МЧС России
Федеральное государственное бюджетное образовательное учреждение высшего образования
«Санкт-Петербургский университет Государственной противопожарной службы Министерства Российской Федерации по делам гражданской обороны, чрезвычайным ситуациям и ликвидации последствий стихийных бедствий»
Кафедра прикладной математики и информационных технологий

Эконометрика

Сборник статей

Санкт-Петербург
2022

УДК 681/518(075/8)
ББК 65ф.я73
Э40

Эконометрика: Сб. статей. / Под ред. Т.Н. Антошиной, А.А. Кабанова, А.В. Матвеева. – СПб.: ФГБОУ ВО СПБУ ГПС МЧС России, 2022. – 43 с.

В сборнике кратко рассматриваются актуальные вопросы эконометрики, объёмом не более одной страницы на вопрос. В него вошли статьи студентов 3 курса факультета обеспечения безопасности жизнедеятельности Санкт-Петербургского университета ГПС МЧС России набора 2018 года, обучающихся по специальности 38.05.01 «Экономическая безопасность». Вступительная статья написана составителем сборника А.А. Кабановым, заключительная статья – Т.Н. Антошиной. Замечания и предложения по сборнику просим присылать по *e-mail:* ***akabanov@inbox.ru.***

За последние три столетия
экономика – как предмет
исследования учёных – неоднократно
качественно изменялась.
Евгений Филиппович Борисов.[1]

Эконометрика как наука

А.А. Кабанов, доцент кафедры
прикладной математики и информационных технологий,
кандидат юридических наук, доцент

Эконометрика – это наука, предметом изучения которой является количественное выражение взаимосвязей экономических явлений и процессов. Её цель состоит в описании конкретных количественных взаимосвязей, обусловленных общими качественными закономерностями, выявленными в экономической теории. Предмет исследования эконометрики – это массовые экономические явления и процессы. Экономические санкции возможны только со стороны стран с более современными технологиями, поэтому лучшим ответом на них является разработка своих уникальных, принципиально новых технологий. Этому процессу способствует самостоятельный поиск ответов на актуальные вопросы, имеющий место среди студентов, обучающихся по специальности «Экономическая безопасность». Каждый студент самостоятельно выбирал тему исследования, искал информацию, известную по этому вопросу, самостоятельно выбирал и формулировал то, что на его взгляд, представляет более важное. Навыки подобного рода исследований пригодятся и в других областях деятельности. Некоторые статьи носят компилятивный характер. Но имеются также статьи, представляющие творческое переосмысление известной информации.

Для роста экономики главное – это человеческий капитал. Навыки, приобретённые студентами при работе над статьями для сборника, – важный шаг в формировании человеческого капитала в нашей стране. Затянувшаяся стагнация экономики не может продолжаться долго. Человеческий капитал пробьёт себе дорогу, не смотря на препятствия, создаваемые отдельными недальновидными лицами. Активность студентов – очевидное тому доказательство.

[1] Хрестоматия по экономической теории / Сост. Е.Ф. Борисов. – М.: Юристъ, 2000. – С. 5.

Понятие, предмет и задачи эконометрики

К.Д. Асриева, студентка учебной группы 1831эб;
Е.Е. Грищенко, студентка учебной группы 1832эб

Эконометрика – это инструментальная наука, которая позволяет изучать количественные взаимосвязи различных хозяйственных процессов и объектов с помощью математических и статистических методов и моделей. Эконометрика, чаще всего, занимается разработкой и применением на практике статистических методов для измерения взаимосвязей между хозяйственными переменными.

С одной стороны, она изучает экономические процессы (следовательно, базируется на экономической теории), с другой стороны – призвана характеризовать их с математической точностью (в виде математических моделей), и при этом – имеет дело не с функциональными (конкретно проявляющимися), а со случайными, т.е. стохастическими связями.

Эконометрика – одна из базовых дисциплин экономического образования во всем мире.

Предмет изучения эконометрики – это количественные характеристики хозяйственных процессов и явлений.

Задачи эконометрики:

- Спецификация модели – это построение эконометрических моделей для эмпирического анализа данных.
- Параметризация модели – оценка параметров построения модели, делающих выбранную модель наиболее адекватной реальным данным.
- Верификация модели – проверка качества параметров модели и всей модели в целом.
- Прогнозирование модели – составление прогноза и рекомендация для конкретных эконометрических явлений по результатам моделирования.
- Применение построенных моделей с целью объяснения действия исследуемых экономических показателей, прогнозирования, осознанного выполнения экономической политики.

Составляющие эконометрики и их взаимосвязь. Основные разделы эконометрики

К.Д. Асриева, студентка учебной группы 1831эб

1) «Эконометрика – довольно схожа с экономической статистикой, но они не идентичны. Она не идентична и дисциплине, которую мы называем экономической теорией. Главные три отправных точки – статистика, экономическая теория и математика. Это единство всех трёх составляющих. И это единство образует эконометрику», т.к. в этой науке необходим статистический анализ, в котором используются математические данные и базовые знания терминологии из экономической теории. (Р. Фриш)[1].

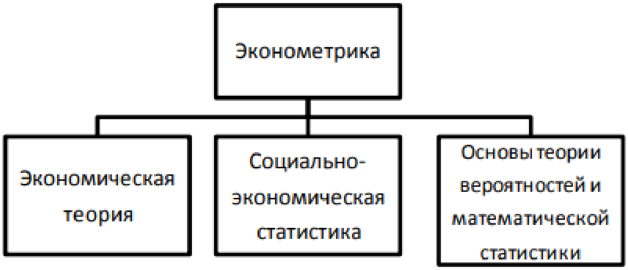

Рис.1 Составляющие эконометрики

2) в эконометрике существуют два раздела: корреляционный и регрессионный анализ.

Основная задача корреляционного анализа – это установление существования взаимосвязи между признаками, определение тесноты этой связи.

Регрессионный анализ изучает вопросы установления формы связи, вида функции, устанавливающей зависимость одного фактора от одного или нескольких других факторов.

Корреляционный и регрессионный анализ могут подразделяться на парный и множественный, линейный и нелинейный анализ.

[1] Цит. по: Эконометрика. Курс лекций. – Екатеринбург: ГОУ ВПО УрГУ, 2007. – С. 5 URL: https://docviewer.yandex.ru/ (дата обращения: 09.02.2021)

Понятие эконометрической модели, основные типы моделей

К.А. Датова, студентка учебной группы 1831эб

Эконометрическая модель является главным элементом эконометрики. Она представляет собой модель, которая описывает функционирование экономической системы.

Эконометрическая модель – это уравнение (или система уравнений), которое с математической точки зрения описывает основные зависимости между экономическими явлениями и процессами.

Эконометрическая модель включает в себя эндогенные и экзогенные переменные. Эндогенные переменные также принято называть внутренними, так как их значения определяются внутри модели и внутренней структурой изучаемого процесса. Такие переменные обозначаются как ***Y.*** Примером эндогенных переменных является: выпуск продукции, объём потребления и т.д.

Экзогенные переменные в отличие от эндогенных независимы от структуры процесса или явления. Эти переменные считаются заданными. Соответственно экзогенные переменные обозначаются как ***X.*** В качестве экзогенных переменных могут использоваться объём инвестиций, поставка ресурсов и прочее.

Говоря о типах эконометрических моделей, можно выделить следующие: 1) однофакторные регрессионные модели; 2) многофакторные регрессионные модели; 3) модели временных рядов.

Однофакторные регрессионные модели связывают **X** и **Y**, то есть экзогенную и эндогенную переменные. Такая связь записывается в виде уравнения. В зависимости от вида функции модели могут делиться на линейные и нелинейные модели. *Многофакторные* модели описывают линейную и нелинейную зависимость переменной от нескольких факторов. При построении эконометрических моделей могут быть использованы два типа данных: пространственные данные и временные ряды. Пространственными данными можно назвать набор значений экономических переменных, которые были получены в какой-то момент времени. Например, число работников, стаж работы и прочее. Модели временны́х рядов описывают зависимость среднего значения зависимой переменной от времени.

К *моделям временны́х рядов*, в которых результативный признак зависит от времени, относятся: модель тренда (модель зависимости результативного признака от трендовой компоненты); модель сезонности (модель зависимости результативного признака от сезонной компоненты); модель тренда и сезонности. К временны́м данным может относиться, например, ежедневный курс валюты.

Теоретическая и выборочная регрессионные модели. Основные этапы эконометрического моделирования

Е.Е. Грищенко, студентка учебной группы 1832эб

Проанализируем теоретическую регрессионную модель зависимости переменной **Y** от переменной **X:**

$$Y = f(X, \beta) + \varepsilon$$

Обе переменные можно рассматривать как случайные величины. Для конкретных значений этих величин модель примет вид выборочной регрессионной модели:

$$y = f(x, \beta) + \varepsilon$$

Если функция $f(x)$ определяет условное математическое ожидание величины **Y** при условии, что случайная величина **X** примет значение x, т.е.

$$f(x) = M(Y/X=x),$$

то она называется теоретической функцией парной регрессии **Y** на **X,** определяется теоретическим уравнением парной регрессии и обозначается $\hat{y} = f(x)$. При этом сама модель называется теоретической моделью парной регрессии.

Таким же образом можно определить теоретическую функцию множественной регрессии, если переменная **Y** зависит от нескольких переменных $X_1, X_2, \ldots, X_n$:

$$f(x_1, x_2, \ldots, x_n) = M(Y/X_1=x_1, X_2=x_2, \ldots, X_n= x_n)$$

В этом случае, теоретическое уравнение множественной регрессии $\hat{y} = f(x_1, x_2, \ldots, x_n)$ определяет теоретическую модель множественной регрессии:

$$Y=f(X_1, X_2, \ldots, X_n, \beta) + \varepsilon$$

Этапы эконометрического моделирования:

I. Постановочный – определение целей и задач модели.
II. Априорный – подготовительный анализ ситуации.
III. Спецификация модели – выбор типа модели, состава переменных и формы математической связи между ними.
IV. Информационный – сбор первичных данных.
V. Идентификация модели – оценивание параметров модели.
VI. Верификация модели – контроль качества модели в целом и её параметров.
VII. Интерпретация результатов – формулирование заключений (выводов) и советов на основе построенной модели.

Выбор формы уравнения регрессии

К.Е. Скоробогатько, студентка учебной группы 1832эб;
А.А. Кабанов, доцент кафедры
прикладной математики и информационных технологий,
кандидат юридических наук, доцент

Для определения формы исходного уравнения регрессии чаще всего используют метод перебора различных уравнений.

Данный метод заключается в возможности ЭВМ реализовывать большое количество уравнений, выбранных для описания взаимосвязей социально-экономического явления или процесса, с помощью специально разработанного алгоритма перебора с последующей статистической проверкой на основе критериев Стьюдента или Фишера. Этот способ достаточно трудоёмок и связан с большим объёмом вычислительных работ. Поэтому выбор осуществляется с помощью компьютерных программ (например, на Excel).

Основные виды множественной регрессии[1]:

1) линейная: $\hat{Y} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n$ **;**

2) степенная: $\hat{Y} = \beta_0 X_1^{\beta 1} \bullet X_2^{\beta 2} \bullet \ldots \bullet X_n^{\beta n}$ **;**

3) показательная: $\hat{Y} = e^{\beta 0 + \beta 1 X 1 + \beta 2 X 2 + \ldots + \beta n X n}$ **;**

4) параболическая: $\hat{Y} = \beta_0 + \beta_1 X_1^2 + \beta_2 X_2^2 + \ldots + \beta_n X_n^2$ **;**

5) гиперболическая: $\hat{Y} = \beta_0 + \beta_1/X_1 + \beta_2/X_2 + \ldots + \beta_n/X_n$ **.**

Наиболее простые и широко используемые – линейные модели.

После выбора общего вида уравнения регрессии, решают задачу отбора в регрессионную модель факторов, оказывающих наибольшее влияние на результирующую переменную.

[1] См., напр.: Иванов А.Н., Подружкина Т.А. Эконометрика: курс лекций для курсантов, студентов и слушателей (Экономическая безопасность). – СПб.: Санкт-Петербургский университет ГПС МЧС России, 2017. – 68 с. ***Режим доступа:*** http://elib.igps.ru/?4&type=card&cid=ALSFR-9945d866-5f4d-4a9e-887e-d1661fa8b1dd&remote=false

Корреляционное поле

К.Е. Скоробогатько, студентка учебной группы 1832эб;
Д.К. Палилова, студентка учебной группы 1831эб

Корреляционное поле (диаграмма рассеивания) – графическое изображение выборочных статистических данных в виде точек с координатами $(x_i, y_i), i = 1, 2, \ldots, n$ в декартовой системе координат[1].

Расположение точек корреляционного поля определяет вид функции регрессии. Точки корреляционного поля, располагающиеся недалеко от некоторой воображаемой прямой, говорят о появлении парной линейной регрессии.

Воображаемая прямая может располагаться под разными углами к положительному направлению оси $0x$. В соответствии с этим регрессия может быть прямой или обратной.

Между признаками может существовать прямая линейная зависимость в том случае, если расположение точек определяет прямую, образующую острый угол с положительным направлением оси абсцисс.

Таким образом, визуальный анализ корреляционного поля помогает выявлять наличие линейной или нелинейной зависимости между признаками, её тесноту и форму.

Известен способ построения корреляционного поля с помощью MS Excel:

«Для построения поля корреляции (или диаграммы рассеивания) в MS Excel используем **Мастер диаграмм.**

В диалоговом окне выбираем **Точечная**

После вставки диаграммы можно добавить линию регрессии. Для этого нажимаем на одной из точек правую кнопку мыши и выбираем команду **Добавить линию тренда**

Выбираем тип – **Линейная**, Параметры – **Показывать уравнение на диаграмме»**[2]**.**

---

[1] См., напр.: Иванов А.Н., Подружкина Т.А. Эконометрика: курс лекций для курсантов, студентов и слушателей (Экономическая безопасность). – СПб.: Санкт-Петербургский университет ГПС МЧС России, 2017. – 68 с. ***Режим доступа:*** http://elib.igps.ru/?4&type=card&cid=ALSFR-9945d866-5f4d-4a9e-887e-d1661fa8b1dd&remote=false

[2] https://math.semestr.ru/corel/scatter-chart.php (дата обращения: 17.02.2021)

Выборочное уравнение парной линейной регрессии

А.Г. Парк, студентка учебной группы 1831эб;
А.А. Кабанов, доцент кафедры
прикладной математики и информационных технологий,
кандидат юридических наук, доцент

Рассмотрим требования к теории и правила использования корреляционного и регрессионного анализа на наглядном примере, показав влияние только одного фактора на эффективный фактор в сочетании с линейной зависимостью. Парную линейную зависимость можно выразить уравнением:

$$y = a + b \cdot x$$

где y – зависимая переменная; a – свободный фактор уравнения; b – коэффициент регрессии; x – независимая переменная.

Выборочное уравнение линейной регрессии Y на X:

$$y = a + b \cdot x$$

где числовые параметры:

a – свободный член;

b – коэффициент регрессии;

$$a = \overline{y} - b \cdot \overline{x}, \qquad b = \frac{\operatorname{cov}(x, y)}{\sigma_x^2},$$

$\overline{y}$, $\overline{x}$ – выборочные средние случайных величин X и Y;

$\sigma_x^2 = \overline{x^2} - \overline{x}^2$ – выборочная дисперсия случайной величины X;

$\operatorname{cov}(x, y) = \overline{y \cdot x} - \overline{y} \cdot \overline{x}$ – выборочная ковариация X и Y.

Таким образом, мы рассмотрели выборочное уравнение парной линейной регрессии.

Выборочная модель и выборочное уравнение линейной регрессии

К.Е. Скоробогатько, студентка учебной группы 1832эб

В случае если корреляционное поле, построенное по некоторой *выборке*, показало существование линейной зависимости между признаками Y и X, теоретическая модель (1) и теоретическое уравнение (2) парной линейной регрессии Y на X имеют вид:

$$y = \beta_0 + \beta_1 x_1 + \varepsilon \qquad (1)$$

и

$$\hat{y} = \beta_0 + \beta_1 x \qquad (2),$$

где β_0 и β_1 неизвестные теоретические коэффициенты, а ε – теоретическое случайное отклонение или случайный элемент.

Истинные значения β_0 и β_1 практически найти невозможно, потому, что нельзя получить все значения признаков Y и X, которых может быть бесконечное множество. Поэтому из генеральной совокупности выбирают парную выборку ограниченного объёма (xi, yi), $i = 1, 2, \ldots, n$ и по ней определяют приближённые значения теоретических коэффициентов.

Иначе говоря, по выборочным данным строят выборочную (эмпирическую) модель регрессии:

$$y = b_0 + b_1 x + e \qquad (3)$$

и выборочное уравнение регрессии:

$$\hat{y} = b_0 + b_1 x \qquad (4),$$

где b_0 и b_1 – приближённые значения неизвестных коэффициентов β_0 и β_1, называемые оценками коэффициентов, выборочными коэффициентами регрессии или параметрами, а величина e – оценка теоретического случайного элемента ε, называемая остатком.

Уравнения (3) и (4) для конкретных значений выборки (x_i, y_i), $i = 1, 2, \ldots, n$ имеют вид:

$$y_i = b_0 + b_1 x_i + e_i \qquad (5)$$

и

$$\hat{y}_i = b_0 + b_1 x_i \qquad (6).$$

По конкретной выборке можно найти оценки неизвестных коэффициентов так, чтобы построенная линия регрессии была лучшей среди всех остальных прямых.

Метод наименьших квадратов, его суть и применение для расчёта оценок коэффициентов выборочного уравнения регрессии

К.Е. Скоробогатько, студентка учебной группы 1832эб;
А.А. Кабанов, доцент кафедры
прикладной математики и информационных технологий,
кандидат юридических наук, доцент

Для расчёта оценок коэффициентов $\boldsymbol{b_0}$ и $\boldsymbol{b_1}$ выборочного уравнения регрессии

$$y = b_0 + b_1 x$$

наиболее распространённым является метод наименьших квадратов.

Метод наименьших квадратов (МНК) – это метод определения оценок, который использует условие минимизации суммы квадратов оценок случайных отклонений значения функции.

Задача МНК – нахождение лучшей, оптимальной модели.

Согласно МНК коэффициенты выбираются таким образом, чтобы минимизировать сумму квадратов отклонений выборочных значений от расчётных, то есть минимизировать функцию:

$$L = \sum_{i=1}^{n} e_i^{\,2} = \sum_{i=1}^{n} (y_i - y_i)^2 \to \min$$

$$\text{или } L = \sum_{i=1}^{n} \left(y_i - b_0 - b_1 x_i \right)^2 \to \min$$

Решением данной задачи являются такие значения параметров $\boldsymbol{b_0}$ и $\boldsymbol{b_1}$, для которых функция $\boldsymbol{L}$ минимальна.[1]

[1] См., напр.: Иванов А.Н., Подружкина Т.А. Эконометрика: курс лекций для курсантов, студентов и слушателей (Экономическая безопасность). – СПб.: Санкт-Петербургский университет ГПС МЧС России, 2017. – 68 с. ***Режим доступа:*** http://elib.igps.ru/?4&type=card&cid=ALSFR-9945d866-5f4d-4a9e-887e-d1661fa8b1dd&remote=false

Проверка качества уравнения регрессии. Средняя ошибка аппроксимации

Д.Л. Дойкова, студентка учебной группы 1832эб

Качество уравнения регрессии определяется по таким критериям, как: оценки значимости коэффициентов уравнения регрессии или оценки значимости отдельных коэффициентов уравнения регрессии.

Значение средней ошибки аппроксимации указывает на правильность выбора уравнения регрессии.

Дисперсный анализ, то есть анализ величин разброса значений, позволяет проверить значимость уравнения регрессии.

Рассмотрим выражение между значением $\boldsymbol{y_i}$ переменной $\boldsymbol{Y}$ и средним значением $\boldsymbol{\bar{y}}$: $\boldsymbol{y_i - \bar{y}}$

Далее приведём разность к виду:

$$yi-\bar{y} = y_i-\bar{y}+\hat{y}_i-\hat{y}_i = (\hat{y}_i - \bar{y})+(y_i - \hat{y}_i)$$

На отклонение значения $\boldsymbol{y_i}$ влияет определение уравнения регрессии, а также случайная составляющая.

Средняя ошибка аппроксимации.

Средняя ошибка аппроксимации – это несоответствие расчётных показателей к фактическим.

Средняя ошибка аппроксимации определяется по следующей формуле:

$$A = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - y_i}{y_i}\right| \cdot 100\%$$

В случае, если значение средней ошибки аппроксимации находится в пределах 5-7%, то это говорит о хорошем подборе значения, приближённого к исходным данным.

Если же значение аппроксимации находится в границах 8-10%, то это свидетельствует о повышенной, но разрешимой ошибке аппроксимации.

Парный линейный коэффициент корреляции, его свойства, оценка и значимость, его связь с коэффициентом регрессии

Д.Л. Дойкова, студентка учебной группы 1832эб

Наилучшим показателем степени тесноты связи считается линейный коэффициент парной корреляции, характеризующий связь между ***X*** и ***Y***. Его значение колеблется от 0 до ±1. Коэффициент парной корреляции рассчитывается по формуле[1]:

$$r_{\text{ген.}} = \frac{M\left((X - M(X)(Y - M(Y)\right)}{\sigma(X)\sigma(Y)}$$

Свойства линейного парного коэффициента корреляции:

1. $r_{\text{ген.}}$ находится в интервале от -1 до +1.
2. $r_{\text{ген.}}$ не имеет единиц измерения.
3. Если $r_{\text{ген.}} = 0$, значит между ***X*** и ***Y*** нет линейной зависимости.
4. Если $r_{\text{ген.}} = \pm 1$, значит между ***X*** и ***Y*** есть линейная зависимость.
5. Если $r_{\text{ген.}}$ стремится к 0, то зависимость между ***X*** и ***Y*** слабее.
6. Если $r_{\text{ген.}}$ стремится к -1 или +1, то линейная зависимость между ***X*** и ***Y*** сильнее.

Оценка значимости коэффициента корреляции определяется с помощью t-критерия Стьюдента. Фактическое значение определяется по формуле:

$$T_{\text{в}} = \frac{r_{\text{в}}}{\sqrt{1 - \left(r_{\text{в}}\right)^2}} \sqrt{n-2}$$

Если $\left|T_{\text{в}}\right| > t_{\text{кр.}}$, то полученное значение коэффициента корреляции признаётся значимым, между переменными есть тесная взаимосвязь. Если $\left|T_{\text{в}}\right| < t_{\text{кр.}}$, то полученное значение коэффициента корреляции не признаётся значимым, между переменными нет связи.

Парный линейный коэффициент корреляции связан с коэффициентом регрессии следующим образом:

1. Если коэффициент регрессии $b_1 > 0$, то коэффициент корреляции имеет интервал $0 \leq r_{\text{в}} \leq 1$.

2. Если коэффициент регрессии $b_1 < 0$, то коэффициент корреляции имеет интервал $-1 \leq r_{\text{в}} \leq 0$.

---

[1] См., напр.: Иванов А.Н., Подружкина Т.А. Указ. соч.

Коэффициент детерминации и его смысловое значение

А.К. Травина, студентка учебной группы 1831эб;
Т.Н. Антошина, доцент кафедры
прикладной математики и информационных технологий,
кандидат педагогических наук

Коэффициент детерминации – основной показатель, который показывает меру качества регрессионной модели, описывающей связь между зависимой и независимыми переменными модели.

Коэффициент детерминации – это доля объяснённой дисперсии отклонений зависимой переменной от её среднего значения.

Таким образом, коэффициент детерминации показывает, какая доля вариации объясняемой переменной $\boldsymbol{y}$ учтена в модели и обусловлена влиянием на неё факторов, включённых в модель:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \widehat{y_i})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$$

где $\boldsymbol{y_i}$ – фактические значения наблюдаемой переменной, $\bar{\boldsymbol{y}}$– среднее значение по наблюдаемым данным, $\widehat{\boldsymbol{y_i}}$ – модельные (расчётные) значения, построенные по оцененным параметрам.

Этот показатель может находиться в пределах от 0 до 1.

Чем ближе к 1, тем сильнее зависимость. Приемлемой считается 0,5, хорошей – 0,8. При $R^2 = 1$ между переменными существует линейная функциональная зависимость. При $R^2 = 0$ линия регрессии параллельна оси $\boldsymbol{0x}$.

В этом и состоит его смысловое значение.

Основные типы нелинейной парной зависимости между признаками

А.К. Травина, студентка учебной группы 1831эб;
Д.Л. Дойкова, студентка учебной группы 1832эб

Для того, чтобы описать нелинейную взаимосвязь между признаками ***X*** и ***Y***, можно использовать элементарные нелинейные функции математики. Название регрессии, которая описывает нелинейную зависимость, относится к видам нелинейной функции. Пример:

$\hat{Y} = \beta_0 + \beta_1 X + \beta_2 X^2$ – параболическая регрессия;

$\hat{Y} = \beta_0 + \frac{\beta_1}{X}$ – гиперболическая регрессия;

$\hat{Y} = \beta_0 X^{\beta_1}$ – степенная регрессия;

$Y = \beta_0 \cdot \beta_1^X$ – показательная регрессия;

$Y = \beta_0 \cdot e^{\beta_1 X + \beta_2}$ – экспоненциальная регрессия.

Для того, чтобы решить задачу нелинейной регрессии нужно преобразовать нелинейную модель в линейную. Данное преобразование происходит путём линеаризации.

Линеаризация – это процесс, который позволяет преобразовывать нелинейную модель к линейному виду.

Стоит выделить два вида нелинейной регрессии:

1. Первый вид – регрессии, нелинейные касательно объясняющих переменных, но линейные по оценке параметров. Например, гиперболическая и параболическая регрессии.

2. Второй вид – регрессии, нелинейные по оценке параметров. Например, показательная, степенная, экспоненциальная регрессии.

Нелинейные модели можно разделить на два типа: модели внутренне линейные и модели внутренне нелинейные.

Если нелинейная модель внутренне линейна, то она может быть интерпретирована к линейному виду. Но при этом нелинейные функции, которые внутренне не линейны, не могут быть сведены к линейным.

Понятие процесса линеаризации

В.А. Остапчук, студентка учебной группы 1832эб

Линеаризацией называется процесс преобразования нелинейных уравнений в линейные.

Благодаря линеаризации можно выявить качественные и количественные свойства нелинейной системы.

Для линеаризации модели могут использоваться модели нелинейные по переменным и нелинейные по параметрам.

Существует три метода линеаризации.

Методы линеаризации:

1) Метод логарифмирования — применяется к степенным функциям;

Логарифмирование — это переход от уравнения f(x) = g(x) к уравнению $log_a f(x) = log_a g(x)$

2) Метод обратного преобразования — используется для дробных функций;

Метод обратного преобразования – это способ генерации случайных величин с заданной функцией распределения, путём модификации работы генератора равномерно распределённых чисел.

3) Комплексный метод — предназначен для дробных и степенных функций.

Регрессии, нелинейные по переменным и линейные по оцениваемым параметрам

Д.А. Смольская, студентка учебной группы 1831эб

Группа регрессий, которые не линейны по независимым переменным, но линейны по оценочным параметрам, содержит уравнения, в которых зависимая переменная линейно сопряжена с параметрами. Примером подобных регрессий могут быть:

- полиномы различных степеней

$$y_i = a_0 + a_1 x_i + a_2 x_i^2 + a_3 x_i^3 + \ldots + a_k x_i^k + \varepsilon_i$$

(полином k-й степени);

- равносторонняя гипербола

$$y_i = a + \frac{b}{x_i} + \varepsilon_i$$

При оценке характеристик регрессий, нелинейных по объясняющим переменным, применяется способ замены переменных. Сущность его заключается в замене нелинейных объясняющих переменных новыми, линейными переменными, в следствии чего нелинейная регрессия сводится к линейной. К новой, преобразованной регрессии может быть использован обыкновенный метод наименьших квадратов.[1]

[1] https://pandia.ru/text/77/203/77731.php (дата обращения: 05.08.2021)

Регрессии, нелинейные по оцениваемым параметрам, их виды

В.А. Остапчук, студентка учебной группы 1832эб;
Т.Н. Антошина, доцент кафедры
прикладной математики и информационных технологий,
кандидат педагогических наук

К классу регрессий**,** нелинейных по оцениваемым параметрам, относятся уравнения, в которых результирующая переменная – ***y*** нелинейно связана с параметрами ***x***.

Регрессии, нелинейные по оцениваемым параметрам:

1) степенная функция:

$$y = a * x^{b} * \varepsilon$$

2) показательная функция:

$$y = a * b^{x} * \varepsilon$$

3) экспоненциальная функция:

$$y = e^{a+b*x} * \varepsilon$$

Нелинейные модели данного класса подразделяется на 2 типа:

- нелинейные модели внутренне линейные, которые могут быть приведены к линейному виду с помощью таких преобразований, как замена переменных или логарифмирование;

- нелинейные модели внутренне нелинейные, которые не могут быть приведены к линейной функции.

Расчёт оценок параметров для некоторых видов нелинейной регрессии

В.А. Остапчук, студентка учебной группы 1832эб;
Т.Н. Антошина, доцент кафедры
прикладной математики и информационных технологий,
кандидат педагогических наук

Оценка параметров нелинейной регрессии по объясняющим переменным проводится методом наименьших квадратов, так как эти функции линейны по параметрам.

Например, для полинома k-го порядка:

$$y_i = a + b_1x + b_2x^2 + \dots + b_kx^k + \varepsilon$$

1) проводим замену переменных $x_1 = x$; $x_2 = x^2$; $x_k = x^k$

2) получаем линейную модель множественной регрессии

$$y = a + b_1x_1 + b_2x_2 + \dots + b_kx_k + \varepsilon$$

3) можно сделать вывод, что полином k-го порядка сводится к линейной регрессии с её методами оценивания и проверки гипотез.

Степенная функция $y = a * x^b * \varepsilon$ не линейна относительно параметров $\boldsymbol{a}$ и $\boldsymbol{b}$, но внутренне линейна, т.к. прологарифмировав, её можно привести к линейному виду:

$$ln\ y = ln\ a + b{*}ln\ x + ln\ \varepsilon$$

или производя обозначения:

$$y_1 = a_1 + b{*}\ x_1 + \varepsilon_1 \text{, где } y_1 = ln\ y;\ a_1 = ln\ a;\ x_1 = ln\ x;\ \varepsilon_1 = ln\ \varepsilon\ .$$

Параметры данной функции можно найти благодаря МНК.[1]

$$\sum y_1 = n{*}\ a_1 + b\ {*}\sum x$$

$$\sum y_1 * x1 = a_1 \sum x_1 + b * \sum x_1{}^2\ ,$$

или

$$\sum ln\ y = n{*}\ ln\ a + b{*}\sum ln\ x$$
$$\sum ln\ y * ln\ x = ln\ a \sum ln\ x + b{*}\sum (ln\ x)^2$$

Параметр $\boldsymbol{b}$ определяется из системы, а параметр $\boldsymbol{a}$ – косвенным путём:

$$\boldsymbol{a} = e^{a}{}_1$$

[1] См., напр.: Иванов А.Н., Подружкина Т.А. Эконометрика: Курс лекций для курсантов, студентов и слушателей, обучающихся по направлению подготовки 38.03.01 «Экономика» и специальности 38.05.01 «Экономическая безопасность» / Под. ред. Э.Н. Чижикова. СПб.: Санкт-Петербургский университет ГПС МЧС России, 2016.С. 39.

Приближённый метод для оценок коэффициентов некоторых моделей нелинейной регрессии

А.В. Яманова, студентка учебной группы 1832эб;
М.В. Зубец, студентка учебной группы 1831эб

Суть приближённого метода состоит в исследовании внутренне нелинейной регрессионной модели ***Y***, которая зависит от признака ***X*** вида

$$\hat{Y} = \beta_0 X^{\beta_1} + \beta_2$$

Данная модель не может быть сведена к линейной, а также для такой модели метод наименьших квадратов не может быть применим.

Приближённый метод нахождения параметров $b_0, b_1, b_2 -$ оценок коэффициентов $\beta_0, \beta_1, \beta_2$ состоит из следующих этапов:

1. Задаём параметрам $b_0,\ b_1$ и b_2 произвольные значения
2. Находим для данных параметров теоретические значения из уравнения регрессии в точках

$$x_i, i = 1,2, \dots, n:$$
$$\hat{y}_i = b_0(x_i)^{b_1} + b_2$$

3. Найдём сумму квадратов остатков

$$S_1 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} e_i{}^2$$

4. Заменяем значение любого из параметров b_0, b_1 и b_2 уменьшив его
5. Найдём для нового значения параметра теоретические значения из уравнения регрессии и вторую сумму квадратов остатков S_2
6. Сравним значение S_1 и S_2 и преобразовываем параметр.

Сравниваем полученную сумму остатков по следующему принципу:

- Если $S_1 > S_2$, то можно уменьшить значение параметра и затем полученную сумму сравнить с предыдущей суммой S_2
- Если $S_1 < S_2$, то параметр нужно не уменьшать, а увеличивать и полученную сумму остатков сравнить с S_1

7. Выбираем иной параметр и заново применяем условия пунктов 4-6 до тех пор, пока станет невозможно внести изменения, которые уменьшат S.
8. Делаем вывод о минимизации S. Последние оценки принимаются за оценки параметров.

Индексы корреляции и детерминации для парной нелинейной регрессии, их смысловое значение, оценки и значимость

А.В. Яманова, студентка учебной группы 1832эб

Индекс корреляции – это показатель тесноты связи, который не является коэффициентом для нелинейной зависимости.

Индекс корреляции определяется по формуле:

$$\mathrm{RI} = \sqrt{1 - \frac{\sum_{i=1}^{n}\left(y_i - \widehat{y_i}\right)^2}{\sum_{i=1}^{n}\left(y_i - \bar{y}\right)^2}}$$

Индекс корреляции лежит в интервале $0 \leq RI \leq 1$. Стоит отметить, что чем ближе к единице значение индекса, тем теснее связь признаков, а также это говорит о надёжности уравнения регрессии.

Оценка статистической значимости индекса корреляции проводится так же, как и оценка значимости коэффициента корреляции.

Индекс детерминации – это квадрат индекса корреляции нелинейных связей. Для того, чтобы найти индекс детерминации необходимо возвести индекс корреляции в квадрат:

$$\mathrm{RI}^2 = 1 - \frac{\sum_{i=1}^{n}\left(y_i - \widehat{y_i}\right)^2}{\sum_{i=1}^{n}\left(y_i - \bar{y}\right)^2}$$

Индекс детерминации RI^2 можно сравнивать с коэффициентом детерминации r^2 для обоснования возможности применения линейной функции.

На основе коэффициента определяется индекс детерминации для подсчёта производных бета и альфа в процентном соотношении, если процент ниже установленного минимума (может измеряться в пределах 75%), то установленные значения будут некорректными.

Спецификация модели. Модель линейной множественной регрессии

А.Г. Парк, студентка учебной группы 1831эб

Спецификация модели– это математическая форма записи уравнения зависимости результирующей переменной от одного или же нескольких факторов. Эконометрические модели делятся на несколько видов в зависимости от спецификации.

Регрессия является зависимостью среднего значения величины $\boldsymbol{y}$ от другой величины $\boldsymbol{x}$ или же нескольких величин $\boldsymbol{x_i}$.

Количество факторов, включённых в регрессионное уравнение, определяет тип регрессии, которая может быть простой и множественной. Последнюю рассмотрим подробнее.

Множественная регрессия представляет собой модель, которая имеет среднее значение объясняемой переменной $\boldsymbol{y}$ – это функция нескольких объясняющих переменных $\boldsymbol{x_1, x_2, \ldots, x_n}$.

Множественная регрессия в неявном виде – это модель типа: $\boldsymbol{\hat{y}=f(x_1, x_2, \ldots, x_n)}$

В явном виде: $\boldsymbol{\hat{y}=a+b_1x_1+b_2x_2+\ldots+b_nx_n}$.

Примером модели множественной регрессии может служить зависимость заработной платы сотрудников от возраста, уровня знаний, квалификации, стажа работы, и т.д.

Основные модели множественной регрессии, следующие:

1) линейная;
2) степенная;
3) показательная;
4) параболическая;
5) гиперболическая.

Безусловно считается, что линейные модели – самые простые и часто используемые.

Множественная линейная регрессия – это прямая зависимость среднего значения $\boldsymbol{Y}$ от двух или более других значений $\boldsymbol{X_1, X_2, \ldots, X_n}$. Значение $\boldsymbol{Y}$ обычно называется зависимой или результирующей переменной, а значения $\boldsymbol{X_1, \ldots, X_n}$ называются независимыми переменными.

Требования к факторам, включаемым во множественную регрессию

А.Г. Парк, студентка учебной группы 1831эб

При рассмотрении экономических процессов часто необходимо обращаться к моделям, которые содержат не один фактор-признак, а несколько. Следовательно, в модель следует включать не один фактор, а строить уравнение множественной регрессии.

Уравнение множественной регрессии представляет собой функцию k переменных: $y=f(x_1, x_2, \ldots, x_k)$.

Существуют следующие требования, предъявляемые к факторам, для включения в модель:

1. Факторы должны поддаваться количественной оценке. При условии, что необходимо включить в модель качественный фактор, который не имеет количественного измерения, его необходимо определить количественно.

2. Каждому из факторов должно быть своё теоретическое обоснование.

3. Факторы не должны быть линейно зависимыми и, прежде всего, не должны находиться в точной функциональной зависимости.

4. Модель должна содержать только наиболее важные факторы, которые имеют существенное влияние на исследуемые показатели.

Подбор факторов обычно проводится в два этапа:

- факторы подбираются исходя из характера проблемы;
- на основе матрицы показателей корреляции определяется t-статистика для параметров регрессии.

Сложность и взаимопроникновение отдельных факторов, определяющих анализируемое экономическое явление, может проявляться в так называемой мультиколлинеарности.

Под мультиколлинеарностью понимается тесная взаимосвязь между факторными признаками, включёнными в модель.

Понятие мультиколлинеарности факторов, её выявление и устранение

Я.А. Коцеруба, студентка учебной группы 1832эб

Мультиколлинеарность – нарушение одного из основных условий, лежащих в основе построения линейной модели множественной регрессии.

Правила выявления мультиколлинеарных факторов:

1) если в корреляционной матрице факторных переменных присутствуют коэффициенты парной корреляции по абсолютной величине большие 0,8, то можно сделать вывод, что в данной модели множественной регрессии существует мультиколлинеарность;

2) вычислить собственные числа корреляционной матрицы факторных переменных λ_{min} и λ_{max}.

Если $\lambda_{min} < 10^{-5}$, то в модели регрессии присутствует мультиколлинеарность.

Если отношение

$$\frac{\lambda_{\min}}{\lambda_{\max}} < 10^{-5},$$

то оно свидетельствует о наличии мультиколлинеарных факторных переменных;

3) вычислить определитель корреляционной матрицы факторных переменных. Если его величина очень мала, то в модели регрессии присутствует мультиколлинеарность.

Основные способы устранения мультиколлинеарности:

1) один из наиболее простых способов устранения мультиколлинеарности состоит в получении дополнительных данных. Однако на практике в некоторых случаях реализация данного метода может быть весьма затруднительна;

2) способ преобразования переменных, например, вместо значений всех переменных, участвующих в модели (и результативной в том числе) можно взять их логарифмы:

$$\ln y = \beta_0 + \beta_1 \ln x_1 + \beta_2 \ln x_2 + \varepsilon.$$

Однако данный способ также не способен гарантировать полного устранения мультиколлинеарности факторов.

Если рассмотренные способы не помогли устранить мультиколлинеарность факторов, то переходят к использованию смещённых методов оценки неизвестных параметров модели регрессии, или методов исключения переменных из модели множественной регрессии.

Понятие и применение фиктивных переменных

П.В. Кучина, студентка учебной группы 1831эб

Термин «фиктивная переменная» обозначает искусственно введённую переменную, которая может принимать только два значения 0 и 1, и использоваться для учёта качественных признаков.

Это могут быть различные признаки, например: пол, образование, принадлежность к определённому региону. Чтобы представить эти переменные в регрессионную модель, им должны быть присвоены специальные цифровые метки, т. е. качественные переменные преобразованы в количественные. Такого вида спроектированные переменные в эконометрике называют фиктивными переменными.

Применение фиктивных переменных

Общее теоретическое регрессионное уравнение для значений признаков имеет вид:

$$\hat{y} = \beta_0 + \beta_1 x + \alpha z$$

Исследуем зависимость потребления кофе ***Y*** от цены ***X***.

Z – пол работника.

Будем считать, что значение качественного признака ***Z***=1, если работник мужского пола и ***Z***=0, если работник женского пола.

Получим регрессионное уравнение для женщин $\hat{y} = b_0 + b_1 x$ и для мужчин $\hat{y} = b_0 + b_1 x + \alpha = (b_0 + \alpha) + b_1 x$.

Таким образом, для данного случая сила влияния цены на потребление кофе (параметр b1) для мужчин и женщин одинакова, а свободные члены в моделях отличаются на величину $\boldsymbol{\alpha}$. Ясно, что при $\boldsymbol{\alpha>0}$ предпочтение отдают мужчинам, а иначе – женщинам.

Преимущества использования фиктивных переменных

1) интервалы между наблюдениями могут различаться;

2) коэффициенты при фиктивных переменных легко объясняемы, они наглядно представляют структуру динамического процесса.

Понятие и смысловое значение парных и частных коэффициентов множественной корреляции

Я.А. Коцеруба, студентка учебной группы 1832эб

Парный коэффициент корреляции характеризует тесноту линейной зависимости между двумя переменными на фоне действия всех остальных переменных, входящих в модель.

Парный коэффициент корреляции $\boldsymbol{\rho}$ в силу своих свойств является одним из самых распространённых способов измерения связи между случайными величинами в генеральной совокупности. Для выборочных данных используется эмпирическая мера связи.

Коэффициент корреляции не имеет размерности, следовательно, его можно сопоставлять для разных статистических рядов. Его величина находится в пределах от –1 до +1. Если коэффициент равен нулю, то линейная связь между $\boldsymbol{x}$ и $\boldsymbol{y}$ отсутствует. В данном случае не отрицается возможность существования иной формы зависимости между переменными. Положительный знак коэффициента корреляции указывает на положительную корреляцию, т.е. все данные наблюдения лежат на прямой с положительным углом наклона в плоскости $\boldsymbol{xy}$ и с увеличением $\boldsymbol{x}$ растёт $\boldsymbol{y}$. Когда $\boldsymbol{x}$ уменьшается, то $\boldsymbol{y}$ уменьшается. Отрицательный знак коэффициента свидетельствует об отрицательной корреляции.

Частный коэффициент корреляции характеризует тесноту линейной зависимости между двумя переменными при исключении влияния всех остальных переменных, входящих в модель.

Частный коэффициент корреляции обладает всеми свойствами парного коэффициента корреляции. Если парный коэффициент корреляции между двумя случайными величинами оказался больше соответствующего частного коэффициента, то можно сделать вывод о том, что фиксирование всех других переменных приводит к усилению взаимосвязи между изучаемыми величинами, т.е. более высокое значение парного коэффициента обусловлено присутствием «третьей величины». Более низкое значение парного коэффициента корреляции в сравнении с соответствующими частными свидетельствует об ослаблении связи между изучаемыми величинами действием фиксируемых величин.

Линейные коэффициенты множественной корреляции и детерминации, их оценки и смысловое значение

П.В. Кучина, студентка учебной группы 1831эб

Множественный коэффициент корреляции R — это положительный квадратный корень из *R*-квадрата. Эта статистика полезна при проведении многомерной регрессии (т.е. использовании нескольких независимых переменных), когда необходимо описать зависимость между переменными.

Множественный коэффициент корреляции характеризует тесноту связи между зависимой переменной и предиктором. Он изменяется в пределах от 0 до 1 и рассчитывается по формуле:

$$r_{y/(x)} = \sqrt{1 - \frac{|R|}{R_{11}}}$$

$|R|$ — определитель корреляционной матрицы;

$\boldsymbol{R_{11}}$ — алгебраическое дополнение элемента матрицы, стоящего в первой строке первого столбца.

Значимость множественного коэффициента корреляции проверяется по таблице F-критерия Фишера.

Коэффициент детерминации (***R²*** – ***R****-квадрат*) — это доля дисперсии зависимой переменной, объясняемая рассматриваемой моделью. Более точно – это единица минус доля необъяснённой дисперсии в дисперсии зависимой переменной. В случае линейной зависимости $\boldsymbol{R^2}$ является квадратом так называемого множественного коэффициента корреляции между зависимой переменной и объясняющими переменными. В частности, для модели линейной регрессии с одним признаком $\boldsymbol{x}$ коэффициент детерминации равен квадрату обычного коэффициента корреляции между $\boldsymbol{y}$ и $\boldsymbol{x}$. Истинный коэффициент детерминации модели зависимости случайной величины $\boldsymbol{y}$ от признаков $\boldsymbol{x}$ определяется следующим образом:

$$R^2 = 1 - \frac{V(y|x)}{V(y)} = 1 - \frac{\sigma^2}{\sigma_y^2},$$

где $V(y|x) = \sigma^2$ — условная (по признакам $\boldsymbol{x}$) дисперсия зависимой переменной (дисперсия случайной ошибки модели).

Проверка статистических гипотез для множественной регрессии и корреляции

А.К. Травина, студентка учебной группы 1831эб

Статистические методы проверки гипотез используют, чтобы проверить, значимы ли параметры, т. е. существенно ли они отличаются от нуля для общей совокупности.

Основной гипотезой выступает гипотеза о несущественном отличии от нуля параметра или статистической характеристики в общей совокупности. Если основная гипотеза неверна, то используется альтернативная гипотеза. В ней говорится о неравенстве нулю параметра или статистической характеристики в общей совокупности. Проверка значимости отдельных коэффициентов множественной регрессии осуществляется с помощью ***t***-критерия Стьюдента.

Найденное значение ***t***-критерия сравнивается с табличным значением, определяемым по таблицам распределения Стьюдента. Табличное значение определяется в зависимости от уровня значимости (***a***) и числа степеней свободы, равное (***n***-2), где n-число наблюдений.

Если фактическое значение ***t***-критерия выше табличного (по модулю), то основную гипотезу отвергают и считают, что с вероятностью (1-***a***) параметр или статистическая характеристика в общей совокупности существенно отличается от нуля.

Если фактическое значение ***t***-критерия по модулю ниже табличного, то используют основную гипотезу, т. е. параметр или статистическая характеристика в общей совокупности незначимо отличается от нуля при уровне значимости ***a***.

Модель временного ряда. Основные элементы и характеристики временного ряда

М.Р. Ахмадова, студентка учебной группы 1832эб

Модели, построенные по данным, характеризующим один объект за ряд определённых последовательных периодов, называются моделями временны́х рядов.

Временной ряд – это совокупность значений какого-либо показателя за несколько последовательных моментов или периодов времени. Каждый уровень временного ряда формируется под воздействием большого числа факторов, которые условно можно подразделить на три группы:

1) факторы, формирующие тенденцию ряда;
2) факторы, формирующие циклические колебания ряда;
3) случайные факторы.

Во многом фактический уровень временного ряда можно показать, как сумму или произведение трендовой, циклической и случайной составляющей.

Модель, в которой временной ряд представлен как сумма перечисленных элементов, называют аддитивной моделью временного ряда. Аддитивная модель имеет вид:

$$Y=T+S+E.$$

Модель, в которой временной ряд представлен как произведение перечисленных компонент, называется мультипликативной моделью временного ряда. Мультипликативная модель имеет вид:

$$Y=T*S*E.$$

Основная задача эконометрического исследования отдельного временного ряда – выявление и придание количественного выражения каждой из перечисленных выше компонент с тем, чтобы использовать полученную информацию для прогнозирования будущих значений ряда или при построении моделей взаимосвязи двух или более временных рядов.

Временные ряды состоят из двух элементов:

- периода времени, за который или по состоянию на который приводятся числовые значения;
- числовых значений того или иного показателя, называемых уровнями ряда.

Систематические элементы временного ряда

Д.А. Смольская, студентка учебной группы 1831эб

Временной ряд — это ряд наблюдаемых значений исследуемого показателя, которые расположены в хронологическом порядке или в порядке возрастания периода. Выделяют 3 главных систематических элемента временного ряда: тренд; сезонность; цикличность. При разных сочетаниях в исследуемом явлении, либо процессе этих явлений зависимость уровней ряда от времени может принимать разнообразные формы. Во-первых, большая часть временных рядов финансовых показателей обладают тенденцией, описывающей совокупное длительное влияние большого количества факторов на динамику исследуемого показателя. Во-вторых, исследуемый показатель способен подвергается циклическим колебаниям. Данные колебания имеют все шансы носить сезонный характер, так как финансовая деятельность ряда отраслей экономики находится в зависимости от времени года. Некоторые временны́е ряды никак не содержат тенденции и циклической компоненты, а любой последующий их уровень образовывается равно как совокупность среднего уровня ряда и определённой (позитивной либо негативной) неожиданной компоненты.

Автокорреляция уровней временного ряда и выявление его структуры

М.Р. Ахмадова, студентка учебной группы 1832эб

Корреляционную зависимость между последовательными уровнями временного ряда называют автокорреляцией уровней этого ряда.

Коэффициент корреляции имеет вид:

$$r_{xy}=\frac{\Sigma(x_j-\bar{x})\cdot(y_j-\bar{y})}{\sqrt{\Sigma(x_j-\bar{x})^2\cdot\Sigma(y_j-\bar{y})^2}}$$

В качестве переменой x рассмотрим ряд $y_2,y_3,\dots y_g$
в качестве переменной y – ряд $y_1,y_2,\dots y_g$.

Тогда коэффициент автокорреляции первого порядка:

$$r_1=\frac{\sum_{t=2}^{n}(y_t-\bar{y}_1)(y_{t-1}-\bar{y}_2)}{\sqrt{\sum_{t=2}^{n}(y_t-\bar{y}_1)^2\sum_{t=2}^{n}(y_{t-1}-\bar{y}_2)^2}},$$

где $\bar{y}_1=\frac{1}{n-1}\sum_{t=2}^{n}y_t,\qquad \bar{y}_2=\frac{1}{n-1}\sum_{t=2}^{n}y_{t-1}.$

Коэффициент автокорреляции первого порядка измеряет зависимость между соседними уровнями ряда t и t-*1*, т.е. при лаге 1. Так, коэффициент автокорреляции второго порядка характеризует тесноту связи между уровнями y_t и y_{t-2}, и определяется по формуле:

$$r_2=\frac{\sum_{t=3}^{n}(y_t-\bar{y}_3)(y_{t-2}-\bar{y}_4)}{\sqrt{\sum_{t=3}^{n}(y_t-\bar{y}_3)^2\sum_{t=3}^{n}(y_{t-2}-\bar{y}_4)^2}},$$

где $\bar{y}_3=\frac{1}{n-2}\sum_{t=3}^{n}y_t,\qquad \bar{y}_4=\frac{1}{n-2}\sum_{t=3}^{n}y_{t-2}.$

Число периодов, по которым рассчитывается коэффициент автокорреляции, называют лагом. С увеличением лага число пар значений, по которым рассчитывается коэффициент автокорреляции, уменьшается. Отметим два свойства коэффициента автокорреляции.

Во-первых, он строится по аналогии с линейным коэффициентом корреляции и характеризует тесноту только *линейной* связи текущего и предыдущего уровней ряда.

Во-вторых, по знаку коэффициента автокорреляции нельзя делать вывод о возрастающей или убывающей тенденции в уровнях ряда.

Выявление тенденции временного ряда

А.А. Бизенкова, студентка учебной группы 1832эб

Основной тенденцией, то есть трендом называют довольно мягкое и устойчивое изменение уровня явления во времени, свободное от случайных колебаний.

Наиболее известными способами выявления основных тенденций ряда динамики выделяют такие методы, как:

1. Укрупнение интервалов и их характеристики средними уровнями.

Данный метод заключается в переходе от маленьких промежутков времени к более продолжительным, например: от суток – к неделям; от недель к месяцам.

2. Скользящей средней.

Суть метода заключается в замене абсолютных данных средними арифметическими за определённые периоды. Расчёт средних показателей ведётся способом скольжения, т.е. постепенным исключением из принятого периода первого уровня и включением следующего.

3. Аналитического выравнивания.

При данном методе уровни ряда динамики описываются в виде временны́х функций.

Автокорреляция остатков и её обнаружение

А.Г. Парк, студентка учебной группы 1831эб

Автокорреляция – это статистическая взаимосвязь между последовательностями значений одного ряда, взятыми с запаздыванием, например, для случайного процесса – с запаздыванием во времени. Этот термин широко используется в эконометрике.

Автокорреляция остатков обычно обнаруживается при регрессионном анализе временны́х рядов, и почти никогда при анализе пространственных выборок.

Автокорреляция остатков наблюдается в тот момент, когда значения предыдущих остатков выше (положительно) или ниже (отрицательно) значения последующих. При положительной автокорреляции остатки изменяются монотонно во время наблюдения, а при отрицательной автокорреляции часто меняется знак остатков.

Основными причинами автокорреляции являются:

• неправильный выбор формы регрессионной зависимости;

• пренебрежение одним или несколькими важными факторами в модели;

• циклические значения экономических переменных при построении модели на основе временных данных.

Для обнаружения автокорреляции используются статистические тесты и построение графиков зависимостей остатков от времени, и визуальное определение наличия автокорреляции остатков. Самым распространенным является тест Да́рбина – Уо́тсона. Этот тест обнаруживает автокорреляцию первого порядка, когда автокорреляция соответствует уравнению:

$$d = \frac{\sum_{t=2}^{n}(\varepsilon_t - \varepsilon_{t-1})^2}{\sum_{t=1}^{n}\varepsilon_t^2}$$

Критерий Да́рбина – Уо́тсона определяется как отношение суммы квадратов разностей последовательных значений остатков к сумме квадратов остатков. Практически во всех задачах по эконометрике значение критерия Дарбина – Уотсона указывается наряду с коэффициентом корреляции, значениями критериев Фи́шера и Стью́дента.

Автокорреляция остатков. Критерии обнаружения автокорреляции

А.А. Бизенкова, студентка учебной группы 1832эб

Автокорреляция остатков – это зависимость между последовательными значениями предыдущих, текущих и последующих данных.

Автокорреляция в остатках имеет свои причины:

1. Она может относиться к исходным данным, если зависит от ошибок измерения в результате.

2. Выводом неправильной спецификации модели.

Данное явление имеет место, когда нарушается предположение о независимости случайных величин, т.е. когда

$$\boldsymbol{cov(u_j, u_i) \neq 0, j \neq 1}$$

В простом виде автокорреляция остатков может быть представлена уравнением:

$$\boldsymbol{u_t = \rho u_{t-1} + \varepsilon_t},$$

где $\rho < 1$ − автокорреляционный коэффициент первого порядка, t – период времени, ε_t – ошибка.

Критерии обнаружения автокорреляции:

1. Графический способ, который может быть представлен в виде:

· анализа графиков зависимости Y от t;

· либо зависимость графиков остатков от времени;

· или обусловленность остатков в текущий момент времени от предыдущего.

2. Использование тест серий Вальда-Вольфовица.

3. Применяя статистику Дарбина-Уотсона.

Методы сглаживания

А.А. Бизенкова, студентка учебной группы 1832эб

Сглаживание временны́х рядов – это выделение основной тенденции $f(t)$ из совокупности динамического ряда, который содержит случайную составляющую $\varepsilon(t)$

Цель сглаживания – это построение модели прогнозирования для последующих периодов, исходя из прошлых наблюдений для выделения неслучайных компонентов.

Самый простой метод сглаживания рядов – это скользящее среднее. Содержание метода заключается в том, что для любого нечётного количества точек последовательности ряда осуществляется замена основной точки на среднее арифметическое остальных точек:

$$\boldsymbol{s_i = \frac{1}{2k+1} \sum_{j=-k}^{k} x_i + j}$$

где x_i – исходный ряд, s_i – сглаженный ряд.

Метод скользящего среднего имеет свои недостатки:

- Скользящее среднее малоэффективно в вычислениях, потому что для каждой точки среднее необходимо вычислять заново.
- Скользящее среднее нельзя продлить на первые и последние точки ряда.
- Скользящее среднее не определено за пределами ряда, и в результате не может использоваться для прогнозирования.

Кроме скользящего среднего используется также экспоненциальное сглаживание, или метод Хольта-Уинтерса.

Экспоненциальное сглаживание заключается в том, что берётся среднее взвешенное значение по всем прошлым результатам, но бо́льшее внимание уделяется все-таки последним наблюдениям.

Задачи прогнозирования с помощью временных рядов

А.А. Бизенкова, студентка учебной группы 1832эб

Временно́й ряд — это множество наблюдений, получаемых последовательно во времени.

Прогнозирование временных рядов заключается в построении модели для прогноза событий в будущем до того, как они будут определены, основываясь на событиях прошлого.

Прогнозирование строится на предположении, что закономерность развития, действующая в прошлом внутри ряда динамики, сохранится и в прогнозируемом будущем.

Главной задачей прогнозирования временных рядов является определение факторов, которые оказывают существенное влияние на каждое значение временно́го ряда. Для этого используется каждая составляющая временно́го ряда, вычисляется её ценность в общую составляющую, а после, на основе этого вычисления, прогнозируются будущие значения временно́го ряда.

Для решения данного вопроса, используется метод декомпозиции временно́го ряда. Согласно этому методу, первоначальный временно́й ряд представляет собой композицию тренда с постоянно изменяющимися элементами. Также для построения прогноза применяется анализ ряда по частям.

Актуальность эконометрики для деятельности специалистов по экономической безопасности

Т.Н. Антошина, доцент кафедры прикладной математики и информационных технологий, кандидат педагогических наук

Развитие и совершенствование эконометрики является в настоящее время актуальной задачей современной России.

Описание эконометрических решений охватывает период с конца девятнадцатого века до середины двадцатого века, показывая, как экономисты впервые научились использовать статистические методы для измерения и проверки «законов» экономики.

В сборнике даны основные понятия эконометрики, раскрыты цели и задачи предмета эконометрики; представлены основные методы по разработке множественного регрессионного и корреляционного анализа, при помощи которого можно доказать различные экономические теории на основе анализа данных временны́х рядов; применения статистики к экономическим данным, начиная с обсуждения законов погрешности путём разработки метода наименьших квадратов.

Несмотря на научный характер, данный сборник изложен очень доступным языком; он не требует высокого уровня предварительных статистических знаний, и будет интересен как практикующим экономистам, так и обучающимся, которые только начинают свой путь в освоении статистики и экономики.

Содержание

Составление, вступительная статья
и компьютерная вёрстка:
Кабанов Андрей Александрович,
кандидат юридических наук, доцент,
e-mail: ***akabanov@inbox.ru***

сайт: ***otvet-akab.ru***

Авторский коллектив:
Антошина Татьяна Николаевна, Асриева Кристина Дмитриевна, Ахмадова Марха Руслановна, Бизенкова Алиса Александровна, Грищенко Ева Евгеньевна, Датова Карина Алексеевна, Дойкóва Дина Леонидовна, Зубец Маргарита Владимировна, Кабáнов Андрей Александрович, Коцеруба Янина Александровна, Кучина Полина Валентиновна, Матвеев Александр Владимирович, Остапчук Валерия Алексеевна, Палилова Диана Константиновна, Парк Ангелина Германовна, Скоробогатько Ксения Евгеньевна, Смольская Диана Андреевна, Травина Анна Константиновна, Яманова Анастасия Валерьевна.

ЭКОНОМЕТРИКА
Сборник статей

Редакционная коллегия:
Т.Н. Антошина, А.А. Кабанов, А.В. Матвеев.
Компьютерная верстка: А.А. Кабанов
Печатается в авторской редакции

Подписано в печать и свет 19.01.2022. Формат 60×84 1/16
Печать офсетная Объём 2,7 п. л. Тираж 66 экз.

Отпечатано в Полиграфическом центре ООО «НПО ПБ АС»
ФГБОУ ВО СПБУ ГПС МЧС России
196105, Санкт-Петербург, Московский пр., д. 149.