменения данного показателя, тогда, учитывая интересы той или иной группы (например, потенциального инвестора может интересовать темп прироста реализации или рентабельность продаж, акционера – чистая прибыль на человека и т. д.), руководство имеет возможность наращивать определенный частный показатель, параллельно отслеживая изменения индикатора I_{11}. При этом остальные составляющие I_{11}, не взятые менеджером в расчет, будут меняться, однако потребуется их «сдерживание» в допустимых для изменения интегрального показателя I_{11} пределах. Причем по выборкам по годам коэффициенты a_i различаются незначительно.

Таким образом, на основе R^2 можно осуществить спецификацию модели, т. е. подобрать оптимальную для целей управления функцию с помощью математической формулы модели.

Контрольные вопросы и задания

1. Насколько обоснованным является применение понятия случайная величина к анализу социально-экономического процесса?
2. Перечислите виды случайных величин и способы задания.
3. Назовите основные индикаторы качества модели.
4. В чем заключается проблема спецификации модели?
5. Насколько однозначно на практике решается проблема спецификации?
6. Проанализируйте алгоритм выбора функции для анализа социально-экономического процесса.

4. Элементы теории принятия статистических решений

4.1. Статистическая проверка гипотез

Проверка гипотез позволяет проверить свойства генеральной совокупности на основе выборочного наблюдения, т. е. свойства реального мира, имея лишь отдельные наблюдения

Целью регрессионного анализа, так же как и любого статистического исследования, не является поиск истинных значений. Мы практически никогда не обладаем генеральной совокупностью наблюдений, а имеем лишь выборку из нее. Поэтому достижение правды, с точки зрения эконометриста, – построение генеральной регрессии (population regression). Но поскольку чаще всего это невозможно, нам приходится довольствоваться теми регрессиями, которые не противоречат построенным нами теориям (моделям)[39].

[39] Аистов А. В., Максимов А. Указ. соч. С. 33.

Математический аппарат проверки гипотез позволяет выявить такие регрессии, а также проверить статистическую значимость регрессии, т. е. отбросить те регрессии, которые по имеющимся в нашем распоряжении наблюдениям не позволяют судить о наличии взаимосвязей между событиями окружающего мира. Иногда построение значимой, но дающей совершенно неожиданные результаты регрессии является толчком для развития экономической теории[40].

Во многих случаях результаты наблюдений используются для проверки предположений (гипотез) относительно тех или иных свойств распределения генеральной совокупности. В частности, такого рода задачи возникают при сравнении различных технологических процессов или методов обработки по определенным измеряемым признакам, например по точности, производительности и т. д. *Оценку генерального параметра* получают на основе выборочного показателя с учетом ошибки репрезентативности. *Ошибка выборки* – это разница между значениями показателя, полученного по выборке, и генеральным параметром. В другом случае в отношении свойств генеральной совокупности выдвигается некоторая гипотеза о величине средней, дисперсии, характере распределения, форме и тесноте связи между переменными. Проверка гипотезы осуществляется на основе выявления согласованности эмпирических данных с гипотетическими (теоретическими). *Если расхождение между сравниваемыми величинами не выходит за пределы случайных ошибок, гипотезу принимают.* При этом не делается никаких заключений относительно правильности самой гипотезы, речь идет лишь *о согласованности сравниваемых данных*. Основой проверки статистических гипотез являются данные случайных выборок. При этом безразлично, оцениваются ли гипотезы в отношении реальной или гипотетической генеральной совокупности. Последнее открывает путь применения этого метода за пределами собственно выборки: при анализе результатов эксперимента данных сплошного наблюдения, но малой численности. В этом случае рекомендуется проверить, не вызвана ли установленная закономерность стечением случай-

[40] Аистов А. В., Максимов А. Указ. соч. С. 33.

ных обстоятельств, насколько она характерна для того комплекса условий, в которых находится изучаемая совокупность.

Пусть X – наблюдаемая дискретная или непрерывная случайная величина. *Статистической гипотезой* (обозначается H) называется произвольное предположение относительно параметров или вида распределения случайной величины X. В общем случае статистической гипотезой называется предположение о свойстве генеральной совокупности, которое проверяется на основе данных выборки. Так может быть выдвинута гипотеза о том, что средняя μ в генеральной совокупности равна некоторой величине a (записывается $H: \mu = a$) или о том, что генеральная средняя больше некоторой величины $H : \mu > в.$

Различают *простые* и *сложные* гипотезы. Гипотеза называется простой, если она однозначно характеризуется параметром распределения случайной величины. Например, $H: \mu = a.$ Гипотеза называется сложной, если она состоит из конечного или бесконечного числа простых гипотез, при этом указывается некоторая область вероятных значений параметра. Например, $H: \mu > в.$ Эта гипотеза состоит из множества простых гипотез $H: \mu = c,$ где c – любое число большее $в.$ Другим примером простой гипотезы является предположение о том, что случайная величина X распределена по нормальному закону $N\ (0;1)$. Если же высказывается предположение, что случайная величина X имеет нормальное распределение $N\ (m;1)$, где $a \le m \le b$, то это сложная гипотеза.

Часто распределение случайной величины X известно, и по выборке наблюдений необходимо проверить предположения о значении параметров этого распределения. Гипотезы о параметрах генеральной совокупности называются *параметрическими*, о распределениях – *непараметрическими*.

Гипотеза о том, что две совокупности, сравниваемые по одному или нескольким признакам, не отличаются, называется нулевой гипотезой или *нуль-гипотезой* (обозначается H_0). При этом предполагается, что действительное различие сравниваемых величин равно нулю, а выявленное по данным отличие от нуля носит случайный характер. Например, H_0: $\mu_1 = \mu_2$ и т. д.

Нулевой гипотезой назначается проверяемая гипотеза. Наряду с гипотезой *Ho* рассматривают одну из альтернативных (конкурирующих) гипотез H_1. Альтернативная гипотеза H_1 может быть сформулирована по-разному в зависимости от того, какие отклонения от гипотетической величины нас особенно беспокоят: положительные, отрицательные либо и те и другие. Соответственно, альтернативные гипотезы могут быть записаны как:

$H_1: \mu > a, H_1: \mu < a, H_1: \mu \neq a.$

Например, если проверяется гипотеза о равенстве параметра θ некоторому заданному значению θ_0 , т. е. $H_0: \theta = \theta_0$,то в качестве альтернативной гипотезы можно рассмотреть одну из следующих гипотез:

- $H_1: \theta > \theta_0$
- $H_1: \theta < \theta_0$
- $H_1: \theta \neq \theta_0$
- $H_1: \theta = \theta_1$, где θ_1 – заданное значение, $\theta_1 \neq \theta_0$.

Выбор альтернативной гипотезы определяется конкретной формулировкой задачи. Правило, по которому принимается решение принять или отклонить гипотезу *Ho*, называется критерием. *Статистическим критерием* называют правило, устанавливающее условия отклонения проверяемой нулевой гипотезы.

Так как решение принимается на основе выборки наблюдений случайной величины *X*, необходимо выбрать подходящую статистику, называемую в этом случае статистикой Z критерия. При проверке простой параметрической гипотезы $H_0: \theta = \theta_0$ в качестве статистики критерия выбирают ту же статистику, что и для оценки параметра незначительно, т. е. $\tilde{\theta}$.

Нулевая гипотеза отвергается тогда, когда по выборке получается результат, который при истинности выдвинутой нулевой гипотезы маловероятен. Перед анализом выборки фиксируется некоторая малая вероятность α, называемая уровнем значимости. Границей невозможного или маловероятного обычно считают α = 0,05 (5 %); 0,01 (1 %);0,001 (10 %). Если ориентироваться на правило «трех сигма», которое дает соотношение: $\sigma = 1/6(x_{max}-x_{min})$, то, так как в нормальном распределении в размахе вариации «укладывается» 6σ (±3σ)), вероятность ошибки α должна быть равна 0,0027. Однако для этого уровня вероятности ошибки значений критериев редко табулируются: как правило, значения критериев в статистико-математических таблицах рассчитаны для вероятностей ошибки 0,05; 0,01; 0,001. Правила работы с таблицами рассмотрены в гл. 4.2.

Проверка статистических гипотез состоит из следующих этапов:

- формулируется в виде статистической гипотезы задача исследования;
- выбирается статистическая характеристика гипотезы;
- выбираются испытуемая и альтернативная гипотезы на основе анализа возможных ошибочных явлений и их последствий;
- определяется область допустимых значений, критическая область, а также критическое значение статистического критерия (t; F; χ^2) по соответствующей таблице;
- вычисляется фактическое значение статистического критерия;
- проверяется гипотеза на основе сравнения фактического и критического значений критерия, и в зависимости от результатов проверки гипотеза либо отклоняется, либо нет.

Методика проверок гипотез (формулировка нулевой и альтернативной гипотез, выбор правила принятия решения, уровня значимости и т. п.) достаточно подробно освещена в литературе. Существуют определенные положения, значительно упрощающие весь алгоритм проверки гипотез.

1. В качестве нулевой гипотезы исследователи обычно выбирают ту гипотезу, которая, скорее всего, будет отвергнута. В этом случае вид альтернативной гипотезы должен быть как можно ближе к тому, что предсказывает теория.

2. Альтернативная гипотеза в форме неравенства (например, H_1: $\beta_j > \beta_0$) тестируется односторонним тестом, а в форме «не равно» (например, H_1: $\beta_j \neq \beta_0$) – двусторонним.

3. В результате проверки нулевая гипотеза может быть либо отвергнута, либо имеющиеся наблюдения не дают оснований ее отвергнуть. Следует отметить, что фразу «не дают основания отвергнуть» некорректно заменять фразой «принимаем нулевую гипотезу».

4. Следующие версии нулевой гипотезы при заданной H_1 тестируются одинаково (односторонним тестом):

$$H_0: \beta_j = 0,\ H_1: \beta_j < 0;$$
$$H_0: \beta_j \geq 0,\ H_1: \beta_j < 0.$$

При проверке гипотез по одному из критериев возможны два ошибочных решения:

1) неправильное отклонение H_0: ошибка 1-го рода;

2) неправильное принятие H_0: ошибка 2-го рода.

В то время как фактически H_0 верна (1) и H_0 не верна (2), принимают 2 ошибочных решения:

- H_0 отклоняется и принимается альтернативная гипотеза;
- H_0 не отклоняется.

Если, например, установлено, что новое минеральное удобрение лучше, хотя на самом деле его действие не отличается от старого, то это ошибка 1-го рода. Если мы решили, что оба вида удобрения одинаковы, то допущена ошибка 2-го рода.

Вероятности, соответствующие неверным решениям, называются риском 1 и риском 2. Риск 1 равен вероятности ошибки α (уровню значимости), риск 2 равен вероятности ошибки β. Поскольку α всегда больше 0, то всегда есть риск ошибки β. Обычно задают значение α и пытаются сделать β возможно малым. Вероятность $1\text{-}\beta$ называется *мощностью критерия*: чем она больше, тем меньше вероятность *ошибки 2-го рода* (табл. 4.1).

Многие экономические показатели имеют нормальный или близкий к нормальному закон распределения. Например, доход населения, прибыль фирм в отрасли, объем потребления и т. д. имеют близкое к нормальному распределение. Нормальное распределение используется при проверке различных гипотез в статистике (о величине математического ожидания при известной дисперсии, о равенстве математических ожиданий и т. д.). Зачастую при моделировании экономических процессов приходится рассматривать случайную величину, которая представляют собой алгебраическую комбинацию нескольких случайных величин СВ. При этом желательно иметь возможность прогнозирования поведения таких СВ. Существенную роль в этом играет ряд специально разработанных теоретических законов распределений. К ним относятся распределения *Стьюдента* и *Фишера*.

Таблица 4.1

Гипотеза H_0	*Решение*	*Вероятность*	*Примечание*
Верна	Принимается	1–α	Доверительная вероятность
	Отвергается	α	Вероятность ошибки первого рода
Неверна	Принимается	β	Вероятность ошибки второго рода
	Отвергается	1–β	Мощность критерия

Пусть случайная величина $U \sim N(0;1)$), случайная величина V – независимая от U величина, распределенная по закону χ^2 с n степенями свободы. Тогда величина

$$T = \frac{U}{\sqrt{V/n}} \qquad (4.1)$$

имеет распределение *Стьюдента (t-распределение)* с n степенями свободы. Из формулы (4.1) очевидно, что распределение Стьюдента определяется только одним параметром n – числом степеней свободы. График функции плотности вероятности случайной величины, имеющей распределение Стьюдента, является симметричной кривой (линия симметрии – ось ординат), $M(T)=0$; $D(T)=n/n\text{-}2$ (см. рис. 4.1).

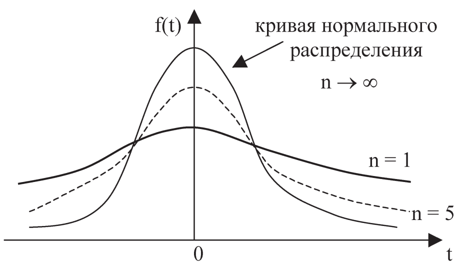


Рис. 4.1. Графики функции плотности вероятности случайной величины с распределением Стьюдента

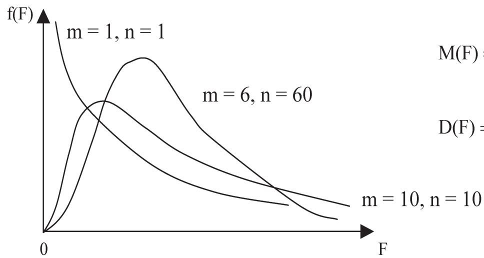
С увеличением числа степеней свободы распределение Стьюдента приближается к стандартизированному нормальному; при $n > 30$ распределение Стьюдента практически заменяется нормальным распределением. Распределение Стьюдента применяется для нахождения интервальных оценок, а также при проверке статистических гипотез. При этом используется таблица критических точек распределения Стьюдента.

Кроме распределения Стьюдента крайне популярным в анализе социально-экономических процессов является F-распределение Фишера.

Пусть V и W – независимые случайные величины, распределенные по закону χ^2 со степенями свободы $k_1 = m$ и $k_2 = n$ соответственно. Тогда величина, рассчитанная по формуле (4.2) имеет распределение Фишера со степенями свободы $k_1 = m$ и $k_2 = n$.

$$F = \frac{V/m}{W/n}. \tag{4.2}$$

Таким образом, распределение Фишера определяется двумя параметрами – числами степеней свободы. Статистики используют термин *степени свободы* (degrees of freedom, df) в качестве доверия результатам тестирования. Хотя на интуитивном уровне необходимость оценки степеней свободы очевидна, формулы, задающей это соотношение, нет. Под числом степеней свободы понимается число вариантов свободного варьирования признака. Оно равно числу наблюдений с учетом корректировки на количество независимых переменных и свободный член. Так, для модели парной регрессии число степеней свободы (n-2).



$$M(F) = \frac{n}{n-2} \quad (n > 2),$$

$$D(F) = \frac{2n^2(m+n-2)}{m(n-2)^2(n-4)} \quad (n > 4).$$

Рис. 4.2. Распределения Фишера
с различным числом степеней свободы

При больших m и n это F-распределение приближается к нормальному. Распределение Фишера используется при проверке статистических гипотез, в дисперсионном и регрессионном анализах. При этом используется таблица критических точек распределения Фишера.

4.2. Оценка значимости уравнения регрессии и отдельных его параметров

После построения уравнения линейной регрессии проводится оценка значимости как уравнения в целом, так и отдельных его параметров. Первая задача решается с помощью *F-критерия* Фишера, вторая – *t-критерия* Стьюдента.

Проверить значимость уравнения регрессии – это установить, соответствует ли математическая модель, выражающая зависимость между переменными, экспериментальным данным и достаточно ли включенных в уравнение объясняющих переменных (одной или нескольких) для описания зависимой переменной.

Проверка статистических гипотез на основе *F-критерия* Фишера и *t-критерия* Стьюдента состоит из следующих этапов:

- формулируется в виде статистической гипотезы задача исследования;
- выбираются испытуемая и альтернативная гипотезы на основе анализа возможных ошибочных явлений и их последствий;
- определяется область допустимых значений, критическая область, а также критическое значение статистического критерия (F, t) по соответствующей таблице;
- вычисляется фактическое значение статистического критерия: *F-статистика* или *t-статистика*;
- проверяется гипотеза на основе сравнения фактического и критического значений критерия, и в зависимости от результатов проверки гипотеза либо отклоняется, либо нет.

Оценка значимости уравнения регрессии в целом дается с помощью *F-критерия* Фишера. При этом выдвигается нулевая гипотеза, что коэффициент регрессии равен нулю, следовательно, фактор x не оказывает влияния на результат y.

Величина *F-отношения* (*F-критерий*) получается при сопоставлении факторной и остаточной дисперсии в расчете на одну степень свободы[41].

$$F = D_{факт} / D_{ост}. \qquad (4.3)$$

F-критерий проверки для нулевой гипотезы

$$H_0: D_{факт} = D_{ост.} \qquad (4.4)$$

Если нулевая гипотеза справедлива, то факторная и остаточная дисперсии не отличаются друг от друга. Для H_0 необходимо опровержение, если факторная дисперсия превышает остаточную в несколько раз.

Английским статистиком Снедекором разработаны таблицы критических значений *F-отношений* при разных уровнях существенности нулевой гипотезы и различном числе степеней свободы. Табличное значение *F-критерия* – это максимальная величина отношения дисперсий, которая может иметь место при случайном их расхождении для данного уровня вероятности наличия нулевой гипотезы.

Вычисленное значение *F-отношения* признается достоверным (отличным от 1), если оно больше табличного. В этом случае нулевая гипотеза об отсутствии связи признаков отклоняется и делается вывод о существенности этой связи: если $F_{факт} > F_{табл}$, то H_0 отклоняется.

Если же величина оказалась меньше табличной $F_{факт} \leq F_{табл}$, то вероятность нулевой гипотезы меньше заданного уровня (например, 0,05) и она не может быть отклонена без серьезного риска сделать неправильный вывод о наличии связи.

Проверка значимости уравнения регрессии производится на основе дисперсионного анализа[42]. В математической статистике дисперсионный анализ рассматривается как самостоятельный инструмент (метод) статистического анализа. В эконометрике же он применяется как вспомогательное средство для изучения качества модели. Центральное место в анализе дисперсии занимает разложение общей суммы квадратов отклонений переменной y от среднего значения $\bar{y}$ на 2 части – «объясненную» и «необъясненную» и может быть представлена следующим образом[43]:

[41] Гмурман В. Е. Теория вероятностей и математическая статистика. М.: Высшая школа, 1977. С. 290.

[42] Там же. С. 349.

[43] Доугерти К. Указ. соч. С. 57.

$$\sum_{i=1}^{n}(y_i-\bar{y})^2=\sum_{i=1}^{n}(\hat{y}_i-\bar{y})^2+\sum_{i=1}^{n}(y_i-\hat{y}_i)^2 \tag{4.5}$$

где $\hat{y}_i = a + bx_i$

$$Q = Q_R + Q_e \tag{4.6}$$

где Q – общая сумма квадратов отклонений; Q_R – сумма квадратов отклонений, обусловленная регрессией; Q_e – остаточная сумма квадратов отклонений.

$$Q=\sum_{i=1}^{n}(y_i-\bar{y})^2=\sum_{i=1}^{n}y_i^2-n\bar{y}^2 \tag{4.7}$$

$$Q_R=\sum_{i=1}^{n}(\hat{y}_i-\bar{y})^2=b^2\sum_{i=1}^{n}(x_i-\bar{x})^2 \tag{4.8}$$

$$Q_e=\sum_{i=1}^{n}(y_i-\hat{y}_i)^2 \tag{4.9}$$

Таблица 4.2

Схема дисперсионного анализа

Компоненты дисперсии	*Сумма квадратов*	*Число степеней свободы*	*Средние квадраты*
Регрессия	$Q_R=\sum_{i=1}^{n}(\hat{y}_i-\bar{y})^2$	$m-1$	$D_{факт}=s_R^2=\dfrac{Q_R}{m-1}$
Остаточная	$Q_e=\sum_{i=1}^{n}(y_i-\hat{y}_i)^2$	$n-m$	$D_{ост}=s^2=\dfrac{Q_R}{n-m}$
Общая	$Q=\sum_{i=1}^{n}(y_i-\bar{y})^2$	$n-1$	

Средние квадраты s_R^2 и s^2 представляют собой несмещенные оценки зависимой переменной, обусловленные соответственно регрессией или объясняющей переменной x и воздействием неучтенных случайных факторов и ошибок; m – число оцениваемых параметров регрессии, n – число наблюдений.

При отсутствии линейной зависимости между зависимой и объясняющей(ими) переменной случайные величины s_R^2 и s^2 имеют χ^2– *распределение* соответственно с (m-*1*) и (n-m) степенями свободы, а их отношение – *F-распределение* с теми же степенями свободы. Поэтому уравнение регрессии значимо на уровне α если фактически наблюдаемое значение статистики больше $F\ \alpha,\ k_1,k_2$:

$$F = \frac{Q_R(n-m)}{Q_e(m-1)} = \frac{s_R^2}{s^2} > F_{\alpha;k1;k2} \qquad (4.10)$$

где $F\ \alpha;\ k_1;\ k_2$ – табличное значение *F-критерия* Фишера, определенное на уровне значимости α при $k1 = m\text{-}1$ и $k2 = n\text{-}m$ числе степеней свободы.

Учитывая смысл величин s_R^2 и s^2, можно сказать, что значение F показывает, в какой мере регрессия лучше оценивает значение зависимой переменной по сравнению с ее средней.

В случае парной линейной регрессии $m = 2$, и уравнение регрессии значимо на уровне α, если

$$F = \frac{Q_R(n-2)}{Q_e} > F_{\alpha;1;n-2} \qquad (4.11)$$

Если известен коэффициент детерминации R^2, то критерий значимости уравнения регрессии или самого коэффициента детерминации может быть записан в виде

$$F = \frac{R^2(n-m)}{(1-R^2)(m-1)} > F_{\alpha;k1;k2} \qquad (4.12)$$

В линейной регрессии обычно оценивается значимость не только уравнения в целом, но и отдельных его параметров. С этой целью по каждому из параметров определяется его стандартная ошибка, называемая стандартной ошибкой коэффициента.

Оценки истинных, но неизвестных значений параметров – это числа, зависящие от количества и состава наблюдений, т. е. от выборки. При различных выборках мы получили бы различные оценки. Если продолжать брать все больше выборок и получать дополнительные оценки, то оценки каждого параметра будут соответствовать некоторому распределению вероятностей, которое может быть суммировано как среднее и мера дисперсии, следовательно, сравниваемые параметры распределены нормально. Нормальное распределение имеет следующее свойство: область, находящаяся в пределах 1,96 стандартного отклонения от его среднего значения составляет 95 % всей области[44]. Учитывая

[44] Доугерти К. Указ. соч. С. 78.

это, можно указать такой интервал вокруг оценки параметра, что с вероятностью 95 % истинное значение параметра лежит внутри этого интервала. Данный интервал, называемый 95 %-ным доверительным интервалом, определяется так:

$$b \pm 1{,}96 \textit{ среднего квадратического отклонения от } b. \quad (4.13)$$

Можно проверить гипотезу о том, что истинное значение параметра равно нулю, изучая ее *t-статистику*, которая определяется следующим образом:

$$t = \frac{b}{\textit{стандартная ошибка } b} \quad (4.14)$$

Областью принятия гипотезы, или *областью допустимых значений*, называется множество возможных значений статистического критерия, при которых основная гипотеза принимается. Если наблюдаемое значение статистического критерия, рассчитанное по данным выборочной совокупности, принадлежит *критической области,* то основная гипотеза отвергается. Если наблюдаемое значение статистического критерия принадлежит области принятия гипотезы, то основная гипотеза принимается.

В ряде прикладных задач требуется оценить значимость коэффициента корреляции r. При этом исходят из того, что при отсутствии корреляционной связи *t-статистика*, найденная по формуле:

$$t = r_{xy}\frac{\sqrt{n-2}}{\sqrt{1-r_{xy}^{2}}} \quad (4.15)$$

имеет *t-распределение* Стьюдента с (n-2) степенями свободы.

Коэффициент корреляции r_{xy} значим на уровне α (иначе гипотеза H_0 о равенстве генерального коэффициента корреляции нулю отвергается), если:

$$|t| = |r_{xy}|\frac{\sqrt{n-2}}{\sqrt{1-r_{xy}^{2}}} > t_{1-\alpha;n-2} \quad (4.16)$$

где $t_{1-\alpha;n-2}$ – табличное значение *t-критерия* Стьюдента, определенное на уровне значимости α при числе степеней свободы (n-2).

Процедура оценивания существенности коэффициента корреляции не отличается от рассмотренной выше для коэффициента регрессии: вычисляется значение *t-критерия*, его величина сравнивается с табличным значением при (*n-2*) степенях свободы.

Проверка гипотез о значимости коэффициентов регрессии и корреляции равносильна проверке гипотезы о существенности линейного уравнения регрессии.

Выполним проверку отдельных параметров уравнения регрессии на основе модели по сельскому хозяйству Ярославской области, приведенной в гл. 3.1. Рассмотрим на практике возможность варианта проверки несколько отклоняющейся от предложенного в теории алгоритма.

Получена модель следующего вида:

$$Y = 72,7 \cdot K^{-0,05} L^{0,84} I^{0,21}, \tag{4.17}$$

$R^2 = 0,99.$

Отметим высокое качество модели, так как на основе R^2 99 % результата объясняется влиянием суммарного изменения факторов. На основе *F-критерия* Фишера получим статистически значимое уравнение. Однако опытный исследователь сразу обратит внимание на то, что, в отличие от моделей промышленности, в модели сельского хозяйства параметр при факторе капитал имеет отрицательную степень.

Оценим надежность и статистическую значимость параметров β и α (табл. 4.3).

Таблица 4.3

Данные статистического анализа

статистики	β	α	*lgA*	γ	*A*
коэффициенты	0,840994	-0,05572	1,861402241	0,214723	72,67788
стандартные ошибки	0,033934	0,104214	0,129396946		
R^2	0,993539	0,0203	#Н/Д		
F	615,1371	8	#Н/Д		
суммы квадратов	0,507003	0,003297	#Н/Д		

Несмотря на то что получено отрицательное значение параметра при факторе «капитал», модель может использоваться на практике, так как следующие допустимые погрешности параметров:

$\Delta(\beta)=0{,}03$; $\Delta(\alpha)=0{,}1$.

Коэффициент регрессии по *t-критерию* Стьюдента будет статистически значим. Значительное превышение погрешности $\Delta(\alpha)=0{,}1$ над полученным значением параметра $\alpha=-0{,}05$ фактически означает, что в рамках данной модели этот параметр равен нулю: $\alpha=0$. Это можно интерпретировать как независимость объёма производства от капитала. В условиях российского сельского хозяйства подобный результат не вызывает значительных возражений и находится в согласии как с исторически сложившимися социально-экономическими реалиями, так и с общецивилизационными соображениями относительно хозяйства Северной Евразии.

Контрольные вопросы и задания

1. Дайте определение статистической гипотезы. Приведите примеры.
2. Что означает уровень значимости критерия? Чем он определяется? Как связан уровень значимости с характеристиками случайной величины?
3. Как определяется доверительный интервал? В чем его экономический смысл?
4. Что общего в методике построения доверительных интервалов и проверки статистических гипотез?
5. Поясните смысл понятий «ошибка первого рода», «ошибка второго рода», «мощность критерия».
6. Приведите алгоритм выполнения статистических критериев.
7. Сформулируйте, в чем различие между *F-критерием* Фишера и *t-критерием* Стьюдента?

Список литературы

1. Аистов, А. В. Эконометрика шаг за шагом: учеб. пособие для вузов / А. В. Аистов, А. Г. Максимов. – М.: Изд. дом ГУВШЭ, 2006.

2. Арженовский, С. В. Статистические методы прогнозирования. учеб. пособие / С. В. Арженовский, И. Н. Молчанов; Рост. гос. экон. ун-т. – Ростов-н/Д., – 2001.

3. Вентцель, Е. С. Теория вероятностей: учебник для вузов / Е. С. Вентцель. – 9-е изд. – М.: Академия, 2003. – 576 с.

4. Гмурман, В. Е. Теория вероятностей и математическая статистика / В. Е. Гмурман. – М.: Высшая школа, 1977.

5. Зеткина О. В. Совершенствование управления экономической устойчивостью промышленного предприятия на основе интегрального показателя: дис. ... канд. экон. наук / О. В. Зеткина. – Ярославль, 2004.

6. Доугерти, К. Введение в эконометрику / К. Доугерти; пер. с англ. – М.: ИНФРА-М, 1999.

7. Леоненков, А. В. Решение задач оптимизации в среде MS Excel / А. В. Леоненков. – СПб.: БХВ-Петербург, 2005.

8. Маленво, Э. Статистические методы эконометрии / Э. Маленво; пер. с фр. – 1976. – Вып. 2.

9. Россия и ВТО: региональный уровень оценки последствий интеграции / под ред. Д. С. Лебедева – Ярославль, 2013.

10. Тюрин, Ю. Н. Анализ данных на компьютере / Ю. Н. Тюрин, А. А. Макаров; под ред. В. Э. Фигурнова. – 3-е изд, перераб. и доп. – М.: Инфра-М, 2003.

11. Уэйн, Л. Винстон. Microsoft Excel: анализ данных и построение бизнес моделей / Л. Винстон Уэйн; пер. с англ. – М.: Русская Редакция, 2005.

12. Шикин, Е. В. Математические методы и модели в управлении: учеб. пособие / Е. В. Шикин, А. Г. Чхартишвили. – 3-е изд. – М.: Дело, 2004. (Серия «Классический университетский учебник».)

13. Frisch, O. R. The Responsibility of the Econometrician // H Econometrica / O. R. Frisch. – 1946. – Vol. 14, № 1. – P. 1–4 (Arrow K. J. The Work of Ragan Frisch, Econometrician // Econometrica. – Vol. 28, № 2).

14. ЭСКО: Электронный журнал энергосервисной компании «Экологические системы». – 2009. – № 2, февраль. – URL:http: // esco.co.ua/journal/2009_2/art067.htm

15. URL:http://www.cpe.une.edu/projects/rlms

16. URL:http://www.cpe.une.edu/projects/rlms

17. URL:http://www.egartech.ru/fields/derivatives/riskfactors

Приложение

Таблица

Личные потребительские расходы населения США (млрд долл., в ценах 1972 г.)

Категория	*Статья*	*Переменная*	*1959*	*1960*	*1961*	*1962*	*1963*	*1964*	*1965*
	Дата	DATE	1959,0	1960,0	1961,0	1962,0	1963,0	1964,0	1965,0
	Время	TIME	1,0	2,0	3,0	4,0	5,0	6,0	7,0
ДОХОД	Личный доход	PI	544,9	559,7	575,4	602,0	622,9	658,0	700,4
	Личный располагаемый доход	DPI	479,7	489,7	503,8	524,9	542,3	580,8	616,3
РАСХОДЫ	Совокупные личные расходы	TPE	440,4	452,0	461,4	482,0	500,5	528,0	557,5
Текущие расходы	Питание	FOOD	99,7	100,9	102,5	103,5	104,6	108,8	113,7
	Одежда	CLOT	36,3	36,6	37,3	38,9	39,6	42,6	44,2
	Бензин	GASO	13,7	14,2	14,3	14,9	15,3	16,0	16,8
	Моторное масло	FUEL	5,2	5,0	4,7	4,7	4,9	5,2	5,5
	Табак	TOB	10,7	10,9	11,2	11,2	11,4	11,3	11,6
	Косметика	COSM	3,1	3,5	3,9	4,2	4,5	4,8	5,3
	Лекарства	PHAR	3,5	3,9	4,3	4,7	4,9	5,1	5,3

Услуги	Жилье	HOUS	60,9	64,0	67,0	70,7	74,0	77,4	81,6
	Газ	GAS	3,9	4,1	4,3	4,7	4,9	5,1	5,3
	Вода	WAT	2,0	2,2	2,3	2,5	2,7	2,8	2,9
	Телефон	TELE	4,7	5,0	5,4	5,7	6,1	6,6	7,3
	Местный транспорт	LOCT	3,9	3,9	3,6	3,6	3,5	3,4	3,3
	Воздушный транспорт	AIR	0,9	0,9	1,0	1,1	1,2	1,4	1,6
	Медицинские услуги	DOC	8,8	9,0	9,1	9,8	10,2	11,9	12,1
	Услуги стоматологов	DENT	3,2	3,2	3,3	3,5	3,4	3,9	4,0
	Отдых	REC	9,6	10,0	10,4	10,9	11,3	11,6	11,9
	Частное образование	PRIV	5,6	6,0	6,3	6,6	7,0	7,4	8,1
Товары длительного пользования	Кухонное оборудование	KIT	4,2	4,2	4,2	4,4	4,6	5,1	5,2
	Посуда	TAB	2,6	2,5	2,5	2,6	2,5	2,8	3,1
	Ювелирные изделия	JEWL	2,2	2,2	2,2	2,3	2,5	2,6	2,9

Продолжение таблицы

Категория	*Статья*	*Переменная*	*1966*	*1967*	*1968*	*1969*	*1970*	*1971*	*1972*
	Дата	DATE	1966,0	1967,0	1968,0	1969,0	1970,0	1971,0	1972,0
	Время	TIME	8,0	9,0	10,0	11,0	12,0	13,0	14,0
ДОХОД	Личный доход	PI	740,6	774,4	816,2	853,5	876,8	900,0	951,4
	Личный располагаемый доход	DPI	646,8	673,5	701,3	722,5	751,6	779,2	810,3
РАСХОДЫ	Совокупные личные расходы	TPE	585,7	602,7	634,4	657,9	672,1	696,8	737,1
Текущие расходы	Питание	FOOD	116,6	118,6	123,4	125,9	129,4	130,0	132,4
	Одежда	CLOT	49,6	46,9	49,0	50,0	49,4	51,8	55,4
	Бензин	GASO	17,8	18,4	19,9	21,4	22,9	24,2	25,4
	Моторное масло	FUEL	5,6	5,6	5,3	5,0	4,7	4,6	5,0
	Табак	TOB	11,7	11,8	11,7	11,4	11,7	11,8	12,2
	Косметика	COSM	5,9	6,3	6,6	6,8	7,0	7,1	7,4
	Лекарства	PHAR	5,5	5,8	6,4	7,0	7,7	8,0	8,7

Услуги	Жилье	HOUS	85,3	89,1	93,5	98,4	102,0	106,4	112,5
	Газ	GAS	5,4	5,7	5,9	6,2	6,3	6,4	6,6
	Вода	WAT	3,0	3,0	3,1	3,3	3,5	3,6	3,9
	Телефон	TELE	8,1	8,7	9,5	10,4	11,2	11,7	12,4
	Местный транспорт	LOCT	3,3	3,2	3,3	3,5	3,4	3,4	3,4
	Воздушный транспорт	AIR	1,7	2,1	2,4	2,8	2,7	2,8	3,1
	Медицинские услуги	DOC	12,1	12,5	12,8	13,6	14,4	14,8	15,7
	Услуги стоматологов	DENT	4,1	4,3	4,4	4,8	5,1	5,1	5,3
	Отдых	REC	12,4	12,7	13,4	14,1	14,6	15,1	15,8
	Частное образование	PRIV	8,8	9,3	10,0	10,6	10,9	11,2	11,7
Товары длительного пользования	Кухонное оборудование	KIT	5,8	6,0	6,6	7,0	7,3	7,9	8,9
	Посуда	TAB	3,5	3,7	3,8	3,8	3,7	3,8	4,0
	Ювелирные изделия	JEWL	3,6	3,9	4,1	4,1	4,1	4,3	4,6

Продолжение таблицы

Категория	*Статья*	*Переменная*	*1973*	*1974*	*1975*	*1976*	*1977*
	Дата	DATE	1973,0	1974,0	1975,0	1976,0	1977,0
	Время	TIME	15,0	16,0	17,0	18,0	19,0
ДОХОД	Личный доход	PI	1007,9	1004,8	1010,8	1056,2	1105,4
	Личный располагаемый доход	DPI	865,3	858,4	875,8	906,8	924,9
РАСХОДЫ	Совокупные личные расходы	TPE	768,5	763,6	780,2	823,1	864,3
Текущие расходы	Питание	FOOD	129,4	128,1	132,3	139,7	145,2
	Одежда	CLOT	59,3	58,7	60,9	63,8	67,5
	Бензин	GASO	26,2	24,8	25,6	26,8	27,7
	Моторное масло	FUEL	5,4	4,2	4,2	4,6	4,4
	Табак	TOB	12,8	13,0	12,9	13,7	13,1
	Косметика	COSM	7,9	7,8	7,4	7,5	7,8
	Лекарства	PHAR	9,3	9,8	9,7	10,0	10,2

Услуги	Жилье	HOUS	118,2	124,2	128,3	134,9	141,3
	Газ	GAS	6,4	6,5	6,6	6,7	6,5
	Вода	WAT	4,1	4,3	4,4	4,3	4,4
	Телефон	TELE	13,7	14,4	15,9	17,1	18,3
	Местный транспорт	LOCT	3,4	3,5	3,5	3,6	3,6
	Воздушный транспорт	AIR	3,4	3,7	3,6	4,0	4,3
	Медицинские услуги	DOC	16,9	17,2	17,8	18,0	19,2
	Услуги стоматологов	DENT	6,1	6,2	6,4	6,9	7,2
	Отдых	REC	16,9	17,6	17,9	19,1	20,4
	Частное образование	PRIV	11,9	11,7	12,1	12,2	12,2
Товары длительного пользования	Кухонное оборудование	KIT	9,9	9,9	9,3	9,7	10,5
	Посуда	TAB	4,2	4,1	3,7	3,9	4,1
	Ювелирные изделия	JEWL	5,2	5,4	5,5	6,1	6,3

Учебное издание

Зеткина Оксана Валерьевна

Эконометрика: основы математического моделирования социально-экономических процессов

(Эконометрические модели
в анализе социально-экономических процессов)

Учебное пособие

Редактор, корректор М. В. Никулина
Правка, верстка Е. Б. Половкова

Подписано в печать 13.12.13. Формат 60×84 1/16.
Усл. печ. л. 7,21. Уч.-изд. л. 5,2.
Тираж 100 экз. Заказ

Оригинал-макет подготовлен
в редакционно-издательском отделе ЯрГУ.

Отпечатано в типографии ООО «Филигрань».
г. Ярославль, ул. Свободы, д. 91.
Тел. (4852) 982705,
pechataet@bk.ru