

ЭКОНОМЕТРИКА ПРАКТИКУМ

Под редакцией профессора **И.А. Кацко**

Допущено Министерством
сельского хозяйства Российской Федерации
в качестве **учебно-практического пособия**
для студентов высших аграрных
учебных заведений, обучающихся
по направлению «Экономика»

BOOK.ru

ЭЛЕКТРОННО-БИБЛИОТЕЧНАЯ СИСТЕМА

КНОРУС • МОСКВА • 2019

УДК 330(075.8)

ББК 65.01я73

Э40

Рецензенты:

Л.И. Ниворожкина, заведующая кафедрой математической статистики, эконометрики и актуарных расчетов Ростовского государственного экономического университета (РИНХ), заслуженный деятель науки РФ, д-р экон. наук, проф.,

С.Г. Чифранов, заведующий кафедрой организации и технологии защиты информации Майкопского государственного технологического университета, д-р экон. наук, проф.

Авторский коллектив:

П.С. Бондаренко, Г.В. Горелова, И.А. Кацко, А.Е. Жминько, Т.В. Соловьёва, С.А. Кацко, С.К. Куижева, А.А. Митус, Н.Б. Паклин, А.Е. Сенникова

Э40 **Эконометрика. Практикум** : учебно-практическое пособие / коллектив авторов ; под ред. И.А. Кацко. — Москва : КНОРУС, 2019. — 218 с. — (Бакалавриат).

ISBN 978-5-406-06368-2

Призвано оказать помощь студентам в овладении приемами и методами эконометрического анализа данных с использованием компьютера. Содержит начальные сведения работы с часто используемыми средствами (MS Excel, Statistica, Gretl, Stata, Deductor), примеры выполнения заданий и задания для самостоятельного решения по основным разделам эконометрики с использованием реальных данных статистической отчетности сельскохозяйственных организаций Краснодарского края. Отражает опыт преподавания авторами соответствующей дисциплины в Кубанском государственном аграрном университете.

Соответствует ФГОС ВО последнего поколения.

Для оказания методической помощи и самостоятельной работы в соответствии с программой курса бакалавров, магистров и специалистов по направлению подготовки «Экономика», а также преподавателей, аспирантов и слушателей курсов повышения квалификации.

УДК 330(075.8)

ББК 65.01я73

ЭКОНОМЕТРИКА

Практикум

Изд. № 11216. Подписано в печать 13.09.2018. Формат 60×90/16.

Гарнитура «Newton». Печать офсетная.

Усл. печ. л. 14,0. Уч.-изд. л. 9,61. Тираж 500 экз.

ООО «Издательство «КноРус».

117218, г. Москва, ул. Кедрова, д. 14, корп. 2.

Тел.: 8-495-741-46-28.

E-mail: office@knorus.ru <http://www.knorus.ru>

Отпечатано в АО «Т8 Издательские Технологии».

109316, г. Москва, Волгоградский проспект, д. 42, корп. 5.

Тел.: 8-495-221-89-80.

ISBN 978-5-406-06368-2

© Коллектив авторов, 2019

© ООО «Издательство «КноРус», 2019

Оглавление

Предисловие	4
Глава 1. Парная регрессия и корреляция (однофакторный корреляционно-регрессионный анализ)	6
Глава 2. Множественный корреляционно-регрессионный анализ	46
Глава 3. Оценивание систем одновременных уравнений.....	82
Глава 4. Фиктивные переменные.....	103
Глава 5. Модели с дискретной зависимой переменной.....	107
Глава 6. Временные ряды.....	141
6.1. Модели временных рядов экономических объектов	141
6.2. Анализ взаимосвязи временных рядов.....	166
Глава 7. Анализ панельных данных	176
Литература	201
Ответы	203
Приложение А Показатели по крупным и средним сельскохозяйственным организациям Краснодарского края.....	204
Приложение Б Статистические данные по крупным и средним сельскохозяйственным организациям Краснодарского края, 2015 г.....	207
Приложение В Данные по производству молока на сельскохозяйственных предприятиях Краснодарского края, 2014 г.....	209
Приложение Г Урожайность озимой пшеницы и количество внесенных минеральных удобрений на 1 га посева на сельскохозяйственных предприятиях.....	211
Приложение Д Урожайность сельскохозяйственных культур в хозяйствах всех категорий Краснодарского края, ц с 1 га посевной площади	212
Приложение Е Урожайность зерновых и зернобобовых культур в Краснодарском крае, ц с 1 га посевной площади.....	213
Приложение Ж Социально-экономические показатели Краснодарского края	215

Предисловие

В последние десятилетия существенно ускорились социально-экономические процессы в обществе, усложнились связи между общественными явлениями, поэтому с 1970-х годов во всем мире считается, что теоретическое изучение экономических процессов и явлений, необходимое для адекватного реагирования (прогнозирования, управления) в банковском деле, финансах, бизнесе (аграрном, промышленном), сфере услуг на уровне организаций и государства должно основываться на эконометрических моделях, полученных по ретроспективным данным.

Эконометрика – это наука, которая позволяет придавать количественное выражение качественным закономерностям и связям в экономических явлениях и процессах. Она базируется на экономической теории, экономической и математической статистике и обычно предполагает вероятностную природу изучаемых данных.

Объектом эконометрики являются социально-экономические явления и процессы, их связи и зависимости.

Эконометрика является одной из ведущих дисциплин при подготовке студентов по направлению «Экономика». В современных условиях развития экономики повышается роль специалистов, обладающих знаниями и навыками использования методов эконометрического анализа при изучении социально-экономических явлений. При подготовке бакалавров-экономистов уделяется большое внимание эконометрическому моделированию связей и закономерностей социально-экономических явлений и процессов.

Государственные образовательные стандарты включают эконометрику, наряду с микроэкономикой и макроэкономикой, в качестве обязательной дисциплины для обучения, как бакалавров, так и магистров.

Разная глубина изучения и степень детализации данной дисциплины позволяет рассматривать эконометрику как на уровне бакалавриата, так и на уровне магистратуры и специалитета.

Цель пособия – оказать помощь студентам в овладении приемами и методами статистического анализа данных с использованием компьютера.

Предлагаемое учебно-практическое пособие в основном посвящено формированию практических навыков эконометрического моделирования с использованием пакетов программ *Excel* и *Statistica 10*, которые вполне пригодны для решения большинства эконометрических задач и, при необходимости, позволяют легко перейти к профессиональным эконометрическим пакетам *Stata*, *Gretl* и др.

Изучение дисциплины направлено на приобретение студентами опыта построения эконометрических моделей, принятия решений об идентификации модели, выбора метода оценки параметров модели, интерпретацию результатов и последующего прогнозирования.

Основное внимание в пособии уделяется построению эконометрических моделей на основе пространственных данных и временных рядов.

Основное внимание в пособии уделяется построению эконометрических моделей на основе пространственных данных и временных рядов.

Большая часть задач составлена так, чтобы обеспечить индивидуальную работу студентов, предусмотрена возможность комбинаций объясняющих переменных, выбор различных зависимых переменных, предлагаются дифференцированные задания.

Учебно-практическое пособие предназначено для выполнения практических (лабораторных) заданий студентами очной и заочной форм обучения.

Задания могут быть использованы при самостоятельном изучении курса.

Следует отметить некоторое отличие классических эконометрических методов от методов, использующихся в бизнес-аналитике, основанной на разведочном анализе данных, нашедшем свое воплощение в информационных технологиях *Data Mining* и *KDD (Knowledge Discovery in Databases)* реализованных, например, в аналитической платформе *Deductor* (пример 5.2).

В настоящее время, в соответствии с классификацией методов анализа данных, можно выделить два подхода к изучению предметной области:

- а) верификации гипотез о процессах в предметной области, которая основывается на использовании традиционных статистических методов;
- б) открытию новых знаний с использованием методов *Data Mining* (машинного обучения), способных автоматически обнаруживать ранее неизвестные связи, скрытые в данных.

Эконометрика постулирует первый подход. В бизнес-аналитике комбинируются и применяются оба подхода, причем жажда прибыли позволяет жертвовать «пониманием» модели, а в эконометрике «понимание» – необходимый атрибут использования модели экономистами на практике (при прогнозировании и управлении), что выводит на первый план идею линейности в малом большинства процессов в окружающем мире и рассмотрении в первую очередь линейных моделей, имеющих простую интерпретацию.

Авторы полагают, что реализованный ими в учебном пособии подход позволит студентам получить необходимые навыки построения и анализа эконометрических моделей социально-экономических явлений и процессов.

Выражаем благодарность авторам всех источников, отраженных в списке литературы, а также сотрудникам кафедры статистики и прикладной математики Кубанского ГАУ в разные годы принимавших участие в формировании взглядов на преподавание эконометрики.

Следует отметить, что пособие предполагает знание основ теории вероятностей и математической статистики (понятия вероятности, выборочного метода и проверки статистических гипотез), например, на уровне нашего пособия [Бондаренко П.С. и др.].

Исходные данные для большинства примеров и заданий можно скачать по адресу <http://kubsau.ru/education/chairs/statistics/publications/>

Глава 1. Парная регрессия и корреляция (однофакторный корреляционно-регрессионный анализ)

Человек наблюдает за явлениями и процессами в окружающем мире в разных масштабах (от квантовой механики до космоса), временных диапазонах (от долей секунды до времени существования вселенной) и в разной степени связанных с социумом (от природных явлений до социально-экономических процессов).

Изучаемые явления (процессы) характеризуются признаками или показателями, которые подразделяются на факторные и результативные. Признаки, оказывающие влияние на другие признаки, называются факторными (независимыми, экзогенными). Признаки, изменяющиеся под влиянием факторных признаков, называются результативными (зависимыми, эндогенными). Если на изменение результативного признака оказывает влияние один, доминирующий фактор, то связь между ними называется парной и тогда строится модель парной регрессии. Неявно предполагается, что этот фактор (x) является причиной изменения результативного признака (y).

Связи между изучаемыми переменными подразделяются на функциональные и статистические. При функциональной связи определенному значению одной переменной величины соответствует строго определенное значение другой переменной. При изменении одной из них на определенную величину, другая переменная изменяется на величину, в соответствии с видом функции, связывающей переменные. Статистической называется связь между переменными или признаками, когда определенному значению факторного признака соответствует несколько различных значений результативного признака. Частным случаем статистической связи является корреляционная, которая проявляется в среднем, в массе наблюдений, как статистическая закономерность. При корреляционной связи с изменением факторного признака на определенную величину изменяется среднее значение результативного признака. Обычно корреляционная зависимость представляется как функциональная зависимость между переменными в виде уравнения регрессии.

Задачами корреляционно-регрессионного анализа являются:

- установление типа уравнения регрессии;
- определение параметров уравнения регрессии и оценка значимости параметров;
- оценка тесноты и направления связи между переменными;
- оценка значимости уравнения регрессии;
- определение прогнозных значений зависимой переменной и оценка полученного прогноза.

Модель парной регрессии – это уравнение связи между двумя переменными x и y , т.е. функция $y = f(x)$, где y – результативный признак или функция отклика, x – факторный признак или независимая переменная. Но так как

связь между переменными в эконометрике является статистической, а не функциональной, то всегда будут различия между фактически наблюдаемыми значениями результативного признака (y) и теоретическими (\hat{y}), т.е. возможными значениями, найденными по уравнению связи. Тогда реальные наблюдения, в силу влияния неучтенных факторов, отличаются от теоретических значений y на величину ε – случайную ошибку или остаток:

$$y = \hat{y} + \varepsilon, \quad y = f(x) + \varepsilon, \quad \text{где } \hat{y} = f(x). \quad (1.1)$$

В основе регрессионного анализа обычно лежит некоторая вероятностно-статистическая модель, соответствующая закону распределения ошибок (ε). Вид функции $f(x)$ при изучении парной регрессии устанавливается на основе теоретического анализа или графического отображения пар наблюдений $(x_i, y_i, i = 1, 2, \dots, n)$ на плоскости. Параметры уравнения регрессии находятся методом наименьших квадратов таким образом, чтобы сумма квадратов остатков была минимальной:

$$S = \sum_{i=1}^n \varepsilon_i^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2. \quad (1.2)$$

Найдя и приравняв нулю частные производные по параметрам уравнения регрессии $\hat{y} = f(x)$, составляется и решается соответствующая система нормальных уравнений. Для наиболее широко используемых уравнений, системы уравнений приведены в таблице 1.1.

Таблица 1.1 – Функции, применяемые в парном регрессионном анализе

№ п/п	Функция	Нормальные уравнения
1	$y = a + bx$ линейная	$an + b \sum x = \sum y$ $a \sum x + b \sum x^2 = \sum (xy)$
2	$y = a + bx + cx^2$ парабола второго порядка	$an + b \sum x + c \sum x^2 = \sum y$ $a \sum x + b \sum x^2 + c \sum x^3 = \sum (xy)$ $a \sum x^2 + b \sum x^3 + c \sum x^4 = \sum (x^2 y)$
3	$y = ab^x$ или $\ln y = \ln a + x \ln b$ показательная	$n \ln a + \ln b \sum x = \sum \ln y$ $\ln a \sum x + \ln b \sum x^2 = \sum (x \ln y)$
4	$y = ax^b$ или $\ln y = \ln a + b \ln x$ степенная	$n \ln a + b \sum \ln x = \sum \ln y$ $\ln a \sum \ln x + b \sum (\ln x)^2 = \sum (\ln x \ln y)$
5	$y = a + b \ln x$ полулогарифмическая	$an + b \sum \ln x = \sum y$ $a \sum \ln x + b \sum (\ln x)^2 = \sum (y \ln x)$
6	$\ln y = a + b \ln x$ или $y = e^{a+b \ln x}$ логлинейная	$an + b \sum \ln x = \sum \ln y$ $a \sum \ln x + b \sum (\ln x)^2 = \sum (\ln x \ln y)$
7	$\ln y = a + bx$ или $y = e^{a+bx}$ экспоненциальная	$an + b \sum x = \sum \ln y$ $a \sum x + b \sum x^2 = \sum (x \ln y)$
8	$y = \frac{a}{1 + be^{-cx}}$ логистическая	решается численно (приближенными методами)

В эконометрике применяются как линейные, так и нелинейные уравнения регрессии. Нелинейные уравнения регрессии можно подразделить на два класса: нелинейные, относительно объясняющих переменных, но линейные по оцениваемым параметрам; нелинейные по оцениваемым параметрам. Чаще всего параметры этих уравнений находят с использованием решения соответствующих систем нормальных уравнений.

Ряд нелинейных зависимостей можно привести к линейному виду путем соответствующих преобразований (таблица 1.2).

Таблица 1.2 – Преобразование функций в линейный вид

№ п/п	Функция	Линеаризующие преобразования			
		преобразования переменных		выражения для величин a и b	
		y'	x'	a'	b'
1	$y = a + b/x$	y	$1/x$	a	b
2	$y = 1/(a+bx)$	$1/y$	x	a	b
3	$y = x/(a+bx)$	x/y	x	a	b
4	$y = ab^x$	$\ln y$	x	$\ln a$	$\ln b$
5	$y = ae^{bx}$	$\ln y$	x	$\ln a$	b
6	$y = 1/(a + be^{-x})$	$1/y$	e^{-x}	a	b
7	$y = ax^b$	$\ln y$	$\ln x$	$\ln a$	b
8	$y = a + b \lg x$	y	$\lg x$	a	b
9	$y = a/(b + x)$	$1/y$	x	b/a	$1/a$
10	$y = ax/(b + x)$	$1/y$	$1/x$	b/a	$1/a$
11	$y = ae^{b/x}$	$\ln y$	$1/x$	$\ln a$	b
12	$y = a + bx^n$	y	x^n	a	b

I. Корреляционно-регрессионный анализ между двумя признаками проводится в следующей последовательности:

1. Исходя из целей и задач исследования зависимости, устанавливается результативный и факторный признаки. По каждой единице совокупности объектов или наблюдений определяются значения результативного и факторного признаков. Совокупность должна быть однородной в отношении изучаемых признаков, а значения признаков должны быть достоверными и объективными.

2. Обосновывается (обычно графическим способом) модель уравнения регрессии (например, пример 1.1, рисунок 1.1). Для этого в прямоугольной декартовой системе координат отображают эмпирические наблюдения, образующие корреляционное поле, на основании изучения которого делаются выводы о направлении и возможной функциональной форме связи между факторным и результативным признаками (прямая или обратная, линейная или нелинейная).

Пример 1.1. Имеются следующие выборочные данные по 15 предприятиям центральной зоны Краснодарского края за 2014 г.

Таблица 1.3 – Среднегодовая стоимость основных фондов и производство продукции на 1 га сельскохозяйственных угодий по совокупности предприятий

№	Среднегодовая стоимость основных фондов на 1 га сельскохозяйственных угодий, тыс. руб. (x)	Валовая продукция на 1 га сельскохозяйственных угодий, тыс. руб. (y)
1	55,9	66,2
2	50,3	63,3
3	82,4	96,9
4	97,7	83,1
5	73,1	69,8
6	66,7	64,3
7	45,6	60,4
8	53,7	74,6
9	52,5	58,7
10	34,0	44,9
11	51,2	51,2
12	59,1	75,8
13	40,5	59,9
14	56,7	53,9
15	69,9	75,8

3. Методом наименьших квадратов определяются параметры уравнения регрессии.

Для линейных и нелинейных уравнений, приводимых к линейному виду, при условии, что зависимость выражается уравнением:

$$y = a + bx,$$

составляется и решается следующая система нормальных уравнений Гаусса:

$$\begin{cases} \Sigma y = na + b\Sigma x, \\ \Sigma yx = a\Sigma x + b\Sigma x^2. \end{cases} \quad (1.3)$$

Параметры уравнения линейной регрессии также можно найти по формулам, вытекающим из системы (1.3):

$$b = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\overline{x^2} - (\bar{x})^2}, \quad a = \bar{y} - b \cdot \bar{x}, \quad (1.4)$$

$$\text{где } \bar{y} = \frac{\Sigma y}{n}, \quad \bar{x} = \frac{\Sigma x}{n}, \quad \overline{xy} = \frac{\Sigma xy}{n}, \quad \overline{x^2} = \frac{\Sigma x^2}{n}.$$

Замечание. Важнейшим методом анализа данных является визуализация (представление данных в виде таблиц, диаграмм, кросс-таблиц, кросс-диаграмм, графиков).

Рассмотрим в *Excel* применение диаграмм рассеяния. Выделим на листе *Excel* диапазон, содержащий данные таблицы 1.3, выполним команду:

Вставка – Точечная – Точечная с маркерами.

В результате получим рисунок 1.1, который показывает возможную зависимость между факторным и результативным признаками.

Достоинством графического представления данных является возможность увидеть ошибки, допущенные при вводе данных (артефакты – объекты, созданные человеком) или аномальные значения признаков – выбросы, явно не принадлежащие изучаемой совокупности.

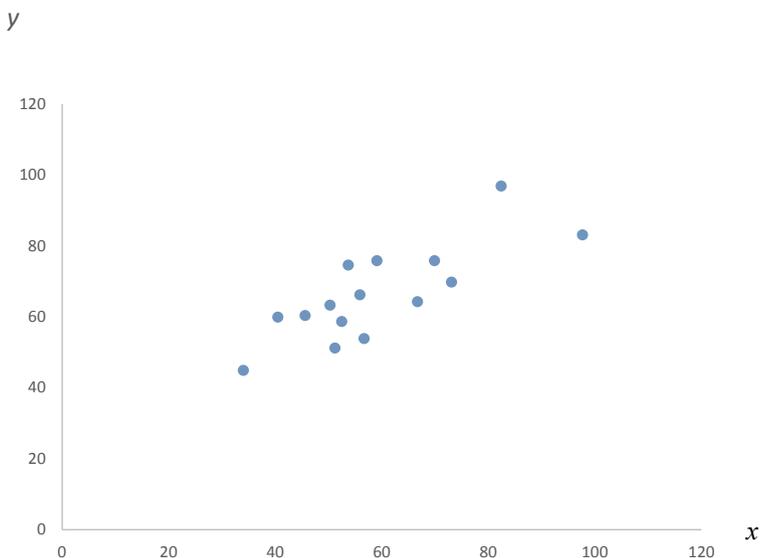


Рисунок 1.1– Корреляционное поле

Например, при вводе исходных данных (таблица 1.3) мы вместо 66,2 ввели 662. Построим соответствующую диаграмму рассеяния (рисунок 1.2) из которой видно, что есть наблюдение, отличающееся от других данных.

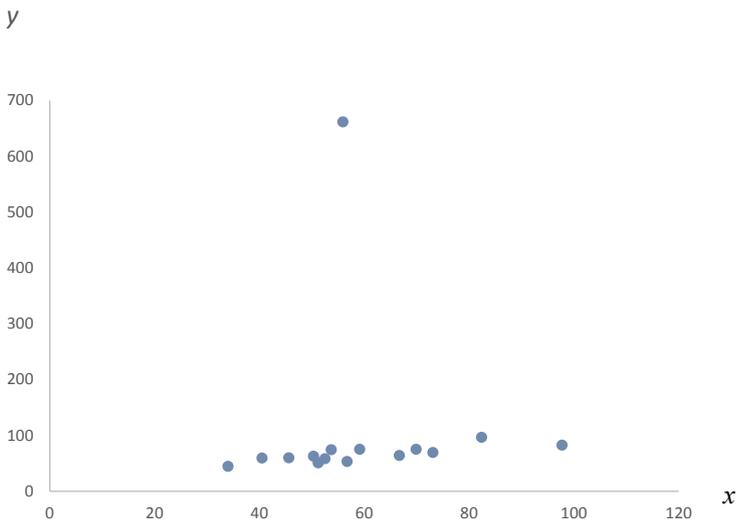


Рисунок 1.2 – Диаграмма рассеяния с артефактом (или выбросом)

Коэффициент регрессии линейного уравнения b показывает на сколько единиц в среднем изменится результирующий признак y при увеличении факторного признака x на единицу.

Для проведения расчетов строится вспомогательная таблица.

4. Качество уравнения регрессии оценивается с помощью средней ошибки аппроксимации:

$$\bar{A} = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \cdot 100. \quad (1.5)$$

Если средняя ошибка аппроксимации не более 10-12%, то качество уравнения регрессии считается хорошим.

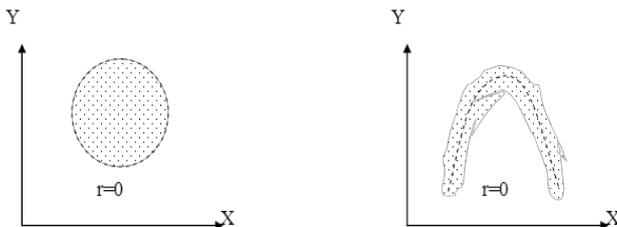
5. При линейной зависимости теснота связи между переменными x и y оценивается с помощью коэффициента корреляции:

$$r = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\sigma_x \cdot \sigma_y}, \quad -1 \leq r \leq 1, \quad (1.6)$$

где σ_x и σ_y – средние квадратические отклонения по x и по y :

$$\sigma_x = \sqrt{\overline{x^2} - (\bar{x})^2}; \quad \sigma_y = \sqrt{\overline{y^2} - (\bar{y})^2}; \quad \bar{y^2} = \frac{\sum y^2}{n}. \quad (1.7)$$

Уравнения регрессии y на x ($y = a + bx$) и x на y ($x = c + dy$), как правило не совпадают, но, $r_{yx} = r_{xy}$.



Отсутствие корреляции



Рисунок 1.3 – Виды корреляционных полей

Чем ближе абсолютное значение коэффициента корреляции к нулю, тем линейная связь между признаками слабее. Чем ближе к единице, тем связь сильнее. Если $r = \pm 1$, то связь линейная функциональная. Если $r = 0$, то связь между признаками отсутствует, т. е. они линейно независимы.

Знак коэффициента корреляции показывает направление связи между признаками. Если $r > 0$, то связь прямая, обе переменные изменяются в одном направлении. Если $r < 0$, то связь обратная, переменные изменяются в противоположных направлениях (рисунок 1.3).

При нелинейной зависимости теснота связи между переменными x и y определяется с помощью индекса корреляции:

$$R_{xy} = \sqrt{1 - \frac{\sigma_{ocm}^2}{\sigma_y^2}} = \sqrt{1 - \frac{\sum (y - \hat{y})^2}{\sum (y - \bar{y})^2}}, \quad \sigma_{ocm}^2 = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n} \quad (1.8)$$

Квадрат коэффициента (индекса) корреляции называется коэффициентом (индексом) детерминации. Коэффициент детерминации R^2 показывает долю влияния фактора x на результативную переменную y , а $(1-R^2)$ – долю

влияния других, неучтенных в модели факторов (толерантность – в англоязычных источниках).

6. Относительное влияние факторного признака на результативный признак определяется с помощью коэффициента эластичности (таблица 1.4)

$$\mathfrak{E} = f'(x) \frac{x}{\hat{y}}. \quad (1.9)$$

Таблица 1.4 – Коэффициенты эластичности по некоторым элементарным функциям

Функция, y	Первая производная, y'_x	Коэффициент эластичности, $\mathfrak{E} = y'_x \cdot \frac{x}{\hat{y}}$
Линейная $y = a + bx$	b	$\mathfrak{E} = \frac{bx}{a+bx}$
Парабола второго порядка $y = a + bx + cx^2$	$b+2cx$	$\mathfrak{E} = \frac{(b+2cx)x}{a+bx+cx^2}$
Гипербола $y = a + \frac{b}{x}$	$-\frac{b}{x^2}$	$\mathfrak{E} = \frac{-b}{ax+b}$
Показательная $y = ab^x$	$ab^x \ln b$	$\mathfrak{E} = x \ln b$
Степенная $y = ax^b$	abx^{b-1}	$\mathfrak{E} = b$
Полулогарифмическая $y = a + b \ln x$	$\frac{b}{x}$	$\mathfrak{E} = \frac{b}{a+b \ln x}$
Логистическая $y = \frac{a}{1 + be^{-cx}}$	$\frac{abce^{-cx}}{(1+be^{-cx})^2}$	$\mathfrak{E} = \frac{cx}{\frac{1}{b}e^{cx} + 1}$

При линейной форме связи средний коэффициент эластичности находится по формуле:

$$\bar{\mathfrak{E}} = b \frac{\bar{x}}{\bar{y}}, \quad (1.10)$$

где \bar{x} и \bar{y} – средние значения признаков; b – коэффициент регрессии.

Средний коэффициент эластичности линейной модели показывает, на сколько процентов в среднем изменится результативный признак при увеличении факторного признака на 1 %.

7. Оценка статистической значимости построенной модели регрессии в целом производится с использованием F -критерия Фишера-Снедекора (далее Фишера). Рассматривается нулевая гипотеза $H_0: r^2 = 0$ ($b = 0$) и альтернативная ей гипотеза $H_1: r^2 \neq 0$ ($b \neq 0$).

Наблюдаемое (фактическое) значение F -критерия находится по формуле:

$$F_n = \frac{r^2}{1-r^2} \cdot \frac{n-m-1}{m}, \quad (1.11)$$

где m – число параметров уравнения регрессии при переменной x ,
 n – число наблюдений.

Для однофакторного линейного уравнения регрессии $m = 1$, тогда наблюдаемое значение критерия рассчитывается по формуле:

$$F_n = \frac{r^2}{1-r^2} \cdot (n-2). \quad (1.12)$$

Для определения критического (табличного) значения используется таблица значений F -критерия Фишера. При уровне значимости $\alpha=0,05$ и числе степеней свободы $k_1 = m$; $k_2 = n - m - 1$:

$$F_{кр} = F_{\alpha=0,05}(k_1; k_2).$$

В случае парной регрессии число степеней свободы $k_1 = 1$, а число степеней свободы $k_2 = n - 2$.

Сравнивается F_n с $F_{кр}$. Если $F_n < F_{кр}$, то нулевая гипотеза принимается и уравнение регрессии статистически незначимо. Если $F_n > F_{кр}$, то нулевая гипотеза отклоняется и принимается альтернативная гипотеза о статистической значимости уравнения регрессии.

8. Так как регрессионный анализ зависимости между признаками проводится по выборочным данным, то проверяется значимость величины выборочного коэффициента корреляции, а также параметров уравнения регрессии a и b с использованием t -критерия Стьюдента при заданном уровне значимости α .

Находится наблюдаемое значение критерия для параметров a и b :

$$t_a = \frac{a}{m_a}, \quad m_a = \frac{s_{ост} \sqrt{\sum x^2}}{n \sigma_x}, \quad (1.13)$$

$$t_b = \frac{b}{m_b}, \quad m_b = \frac{s_{ост}}{\sigma_x \sqrt{n}}, \quad (1.14)$$

где $s_{ост} = \sqrt{\frac{\sum (y - \hat{y})^2}{n-m-1}}$, для парной регрессии $m = 1$.

Критическое значение $t_{кр}$ находится по таблице критических значений Стьюдента при заданном уровне значимости α и числе степеней свободы $k = n - 2$ для двухсторонней критической области.

Если $|t_n| < t_{кр}$, то принимается основная гипотеза о незначимости соответствующего параметра уравнения регрессии. Если $|t_n| > t_{кр}$, то основная гипотеза отвергается и принимается альтернативная гипотеза о статистической значимости соответствующего параметра уравнения регрессии при заданном уровне значимости α .

Для проверки значимости коэффициента корреляции выдвигается нулевая гипотеза $H_0: r_2 = 0$ – коэффициент корреляции в генеральной совокупности равен нулю и изучаемый фактор не оказывает существенного влияния

на результативный признак, альтернативная гипотеза $H_1: r_2 \neq 0$ – коэффициент корреляции в генеральной совокупности значимо отличается от нуля при заданном уровне значимости α .

Для проверки нулевой гипотезы применяется t -критерий Стьюдента и определяется наблюдаемое значение t -критерия:

$$t_n = |r| \cdot \sqrt{\frac{n-2}{1-r^2}}. \quad (1.15)$$

Критическое значение ($t_{кр}$) находится по таблице t -критерия Стьюдента при уровне значимости α и числе степеней свободы $k = n - 2$ для двухсторонней критической области.

Если $|t_n| < t_{кр}$, то основная гипотеза о незначимости коэффициента корреляции принимается. Если $|t_n| > t_{кр}$, то нулевая гипотеза отвергается, считается, что коэффициент корреляции существенно отличается от нуля в генеральной совокупности.

При парной линейной зависимости оценка значимости всего уравнения, коэффициентов корреляции и регрессии дает одинаковые результаты, так как $t_b^2 = t_r^2 = F_n$ (различия объясняются ошибками округлений).

Рассчитывается доверительный интервал для параметров a и b .

Для этого определяется предельная ошибка для каждого параметра:

$$\Delta_a = t \cdot m_a; \quad \Delta_b = t \cdot m_b; \quad (1.16)$$

Доверительные интервалы для параметров a и b определяются по следующим формулам:

$$\begin{aligned} \gamma_{a_{min}} &= a - \Delta_a, & \gamma_{a_{max}} &= a + \Delta_a; \\ \gamma_{b_{min}} &= b - \Delta_b, & \gamma_{b_{max}} &= b + \Delta_b. \end{aligned} \quad (1.17)$$

Если ноль попадает в границы доверительных интервалов, т.е. нижняя граница отрицательна, а верхняя положительна, то оцениваемый параметр с заданной доверительной вероятностью принимается нулевым, т.е. является статистически незначимым. В противном случае принимается статистическая значимость оцениваемого параметра.

9. Прогнозное значение результативного признака определяется путем подстановки в построенное парное линейное уравнение регрессии прогнозного значения факторного признака x_p :

$$\hat{y}_p = a + bx_p. \quad (1.18)$$

10. Точечный прогноз дополняется расчетом стандартной $m_{\hat{y}_p}$ и предельной $\Delta_{\hat{y}_p}$ ошибками, и, соответственно, интервальной оценкой прогнозного значения. Определяется средняя и предельная ошибки прогноза по линейному уравнению регрессии:

Литература

1. Айвазян С.А. Методы эконометрики: Учебник. – М.: Инфра-М, 2010. – 512 с.
2. Бондаренко П.С. Теория вероятностей и математической статистика: учебное пособие / П.С. Бондаренко, Г.В. Горелова, И.А. Кацко; под ред. И.А. Кацко, А.И. Трубилина. – Москва: КНОРУС. 2017. – 390 с.
3. Боровиков В. П. STATISTICA. Искусство анализа данных на компьютере: Для профессионалов. 2-е изд. / В. П. Боровиков – СПб.: Питер, 2003. – 688 с.: ил.
4. Вербик М. Путеводитель по современной эконометрике. – М.: Научная книга, 2008. – 616с.
5. Гмурман В.Е. Руководство к решению задач по теории вероятностей и математической статистике: Учеб. пособие для бакалавров. / В.Е. Гмурман. – М.: Юрайт, 2013. 405 с.
6. Гмурман В.Е. Теория вероятностей и математическая статистика. Учебное пособие. М.: Юрайт, 2013. – 479с.
7. Горелова Г.В., Кацко И.А. Теория вероятностей и математическая статистика в примерах и задачах с применением Excel / Г.В. Горелова, И.А. Кацко Ростов н/Д.: Феникс, 2006. – 475 с.
8. Доугерти К. Введение в эконометрику: Учебник. – М.: Инфра-М, 2010. – 465 с.
9. Дрейпер И. Прикладной регрессионный анализ: В 2-х кн.; пер. с англ. – 2-е изд. перераб. и доп. / И. Дрейпер, Г. Смит – М.: Финансы и статистика, 1986. – Кн. 1. – 366 с., ил; 1987. Кн. 2. – 351 с., ил.
10. Елисеева И. И. Эконометрика: учебник / И. И. Елисеева, С. В. Курьшева, Т. В. Костева и др., под ред. И. И. Елисеевой. – 2-е изд., перераб. и доп. – М.: Финансы и статистика, 2005. – 576 с.: ил.
11. Елисеева И.И. Эконометрика: Уч. для магистров /И.И. Елисеева [и др.]; под ред. И.И. Елисеевой. – М.: Юрайт, 2012. – 453 с.
12. Кацко И. А. Практикум по анализу данных на компьютере / И. А. Кацко, Н. Б. Паклин; под ред. Г. В. Гореловой. – М.: КолосС, 2009. – 278 с.
13. Кремер Н.Ш. Эконометрика / Н.Ш. Кремер, В.А. Путко – М.: ЮНИТИ-ДАНА, 2007. – 573с.
14. Кремер Н.Ш., Путко Б.А. Эконометрика: Учебник для вузов / Под ред. проф. Н.Ш. Кремера. – 3-е изд., перераб. и доп. – М.: Юнити-Дана, 2013. – 328 с.
15. Магнус Я. Р. Эконометрика. Начальный курс: учебник / Я. Р. Магнус, П. К. Катышев, А. А. Пересецкий – М.: Дело, 2004. – 576 с.
16. Магнус Я.Р., Катышев П.К., Пересецкий А.А. Эконометрика. Начальный курс: Учебник. – М.: Дело, 2007. – 504 с.
17. Мхитарян В.С., Архипова М.Ю., Балаш В.А. Эконометрика: Учебник. – М.: Проспект, 2010. – 384с.

18. Общий курс высшей математики для экономистов: Учебник / Под общ. ред. В.И. Ермакова. – М.: ИНФРА-М. 2008. – 656с.
19. Орлов А. И. Организационно-экономическое моделирование: учебник: в 3-х ч.; Ч. 1 Нечисловая статистика / А. И. Орлов. – М.: Изд-во МГТУ им. Н. Э. Баумана, 2009. – 541 с.
20. Орлов А. И. Прикладная статистика / А. И. Орлов. – М.: Экзамен, 2006. – 611 с.
21. Орлов А.И. Эконометрика: учебник / А.И. Орлов – М. : Экзамен, 2009. – 412 с.
22. Паклин, Н. Б. Бизнес-аналитика: от данных к знаниям: учеб. пособие. 2-е изд., перераб. и доп./ Н. Б. Паклин, В. И. Орешков. – СПб.: Питер, 2010. – 704 с.
23. Ратникова Т. А. Анализ панельных данных в пакете «СТАТА» / Т. А. Ратникова // Методические указания к компьютерному практикуму по курсу «Эконометрический анализ панельных данных». – М. : ГУ ВШЭ.– 2004.
24. Ратникова Т. А. Введение в эконометрический анализ панельных данных / Т. А. Ратникова // Экономический журнал ВШЭ. – 2006. – № 2. – С. 267–316.
25. Шанченко Н.И. Лекции по эконометрике: учебное пособие для студентов высших учебных заведений / Н.И. Шанченко. – Ульяновск: УлГТУ, 2008. – 139 с.
26. Шанченко Н.И. Эконометрика: лабораторный практикум: учебное пособие / Н.И. Шанченко. – Ульяновск: УлГТУ, 2011. – 117 с.
27. *Green, W.H. Econometric analysis / W.H. Green. – 6th ed. – Pearson Prentice Hall, 2008. – 1178 p.*