



Конспект лекций к изучению раздела **«Эконометрический анализ»**

УДК 51-7
ББК 22.18

Печатается по решению учебно-методической комиссии экономического отделения Набережночелнинского института (филиала) федерального государственного автономного образовательного учреждения высшего образования «Казанский (Приволжский) федеральный университет» от «05» сентября 2023г. (протокол №5)

Рецензенты:

доктор экономических наук, профессор А.Н. Макаров
кандидат экономических наук, доцент М.Р.Хамидуллин

Розенцвайг А.К., Исавнин А.Г. Конспект лекций к изучению раздела «Эконометрический анализ»: учебно-методическое пособие / А.К. Розенцвайг, А.Г. Исавнин – Набережные Челны: Изд-во Набережночелнинского института КФУ, 2024. – 29 с.

Учебно-методическое пособие содержит последовательное изложение базовых понятий теории в эконометрике. Подробно изложены: основные этапы эконометрического исследования, множественная линейная регрессия, анализ качества уравнения регрессии, уравнение регрессии в стандартизованной форме, использование многофакторной регрессионной модели для экономического анализа, порядок выполнения и оформления типового примера.

Учебно-методическое пособие предназначено для использования в учебном процессе студентами технических направлений в экономике и экономического отделения дневной, заочной и дистанционной форм обучения.

© Розенцвайг А.К., Исавнин А.Г., 2024
© НЧИ КФУ, 2024
© Кафедра бизнес-информатики и математических методов в экономике, 2024 г.

СОДЕРЖАНИЕ

Введение	4
1. Основные этапы эконометрического исследования	5
2. Множественная линейная регрессия.	7
2.1. Анализ множественной линейной регрессии с помощью метода наименьших квадратов (МНК)	7
2.2. Теоретические предпосылки применимости МНК	9
2.3. Оценивание коэффициентов множественной линейной регрессии . . .	10
2.4. Интерпретация параметров и уравнения множественной линейной регрессии	11
3. Анализ качества уравнения регрессии	12
3.1. Дисперсии и стандартные ошибки коэффициентов линейной регрессии	12
3.2. Доверительные интервалы коэффициентов регрессии	13
3.3. Статистическая значимость уравнения регрессии	14
4. Уравнение регрессии в стандартизованной форме.	15
5. Использование многофакторной регрессионной модели для экономического анализа.	16
5.1. Коэффициенты β_i стандартизированной регрессии	16
5.2. Средние и частные коэффициенты эластичности	17
5.3 Линейные коэффициенты парной, частной и множественной корреляции.	17
6. Порядок выполнения и оформления типового примера	20
Литература	28

Введение

Постоянно усложняющиеся процессы и явления реальной экономической жизни привели к необходимости создания и совершенствования особых методов их изучения и анализа. При этом широкое распространение получило использование моделирования и количественного анализа. На базе последних выделилось и сформировалось одно из направлений экономических исследований — эконометрика.

Формально «эконометрика» означает «измерения в экономике». Однако область исследований данной дисциплины гораздо шире. Эконометрика — это наука, в которой на базе реальных статистических данных строятся, анализируются и совершенствуются математические модели реальных экономических явлений. Обоснование и интерпретация самих моделей является задачей экономической теории, а эконометрика позволяет найти количественное подтверждение либо опровержение того или иного существующего экономического закона либо новой теоретической гипотезы. Одним из важнейших направлений эконометрики является построение прогнозов по различным экономическим показателям.

Изучение экономических процессов (взаимосвязей) осуществляется в эконометрике на основе математических (эконометрических) моделей. В этом состоит ее родство с математической экономикой. Но если математическая экономика строит и анализирует обобщенные модели без использования реальных числовых значений, то эконометрика концентрируется на изучении моделей, обоснованных экономической теорией, и сопоставлении их с реальными статистическими данными.

Одной из основных задач экономической статистики является сбор, обработка и представление экономических данных в наглядной форме: в виде таблиц, графиков, диаграмм. Эконометрика также активно пользуется этим инструментарием, но идет дальше, применяя его для анализа теоретически обоснованных экономических взаимосвязей и прогнозирования.

При эконометрическом исследовании имеют место две стороны проблемы обеспечения необходимого качества его результатов – качественная и количественная. Качественная сторона заключается в установлении соответствия между построенной эконометрической моделью и лежащими в ее основе положениями экономической теории. Другая – количественная - состоит в обеспечении наиболее полного соответствия между количественными характеристиками модели и статистической информацией, характерной для поведения изучаемых социально – экономических явлений и процессов в реальных условиях.

1. Основные этапы эконометрического исследования

В эконометрических исследованиях обычно предполагается, что закономерности моделируемого процесса складываются под влиянием ряда других явлений, факторов. Обобщенную форму эконометрической модели, описывающей закономерности развития такого процесса, обозначенного переменной y , в зависимости от уровней, действующих на него внешних явлений, факторов X_i , $i = 1, 2, \dots, p$, можно представить следующим уравнением:

$$y = f(b_0, b_1, \dots, b_p, X_1, X_2, \dots, X_p) + \varepsilon, \quad (1)$$

выражающий вид и структуру взаимосвязей между уровнями переменных y , и X_i ; ε , — случайная ошибка модели, в отношении свойств и характеристик которой, как это будет показано далее, выдвигается ряд дополнительных предположений.

В общем случае процедуру построения эконометрической модели можно разделить на несколько взаимосвязанных между собой этапов. На каждом из них последовательно решаются задачи, которые имеют следующее содержание:

1. *Спецификация модели* - анализ специфических свойств рассматриваемых явлений и процессов (предметной области) и обоснование типа моделей, наиболее подходящих для их описания. Отметим, что в общем случае целями этого этапа являются:

1.1. Обоснование формы модели, выражаемой общим математическим уравнением (системой уравнений), связывающим включенные в модель переменные и содержащим неизвестные параметры (коэффициенты).

1.2. Выбор рационального состава включаемых в модель переменных и определение количественных характеристик, отражающих их уровни в прошлые периоды времени (на однородных объектах некоторой совокупности — территориях, предприятиях и т.п.).

2. *Параметризация модели* - оценка параметров выбранного варианта модели на основании исходных статистических данных, выражающих уровни показателей (переменных) на пространственной совокупности однородных объектов или в различные моменты времени.

3. *Верификация модели* - проверка качества построенной модели и обоснование вывода о целесообразности ее использования в ходе дальнейшего эконометрического исследования, а также для объяснения поведения изучаемых экономических показателей, прогнозирования и предсказания их поведения.

4. При выводе о нецелесообразности использования построенной эконометрической модели в дальнейших исследованиях следует вернуться к первому (или какому-либо другому этапу) и попытаться выбрать более качественную модификацию модели (другой вариант модели).

Выделенные этапы построения моделей достаточно условны и отражают 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. Состав используемых на них процедур, приемов и методов, их очередность зависят от типа разрабатываемой эконометрической модели, особенностей исследуемых экономических процессов, свойств исходных данных и т.п.

2. Множественная линейная регрессия

Множественный регрессионный анализ является расширением парного регрессионного анализа. Он применяется в тех случаях, когда поведение объясняемой, зависимой переменной необходимо связать с влиянием более чем одной факторной, независимой переменной. При оценке влияния каждой независимой переменной необходимо уметь разграничивать ее воздействие на объясняемую переменную от воздействия других независимых переменных. При этом множественный корреляционный анализ сводится к анализу парных, частных корреляций. На практике обычно ограничиваются определением их обобщенных числовых характеристик, таких как частные коэффициенты эластичности, частные коэффициенты корреляции, стандартизованные коэффициенты множественной регрессии.

Затем решаются задачи спецификации регрессионной модели, одна из которых состоит в определении объема и состава совокупности независимых переменных, которые могут оказывать влияние на объясняемую переменную. Хотя это часто делается из априорных соображений или на основании соответствующей экономической (качественной) теории, некоторые переменные могут в силу индивидуальных особенностей изучаемых объектов не подходить для модели. В качестве наиболее характерных из них можно назвать *мультиколлинеарность* или *автокоррелированность* факторных переменных.

2.1. Анализ множественной линейной регрессии с помощью метода наименьших квадратов (МНК)

На любой экономический показатель в реальных условиях обычно оказывает влияние не один, а несколько и не всегда независимых факторов. Например, спрос на некоторый вид товара определяется не только ценой данного товара, но и ценами на замещающие и дополняющие товары, доходом потребителей и многими другими факторами. В этом случае вместо парной

регрессии $M(Y/X=x) = f(x)$ рассматривается множественная регрессия

$$M(Y/X_1 = x_1, X_2 = x_2, \dots, X_p = x_p) = f(x_1, x_2, \dots, x_p) . \quad (2.1)$$

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

$$Y = f(\mathbf{B}, \mathbf{X}) + \sigma , \quad (2.2)$$

где \mathbf{X} — вектор независимых (объясняющих) переменных; \mathbf{B} — вектор параметров уравнения (подлежащих определению); σ - случайная ошибка (отклонение); Y — зависимая (объясняемая) переменная.

Предполагается, что для данной генеральной совокупности именно функция f связывает исследуемую переменную Y с вектором независимых переменных \mathbf{X} . Рассмотрим самую употребляемую и наиболее простую для статистического анализа и экономической интерпретации модель множественной линейной регрессии.

Теоретическое линейное уравнение регрессии имеет вид:

$$Y = b_0 + b_1 X_1 + b_2 X_2 + \dots + b_p X_p + \varepsilon , \quad (2.3)$$

или для индивидуальных наблюдений с номером i :

$$y_i = b_0 + b_1 x_{i1} + b_2 x_{i2} + \dots + b_p x_{ip} + \varepsilon_i , \quad (2.4)$$

где $i = 1, 2, \dots, n$.

Здесь $\mathbf{B} = (b_0, b_1, \dots, b_p)$ — вектор размерности $(p+1)$ неизвестных параметров $b_j, j = 0, 1, 2, \dots, p$, называется j -ым теоретическим коэффициентом регрессии (частичным коэффициентом регрессии). Он характеризует чувствительность величины Y к изменению X_j . Другими словами, он отражает влияние на условное математическое ожидание $M(Y/X_1 = x_1, X_2 = x_2, \dots, X_p = x_p)$ зависимой переменной Y объясняющей переменной X_j при условии, что все другие объясняющие переменные остаются постоянными.

После выбора линейной функции в качестве модели зависимости необходимо оценить параметры регрессии. Характерными особенностями регрессионного анализа в рамках классической линейной многофакторной

модели является выполнимость ряда предпосылок МНК

2.2. Теоретические предпосылки МНК

1°. *Математическое ожидание случайного отклонения ε_i равно нулю для всех наблюдений:*

$$M(\varepsilon_i) = 0, \quad i = 1, 2, \dots, n.$$

2°. *Наличие гомоскедастичности* (постоянство дисперсии случайных отклонений). Дисперсия случайных отклонений ε_i должна быть постоянной:

$$D(\varepsilon_i) = D(\varepsilon_j) = \sigma^2 \text{ для любых наблюдений с номером } i \text{ и } j.$$

3°. *Отсутствие автокорреляции.* Случайные отклонения ε_i и ε_j не должны зависеть друг от друга для всех $i \neq j$.

$$\text{Cov}(\varepsilon_i, \varepsilon_j) = \begin{cases} 0, & \text{если } i \neq j, \\ \sigma^2, & \text{если } i = j. \end{cases}$$

4°. Случайное отклонение должно быть независимым от объясняющих переменных:

$$\text{Cov}(\varepsilon_i, X_{ij}) = 0.$$

5°. Модель эмпирической регрессии должна являться линейной относительно параметров. Это ограничение не распространяется на факторные переменные.

6°. *Отсутствие мультиколлинеарности.* Между объясняющими переменными должна отсутствовать строгая (сильная) линейная зависимость.

7°. *Случайные величины - ошибки ε_i , $i = 1, 2, \dots, n$, должны иметь нормальный закон распределения ($\varepsilon_i \sim N(0)$).*

Выполнимость данной предпосылки важна для проверки статистических гипотез и построения интервальных оценок.

Как и в случае парной регрессии, истинные значения параметров b_j с помощью случайной выборки получить невозможно. В этом случае вместо теоретического уравнения регрессии (2.3) оценивается так называемое *эмпирическое уравнение регрессии*. Эмпирическое уравнение регрессии представим в виде:

$$Y = \hat{b}_0 + \hat{b}_1 X_1 + \hat{b}_2 X_2 + \dots + \hat{b}_p X_p + e. \quad (2.5)$$

Здесь $\hat{b}_0, \hat{b}_1, \dots, \hat{b}_p$ — оценки теоретических значений b_0, b_1, \dots, b_p коэффициентов регрессии (*эмпирические коэффициенты регрессий*); e — эмпирическая оценка неизвестного случайного отклонения.

2.3. Оценивание коэффициентов множественной линейной регрессии

Представим данные наблюдений и соответствующие коэффициенты в матричной форме.

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{pmatrix}, \quad X = \begin{pmatrix} 1 & x_{11} & x_{12} & \dots & x_{1p} \\ 1 & x_{21} & x_{22} & \dots & x_{2p} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_{n1} & x_{n2} & \dots & x_{np} \end{pmatrix}, \quad B = \begin{pmatrix} \hat{b}_0 \\ \hat{b}_1 \\ \dots \\ \hat{b}_p \end{pmatrix}, \quad e = \begin{pmatrix} e_1 \\ e_2 \\ \dots \\ e_n \end{pmatrix}.$$

Здесь Y — n -мерный вектор-столбец наблюдений зависимой переменной Y ; X — матрица размерности $n \times (p + 1)$, в которой i -я строка ($i = 1, 2, \dots, n$) представляет наблюдение вектора значений независимых переменных X_1, X_2, \dots, X_p ; единица соответствует переменной при свободном члене b_0 , B — вектор-столбец размерности $(p+1)$ параметров уравнения регрессии (2.5); e — вектор-столбец отклонений выборочных (реальных) значений y_i зависимой переменной Y от значений \hat{y}_i , размерности n , получаемых из модельного уравнения регрессии:

$$\hat{y}_i = X \cdot \hat{B} = \hat{b}_0 + \hat{b}_1 x_{i1} + \hat{b}_2 x_{i2} + \dots + \hat{b}_p x_{ip}. \quad (2.6)$$

Сумма квадратов отклонений МНК в матричном виде запишется следующим образом:

$$e^T \cdot e \rightarrow \min \quad \text{èèè} \quad S(\hat{B}) = (Y - X \cdot \hat{B})^T (Y - X \cdot \hat{B}) \rightarrow \min.$$

Условие экстремума: $S'(\hat{B}) = 0$. (2.7)

Таким образом, получим в матричной форме оценки параметров линейной регрессии:

$$\hat{B} = (X^T X)^{-1} (X^T Y). \quad (2.8)$$

Для вычисления их не нужно составлять и решать нормальную систему линейных уравнений МНК. Достаточно выполнить указанные в формуле (2.8) алгебраические операции в матричной форме над результатами исходных, выборочных наблюдений X и Y . В частности, при выполнении контрольной работы такие расчеты достаточно легко выполняются с помощью Мастера функций табличного процессора Microsoft Excel.

2.4. Интерпретация оценок параметров и уравнения

множественной линейной регрессии

Интерпретация – содержательное объяснение – результатов анализа экономического явления или объекта, представленного статистическими (выборочными данными), является одной из самых важных задач регрессионного анализа. Так, рассматривая полученные оценки параметров уравнения регрессии, можно сказать, изменение фактора X_i на одну единицу своего измерения ведет к изменению объясняемой переменной Y на \hat{b}_i единиц измерения этой переменной. Направление ее изменения определяется знаком коэффициента перед фактором X_i .

При этом нужно обязательно фиксировать, в каких единицах измерены значения всех переменных, прежде чем заменять слово «единица» конкретными названиями: тонны, рубли и т.п. Отсюда следует, что коэффициенты регрессии перед различными факторами нельзя сравнивать друг с другом.

Все другие показатели характера влияния факторов на объясняемую переменную, не зависящие от масштаба их измерения, такие как стандартизованные коэффициенты регрессии и коэффициенты эластичности, получают на основе этих оценок параметров \hat{b}_i .

При интерпретации модельного уравнения регрессии важно отмечать его следующие характерные особенности. Во-первых, \hat{b}_0 и \hat{b}_i являются только оценками неизвестных констант b_0 и b_i только теоретической, линейной

регрессии, которая не обязательно является таковой. Предпосылки МНК обосновывают условия, в которых можно получить «лучшие» оценки параметров модельной регрессии.

Во-вторых, расчетные значения объясняемой переменной не могут быть детерминированными и их нужно дополнять характеристиками вариации, например, стандартными ошибками или доверительными интервалами.

И, наконец, в-третьих, правильность интерпретации зависит от выбора модельного представления статистической связи и полноты совокупности объясняющих переменных.

3. Анализ качества уравнения регрессии

В общем случае характеристикой «качества» эконометрической модели является несмешенная оценка дисперсии случайных отклонений S_e^2 :

$$S_e^2 = \frac{\sum_{i=1}^n e_i^2}{n - p - 1} = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n - p - 1}, \quad (3.1)$$

где p – число объясняющих переменных, факторов. Корень квадратный из оценки дисперсии обозначается как S_e и называется *стандартной ошибкой регрессии*.

Надежность случайных оценок параметров множественной регрессии устанавливают также с помощью определения оценок их дисперсий (стандартных ошибок). Кроме того, строят доверительные интервалы для теоретических значений при заданном уровне значимости и проверяют статистические гипотезы о значимости различия их эмпирических величин от ожидаемых, теоретических значений.

3.1. Дисперсии и стандартные ошибки параметров линейной регрессии

Ошибка модельной регрессии во многом предопределена тем, что оценки $\hat{b}_0, \hat{b}_1, \dots, \hat{b}_p$ рассчитывают по данным случайных измерений, и они являются

случайными значениями величин b_0, b_1, \dots, b_p - неизвестных коэффициентов регрессии.

Дисперсия многомерной случайной величины \hat{B} определяется с помощью ковариационной матрицы $\Sigma_{\hat{B}}$:

$$\Sigma_{\hat{B}} = \sigma_e^2 (X^T X)^{-1} X^T E_n X (X^T X)^{-1} = \sigma_e^2 (X^T X)^{-1}. \quad (3.2)$$

Неизвестное значение дисперсии случайного отклонения теоретической регрессии σ_e^2 заменяется соответствующей несмешенной выборочной оценкой s_e^2 (3.1). Следовательно, по выборке мы можем определить только выборочные оценки дисперсий коэффициентов эмпирической регрессии $\hat{b}_0, \hat{b}_1, \dots, \hat{b}_p$, которые являются диагональными элементами матрицы $\Sigma_{\hat{B}}$:

$$S_{\hat{b}_j}^2 = S_e^2 (X^T X)_{jj}^{-1}, \quad (3.3)$$

где через $(X^T X)_{jj}^{-1}$ обозначены диагональные элементы обратной матрицы $(X^T X)^{-1}, j = 0, 1, 2, \dots, p$.

Стандартные ошибки коэффициентов регрессии вычисляются по формулам:

$$S_{\hat{b}_j} = S_e \sqrt{(X^T X)_{jj}^{-1}}, \quad (3.4)$$

где $j = 0, 1, \dots, p$.

3.2. Доверительные интервалы коэффициентов регрессии

При заданном уровне значимости α доверительный интервал записывается следующим образом:

$$\left| \frac{\hat{b}_i - b_i}{S_{\hat{b}_i}} \right| = \left| t_{\hat{b}_i} \right| < t_{\alpha, n-p-1}$$

где $t_{\alpha, n-p-1}$ - табличное значение t -критерия Стьюдента

Из данного неравенства следует:

$$\hat{b}_i - S_{\hat{b}_i} \cdot t_{\alpha, n-p-1} < b_i < \hat{b}_i + S_{\hat{b}_i} \cdot t_{\alpha, n-p-1}$$

$$\hat{b}_i - \Delta \hat{b}_i < b_i < \hat{b}_i + \Delta \hat{b}_i , \quad (3.5)$$

где $S_{\hat{b}_i}$ и $\Delta \hat{b}_i = S_{\hat{b}_i} \cdot t_{\alpha, n-p-1}$ - стандартная и предельная ошибки выборочных оценок \hat{b}_i соответственно.

3.3. Статистическая значимость уравнения регрессии

Проверить статистическую значимость уравнения множественной регрессии означает установить, соответствует ли регрессионная модель, принятая для объяснения взаимосвязи между переменными, исходным статистическим данным. Или, другими словами, достаточно ли включенных в уравнение регрессии факторов для описания поведения объясняемой переменной на основе имеющихся выборочных данных.

Проверка значимости уравнения регрессии производится с помощью метода статистического анализа – дисперсионного анализа. Оценивание качества уравнения регрессии с помощью F -критерия Фишера состоит в проверке гипотезы H_0 о статистической значимости уравнения регрессии или показателя тесноты связи. В случае, когда нулевая гипотеза отвергается, влияние включенных в регрессию факторов на объясняемую переменную преобладает над ее изменениями в силу других причин.

Для этого сравнивают фактическое значение критерия $F_{\text{эмп}}$ с критическим, табличным значением $F_{\text{рабл}}$.

$$F_{\text{эмп}} = \frac{Q_{\text{pee}} / p}{Q_e / (n - p - 1)} = \frac{R_{yx1x2}^2}{1 - R_{yx1x2}^2} \cdot \frac{n - p - 1}{p},$$

где p - число объясняющих переменных или факторов, включенных в модель.

$$F_{\text{рабл}} = F_{\alpha, v_1, v_2} = F(\alpha, p, n - p - 1),$$

где α - уровень значимости; v_1 и v_2 - число степеней свободы большей (числителя) и меньшей (знаменателя) дисперсий соответственно.

Если $F_{\text{эмп}} > F_{\text{мабл}}(\alpha, n-p-1)$, то гипотеза H_0 о случайной природе статистической связи отклоняется. С вероятностью ($1-\alpha$) делается заключение о статистической значимости уравнения регрессии в целом и показателя тесноты связи R_{yx1x2} , которые сформировались под совместным воздействием факторов X_1 и X_2 , которое нет оснований считать случайным. В противном случае оснований для отклонения гипотезы H_0 нет и данная статистическая связь статистически незначима.

4. Уравнение регрессии в стандартизованной форме

В эконометрике часто используют стандартизованные переменные, которые имеют нулевую среднюю величину и одинаковую дисперсию, равную единице.

$$\hat{t}_y = \frac{\hat{y} - \bar{Y}}{S_Y}, t_{x1} = \frac{x_1 - \bar{X}_1}{S_{x1}}, \dots, t_{xp} = \frac{x_p - \bar{X}_p}{S_{xp}}.$$

Уравнение регрессии в стандартизованной форме имеет вид:

$$\hat{t}_y = \hat{\beta}_1 \cdot t_{x1} + \hat{\beta}_2 \cdot t_{x2} + \dots + \hat{\beta}_p \cdot t_{xp}, \quad (4.1)$$

где $\hat{\beta}_1 = \frac{\hat{b}_1 S_{x1}}{S_Y}, \hat{\beta}_2 = \frac{\hat{b}_2 S_{x2}}{S_Y}, \dots, \hat{\beta}_p = \frac{\hat{b}_p S_{xp}}{S_Y}$ - стандартизованные коэффициенты регрессии.

Между стандартизованными коэффициентами регрессии $\hat{\beta}_i$ и коэффициентами \hat{b}_i обычной, естественной формы существует простая взаимосвязь:

$$\hat{b}_1 = \frac{\hat{\beta}_1 S_Y}{S_{x1}}, \hat{b}_2 = \frac{\hat{\beta}_2 S_Y}{S_{x2}}, \dots, \hat{b}_p = \frac{\hat{\beta}_p S_Y}{S_{xp}}, \quad (4.2)$$

которая дает другой способ вычисления коэффициентов \hat{b}_i по известным значениям $\hat{\beta}_i$, более удобный в случае, например, двухфакторной регрессионной модели.

Стандартизованные коэффициенты регрессии в частном случае модели с двумя факторами

$$\hat{\beta}_1 = \frac{r_{x1y} - r_{x1x2} \cdot r_{x2y}}{1 - r_{x1x2}^2}, \quad (4.3)$$

$$\hat{\beta}_2 = \frac{r_{x2y} - r_{x1x2} \cdot r_{x1y}}{1 - r_{x1x2}^2}. \quad (4.4)$$

Используя значения коэффициентов парной корреляции в уравнениях (4.3) и (4.4), получим $\hat{\beta}_1$ и $\hat{\beta}_2$. Затем с помощью формул (4.2) нетрудно вычислить оценки \hat{b}_i .

5. Использование многофакторной регрессионной модели для экономического анализа

5.1. Коэффициенты β_i стандартизованной регрессии

Стандартизованные коэффициенты регрессии показывают, на сколько стандартных отклонений S_y изменится в среднем объясняемая переменная Y , если соответствующая объясняющая переменная X_i изменится на величину S_{x_i} одного ее стандартного отклонения при сохранении неизменными значений среднего уровня всех остальных факторов.

В силу того, что в стандартизованной регрессии все переменные заданы как центрированные и нормированные случайные величины, коэффициенты β_i сравнимы между собой. Это позволяет ранжировать соответствующие им факторы X_i по силе воздействия на объясняемую переменную Y .

Эта особенность стандартизованных коэффициентов регрессии позволяет использовать при отсеве наименее значимых факторов X_i с близкими к нулю значениями их выборочных оценок $\hat{\beta}_i$. Решение об исключении их из модельного уравнения линейной регрессии принимается после проверки статистических гипотез о равенстве нулю его средней величины.

5.2. Средние и частные коэффициенты эластичности

В экономических исследованиях для количественного представления результатов множественного анализа также часто применяют коэффициенты эластичности. Коэффициент эластичности характеризует изменение в процентах объясняемой переменной при изменении факторной переменной на 1%..

Поскольку коэффициент эластичности для линейной функции не является постоянной величиной, то вводят средний показатель эластичности:

$$\overline{\mathcal{E}}_{y/x} = b \cdot \frac{\bar{X}}{\bar{Y}}.$$

Для характеристики относительной силы влияния каждого фактора в многофакторной регрессии используют частные коэффициенты эластичности. Для характеристики влияния X_1 и X_2 на Y рассчитывают частные коэффициенты эластичности:

$$\overline{\mathcal{E}}l_{y/x_1} = \hat{b}_1 \cdot \frac{\bar{X}_1}{\bar{Y}}, \quad \overline{\mathcal{E}}l_{y/x_2} = \hat{b}_2 \cdot \frac{\bar{X}_2}{\bar{Y}}.$$

Экономическая интерпретация этой статистической взаимосвязи является следующей. При увеличении фактора X_1 на 1% от его среднего уровня объясняемая переменная Y изменяется на \hat{b}_1/\bar{Y} ,% своего от среднего уровня, а при увеличении фактора X_2 на 1% от его среднего уровня объясняемая переменная Y изменяется на \hat{b}_2/\bar{Y} ,% от своего среднего уровня.

Направление изменения объясняемой переменной зависит от знака коэффициента при соответствующем факторе.

5.3. Линейные коэффициенты парной, частной и множественной корреляции

Одной из основных характеристик статистической связи является выборочный коэффициент линейной корреляции. При исследовании двухфакторной модели $\hat{Y} = \hat{b}_0 + \hat{b}_1 \cdot X_1 + \hat{b}_2 \cdot X_2$ линейные коэффициенты парной корреляции определяются следующим:

$$r_{x_1y} = \frac{\overline{x_1y} - \overline{x_1} \cdot \overline{y}}{S_{x_1} \cdot S_y},$$

$$r_{x_2y} = \frac{\overline{x_2y} - \overline{x_2} \cdot \overline{y}}{S_{x_2} \cdot S_y},$$

$$r_{x_1x_2} = \frac{\overline{x_1x_2} - \overline{x_1} \cdot \overline{x_2}}{S_{x_1} \cdot S_{x_2}}.$$

Анализируя значения коэффициентов парной корреляции, приходим к выводам относительно силы и направления взаимосвязи между каждой парой переменных. Однако очевидно, что влияние каждой из факторных переменных на объясняемую переменную не может выражаться также просто, как в случае парной регрессии, в силу их взаимосвязи между собой.

Для характеристики влияния каждого фактора множественной регрессии рассчитывают частные коэффициенты корреляции. Они также используются при отборе факторов: целесообразность включения того или иного фактора в модель можно обосновывать величиной показателя частной корреляции.

Частные коэффициенты корреляции характеризуют влияние фактора X_i на объясняемую переменную Y при устраниении влияния других факторов, включенных в уравнение регрессии. В зависимости от их числа вычисляют частные корреляции различного порядка.

В случае двухфакторной модели частные коэффициенты корреляции первого порядка вычисляются непосредственно с помощью парных коэффициентов корреляции по формулам:

$$r_{x_1y/x_2} = \frac{r_{x_1y} - r_{x_1x_2} \cdot r_{x_2y}}{\sqrt{(1 - r_{x_1x_2}^2)(1 - r_{x_1y}^2)}},$$

$$r_{x_2y/x_1} = \frac{r_{x_2y} - r_{x_1x_2} \cdot r_{x_1y}}{\sqrt{(1 - r_{x_1x_2}^2)(1 - r_{x_2y}^2)}},$$

$$r_{x_1x_2/y} = \frac{r_{x_1x_2} - r_{x_1y} \cdot r_{x_2y}}{\sqrt{(1 - r_{x_1y}^2)(1 - r_{x_2y}^2)}}.$$

Если сравнить значения коэффициентов парной и частной корреляции, то приходим к выводу, что из-за сильной связи x_1 и y , умеренной связи x_2 и y коэффициенты r_{x_1y} и r_{x_1y/x_2} отличаются незначительно, r_{x_2y} и r_{x_2y/x_1} - значительно. Выводы о направлении связи на основе этих коэффициентов

совпадают.

Практическая значимость эмпирического уравнения множественной регрессии оценивается помошью коэффициента множественной корреляции $R_{YX_1X_2...X_p}$ или его квадрата $R^2_{YX_1X_2...X_p}$ – коэффициента множественной детерминации.

Коэффициент множественной корреляции характеризует тесноту связи рассматриваемого набора факторов с объясняемой переменной, или оценивает тесноту совместного влияния факторов на результат. Его величина не может быть меньше, чем величина наибольшего коэффициента парной корреляции:

$$R_{YX_1X_2...X_p} \geq \max_{i=1,2,...,p} (r_{YX_i}).$$

Коэффициент множественной корреляции должен существенно отличаться от коэффициентов парной корреляции. Целесообразность включения в состав регрессии нового фактора можно обосновать, сравнивая коэффициенты множественной и парной корреляции с его участием.

Коэффициент множественной корреляции можно рассчитать несколькими способами, это полезно делать для контроля правильности вычислений:

1) наиболее просто это сделать для уравнения линейной регрессии в стандартизованной форме:

$$R_{YX_1X_2...X_p} = \sqrt{\beta_1 \cdot r_{X_1Y} + \beta_2 \cdot r_{X_2Y} + \dots + \beta_p r_{X_pY}}.$$

2) как корень квадратный из коэффициента множественной детерминации:

$$\begin{aligned} R_{YX_1X_2...X_p} &= \sqrt{R^2_{YX_1X_2...X_p}} = \sqrt{\frac{S_y^2}{S_{\hat{y}}^2}} = \sqrt{\frac{S_y^2 - S_e^2}{S_y^2}} = \\ &= \sqrt{1 - \frac{S_e^2}{S_y^2}} = \sqrt{1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_{(x_1x_2...x_p)_i})^2}{\sum_{i=1}^n (y_i - \bar{y})^2}}. \end{aligned}$$

На основании значения коэффициента множественной регрессии делается вывод о том, что зависимостью переменной y от факторных переменных объясняется $R^2 * 100\%$ вариации, представленной исходными данными. Поэтому факторы, не включенные в модель, составляют соответственно $(1 - R^2) * 100\%$ от

общей вариации объясняемой переменной y .

6. Порядок выполнения и оформления типового примера.

Исследуем зависимость объясняемой переменной Y от факторных переменных X_1 и X_2 , используя множественный регрессионный анализ. Для этого необходимо выполнить расчеты:

1. Вычисление осредненных характеристик выборки (допускается выполнение на компьютере, остальные разделы – с помощью калькулятора).
2. Вычисление коэффициентов парной корреляции.
3. Вычисление стандартизованных коэффициентов множественной линейной регрессии и ранжирование с их помощью факторных переменных.
4. Обоснование формы и оценка параметров линейной множественной регрессии.
5. Построение множественной линейной регрессии в естественной форме.
6. Вычисление стандартной ошибки регрессии.
7. Вычисление частных коэффициентов корреляции и частных коэффициентов эластичности.
8. Вычисление коэффициента множественной регрессии и индекса множественной корреляции.
9. Проверка гипотезы о статистической значимости полученного уравнения множественной регрессии
10. Вычисление доверительных интервалов параметров регрессии при уровне значимости $\alpha = 0,1$ или $0,05$.

Каждый из них должен включать:

- а) изложение теоретических предпосылок и обоснование расчетных формул;
- б) выполнение расчетов;
- в) обсуждение полученных результатов.

Имеются следующие выборочные данные:

№ п/п	X₁	X₂
1	28	-69
2	30	-67
3	29	-62
4	30	-77
5	31	-78
6	29	-63
7	30	-72
8	30	-79
9	29	-74
10	29	-68
11	30	-67
12	29	-72
13	30	-73
14	29	-66
15	29	-71
16	29	-66
17	29	-76
18	29	-73
19	30	-67
20	29	-69
21	29	-59
22	30	-76
23	31	-78
24	32	-82

№ п/п	X₁	X₂	Y
25	28	-69	298
26	30	-67	310
27	29	-62	295
28	30	-77	316
29	31	-78	311
30	29	-63	308
31	30	-72	309
32	30	-79	306
33	29	-74	307
34	29	-68	293
35	30	-67	303
36	29	-72	300
37	30	-73	311
38	29	-66	302
39	29	-71	299
40	29	-66	304
41	29	-76	309
42	29	-73	297
43	30	-67	315
44	29	-69	297
45	29	-59	308
46	30	-76	304
47	31	-78	318
48	32	-82	337

где X_1 и X_2 - объясняющие или факторные переменные; Y - объясняемая переменная; $n = 48$ - число выполненных измерений или объем выборки.

Исследование должно включать в себя наряду с получением оценок статистических характеристик многомерной пространственной выборки обоснование применяемых методов их вычисления, а также элементы экономического анализа.

Выполнение численных расчетов необходимо сопровождать обсуждением экономического содержания полученных результатов. В конце работы на отдельном листе прилагается список использованной методической и учебной литературы.

Вычисления всех средних величин, которые необходимы для проведения последующего анализа, можно оформить в форме таблицы Microsoft Excel:

На основе их получают следующие показатели:

1. Статистические дисперсии:

$$S_Y^2 = \bar{Y}^2 - \bar{Y}^2 = 93953,00 - 306,21^2 = 189,457;$$

$$S_{X_1}^2 = 1,423; S_{X_2}^2 = 38,177;$$

Стандартные (среднеквадратические) отклонения:

$$S_{X_1} = \sqrt{1,423} = 1,19; S_{X_2} = 6,18; S_Y = 13,76.$$

Ковариации:

$$S_{YX_1} = \bar{Y} \cdot \bar{X}_1 - \bar{Y} \cdot \bar{X}_1 = 9105,23 - 29,69 \cdot 306,21 = 14,67;$$

$$S_{YX_2} = -50,83; S_{X_1 X_2} = -4,51.$$

2. Линейные коэффициенты парной корреляции:

$$r_{YX_1} = 0,8934; r_{YX_2} = -0,5977; r_{X_1 X_2} = -0,6177.$$

Анализируя значения коэффициентов парной корреляции, приходим к следующему выводу:

Связанные переменные	Теснота связи	Направление связи
x_1 и y	сильная	прямая
x_2 и y	умеренная	обратная
x_1 и x_2	умеренная	обратная

3. Стандартизованные коэффициенты регрессии:

$$\hat{\beta}_1 = 0,8433; \hat{\beta}_2 = -0,0819.$$

Сравнивая модули значений $\hat{\beta}_1$ и $\hat{\beta}_2$, приходим к выводу, что сила влияния фактора X_1 на объясняемую переменную Y намного больше, чем сила влияния фактора X_2 .

4. Выборочные оценки значений коэффициентов множественной линейной

регрессии:

$$\hat{b}_1 = 9,703; \quad \hat{b}_2 = -0,182;$$

$$\hat{b}_0 = \bar{Y} - \hat{b}_1 \bar{X}_1 - \hat{b}_2 \bar{X}_2 = 306,21 - 9,73 \cdot 29,69 - 0,18 \cdot 69,90 = 4,614.$$

5. Уравнение множественной линейной регрессии в естественной форме:

$$\hat{Y} = \hat{b}_0 + \hat{b}_1 \cdot X_1 + \hat{b}_2 \cdot X_2 = 4,614 + 9,730 \cdot X_1 - 0,182 \cdot X_2$$

6. Вычисление стандартной ошибки регрессии S_e :

Таблица 2

№ п/п	x_{i1}	x_{i2}	Y_i	\hat{Y}_i	$e_i = Y_i - \hat{Y}_i$	e_i^2
1	28	-69	298	289,63	8,37	70,12
2	30	-67	310	308,72	1,28	1,64
3	29	-62	295	298,08	-3,08	9,48
4	30	-77	316	310,54	5,46	29,76
5	31	-78	311	320,46	-9,46	89,42
6	29	-63	308	298,26	9,74	94,84
7	30	-72	309	309,63	-0,63	0,40
8	30	-79	306	310,91	-4,91	24,10
9	29	-74	307	300,27	6,73	45,32
10	29	-68	293	299,17	-6,17	38,11
11	30	-67	303	308,72	-5,72	32,73
12	29	-72	300	299,90	0,10	0,01
13	30	-73	311	309,81	1,19	1,40
14	29	-66	302	298,81	3,19	10,18
15	29	-71	299	299,72	-0,72	0,52
16	29	-66	304	298,81	5,19	26,95
17	29	-76	309	300,63	8,37	70,02
18	29	-73	297	300,09	-3,09	9,52
19	30	-67	315	308,72	6,28	39,43
20	29	-69	297	299,36	-2,36	5,55
21	29	-59	308	297,53	10,47	109,58
22	30	-76	304	310,36	-6,36	40,48
23	31	-78	318	320,46	-2,46	6,03
24	32	-82	337	330,92	6,08	37,02
25	30	-79	316	310,91	5,09	25,92
26	30	-64	302	308,17	-6,17	38,11
27	31	-77	320	320,27	-0,27	0,08

№ п/п	x_{i1}	x_{i2}	Y_i	\hat{Y}_i	$e_i = Y_i - \hat{Y}_i$	e_i^2
28	29	-64	291	298,44	-7,44	55,41
29	29	-71	290	299,72	-9,72	94,49
30	32	-74	324	329,46	-5,46	29,77
31	29	-63	304	298,26	5,74	32,93
32	29	-68	292	299,17	-7,17	51,46
33	29	-63	309	298,26	10,74	115,31
34	30	-61	303	307,63	-4,63	21,40
35	30	-67	307	308,72	-1,72	2,96
36	31	-76	329	320,09	8,91	79,36
37	32	-73	336	329,27	6,73	45,24
38	31	-77	310	320,27	-10,27	105,56
39	30	-64	301	308,17	-7,17	51,46
40	27	-63	270	278,80	-8,80	77,48
41	32	-73	338	329,27	8,73	76,14
42	28	-60	283	287,98	-4,98	24,85
43	30	-63	305	307,99	-2,99	8,95
44	32	-82	332	330,92	1,08	1,18
45	30	-70	304	309,27	-5,27	27,75
46	28	-60	293	287,98	5,02	25,15
47	29	-73	299	300,09	-1,09	1,18
48	27	-66	283	279,35	3,65	13,33
Σ	1425	-3355	14698	14698,00	0,00	1798,07
Σ/n	29,69	-69,9	306,21	306,21	0,00	37,46

$$S_e^2 = \bar{e^2} - \bar{e}^2 = 37,46 - 0 = 37,46; \quad S_e = \sqrt{37,46} = 6,12.$$

7. Частные коэффициенты линейной корреляции:

$$r_{x1y/x2} = 0,832; \quad r_{x2y/x1} = -0,144; \quad r_{x1x2/y} = -0,216$$

Сравнение с парными коэффициентами корреляции показывает, что существует слабая взаимосвязь между факторными переменными. Но она сильно проявилась на связи Y и X_2 , уменьшив характеристику связи с -0,598 до -0,144. На связь Y и X_1 она повлияла незначительно (0,893 и 0,832 соответственно).

Средние частные коэффициенты эластичности:

$$\bar{\mathcal{E}}l_{y/x_1} = 90,944, \quad \bar{\mathcal{E}}l_{y/x_2} = 0,041.$$

При увеличении фактора X_1 на 1% от его среднего уровня объясняемая

переменная Y возрастает на 0,944 % от среднего уровня, а при увеличении фактора X_2 на 1% от его среднего уровня объясняемая переменная Y возрастает на 0,041 % от своего среднего уровня.

8. Характеристика совокупного влияния всех факторов на объясняемую переменную - коэффициент множественной корреляции R_{yx1x2} :

- для уравнения регрессии в стандартизованной форме:

$$R_{yx1x2} = \sqrt{\beta_1 \cdot r_{x1y} + \beta_2 \cdot r_{x2y}} = 0,8957$$

- как корень квадратный из коэффициента множественной детерминации:

$$R_{yx1x2} = \sqrt{R_{yx1x2}^2} = \sqrt{\frac{S_{\text{общ}}^2}{S_y^2}} = \sqrt{\frac{S_y^2 - S_e^2}{S_y^2}} = \sqrt{1 - \frac{S_e^2}{S_y^2}} = 0,8957.$$

На основании значения коэффициента множественной регрессии делается вывод, что зависимость объясняемой переменной y от факторов x_1 и x_2 характеризуется как тесная, в которой 80,2% вариации y определяются вариацией данных факторов. Прочие факторы, не включенные в модель, составляют соответственно только 19,8 % от общей вариации y .

9. Оценивание качества уравнения регрессии с помощью F -критерия Фишера состоит в проверке гипотезы H_0 о статистической значимости уравнения регрессии или показателя тесноты связи. Для этого сравнивают фактическое значение критерия $F_{\text{эмп}}$ с критическим, табличным значением $F_{\text{мабл}}$.

$$F_{\text{эмп}} = \frac{R_{yx1x2}^2}{1 - R_{yx1x2}^2} \cdot \frac{n - p - 1}{p} = \frac{0,8957}{1 - 0,8957} \cdot \frac{48 - 2 - 1}{2} = 193,22,$$

где p - число объясняющих переменных, или факторов, включенных в модель.

$$F_{\text{мабл}} = F_{\alpha, v_1, v_2} = F(0,05, 2, 45) = 3,25,$$

где $\alpha = 0,05$ - уровень значимости; $v_1 = 2$ и $v_2 = 45$ - число степеней свободы большей (числителя) и меньшей (знаменателя) дисперсий соответственно.

Так как в нашем случае $F_{\text{эмп}} > F_{\text{мабл}} (193,22 > 3,25)$, то гипотеза H_0 о

случайной природе статистической связи отклоняется. Имеющиеся статистические данные свидетельствуют о том, что в 95 % случаев связь обусловлена влиянием факторов регрессии, а остальные не включенные в нее факторы являются статистически не значимыми. С вероятностью 0,95 делаем заключение о статистической значимости уравнения регрессии в целом и показателя тесноты связи R_{yx1x2} , которые сформировались под неслучайным воздействием факторов X_1 и X_2 .

10. Найдем доверительные интервалы для коэффициентов регрессии \hat{b}_0 , \hat{b}_1 и \hat{b}_2 .

$$(X^T \cdot X)^{-1} = \begin{pmatrix} 13,4317 & -0,5013 & -0,021072 \\ 0,5014 & 0,0234 & 0,002763 \\ 0,0211 & 0,0028 & 0,000872 \end{pmatrix}.$$

Табличная величина критерия Стьюдента:

$$t_{\alpha,n-p-1} = t(0,05, 45) = 2,027.$$

Дисперсия стандартной ошибки оценки коэффициентов \hat{b}_i :

$$1) S_{\hat{b}_0}^2 = S_e^2 \cdot ((X^T \cdot X)^{-1})_{11} = 37,46 \cdot 13,4317 = 503,151.$$

Предельная ошибка выборки оценки коэффициента \hat{b}_0 :

$$\Delta \hat{b}_0 = S_{\hat{b}_0} \cdot t_{\alpha,n-p-1} = 22,431 \cdot 2,027 = 45,468$$

Таким образом, интервальная оценка коэффициента \hat{b}_0 :

$$4,614 - 45,468 < \hat{b}_0 < 4,614 + 45,468.$$

$$-40,854 < \hat{b}_0 < 50,082$$

$$2) S_{\hat{b}_1}^2 = S_e^2 \cdot ((X^T \cdot X)^{-1})_{22} = 37,46 \cdot 0,0234 = 0,877,$$

$$\Delta \hat{b}_1 = S_{\hat{b}_1} \cdot t_{\alpha,n-p-1} = 0,936 \cdot 2,027 = 1,898.$$

Таким образом: $9,730 - 1,898 < \hat{b}_1 < 9,730 + 1,898$,

$$7,832 < b_1 < 11,628.$$

$$3) S_{\hat{b}_2}^2 = S_e^2 \cdot ((X^T \cdot X)^{-1})_{33} = 37,46 \cdot 0,00087 = 0,0326$$

$$\Delta \hat{b}_2 = S_{\hat{b}_2} \cdot t_{\alpha, n-p-1} = 0,181 \cdot 2,027 = 0,366$$

Таким образом: $-0,182 - 0,366 < \hat{b}_2 < -0,182 + 0,366$

$$-0,548 < \hat{b}_2 < 0,184.$$

ЛИТЕРАТУРА

1. Greene, XV. H. Econometric analysis / W. H. Greene. — 7th ed. — Prentice Hall, 2012.
2. Wooldridge, J. M. Introductory econometrics: A modern approach /J. M. Wooldridge. — 5th cd. — Cengage Learning, 2012.
3. Айвазян, Д. Фантаццини. — М.: Магистр ; ИНФРА-М, 2014.
4. Баум, К. Ф. Эконометрика. Применение пакета stata : учебник и практикум для вузов / К. Ф. Баум ; под ред. С. А. Айвазяна, Г. И. Пеникаса. — М.: Издательство Юрайт, 2016.
5. Бернхт, Э. Р. Практика эконометрики. Классика и современность / Э. Р. Бернхт ; пер. с англ. С. Л. Айвазяна. — М.: ЮНИТИ-ДАНА, 2012.
6. Исследование операций в экономике : учебник для академического бакалавриата / отв. ред. Н. Ш. Кремер. — 3-е изд., перераб. и доп. — М. : Издательство Юрайт, 2016.
7. Костюнин, В. И. Эконометрика : учебник и практикум для прикладного бакалавриата / В. И. Костюнин. — М.: Издательство Юрайт, 2016.
8. Мардас, А. Н. Эконометрика : учебник и практикум для академического бакалавриата / А. Н. Мардас. — 2-е изд., испр. и доп. — М. : Издательство Юрайт, 2016.
9. Подкорытова, О. А. Анализ временных рядов : учеб, пособие для бакалавриата и магистратуры / О. А. Подкорытова, М. В. Соколов. — М. : Издательство Юрайт, 2016.
10. Тимофеев, В. С. Эконометрика : учебник для академического бакалавриата / В. С. Тимофеев, Л. В. Фаддеенков, В. Ю. Щеколдин. — 2-е изд., перераб. и доп. — М.: Издательство Юрайт, 2016.
11. Эконометрика: учебник для бакалавриата и магистратуры / С. В. Курышева [и др.]; под ред. И. И. Елисеевой. — М.: Издательство Юрайт, 2016.

Розенцвайг Александр Куртович,
Исавнин Алексей Геннадьевич

**Конспект лекций к изучению раздела
«Эконометрический анализ»**

учебно-методическое пособие

Подписано в печать 18.01.2024.

Формат 60x84/16. Печать ризографическая.

Бумага офсетная. Гарнитура «Times New Roman».

Усл.п.л. 1,875 Уч.-изд. л. 1,8

Тираж 100 экз. Заказ № 1236

Отпечатано в Издательско-полиграфическом центре
Набережночелнского института
Казанского (Приволжского) федерального университета

423810, г. Набережные Челны, Новый город, пр.Мира, 68/19
тел./факс (8552) 39-65-99 e-mail: ic-nchi-kpfu@mail.ru