Э К О Н О М Е Т Р И К А

САРАТОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИМ. Н.Г. ЧЕРНЫШЕВСКОГО

Экономический факультет

Кафедра менеджмента и маркетинга

Леванова Л.Н.

Учебно-методическое пособие по курсу

ЭКОНОМЕТРИКА

Саратов 2007

УДК 330.43(072.8)
ББК 65в6я73

Л-34

Л-34 Леванова Л.Н.
Эконометрика. Учебно-методическое пособие. Саратов: ООО Издательский центр «Наука», 2007. с. п.л.

ISBN

Учебно-методическое пособие подготовлено в соответствии с положениями и требованиями Государственного образовательного стандарта высшего профессионального образования, включает основные вопросы лекций, терминологический словарь и основные теоретические положения эконометрики, планы семинарских занятий, методические указания по выполнению всех практических и самостоятельных работ, текущие семестровые задания, тестовые задания для оценки остаточных знаний.

Для студентов и преподавателей экономических специальностей.

Составитель

Кандидат экономических наук, доцент Л.Н. Леванова

Рекомендуют к печати:

Кафедра менеджмента и маркетинга Саратовского государственного университета им. Н.Г. Чернышевского

Кандидат экономических наук, профессор Л.И. Дорофеева

ISBN

УДК 330.43(072.8)
ББК 65в6я73

СОДЕРЖАНИЕ

Тема 1. Введение в эконометрику.

1. Предмет и задачи эконометрики.
2. Модель и ее свойства. Сущность эконометрического моделирования.
3. Переменные в моделях и их типы.
4. Экономические показатели как случайные величины.
5. Оценки и их свойства.

Ключевые слова

Эконометрика. Модель. Моделирование. Адекватность модели. Логические, геометрические и математические модели. Экономические и эконометрические модели. Модели микроэкономики, мезоэкономики и макроэкономики. Статические и динамические модели. Этапы эконометрического моделирования. Экзогенные переменные. Эндогенные переменные. Предопределенные переменные. Лаговые эндогенные переменные. Случайная величина. Испытание. Событие. Пространство элементарных событий. Генеральная совокупность. Выборка. Вероятность случайной величины. Дискретная и непрерывная случайные величины. Математическое ожидание. Дисперсия. Теоретическое стандартное отклонение. Функция плотности вероятности. Функция распределения случайной величины. Равномерное распределение. Теорема Ляпунова. Нормальное распределение. Стандартное нормальное распределение. Степень свободы. Распределение Стьюдента. Несмещенность. Эффективность. Состоятельность.

Основные теоретические аспекты темы:

Эконометрика – «экономика» + «метрика». Это наука об измерении и анализе экономических явлений, о количественных выражениях тех связей и соотношений, которые раскрыты и обоснованы экономической теорией. Это сплав четырех компонент: экономической теории, статистических и математических методов, компьютерных вычислений.

Модель – объект любой природы, который создается исследователем с целью получения новых знаний об объекте-оригинале и отражает только существенные (с точки зрения разработчика) свойства оригинала.

Моделирование - процесс построения, изучения и применения моделей.

Модель адекватна объекту-оригиналу – если она с достаточной степенью точности приближения отражает закономерности процесса функционирования реального объекта.

Математическая модель – это абстракция реального мира, в которой интересующие исследователя отношения между реальными элементами заменены подходящими отношениями между математическими категориями. Эти отношения, как правило, представлены в форме уравнений или неравенств.

Вероятностная математическая модель – модель, имитирующая механизм функционирования гипотетического (не конкретного) явления или системы стохастической природы.

Вероятностно - статистическая математическая модель – модель, в которой значения отдельных характеристик (параметров) оцениваются по результатам наблюдений (исходным статистическим данным), то есть это модель, характеризующая механизм функционирования конкретного явления или системы стохастической природы

Экономическая модель - вероятностная математическая модель, описывающая механизм функционирования экономического явления, экономической или социально –экономической системы.

Эконометрическая модель – вероятностно – статистическая модель, описывающая механизм функционирования экономического явления, экономической или социально –экономической системы.

Экзогенные переменные в модели – переменные, задаваемые «извне», автономно от модели, управляемые и планируемые.

Эндогенные переменные в модели – переменные, значения которых формируются в процессе и внутри функционирования анализируемой социально – экономической системы в существенной мере под воздействием экзогенных переменных и во взаимодействии друг с другом. В эконометрическом моделировании являются предметом объяснения.

Предопределенные переменные модели – все экзогенные переменные модели и лаговые эндогенные переменные.

Случайная величина – действительная переменная, которая в зависимости от исхода опыта, т.е. в зависимости от случая принимает различные значения, (значение не может быть точно предсказано).
$X = x_1$; $X = x_2$.

Испытание – реализация определенного комплекса условий, который может воспроизводиться неограниченное число раз.

Событие – результат испытания: достоверное (всегда происходит), невозможное (никогда не происходит), случайное (может произойти или не произойти).

Пространство элементарных событий – совокупность всех возможных, конкретных исходов.

Генеральная совокупность – совокупность всех возможных значений наблюдений интересующего показателя.

Выборка – множество наблюдений, составляющих часть генеральной совокупности.

Под вероятностью некоторого события (события, состоящего в том, что случайная переменная приняла определенное значение) понимается доля числа исходов, благоприятствующих данному событию, в общем числе возможных исходов.

$P\ X = x_1$

Дискретная случайная величина – случайная величина, принимающая в результате испытания значение из конечного либо счетного множества возможных чисел.

Математическое ожидание дискретной случайной величины – это взвешенное среднее всех ее возможных значений, где в качестве весового коэффициента берется вероятность соответствующего исхода.

E(x) = $\mu = \sum_{i=1}^{n} x_i * P_i$.

Теоретическая дисперсия дискретной случайной величины – мера разброса случайной величины вокруг среднего значения, определяется как математическое ожидание квадрата разности между величиной x и ее средним.

$$\sigma_x^2 = \sum_{i=1}^{n} (x-\mu)^2 * P_i = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n}.$$

Теоретическое стандартное отклонение – мера разброса случайной величины вокруг среднего значения, имеющая размерность данной случайной величины. Это среднее квадратическое разброса случайной величины.

m.c.o σ_x

Коэффициент вариации случайной величины – мера относительного разброса случайной величины, показывает какую долю среднего значения случайной величины составляет ее средний разброс.

$$V_x = \frac{\sigma_x}{\mu}.$$

Непрерывная случайная величина – случайная величина, принимающая значение из непрерывного диапазона значений (их нельзя пересчитать, ставя в соответствие им натуральные числа).

Функция плотности вероятности h=f(x) - функция значений случайной переменной, показывает вероятность нахождения случайной переменной внутри единичного интервала вокруг данной точки (вероятность попадания случайной величины x в бесконечно малый интервал).

Функция распределения случайной величины $F_x(x)$- вероятность того, что случайная величина принадлежит интервалу [a,b].

$$F_x(x) = P\{X \in [a,b]\}$$

Математическое ожидание непрерывной случайной величины -

$$\mu = \int_a^b x * f(x)dx.$$

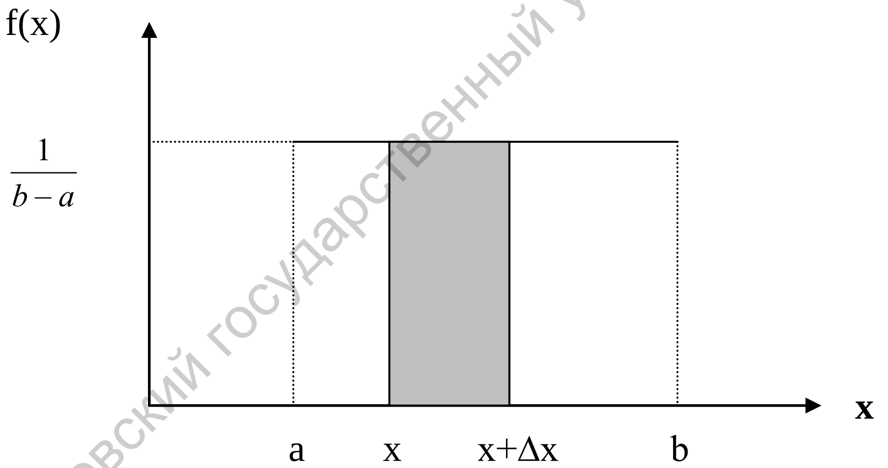
Теоретическая дисперсия непрерывной случайной величины -

$$\sigma_x^2 = \int_a^b (x-\mu)^2 * f(x)dx.$$

Равномерное распределение – это такое распределение вероятности, плотность которого постоянна в заданном интервале изменения случайной величины X. $a \le X \le b$.

Плотность вероятности на [a,b]: $f(x) = \begin{cases} \dfrac{1}{b-a}, & a \le x \le b \\ 0 & x < a, x > b \end{cases}$

Функция распределения: $F(x) = \begin{cases} 0, & x \le a \\ \dfrac{x-a}{b-a}, & a < x \le b \\ 1, & x > b \end{cases}$



Вероятность попадания равномерно распределенной случайной величины в заданный интервал равна произведению плотности вероятности на длину интервала.

$$\Delta P = \frac{\Delta x}{b-a}.$$

Теорема Ляпунова (центральная предельная теорема).

Распределение суммы n произвольно распределенных и взаимно независимых случайных величин при n→∞ стремится к нормальному распределению, если вклад отдельных слагаемых в сумму равномерно мал.

То есть если случайная величина является общим результатом взаимодействия большого числа других случайных величин, ни одна из которых не является доминирующей, то она будет иметь приблизительно нормальное распределение, даже если отдельные составляющие не имеют нормального распределения.

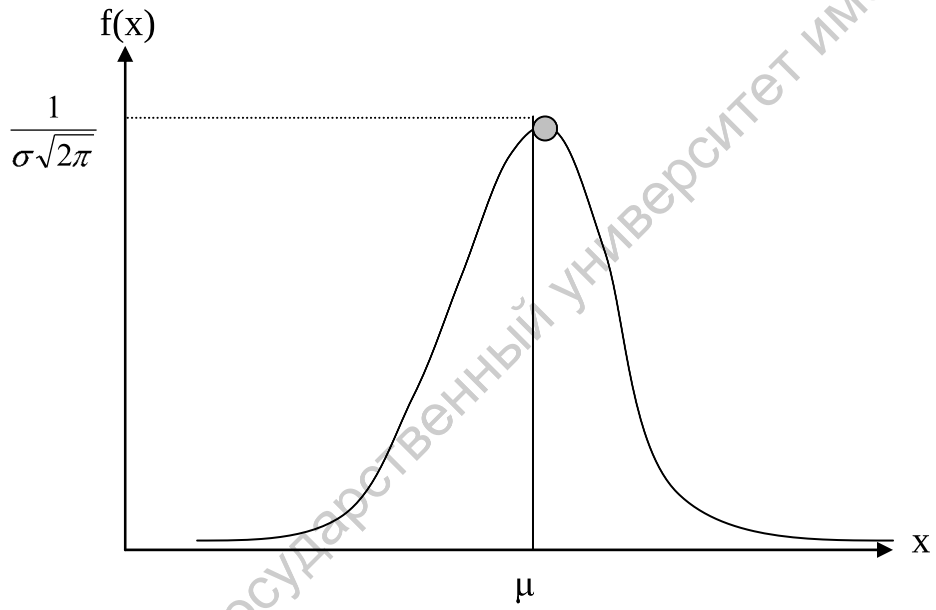
Нормальное распределение случайной величины X характеризуется лишь двумя параметрами: средним значением μ и стандартным отклонением σ.
X=N(μ, σ).

Плотность вероятности нормального распределения на [-∞,+∞]:

$$f_N(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$$

Функция распределения:

$$F_N(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{\frac{-(t-\mu)^2}{2\sigma^2}} dt$$



Стандартное нормальное распределение (z – распределение)– это нормальное распределение случайной величины $z = \frac{x-\mu}{\sigma}$ с параметрами μ =0 и σ=1. Ее распределение может быть протабулировано. (Таблица А1)
Z=N(0, 1).

Распределение Стьюдента (t -статистика) – выборочный аналог нормального распределения, переходящее в него при бесконечно большом числе наблюдений, это распределение случайной величины $t = \frac{x - \bar{x}}{s}$; где s – выборочное стандартное отклонение, определяемое по данным наблюдений и меняющееся от выборки к выборке.

На практике обычно используют не таблицы функции распределения Стьюдента, а таблицы критических точек, то есть точек с заданной вероятностью попадания в начинающиеся от них хвосты распределения. (Таблица А2)

Степени свободы в наборе данных – определяют число единиц данных, независимых друг от друга, которые могут нести отдельные элементы информации.

Методические указания по работе с таблицами t- распределения Стьюдента:

В таблице функции распределения Стьюдента приводятся для различных чисел степеней свободы ν критические точки, соответствующие приведенным в верхней строке таблицы вероятностям α попадания в правый «хвост» распределения.

α – это вероятность превышения t – статистикой приведенного в таблице критического значения при соответствующем числе степеней свободы ν.

Характеристики генеральной совокупности	Формулы оценивания
Математическое ожидание $\mu = \sum_{i=1}^{n} x_i * P_i$	*Выборочное среднее* $\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$
Теоретическая дисперсия $\sigma_x^2 = \sum_{i=1}^{n} (x-\mu)^2 * P_i$	*Выборочная дисперсия* $s^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}$

Оценка называется несмещенной если математическое ожидание оценки равнялось соответствующей характеристике генеральной совокупности. Если это не так, то оценка называется смещенной, и разница между ее математическим ожиданием и соответствующей характеристикой генеральной совокупности называется **смещением**.

Эффективная оценка – эта та оценка, у которой дисперсия минимальна.

Состоятельная оценка - если предел оценки по вероятности равен истинному значению характеристики генеральной совокупности, то есть оценка, которая дает точное значение для большой выборки независимо от входящих в нее конкретных наблюдений. Оценка называется **несостоятельной**, если распределение вероятностей при увеличении размера выборки либо вообще не концентрируется, либо концентрируется около точки, отличной от истинного значения.

Вопросы для обсуждения:

1. Назовите плюсы и минусы моделирования как инструмента исследования экономических процессов и явлений.
2. Может ли выходная переменная модели быть одновременно и входной переменной? Если да, то в каких случаях?
3. Как Вы считаете, если результаты (эндогенные переменные, выходные параметры) модели явно неверные, в чем может быть причина неудачного моделирования?
4. Как Вы считаете, каковы минусы агрегирования при макроэкономическом моделировании?
5. В чем заключается специфичность определения точности измерений социально-экономических явлений?
6. Как Вы считаете, почему изучаемая дисциплина появилась в российском образовании сравнительно недавно?
7. Почему экономические показатели, рассчитанные на данных современных экономик, носят случайный характер?
8. Приведите примеры экономических показателей, к которым можно применить теорему Ляпунова.
9. Приведите примеры экономических показателей, которые можно рассматривать как дискретные и непрерывные случайные величины.
10. Объясните влияние количества наблюдений и σ на график нормального распредления.
11. Почему необходимо рассчитывать z – статистику?

Задания для самостоятельной работы:

Задание 1.

В таблице приведены данные чистого дохода как процента от стоимости акционерного капитала для 42 - х компаний. Рассчитайте выборочные среднюю и дисперсию для приведенных данных компаний**.**

17	14	15	14	11	12	9	18	14	7	17	14	15	20	12	14	9	1	18	27	11
11	23	36	25	10	18	14	23	13	2	6	15	14	10	7	13	8	11	16	44	1

Задание 2.

Рассчитайте вероятность попадания величины z в конечный интервал [0,32; 2,27].

Задание 3.

Найдите интервал, в который попадает случайная величина z с вероятностью 0,5557.

Тема 2. Модель парной регрессии.

1. Спецификация модели парной регрессии: понятие и способы задания функций.
2. Параметризация модели: оценка параметров уравнения линейной регрессии. Метод наименьших квадратов.
3. Интерпретация уравнения парной регрессии: экономический смысл параметров регрессии. Применение модели парной регрессии в микро и макроэкономике.
4. Эксперимент Монте – Карло. Свойства коэффициентов регрессии.
5. Оценка значимости коэффициентов линейной регрессии: проверка гипотез, относящихся к коэффициентам регрессии.
6. Качество оценки: коэффициент детерминации. F – критерий Фишера для проверки качества оценивания.
7. Прогнозирование на основе линейного уравнения регрессии. Интервальный прогноз.

Ключевые слова:

Спецификация модели. Результативный признак, признак-фактор и стохастическая переменная в модели. Графический, аналитический и экспериментальный способы задания функции. Коэффициент вариации случайной величины. Коэффициент корреляции. Метод наименьших квадратов**.** Стандартная ошибка коэффициентов регрессии. t критическое. Доверительный интервал. Коэффициент детерминации. Общая сумма квадратов отклонений. Факторная сумма квадратов отклонений. Остаточная сумма квадратов отклонений. Дисперсии на одну степень свободы. F- критерий Фишера. Средняя ошибка аппроксимации Прогнозное значение. Интервалы прогноза.

Основные теоретические аспекты темы.

Парная регрессия – регрессия между двумя переменными y и x, то есть модель вида: $y=f(x)+\varepsilon$, где:

y - зависимая переменная, фактическое значение результативного признака,

x – независимая, объясняющая, переменная, признак- фактор,

ε - возмущение, случайная, стохастическая переменная.

Причины существования ε:

1. Невключение объясняющих переменных.
2. Агрегирование переменных.
3. Неправильное описание структуры модели.
4. Неправильная функциональная спецификация.
5. Ошибки измерения

Основные типы функций, используемые при количественной оценке связей:

Линейная функция: $y=a+bx$;

Нелинейные функции: y= a+b/x - гипербола;

$y=a+bx+cx^2$ – парабола;

$y=a+bx+cx^2+dx^3$-кубический многочлен;

$y=ax^b$ –степенная функция;

$y=ab^x$-показательная функция;

y=a+blgx-логарифмическая функция;

y= 1/(a+bx);

y=a+bx+c(1/x);

$y=1/(a+bx+cx^2)$;

$\lg y=a+bx+cx^2$

$y=a/(1+be^{-cx})$.

Расчет параметров уравнения линейной регрессии y=a+bx+ε:

Дано: $y=\alpha+\beta x+u$ - уравнение парной регрессии для генеральной совокупности, где:

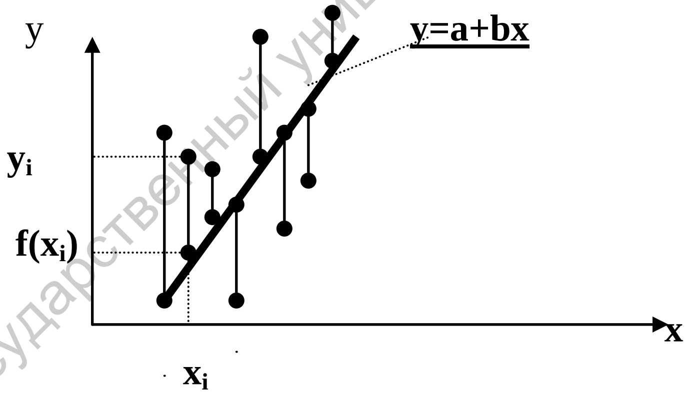
α, β – истинные коэффициенты модели;

u– стохастическая переменная модели.

Необходимо по выборке, состоящей из n статистических значений независимой, объясняющей переменной x: $x_1, \ldots x_n$ и n статистических значений зависимой переменной y: $y_1, \ldots y_n$ построить уравнение парной регрессии y=a+bx + ε, где**:**

a,b - оценки α, β соответствующие генеральной совокупности,

ε - оценка u.



Коэффициент корреляции величин x и y (r_{xy})– свидетельствует о наличии или отсутствии линейной связи между переменными:

$$\mathbf{r_{xy}}=\frac{cov(x,y)}{s_x s_y}=\frac{\overline{yx}-\overline{y}*\overline{x}}{s_x\ s_y}. \qquad r_{xy}\in[-1;1].$$

Если: r_{xy} = -1, то наблюдается строгая отрицательная связь;

r_{xy} = 1, то наблюдается строгая положительная связь;

r_{xy} = 0, то линейная связь отсутствует.

Стандартное отклонение случайной величины x (выборочная дисперсия s_x) - мера разброса случайной величины вокруг среднего значения выборки.

$$s_x=\sqrt{\overline{x^2}-\overline{x}^2}.$$

Коэффициент вариации случайной величины x (V_x.) – мера относительного разброса случайной величины. Показывает, какую долю среднего значения случайной величины составляет ее средний разброс.

$$\mathbf{V_x} = \frac{s_x}{\bar{x}}.$$

Метод наименьших квадратов – метод оценивания параметров линейной регрессии, минимизирующий сумму квадратов отклонений наблюдений зависимой переменной от искомой линейной функции.

$$RSS = \sum_{i=1}^{n} e_i^2 = \sum_i (y_i - f(x_i))^2 = \sum_{i=1}^{n} (y_i - (a + bx_i))^2 \to \min$$

где: y_i – статистические значения зависимой переменной;
$y_x = f(x_i)$ - теоретические значения зависимой переменной, рассчитанные с помощью уравнения регрессии.

$$\begin{cases} \dfrac{dRSS}{da} = 0 \\ \dfrac{dRSS}{db} = 0 \end{cases};$$

Используя метод наименьших квадратов, параметры уравнения линейной регрессии можно вычислить следующим образом $b = \dfrac{\overline{yx} - \bar{y} * \bar{x}}{\overline{x^2} - \bar{x}^2}$; $a = \bar{y} - b * \bar{x}$. ;

где $\overline{yx}$ - среднее значение y*x; $\bar{y}$ - среднее значение y; $\bar{x}$ - среднее значение x.

Экономический смысл параметров уравнения линейной парной регрессии:

Параметр b показывает среднее изменение результата y с изменением фактора x на единицу.

Параметр a=y, когда x=0. Если x не может быть равен 0, то a не имеет экономического смысла. Интерпретировать можно только знак при a: если a>0, то относительное изменение результата происходит медленнее, чем изменение фактора, то есть вариация результата меньше вариации фактора: $V_y < V_x$, и наоборот.

Эксперимент Монте – Карло – это искусственный контролируемый эксперимент, дающий возможность проверки хорошие или плохие оценки дает метод наименьших квадратов.

Этап 1:

1. Выбираются истинные значения α и β;
2. В каждом наблюдении выбирается значение x;
3. Используется процесс генерирования случайных чисел для получения значений случайного фактора u

Этап 2:

В каждом наблюдении генерируется значение y с использованием соотношения $y = \alpha + \beta x + u$.

Этап 3:
Применяется регрессионный анализ для получения оценок a и b с использованием полученных значений y и x.

Вывод по эксперименту:

Свойства коэффициентов регрессии существенным образом зависят от свойств случайной переменной u. Чтобы регрессионный анализ, основанный на обычном методе наименьших квадратов, давал наилучшие из всех возможных результаты, случайная переменная должна удовлетворять четырем условиям, известным как условия Гаусса – Маркова. Если эти условия не выполнены, исследователь должен это осознавать. Если корректирующие действия возможны, то аналитик должен быть в состоянии их выполнить. Если ситуацию исследовать невозможно, исследователь должен быть способен судить, насколько серьезно это может повлиять на результаты.

1 – е условие Гаусса – Маркова: $E(u_i)=0$ для всех наблюдений.

Математическое ожидание случайной переменной в любом наблюдении должно быть равно нулю. Иногда случайная переменная будет положительной, иногда отрицательной, но он не должен иметь систематического смещения ни в каком из двух возможных направлений.

Фактически, если уравнение регрессии включает постоянный член, то обычно бывает разумно предположить, что это условие выполняется автоматически, так как роль постоянного члена состоит в отражении любой систематической, но постоянной составляющей в y, которую не учитывают объясняющие переменные.

2 – е условие Гаусса – Маркова: $\sigma^2(u_i) = \sigma^2_u$ для всех наблюдений.

Теоретическая дисперсия случайной переменной постоянна для всех наблюдений. Не должно быть априорной причины для того, чтобы случайная переменная порождала большую ошибку в одних наблюдениях, чем в других.

Величина σ^2_u, конечно неизвестна. Одна из задач регрессионного анализа состоит в оценке стандартного отклонения случайного члена. Если рассматриваемое условие не выполняется, то коэффициенты регрессии, найденные по МНК, будут неэффективны.

3 – е условие Гаусса – Маркова: $cov(u_i, u_j) =0$.

Отсутствие систематической связи между значениями случайной переменной в любых двух наблюдениях**.** То есть**,** если случайная переменная велика и положительна в одном наблюдении, это не должно обусловливать систематическую тенденцию к тому, что она будет большой и положительной в следующем наблюдении (или большой и отрицательной, или малой и положительной, или малой и отрицательной).

Значения случайной переменной должны быть абсолютно независимой друг от друга. Если рассматриваемое условие не выполняется, то коэффициенты регрессии, найденные по МНК, будут неэффективны.

4 – е условие Гаусса – Маркова: cov(u_i, x_j) =0

Случайная переменная должна быть распределена независимо от объясняющих переменных.

Предположение о нормальности

Наряду с условиями Гаусса – Маркова обычно также предполагается нормальность распределения случайной переменной. Если случайная переменная u нормально распределена, то нормально будут распределены и коэффициенты регрессии.

Теорема Гаусса – Маркова

Если условия Гаусса – Маркова для случайной переменной выполнены, то коэффициенты регрессии, построенные обычным МНК, будут линейными несмещенными (E(a)= α; E(b) = β) наилучшими (так как они являются наиболее эффективными в классе всех несмещенных линейных оценок) оценками.

Стандартные ошибки коэффициентов регрессии:

$$\text{с. о. (a)} = \sqrt{\frac{s_\varepsilon^2}{(n-2)}(1+\frac{\overline{x^2}}{s_x^2})} = m_a = \sqrt{\frac{\sum_{i=1}^{n}(y_i - y_x)^2 * \sum_i x^2}{(n-2) * n * \sum_i (x_i - \overline{x})^2}};$$

$$\text{с. о. (b)} = \sqrt{\frac{s_\varepsilon^2}{(n-2)s_x^2}} = m_b = \sqrt{\frac{\sum_{i=1}^{n}(y_i - y_x)^2 / (n-2)}{\sum_i (x_i - \overline{x})^2}};$$

Нулевая гипотеза: H_0: $\beta = \beta_0$;

Альтернативная гипотеза: H_1: $\beta \neq \beta_0$.

Фактическое значение t-критерия Стьюдента:

$$t = \frac{Разница\_между\_оценкой\_регрессии\_и\_гипотетическим\_значением}{Стандартноеотклонение\_(стандартнаяошибка)\_оценки}$$

$$t_a = \frac{a - \alpha_0}{m_a}; \qquad t_b = \frac{b - \beta_0}{m_b}.$$

Критические значения $t_{крит}$

$t_{крит}$ определяются величиной имеющее t распределение (Таблица А2).

В таблице $t_{крит}$ сгруппированы по уровням значимости и числу степеней свободы. Оценивание каждого параметра в уравнении регрессии поглощает одну степень свободы в выборке. Следовательно, число степеней свободы равно количеству наблюдений в выборке минус количество оцениваемых параметров.

Условие того, что оценка регрессии не должна приводить к отказу от нулевой гипотезы H_0: $\beta = \beta_0$: $-t_{крит} \le \frac{b-\beta_0}{m_b} \le t_{крит}$; или от

гипотезы H_0: $\alpha = \alpha_0$: $-t_{крит} \le \frac{a-\alpha_0}{m_a} \le t_{крит}$.

Ошибки I и II рода:

Ошибка I рода имеет место в том случае, когда отвергается истинная нулевая гипотеза.

Ошибка II рода возникает, когда не отвергается ложная гипотеза.

Чем ниже критическая вероятность, тем меньше риск получения ошибок I рода. Если используйте уровень значимости в 5%, то отвергается истинная гипотеза в 5% случаев. Если уровень значимости составляет 1%, то ошибка I рода совершается в 1% случаев. Следовательно, 1% уровень значимости более надежен.

Если нулевая гипотеза ложна, то чем выше уровень значимости, тем шире область принятия гипотезы, тем выше вероятность того, что вы не отвергаете ее, тем выше риск допущения ошибки II рода.

Если настаивайте на очень высоком уровне значимости, то столкнетесь с относительно высоким риском допущения ошибки II рода, если гипотеза окажется ложной. Если выбираете низкий уровень значимости, то оказываетесь перед относительно высоким риском допущения ошибки I рода, если гипотеза истинна.

Большинство людей выбирают достаточно простую форму обеспечения гарантий и осуществляют проверку на обоих уровнях значимости, представляя результаты каждой такой проверки.

Часто нет необходимости непосредственно ссылаться на оба результата. Если вы отклоняете гипотезу при 1% уровне значимости, то автоматически отклоняете и при уровне значимости в 5%. Если вы не отвергаете гипотезу при уровне значимости в 5%, то из этого следует, что вы не отвергаете ее и при 1% уровне значимости.

Только в одном случае вы должны представить оба результата: если гипотеза отвергается на 5%, но не на 1% уровне значимости.

Доверительные интервалы :

$b - t^*m_b \le \beta \le b + t^* m_b$ – любое гипотетическое значение для β, которое удовлетворяет данному соотношению, будет автоматически совместимо с оценкой b.

$a - t^* m_a \le \alpha \le a + t^* m_a$. - любое гипотетическое значение для α, которое удовлетворяет данному соотношению, будет автоматически совместимо с оценкой a.

Нулевая гипотеза: H_0: $\beta = 0$;

Альтернативная гипотеза: H_1: $\beta \ne 0$.

Общая сумма квадратов отклонений (TSS):

$$\text{TSS} = \sum_i (y_i - \bar{y})^2 = \sum_i y^2 - n * \bar{y}^2 .$$

Факторная сумма квадратов отклонений (ESS):

$$\text{ESS} = \sum_i (y_x - \bar{y})^2 = b^2 * (\sum_i x^2 - n * \bar{x}^2) .$$

Остаточная сумма квадратов отклонений (RSS):

$$\text{RSS} = \sum_{i=1}^{n} (y_i - y_x)^2$$

TSS = ESS + RSS

Коэффициент детерминации (R^2_{xy} , r_{xy}^2) – характеризует долю дисперсии результативного признака у, объясняемую регрессией, в общей дисперсии результативного признака. Чем ближе r_{xy}^2 к 1, тем качественнее регрессионная модель, то есть исходная модель хорошо аппроксимирует исходные данные.

$$R^2_{xy} = r_{xy}^2 = \frac{ESS}{TSS};$$

Дисперсии на одну степень свободы:

$$TMS = \frac{TSS}{n-1}; \; EMS = \frac{ESS}{1}; \; RMS = \frac{RSS}{n-1-1};$$

где $v_1 = df_{ESS} = 1$- степень свободы факторной суммы квадратов отклонений $v_2 = d_{fRSS} = n-1-1$ - степень свободы остаточной суммы квадратов отклонений;

F – критерий Фишера:

$$F_{вычисляемое} = \frac{EMS}{RMS}$$

Английским статистиком Снедекером разработаны таблицы критических значений F – отношений при разных уровнях существенности нулевой гипотезы и различном числе степеней свободы $v_1 = df_{ESS}$; $v_2 = df_{RSS}$ (Таблица А3).

Табличное значение F – критерия – это максимальная величина отношения дисперсий, которая может иметь место при случайном расхождении их для данного уровня вероятности наличия нулевой гипотезы.

Если $F_{\textbf{вычисляемое}} > F_{\text{табличное,}}$ то уравнение регрессии значимо, связь существенна, H_0 – отклоняется.

Средняя ошибка аппроксимации:

$$A = \frac{1}{n} * \sum_{i=1}^{n} \left| \frac{(y_i - y_x)^2}{y_i} \right|; \; A = \frac{100}{\bar{y}} * \sqrt{\frac{\sum_{i=1}^{n} (y_i - y_x)^2}{n}} .$$

Ошибка аппроксимации в пределах 5-7% свидетельствует о хорошем подборе данных.

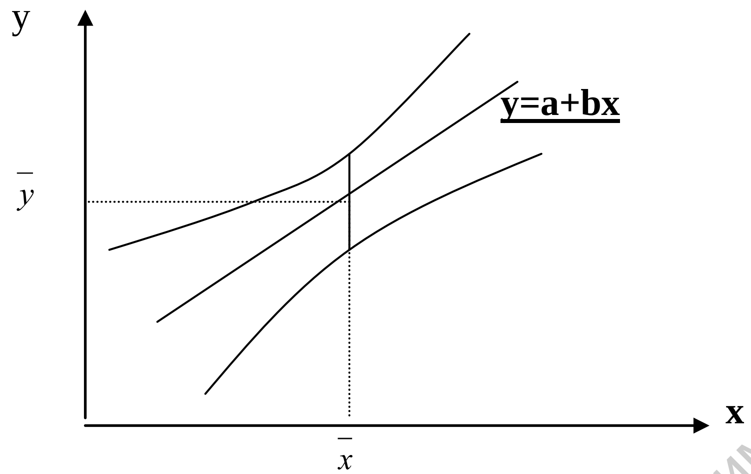
Прогнозное значение $y_P = a + b * x_л$ **;**

Стандартная ошибка $y_x(x_P)$**:**

$$m_y = \sqrt{RMS * (1 + \frac{1}{n} + \frac{(x_л - \overline{x})^2}{\sum (x_i - \overline{x})^2})}$$ **;** $x_P = x_k$

Интервальная оценка прогнозного значения:

$y_P - t_{крит} * m_y \le y^* \le y_P + t_{крит} * m_y$



Методические указания построения модели парной регрессии с помощью ППП Excel:

1. **Встроенная статистическая функция ЛИНЕЙН.**
   1. Введите исходные данные x и y.
   2. Выделите область пустых ячеек (5 строк, 2 столбца) для вывода результатов регрессионной статистики или область (1 строка, 2 столбца) для получения только оценок коэффициентов регрессии.
   3. На панели инструментов найдите значок «функция» «f».
   4. В окне «Категория» выберите «Статистические», в окне «Функция» выберите «ЛИНЕЙН»
   5. Заполните аргументы функции:
      - *Известные значения y* – диапазон, содержащий данные результативного признака;
      - *Известные значения x* – диапазон, содержащий данные факторов независимого признака;
      - *Константа* – логическое значение, которое указывает на наличие или на отсутствие свободного члена в уравнении. Если *константа* =1, то свободный член рассчитывается обычным образом. Если *константа* = 0, то свободный член =0.
      - *Статистика* – логическое значение, которое указывает, выводить дополнительную информацию по регрессионному анализу или нет. Если *Статистика* =1, то дополнительная информация выводится, если *статистика* = 0, то выводятся только оценки параметров уравнения.

6. Нажмите комбинацию клавиш: <CTRL>+<SHIFT>+<ENTER>. Дополнительная регрессионная статистика будет выводиться в порядке, указанном в таблице:

Значение коэффициента b	*Значение коэффициента a*
Стандартная ошибка m_b	*Стандартная ошибка* m_a
Коэффициент детерминации R^2_{xy}	*Стандартная ошибка* $m_y\ x_л = \bar{x}$
F - статистика	*Число степеней свободы (n-m-1)*
ESS	*RSS*

2. **Инструмент анализа данных Регрессия)**
    1. Проверьте доступ к пакету анализа. В главном меню последовательно выберите Сервис/Надстройки. Установите флажок *Пакет анализа*.
    2. В главном меню выберите Сервис/Анализ данных/Регрессия. Щелкните по кнопке ОК.
    3. Заполните диалоговое окно ввода данных и параметров вывода.
        - *Входной интервал y* – диапазон, содержащий данные результативного признака;
        - *Входной интервал x* – диапазон, содержащий данные факторов независимого признака;
        - *Метки* – флажок, который указывает, содержит ли первая строка названия столбцов или нет;
        - *Константа* – ноль – флажок, указывающий на наличие или отсутствие свободного члена в уравнении;
        - *Новый рабочий лист* – данные будут занесены на новый рабочий лист;
        - *Если необходимо получить информацию и графики остатков.* Установите соответствующие флажки в диалоговом окне.
    4. Щелкните по кнопке ОК.

Пример построения модели парной регрессии с помощью пакета Excel и оценка ее значимости.

Задание

Даны статистические данные, описывающие зависимость удельного веса бракованной продукции от удельного веса рабочих со специальной подготовкой на предприятиях. (Приложение В1).

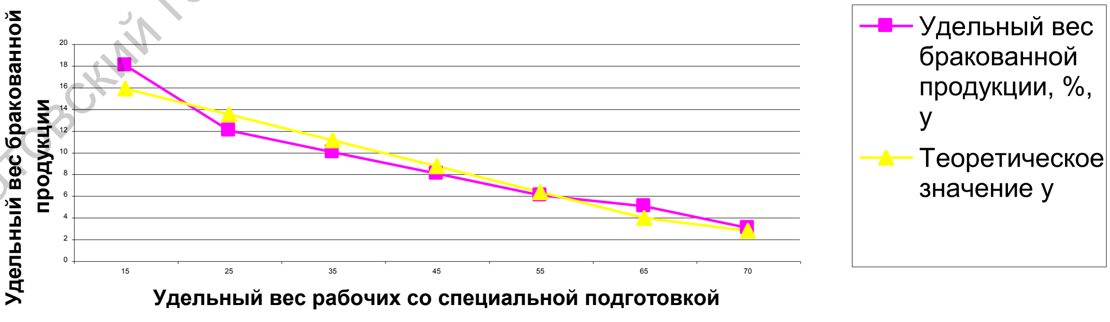
1. Постройте уравнение парной регрессии - дайте интерпретацию модели.
2. Оцените значимость коэффициентов линейной регрессии. Постройте доверительные интервалы
3. Рассчитайте F – критерий Фишера для проверки качества оценивания.
4. Спрогнозируйте значение y для какого – либо x_k., осуществив интервальный прогноз.
5. Постройте графики статистических и теоретических значений y.

№	Удельный вес рабочих со специальной подготовкой, %, x	Удельный вес бракованной продукции, %, y	Теоретическое значение y
1	15	18	15,83427762
2	25	12	13,45184136
3	35	10	11,0694051
4	45	8	8,686968839
5	55	6	6,304532578
6	65	5	3,922096317
7	70	3	2,730878187

Расчет параметров уравнения с помощью встроенной функции ЛИНЕЙН:

-0,2382	19,40793201	Параметр a= 19,41
0,0278	1,339264901	
0,93628	1,395764536	Параметр b = -0,24
73,4624	5	
143,116	9,740793201	Уравнение парной регрессии y=19,41-0,24*x.

Коэффициент корреляции rxy=0,9676156
Коэффициент детерминации rxy2=0,93628
Стандартная ошибка параметра a = 1,34
t- критерий Стьюдента ta = 19,41/1,34=14,5
Стандартная ошибка параметра b =0,028
t - критерий Стьюдента tb =-0,24 / 0,028=8,5
Табличное значение t - критерия Стьюдента для 5% уровня значимости tтабл = 2,57
Вывод: параметры регрессии значимы при 5% - ом уровне значимости.
Факторная сумма квадратов ESS = 143,116
Остаточная сумма квадратов RSS = 9,74079
Общая сумма квадратов TSS = 152,857
Факторная дисперсия на одну степень свободы EMS = 143,116
Остаточная дисперсия на одну степень свободы RMS = 1,948
Вычисляемый F - критерий Фишера F выч = 73,4624
Табличное значение F - критерия Фишера при уровне значимости 5% Fтабл = 6,61
Табличное значение F - критерия Фишера при уровне значимости 1% F табл = 16,26
Вывод: F выч > F табл при обоих уровнях значимости, следовательно модель парной линейной регрессии значима и ее можно использовать для прогнозов.



Вопросы для обсуждения:

1. Объясните, чем вызвано появление в модели регрессии стохастической переменной ε?
2. Почему перед построением модели парной линейной регрессии необходимо рассчитывать коэффициент корреляции?
3. Объясните смысл понятия «число степеней свободы».
4. По каким вычислениям можно судить о значимости модели в целом?
5. Зачем необходимо рассчитывать t-критерий Стьюдента?
6. Зачем необходимо оценивать интервалы прогноза по линейному уравнению регрессии?
7. В каких пределах должна находиться ошибка аппроксимации, чтобы можно было сделать вывод о хорошем подборе модели к исходным данным?
8. Что происходит с интервалами прогноза по мере удаления от среднего значения выборки?
9. Объясните экономический смысл TSS, ESS, RSS.
10. Когда необходимо оценивать значимость модели и параметров регрессии как при 1%, так и при 5% уровне значимости.

Задания для самостоятельной работы:

Задание I

Даны статистические данные, описывающие зависимость y от x.

1. Постройте уравнение парной регрессии - дайте интерпретацию модели.
2. Оцените значимость коэффициентов линейной регрессии. Постройте доверительные интервалы
3. Рассчитайте F – критерий Фишера для проверки качества оценивания.
4. Спрогнозируйте значение y для какого – либо x_k., осуществив интервальный прогноз.
5. Постройте графики статистических и теоретических значений y.

№1.

Статистические данные, описывающие зависимость уровня рентабельности на предприятии от скорости товарооборота.

№	1	2	3	4	5	6	7	8
Число оборотов	5,49	4,68	4,67	4,54	4,56	6,02	5,72	5,43
Уровень рентабельности, %	7,8	3,8	2,1	5,1	9,5	10,5	8,3	9,8

№2.

Статистические данным, описывающим зависимость индекса Лернера от рыночной доли фирмы.

	1	2	3	4	5	6	7
Рыночная доля фирмы, s_i	0,064	0,223	0,273	0,182	0,073	0,05	0,04
Индекс Лернера L	0,1	0,2	0,35	0,15	0,11	0,045	0,038

№3.

Статистические данные, описывающие зависимость уровня рентабельности на предприятии от удельного веса продовольственных товаров в товарообороте.

№	1	2	3	4	5	6	7
Удельный вес продовольственных товаров в товарообороте, %.	74,2	73,5	77	84,3	67,3	70,1	83,1
Уровень рентабельности, %	3,62	3,8	2,77	2,12	4,33	4,01	2,01

№4.

Статистические данные, описывающие зависимость объема спроса на товар от его цены.

№	1	2	3	4	5	6	7	8	9	10
Цена товара, руб.	99	82	77	69	52	44	31	29	28	27,5
Спрос на товар, шт.	100	115	210	270	323	478	544	564	570	574

№5.

В таблице приведены значения выручки от экспорта 1 тонны синтетического каучука за 10 кварталов и цены его на внутреннем рынке.

Период	Выручка от экспорта 1 тонны, долл.	Цена внутреннего рынка, долл. за 1 тонну
1–й квартал	1090	1090
2-й квартал	1190	1550
3-й квартал	1320	2180
4-й квартал	1430	2370
5-й квартал	1470	2440
6-й квартал	1510	2560
7-й квартал	1535	2570
8-й квартал	1570	2700
9-й квартал	1600	2759
10-й квартал	1615	2820

№6.

В таблице представлены средние расходы на потребление Y и агрегированный располагаемый доход X в некоторой национальной экономике в течение 12 лет

Год	Y_t, $	X_t ,$
1986	152	170
1987	159	179
1988	162	187
1989	165	189
1990	170	193
1991	172	199
1992	177	200
1993	179	207
1994	184	215
1995	186	216
1996	190	220
1997	191	225

№7.

В таблице приведены данные о реальной стоимости нескольких конструкторских проектов (млн. $) конструкторской фирмы и ранее выполненные оценки данных проектов. Определите функциональную связь между приведенными данными и оцените ее значимость.

Какова будет действительная стоимость проекта, оценочная стоимость которого 35, 200 млн. $.

№	1	2	3	4	5	6	7	8	9	10
Действительная стоимость проекта	0,918	7,214	14,577	30,028	38,173	15,320	14,837	51,284	34,100	2,003
Оценка стоимости проекта	0,575	6,127	11,215	28,195	30,100	21,091	8,659	40,630	37,800	1,803

№8.

Статистические данные о годовых расходах на техническое обслуживание автобусов и возраста автобусов. Спрогнозируйте расходы на содержание автобуса возрастом 10 лет.

№ автобуса	1	2	3	4	5	6	7	8	9
Расходы на содержание ($)	859	682	471	708	1094	224	320	651	1049
Возраст (годы)	8	5	3	9	11	2	1	8	12

№9.

Статистические данные количества проданных книг в сети книжных магазинов и объем демонстрационного пространства.

№	1	2	3	4	5	6	7	8	9	10
Количество проданных книг, шт	275	142	168	197	215	188	241	295	125	266
Объем демонстрационного пространства (в кв.м)	68	33	41	42	48	39	49	77	31	59

№10.

Статистические данные количества заказов на товары по почте и количества распространенных каталогов.

Город	1	2	3	4	5	6	7	8	9	10	11	12
Количество заказов на товары в каждом городе	24	16	23	15	32	25	18	18	35	34	15	32
Количество распространенных каталогов	6	2	5	1	10	7	15	3	11	13	2	12

№11.

Статистические данные количества выданных инвестиционных кредитов в банке и банковской учетной ставки.

	1	2	3	4	5	6	7	8	9
Количество выданных инвестиционных кредитов, шт	786	494	289	892	343	888	509	987	187
Банковская учетная ставка	10,2	12,6	13,5	9,7	10,8	9,5	10,9	9,2	14,2

№12.

Статистические данные цены товара компании АВС и цены товара конкурента.

	1	2	3	4	5	6	7	8	9	10	11	12	13
Цена товара компании АВС	99	104	99,5	99,9	98,9	101	100	98,2	93,8	99,4	99,7	104	99,5
Цена товара конкурента	100	105	99,5	95,9	98,8	101,5	101,2	99,1	94,8	100	99,5	103,8	99,3

№13.

Для 14 однотипных предприятий имеются данные за год о производительности труда и уровне механизации работ, %.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Производительность труда, %	20	24	28	30	31	33	34	37	38	40	41	43	45	48
Уровень механизации, %	32	30	36	40	41	47	56	54	60	55	61	67	69	76

Задание 14.

По территориям некоторых регионов страны известны данные за год по среднедневной заработной плате у и среднедушевому прожиточному минимуму х.

№	1	2	3	4	5	6	7	8	9	10
Среднедневная заработная плата, руб.	162	151	190	178	161	175	144	191	160	161
Среднедушевой прожиточный минимум, руб.	95	107	125	111	89	97	95	131	92	102

Задание II

По самостоятельно собранным реальным статистическим данным выполните все пункты, указанные в задании I.

Тема 3. Модель множественной регрессии.

1. Спецификация модели множественной регрессии.
2. Отбор факторов при построении модели множественной регрессии. Мультиколлинеарность и методы ее преодоления.
3. Параметризация модели множественной линейной регрессии. Метод наименьших квадратов для модели множественной регрессии.
4. Интерпретация уравнения множественной линейной регрессии: экономический смысл параметров регрессии. Применение модели множественной регрессии в экономике. Стандартизованное уравнение множественной регрессии.
5. Свойства коэффициентов множественной регрессии. Оценка значимости коэффициентов множественной регрессии: проверка гипотез, относящихся к коэффициентам регрессии.
6. Качество оценки: коэффициент детерминации. F – критерий Фишера для проверки качества оценивания модели множественной регрессии.
7. Фиктивные переменные в модели множественной регрессии.
8. Частные уравнения регрессии, частные коэффициенты эластичности, частная корреляция.

Ключевые слова:

Спецификация модели. Результативный признак, признак-факторы и стохастическая переменная в модели Параметры регрессии. Интеркорреляция факторов модели. Мультиколлинеарность факторов. Определитель матрицы межфакторной корреляции. Индекс множественной корреляции. Метод наименьших квадратов. Стандартизованное уравнение. Стандартизованные коэффициенты. Стандартная ошибка коэффициента множественной регрессии. F- критерий Фишера модели множественной регрессии. Скорректированные индексы множественной корреляции и детерминации. Фиктивные переменные.
Частные уравнения регрессии Частные коэффициенты эластичности Коэффициент частной корреляции. Частный F- критерий Фишера.

Основные теоретические аспекты темы.

Множественная регрессия – регрессия между переменными y и $x_1, x_2, \ldots, x_m$ то есть модель вида: $\mathbf{y}=f(x_1, x_2, \ldots, x_m)+\varepsilon$, где:
y - зависимая переменная, фактическое значение результативного признака,
$x_1, x_2, \ldots, x_m$ – независимые, объясняющие переменные, признак- факторы,
ε - возмущение, случайная, стохастическая переменная.

Причины существования ε:

1. Невключение объясняющих переменных.
2. Агрегирование переменных.
3. Неправильное описание структуры модели.
4. Неправильная функциональная спецификация.
5. Ошибки измерения

Основные типы функций, используемые при количественной оценке связей:

Линейная функция: $y=a_0+a_1 x_1+a_2 x_2+, \ldots, +a_m x_m$; или

$y=a+b_1 x_1+b_2 x_2+, \ldots, +b_m x_m$

Нелинейные функции: $y=a\, x_1^{b1} x_2^{b2}, \ldots, x_m^{bm}$ – степенная функция;

$y= 1/(a_0+a_1 x_1+a_2 x_2+, \ldots, +a_m x_m)$ – гипербола;

$y= e^{a0+a1\, x1+a2\, x2+, \ldots, +am\, xm}$ – экспонента.

Отбор факторов:

Факторы, включаемые во множественную регрессию, должны отвечать следующим требованиям:

1. Быть количественно измеримы. Если необходимо включить в модель качественный фактор, не имеющий количественного измерения, то нужно придать ему количественную определенность.
2. Не должны быть коррелированы между собой и тем более находиться в точной функциональной связи.

При дополнительном включении в регрессию (m+1) фактора коэффициент детерминации должен возрастать, а остаточная дисперсия уменьшаться. То есть $R^2_{p+1} \geq R^2_p$; $RSS_{p+1} \leq RSS_p$

Две переменные явно коллинеарны, т. е. находятся между собой в линейной зависимости, если $r_{xixj} \geq 0{,}7$.

Мультиколлинеарность факторов – когда, более чем два фактора связаны между собой линейной зависимостью, то есть имеет место совокупное воздействие факторов друг на друга. То есть некоторые факторы будут действовать в унисон.

Следствие мультиколлинеарности факторов:

1. Затрудняется интерпретация параметров множественной регрессии как характеристик действия факторов в «чистом» виде, ибо факторы коррелированы. Параметры линейной регрессии теряют экономический смысл.
2. Оценки параметров ненадежны, обнаруживают большие стандартные ошибки и меняются с изменением объема наблюдений (не только по величине, но и по знаку), что делает модель непригодной для анализа и прогнозирования.

Замечание:

Высокая коррелированность не всегда ведет к несостоятельным оценкам. Если все остальные факторы, определяющие дисперсию коэффициентов регрессии, благоприятствуют, т.е. число наблюдений и выборочные дисперсии объясняющих переменных велики и дисперсия случайной переменной мала, можно получить, тем не менее, хорошие оценки.

Мультиколлинеарность должна вызываться сочетанием высокой коррелированности и других неблагоприятных условий. Это вопрос степени выраженности явления, а не его сущности. Любая регрессия будет «страдать» от нее в определенной степени, если только все независимые переменные не будут абсолютно некоррелированными.

Правило: Число включаемых факторов обычно в 6-7 раз меньше объема совокупности, по которой строится регрессия. Если это нарушается, то df RSS очень мало, и параметры будут незначимыми, а F-критерий меньше табличного.

Оценка мультиколлинеарности факторов:

1. Расчет определителя матрицы межфакторной корреляции (на примере модели $y=a_0+a_1 x_1 +a_2 x_2 +a_3 x_3 +\varepsilon$)

$$Det|R| = \begin{vmatrix} r_{x1x1} & r_{x2x1} & r_{x3x1} \\ r_{x1x2} & r_{x2x2} & r_{x3x2} \\ r_{x1x3} & r_{x2x3} & r_{x3x3} \end{vmatrix}.$$

Чем ближе к нулю определитель матрицы межфакторной корреляции, тем сильнее мультиколлинеарность факторов и надежнее результаты множественной регрессии. И, наоборот, чем ближе к единице определитель матрицы межфакторной корреляции, тем меньше мультиколлинеарность факторов.

$Det|R| = \begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{vmatrix} = 0$ - полная мультиколлинеарность факторв;

$Det|R| = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 1$ - полное отсутствие мультиколлинеарности.

Оценка значимости мультиколлинеарности факторов может быть проведена методом испытания гипотезы о независимости переменных $H_0 : Det|R| = 1$.

Доказано, что величина $\left[n-1-\frac{1}{6}(2m+5)\lg DetR \right]$ имеет приближенное распределение χ^2 с $df = \frac{1}{2}m(m-1)$ степенями свободы. Если фактическое, рассчитываемое значение χ^2 превосходит табличное (критическое) (Таблица А4), то гипотеза H_0 отклоняется. Это означает, что $Det|R| \neq 1$, недиагональные ненулевые коэффициенты корреляции указывают на коллинеарность факторов. Мультиколлинеарность считается доказанной.

2. Расчет коэффициентов множественной детерминации. Можно найти переменные, ответственные за мультиколлинеарность факторов. Для этого в качестве зависимой переменной рассматривается каждый из факторов.

Чем ближе значение коэффициента множественной детерминации к единице, тем сильнее проявляется мультиколлинеарность факторов. Сравнивая между собой коэффициенты множественной детерминации факторов $R^2_{x1|x2x3..xm\rangle}$; $R^2_{x2|x1x3..xm\rangle}$; и т.д., можно определить переменные, ответственные за мультиколлинеарность, оставляя в уравнении факторы с минимальной величиной коэффициента множественной детерминации.

Методы преодоления сильной межфакторной корреляции:

I. Прямые: улучшение условий Гаусса – Маркова.

- Уменьшение дисперсии случайного параметра.
- Увеличение числа наблюдений.
- Увеличение дисперсии объясняющих переменных.
- Выбор переменных, наименее связанных между собой (начальный этап моделирования).

II. Косвенные:

1. Исключение из модели одного или несколько факторов.
2. Преобразование факторов, при котором уменьшается корреляция между ними.
   - Если коррелированные переменные связаны между собой концептуально, то может быть рационально объединить их в совокупный индекс (взвешенное среднее).
   - При построении модели на основе рядов динамики переходят от первоначальных данных к первым разностям уровней ряда $\Delta y = y_t - y_{t-1}$, чтобы исключить влияние тенденции.
   - Переход к уравнениям приведенной формы: $y = a + b_1x_1 + b_2x_2 + \varepsilon$; $x_1 = c + dx_2$; $y = a + b_1(c + dx_2) + b_2x_2 + \varepsilon$.
   - Переход к совмещенным уравнениям регрессии, т.е. к уравнениям, которые отражают не только влияние факторов, но и их взаимодействие. $y = a + b_1x_1 + b_2x_2 + b_3x_3 + b_1b_2x_1x_2 + b_1b_3x_1x_3 + b_2b_3x_2x_3 + \varepsilon$

Индекс множественной корреляции – характеризует тесноту связи рассматриваемого набора факторов с исследуемым результативным признаком или оценивает тесноту совместного влияния факторов на результат.Чем ближе R к 1, тем теснее связь результативного признака со всем набором исследуемых факторов.

$$R = \sqrt{1 - \frac{\sum_{i=1}^{n}(y_i - y_x)^2}{\sum_{i}(y_i - \overline{y})^2}} = \sqrt{1 - \frac{RSS}{TSS}} = \sqrt{\frac{ESS}{TSS}}\text{; } R \in [0;1]$$

Расчет параметров уравнения множественной линейной регрессии $y=a+b_1 x_1 +b_2 x_2 + \dots +b_m x_m + \varepsilon$

Дано: $y = \alpha + \beta_1 x + \beta_2 x_2 + ... + \beta_m x_m + u$ - уравнение множественной регрессии для генеральной совокупности, где:

$\alpha, \beta_1 \dots \beta_m$– истинные коэффициенты модели;

u– стохастическая переменная модели.

Необходимо по выборке, состоящей из m факторов, каждый из которых имеет n статистических значений: $x_{1j}, \dots x_{nj}$ и n статистических значений зависимой переменной y: $y_1, \dots y_n$ построить уравнение множественной регрессии $y=a+b_1 x_1 +b_2 x_2 + \dots +b_m x_m + \varepsilon$, где:

$a, b_1, \dots b_m$ - оценки $\alpha, \beta_1 \dots \beta_m$, соответствующие генеральной совокупности,

ε - оценка u.

Метод наименьших квадратов для уравнения множественной регрессии - метод оценивания параметров множественной линейной регрессии, минимизирующий сумму квадратов отклонений наблюдений зависимой переменной от искомой линейной функции.

$$RSS = \sum_{i=1}^{n} e_i^2 = \sum_i (y_i - f(x_{ij}))^2 = \sum_{i=1}^{n} (y_i - (a + b_1 x_i + b_2 x_2 + ... + b_m x_m))^2 \to \min$$

где: y_i – статистические значения зависимой переменной;

$f(x_{ij})$ - теоретические значения зависимой переменной, рассчитанные с помощью уравнения множественной регрессии.

$$\begin{cases} \frac{dRSS}{da} \\ \frac{dRSS}{db_1} \\ \frac{dRSS}{db_2} = 0; \\ . \\ . \\ . \\ \frac{dRSS}{db_m} \end{cases} \qquad Y=XA+E;\ X = \begin{pmatrix} 1 & x_{11} & x_{12} & x_{1m} \\ 1 & x_{21} & x_{22} & x_{2m} \\ . & . & . & . \\ 1 & x_{n1} & x_{n2} & x_{nm} \end{pmatrix};\ Y = \begin{pmatrix} y_1 \\ y_2 \\ . \\ y_n \end{pmatrix};\ E = \begin{pmatrix} E_1 \\ E_2 \\ . \\ E_n \end{pmatrix};$$

$$A = (a, b_1, \dots, b_m) = (X^T X)^{-1} X^T Y$$

Экономический смысл параметров уравнения множественной линейной регрессии:

Параметры a, b_1, b_2, … b_m , (a_0, a_1, a_2, … a_m) называются **коэффициентами «чистой» регрессии** и характеризуют среднее изменение результата с изменением соответствующего фактора на единицу при неизменном значении других факторов, закрепленных на среднем уровне.

Стандартизованное уравнение множественной регрессии:

$$t_y = \eta_1 x_1 + \eta_2 x_2 + ... + \eta_m x_m; \qquad \eta_j = b_j * \frac{s_{xj}}{s_{yj}}.$$

η_j - стандартизованные коэффициенты множественной регрессии, показывают, что с изменением фактора x_j на одно выборочное стандартное отклонение s_j при неизменных остальных факторах y изменяется на η_j выборочных стандартных отклонений.

Свойства коэффициентов модели множественной регрессии:

Коэффициенты множественной регрессии определяются значениями независимых переменных и случайным параметром, и их свойства существенно зависят от последнего. Продолжаем считать, что выполняются условия Гаусса – Маркова, а именно:

1. Математическое ожидание случайной переменной в любом наблюдении должно быть равно нулю.
2. Теоретическая дисперсия случайной переменной постоянна для всех наблюдений.
3. Отсутствие систематической связи между значениями случайной переменной в любых двух наблюдениях.
4. Случайная переменная должна быть распределена независимо от объясняющих переменных.

Но существует еще два практических требования:

1. Необходимо иметь достаточное количество данных.
2. Отсутствие строгой линейной зависимости между признак – факторами.

В теореме Гаусса – Маркова для множественной регрессии доказывается, что как и для парной регрессии, обычный МНК дает наиболее эффективные линейные оценки в том смысле, что на основе той же самой выборочной информации невозможно найти другие несмещенные оценки с меньшими дисперсиями при выполнении условий Гаусса – Маркова.

Коэффициенты регрессии будут более точными:

- чем больше число наблюдений в выборке;
- чем больше дисперсия объясняющих переменных;
- чем меньше теоретическая дисперсия случайного параметра
- чем меньше связаны между собой объясняющие переменные.

Нулевая гипотеза: H_0: $\beta_j = \beta_0$;

Альтернативная гипотеза: H_1: $\beta_j \neq \beta_0$.

Стандартные ошибки коэффициентов множественной регрессии:

$$m_{bj} = \frac{s_{yj}\sqrt{1-R^2_{yx1..xm}}}{s_{xj}\sqrt{1-R^2_{xjx1..xm}}} * \frac{1}{\sqrt{n-m-1}}$$; где :

s_y –среднее кавдратическое отклонение для признака у;
s_x – среднее квадратическое отклонение для признака x_j;
$R^2_{yx1....xm}$ – коэффициент детерминации для уравнения множественной регрессии;
$R^2_{xj\ x1....xm}$ – коэффициент детерминации для зависимости фактора x_j со всеми другими факторами уравнения множественной регрессии;
n-m-1 – число степеней свободы для остаточной суммы квадратов отклонений

Фактическое значение t-критерия Стьюдента:

$t_a = \dfrac{a-\alpha_0}{m_a}$ **;** $t_b = \dfrac{b_j-\beta_{0j}}{m_{bj}}$**.**

Критические значения $t_{крит}$ определяются аналогично $t_{крит}$ модели парной регрессии величиной имеющее t распределение (Таблица А2).
Условие того, что оценка регрессии не должна приводить к отказу от нулевой гипотезы H_0: $\beta_j = \beta_{0j}$ **:** $-t_{крит} \le \dfrac{b_j-\beta_{0j}}{m_{bj}} \le t_{крит}$**.**

Доверительные интервалы :
b_j - $t^*m_{bj} \le \beta_j \le b_j + t^*\ m_{bj}$ – любое гипотетическое значение для β_j, которое удовлетворяет данному соотношению, будет автоматически совместимо с оценкой b_j.

Коэффициент множественной детерминации $R^2_{xy} = \dfrac{ESS}{TSS}$;

F – критерий Фишера для модели множественной регрессии:

$TMS = \dfrac{TSS}{n-1}$; $EMS = \dfrac{ESS}{m}$; $RMS = \dfrac{RSS}{n-1-m}$; где:

где: n – число наблюдений;
m – количество факторов в уравнении регрессии;
$v_1=df_{ESS}=m$- степень свободы факторной суммы квадратов отклонений
$v_2=d_{fRSS}$ = n-1-m - степень свободы остаточной суммы квадратов отклонений;

F – критерий Фишера:

$$F_{вычисляемое} = \frac{EMS}{RMS} = \frac{R^2/m}{(1-R^2)/(n-1-m)};$$

Критические значения F – отношений при разных уровнях существенности нулевой гипотезы и различном числе степеней свободы $v_1=df_{ESS}$; $v_2=df_{RSS}$ определяются аналогично критическим значениям F – отношений модели парной регрессии (Таблица А).

Скорректированный индекс множественной корреляции:

В модели множественной регрессии в расчете показателя множественной корреляции используется остаточная дисперсия, которая имеет систематическую ошибку в сторону преуменьшения, тем более значительную, чем больше параметров определяется в уравнении регрессии при заданном объеме n. Если число параметров приближается к объему наблюдений, то остаточная дисперсия будет близка к нулю и индекс корреляции приблизится к единице даже при слабой связи факторов с результатом.

Для того, чтобы не допустить возможного преувеличения тесноты связи, применяется скорректированный индекс (коэффициент) множественной корреляции. Скорректированный индекс множественной корреляции содержит поправку на число степеней свободы, а именно остаточная сумма квадратов RSS делится на число степеней свободы остаточной вариации (n-m-1) , а общая сумма квадратов отклонений – на число степеней свободы в целом по совокупности (n-1).

$$\bar{R} = \sqrt{1 - \frac{\sum_{i=1}^{n} (y_i - y_x)^2 / (n - m - 1)}{\sum_i (y_i - \overline{y})^2 / (n - 1)}};$$

Скорректированный индекс множественной детерминации:

$$\bar{R}^2 = 1 - (1 - R^2) * \frac{(n-1)}{n - m - 1};$$

Чем больше величина m, тем сильнее различия $\overline{R^2}$ и R^2.

Фиктивные переменные (структурные переменные) – переменные, отражающие атрибутивные признаки (пол, профессия, образование, климатические условия, принадлежность к региону и др.) Чтобы ввести такие переменные в регрессионную модель, им должны быть присвоены цифровые метки, то есть качественные переменные необходимо преобразовать в количественные.

Частным уравнением регрессии модели $y=a_0+a_1 x_1 +a_2 x_2 +, \dots , +a_m x_m + \varepsilon$ называется уравнение вида $y_{xi\ x1,\dots,x(i-1),\ x(i+1),\dots,xm} = f(\bar{x}_1,\dots,\bar{x}_{i-1},x_i,\bar{x}_{ii+1},\dots\bar{x}_m)$, то есть уравнения регрессии, которые связывают результативный признак с соответствующими факторами x при закреплении других учитываемых в уравнении регрессии факторов на среднем уровне.

Коэффициент частной корреляции – измеряет влияние на результат фактора x_i при неизменном уровне других факторов**:**

$$r_{yxix1x2\ldots x(i-1)x(i+1)\ldots xm} = \sqrt{1 - \frac{1 - R^2_{yx1x2\ldots xi\ldots xm}}{1 - R_{yx1x2\ldots x(i-1)x(i+1)\ldots xm}}};$$

где: $R^2_{yx1x2\ldots xi\ldots xm}$ – множественный коэффициент детерминации всего комплекса m факторов с результатом;

$R^2_{yx1x2\ldots x(i-1)\ x(i+1)\ \ldots xm}$ – тот же показатель детерминации, но без введения в модель фактора x_i.

Частный F-критерий Фишера модели множественной регрессии для фактора x_i:

$$F_{xi} = \frac{R^2_{yx1..xi..xm} - R^2_{yx1..x(i-1)x(i+1)..xm}}{1 - R^2_{yx1..xi..xm}} * \frac{n-m-1}{1};$$

где в числителе показан прирост доли объясненной вариации у за счет дополнительного включения в модель соответствующего фактора. А в знаменателе – доля остаточной вариации по регрессионной модели, включающей полный набор факторов.

Вопросы для обсуждения:

1. Почему необходимо часто строить модель множественной регрессии; приведите примеры экономических процессов и явлений, в которых Вы бы применяли данную модель?
2. В чем отличие целей построения модели парной регрессии и модели множественной регрессии?
3. Объясните, почему в эконометрическом моделировании возникает проблема мультиколлинеарности?
4. Каким свойствам должны отвечать параметры модели множественной регрессии и почему?
5. Почему параметры «чистой» регрессии не сопоставимы между собой? Каковы методы разрешения данной проблемы?
6. Как должны соотноситься коэффициенты детерминации для m и m+1 факторов модели?
7. В чем заключается смысл расчета скорректированного индекса корреляции, и какова связь его с индексом корреляции при различных количествах вводимых в модель факторах?
8. Каков экономический смысл применения фиктивных переменных в модели множественной регрессии?
9. Сколько фиктивных переменных должно быть в модели, если необходимо оценить влияние двух качественных признаков, один из которых принимает три характеристики, а другой две?
10. Объясните практическое применение в экономике частных коэффициентов эластичности.

Пример построения модели парной регрессии с помощью пакета Excel и оценка ее значимости.

Задание

Даны статистические данные, описывающие зависимость цены от потребительских свойств станка (модель Ланкастера) (Приложение В2):

P – основной размер станка;
N – мощность главного привода;
n - максимальная частота вращения шпинделя;
УА – уровень автоматизации;
Т – класс точности
Y – цена станка.

1. Постройте уравнение множественной регрессии - дайте интерпретацию модели.
2. Проведите анализ факторов на предмет мультиколлинеарности.
3. Постройте стандартизованное уравнение регрессии. Какие выводы можно сделать из построенного стандартизованного уравнения?
4. Оцените значимость коэффициентов линейной регрессии. Постройте доверительные интервалы
5. Рассчитайте F – критерий Фишера для проверки качества оценивания.
6. Спрогнозируйте значение y для какого – либо набора x_{j}..

№	P, мм Размер	N, кВТ Мощность	n, об/мин Скорость вращения шпинделя	УА Уровень автоматизации	Т, Точность	Y Цена станка, тыс.руб.
1	400	11	2000	3	1	53,6
2	400	10	1500	1	1	43,6
3	400	8	2000	1	1	35
4	400	8	2500	1	1,6	39
5	320	8	2500	1	1	27
6	320	8	3000	3	1,6	44
7	320	6,3	2000	1	1	26
8	250	8	3000	3	1,6	32,7
9	250	6,3	3000	1	1,6	22,5
10	250	5,5	2500	1	1	20
11	320	8	2500	1	1,6	29,8
12	400	8	2000	1	1,6	37,5

Задания для самостоятельной работы:

Задание I

Даны статистические данные, описывающие зависимость y от $x_1, \ldots x_m$.

1. Постройте уравнение множественной регрессии - дайте интерпретацию модели.
2. Проведите анализ факторов на предмет мультиколлинеарности.
3. Постройте стандартизованное уравнение регрессии. Какие выводы можно сделать из построенного стандартизованного уравнения?
4. Оцените значимость коэффициентов линейной регрессии. Постройте доверительные интервалы. Если коэффициенты окажутся статистически незначимыми, какова причина данного результата?
7. Рассчитайте F – критерий Фишера для проверки качества оценивания.
8. Спрогнозируйте значение y для какого – либо набора x_{j}..

№1.

№ торговых предприятий	Факторы		Уровень рентабельности, %
	Удельный вес продовольственных товаров в товарообороте, %	Среднемесячная оплата труда, руб.	
1	74,2	1560	3,62
2	73,5	1620	3,8
3	77	1490	2,77
4	84,3	1330	2,12
5	67,3	1970	4,33
6	70,1	1820	4,01
7	83,1	1270	2,01

№2.

№	Факторы					Уровень рента-бельности, %
	Среднеме-сячный товарооборот в расчете на душу населения	Удельный вес продоволь-ственных товаров в товарообо-роте, %	Время обращения товаров, дней	Среднеме-сячная оплата труда	Трудоемкость товаро-оборота (численность работников на 100000 ед. товарооборота)	
1	27	74,2	35	1560	11	3,62
2	29	73,5	32	1620	12	3,8
3	28	77	33	1490	13	2,77
4	21	84,3	41	1330	17	2,12
5	35	67,3	29	1970	9	4,33
6	33	70,1	31	1820	10	4,01
7	21	83,1	39	1270	18	2,01

№3.

№ завода	Факторы		Производительность труда
	Удельный вес рабочих с технической подготовкой, %	Удельный вес механизированных работ, %	
1	64	84	4300
2	61	83	4150
3	47	67	3000
4	46	63	3420
5	49	69	3300
6	54	70	3400
7	53	73	3420
8	61	81	4100
9	57	77	3700
10	54	72	3500
11	60	80	4000
12	67	85	4450
13	63	83	4270
14	50	70	3300
15	67	87	4500

№4.

Статистические данным, описывающие зависимость накопления пяти случайно выбранных семей от дохода и размера имущества. Спрогнозируйте накопления семьи, имеющей доход 40 тыс. руб. и имущество стоимостью 25 тыс. руб.

Семья	Накопления, S	Доход, Y	Имущество, W
1	3,0	40	60
2	6,0	55	36
3	5,0	45	36
4	3,5	30	15
5	1,5	30	90

№5.

В кейнсианской теории спрос на деньги зависит от доходов и процентных ставок. Рассмотрим модель: $m_t=a_0+ a_1y_t+a_2i_t+\varepsilon_t$; где

m_t – агрегат денежной массы М1 (млрд. долл.),

y_t- валовой национальный доход (млрд. долл.);

i_t – процентные ставки по государственным облигациям (%).

Оцените спрос на деньги при 1) ВНД = 1000 и i = 10%; 2) ВНД= 2500 и i = 5%..

t	1	2	3	4	5	6	7	8	9	10	11	12	13	14
m_t	141	146	149	154	161	173	199	216	251	277	310	363	414	478
y_t	506	524	565	597	637	756	873	992	1185	1434	1718	2164	2631	3073
i_t	3,3	2,6	2,9	3,2	3,7	5	5,5	6,6	4,5	7,9	5,3	7,6	11,4	11,1

№ 7.

Статистические данные реального дохода на душу населения y (тыс. долл.) процента рабочей силы, занятой в сельском хозяйстве – x_1, и среднего уровня образования населения в возрасте после 25 лет x_2 (число лет, проведенных в учебных заведениях) для 15 развитых стран с 1983 г.

Страна	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
y	7	9	9	8	8	14	9	8	10	11	11	12	9	10	12
x_1	8	9	7	6	10	4	5	5	6	7	6	4	8	5	8
x_2	9	13	11	11	12	16	11	11	12	14	11	15	15	10	13

№8

Данные о величинах объема реализации продукции y некоторой фирмы, зависящие от цены x_1 и расходов на рекламу x_2.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Y, шт.	23	18	27	29	43	23	55	47	35	38	14	51	20	39	35
x_1, у.е.	37	33	15	36	26	24	15	33	44	34	63	8	44	43	31
x_2,у.е.	39	40	35	48	53	42	54	54	50	53	46	50	43	55	51

№9.

Данные о зависимости цены товара y, руб. от дальности его перевозки x_1, км и расходов на рекламу x_2., тыс руб.

	1	2	3	4	5	6	7
Y, руб	46,72	51,01	49,39	71,71	65,16	67,27	40,09
x_1, км.	10	17	15	25	19	20	8
x_2, тыс. руб	9	13	9	10	5	6	11

№10.

Зависимость выработки продукции на одного работника y, тыс. руб. от ввода в действие новых основных фондов x_1, % от стоимости фондов на конец года и от удельного веса рабочих высокой квалификации в общей численности рабочих x_2, %.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
y	23,1	27,5	23,8	28,9	28,1	33,5	31,7	32,8	32,4	39,2	34,7	36,8	38,7	41,3	46,8	51,7
X_1	5,6	4,4	2,2	7	3,7	6,2	7,5	5,3	5,2	6,7	5,9	7,3	7,6	5,9	7,2	9,1
X_2	10	14	15	16	17	19	19	20	20	20	21	22	22	25	28	29

№11.

По 20 крупнейшим компаниям имеются данные за год. Y – чистая прибыль, млрд. долл., x1 – оборот капитала, млрд. долл., x2 – использованный капитал, млрд. долл., x3 – численность служащих, тыс.чел.

№	Чистая прибыль y	Оборот капитала	Использованный капитал	Численность служащих
1	7,9	6,9	83,6	222
2	2,4	18	6,5	32
3	9,2	107	50,4	82
4	2,8	16,7	15,4	45,2
5	4,7	79,6	29,6	299,3
6	1,9	16,2	13,3	41,6
7	0,5	5,9	5,9	17,8
8	4,5	53,1	27,1	151
9	3,2	18,8	11,2	82,3
10	3,8	35,3	16,4	103
11	4	71,9	32,5	225,4
12	1,6	93,6	25,4	675
13	1,7	10	6,4	43,8
14	3,7	31,5	12,5	102,3
15	2,9	36,7	14,3	105
16	2,7	13,8	6,5	49,1
17	4,1	64,8	22,7	50,4
18	2,1	30,4	15,8	480
19	2,3	12,1	9,3	71
20	4	31,3	18,9	43

№12.

Менеджер отдела кадров компании заинтересован в получении обоснованного прогноза, сможет ли определенный кандидат стать хорошим продавцом. Для этой цели в качестве зависимой переменной y он выбрал данные об объеме продаж за месяц работы и решил принять к рассмотрению следующие независимые переменные:

X1 – результат теста способностей к продаже;
X2 – возраст;
X3 – результат теста тревожности;
X4 - опыт работы;
X5 – средний балл школьного аттестата.

№	Объем продаж за месяц, шт.	Результат теста способностей	Возраст	Результат теста тревожности	Опыт работы	Средний балл школьного аттестата
1	44	10	22,1	4,9	0	2,4
2	47	19	22,5	3	1	2,6
3	60	27	23,1	1,5	0	2,8
4	71	31	24	0,6	3	2,7
5	61	64	22,6	1,8	2	2
6	60	81	21,7	3,3	1	2,5
7	58	42	23,8	3,2	0	2,5
8	56	67	22	2,1	0	2,3
9	66	48	22,4	6	1	2,8
10	61	64	22,6	1,8	1	3,4
11	51	57	21,1	3,8	0	3
12	47	10	22,5	4,5	1	2,7
13	53	48	22,2	4,5	0	2,8
14	74	96	24,8	0,1	3	3,8
15	65	75	22,6	0,9	0	3,7
16	33	12	20,5	4,8	0	2,1
17	54	47	21,9	2,3	1	1,8
18	39	20	20,5	3	2	1,5
19	52	73	20,8	0,3	2	1,9
20	30	4	20	2,7	0	2,2
21	58	9	23,3	4,4	1	2,8
22	59	98	21,3	3,9	1	2,9
23	52	27	22,9	1,4	2	3,2
24	56	59	22,3	2,7	1	2,4
25	49	23	22,6	2,7	1	2,4
26	63	90	22,4	2,2	2	2,6
27	61	34	23,8	0,7	1	3,4
28	39	16	20,6	3,1	1	2,3
29	62	32	24,4	0,6	3	4
30	78	94	25	4,6	5	3,6

№13.

Менеджер по продажам фирмы, занимающейся реализацией запчастей к автомобилям, хотел бы найти модель, с помощью которой уже в мае можно было бы спрогнозировать годовой объем продаж в регионе. Если этот объем можно спрогнозировать для каждого региона, то можно будет составить прогноз продаж и для всей компании в целом. Количество пунктов розничной торговли данной компании в регионе и количество автомобилей, зарегистрированных в регионе на 1 мая, - это две независимые переменные.

Регион	1	2	3	4	5	6	7	8	9	10	11
Годовой объем продаж, (млн. долл.)	52,3	26	20,2	16	30	46,2	35	3,5	33,1	25,2	38,2
Количество пунктов обслуживания	2011	2850	650	480	1694	2302	2214	125	1840	1233	1699
Количество, зарегистрированных автомобилей, млн. шт.	24,6	22,1	7,9	12,5	9	11,5	20,5	4,1	8,9	6,1	9,5

№14.

Исследователю необходимо изучить, насколько некоторый тест способностей сможет предсказать будущую производительность труда работника. Восемь женщин и семь мужчин выполнили предусмотренные тестом задания, предназначенные для оценки ловкости рук при работе с мелкими предметами. После этого каждый из протестированных прошел месячный курс подготовки к работе сборщиком электронных плат, а затем в течение месяца выполнял соответствующую работу, после чего его производительность была оценена индексом, принимающим значение от 0 до 10.

Работник	Оценка производительности	Данные теста способностей	Пол
1	5	60	Ж
2	4	55	Ж
3	3	35	Ж
4	10	96	Ж
5	2	35	Ж
6	7	81	Ж
7	6	65	Ж
8	9	85	Ж
9	9	99	М
10	2	43	М
11	8	98	М
12	6	91	М
13	7	95	М
14	3	70	М
15	6	85	М

№15.

Необходимо исследовать зависимость между результатами письменных вступительных y и курсовых экзаменов x по математике. Получены данные о числе решенных задач (задание – 10 задач) на вступительных и курсовых экзаменах 16 студентами, а также распределение этих студентов по категории «пол».

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
y	9	8	2	10	4	4	9	2	10	4	9	3	9	1	8	5
X	10	8	5	10	6	6	10	4	10	6	9	6	10	4	9	8
пол	м	м	ж	м	ж	ж	м	ж	м	ж	м	ж	м	ж	м	ж

№16.

Зависимость средней заработной платы y, долл. от возраста рабочего x, лет.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
y	338	323	299	334	305	327	336	319	340	315	323	321	293	313	352
X	35	55	32	44	33	58	38	50	45	47	38	35	35	48	58
пол	м	ж	ж	м	ж	ж	м	ж	м	ж	м	м	ж	ж	м

Задание II

По самостоятельно собранным реальным статистическим данным выполните все пункты, указанные в задании I.

Тема 4. Нелинейные модели парной и множественной регрессии.

1. Нелинейная модель парной регрессии. Применение нелинейной модели парной регрессии в микроэкономике и макроэкономике. Тест Бокса-Кокса.
2. Нелинейная модель множественной регрессии. Применение нелинейной модели множественной регрессии в микроэкономике и макроэкономике.

Ключевые слова:

Нелинейные модели внутренне линейные. Нелинейные модели внутренне нелинейные. Коэффициенты эластичности. Коэффициенты роста. Уровень насыщения. Замена переменных. Логарифмирование. Потенциирование. Производственная функция Кобба Производительность факторов производства.– Дугласа. Кривые Энгеля. Функции спроса.

Основные теоретические аспекты темы.

I.Основные типы нелинейных, внутренне линейных, функций, используемых при количественной оценке связей в парной регрессии:

Нелинейные функции: $y= a+b/x+\varepsilon$ - гипербола;

$y=a+bx+cx^2 +\varepsilon$– парабола;

$y=a+bx+cx^2+dx^3+\varepsilon$ - кубический многочлен;

$y=ax^b\varepsilon$–степенная функция;

$y=ab^x\varepsilon$-показательная функция;

$y=a+b\lg x+\varepsilon$-логарифмическая функция;

$y= 1/(a+bx)+ \varepsilon$;

$y=a+bx+c(1/x)+ \varepsilon$;

$y=1/(a+bx+cx^2)+ \varepsilon$;

$\lg y=a+bx+cx^2+\varepsilon$

$y=a/(1+be^{-cx})+ \varepsilon$.

II. Нелинейные модели, внутренне нелинейные:

$y=ax^b+\varepsilon$;

$y=a+bx^c+\varepsilon$;

$$y = a(1 - \frac{1}{1 - x^b}) + \varepsilon \ .$$

1.1 Регрессии, нелинейные относительно включенных в анализ объясняющих переменных, но линейные по оцениваемым параметрам (сводятся к линейным путем замены**):**

$y= a+b/x+\varepsilon$ - гипербола;
$y=a+bx+cx^2 +\varepsilon$– парабола;
$y=a+bx+cx^2+dx^3+\varepsilon$ - кубический многочлен
$y= a+b\ln x+\varepsilon$ – логарифмическая функция
$y= 1/(a+bx)+ \varepsilon$;
$\lg y=a+bx+cx^2+\varepsilon$
$y=a+bx+c(1/x)+ \varepsilon$;
$y=1/(a+bx+cx^2)+ \varepsilon$

1.2. Регрессии, нелинейные по оцениваемым параметрам (сводятся к линейным путем логарифмирования**):**

$y=ax^b\varepsilon$–степенная функция;

Экономический смысл коэффициента b: b - называется коэффициентом эластичности**,** показывает на сколько % изменится в среднем результат при изменении соответствующего фактора x на 1%.

$y=ab^x\varepsilon$-показательная функция – функция постоянного роста;

Экономический смысл коэффициента b: b– коэффициент роста.

$y=e^{a+bx}\varepsilon$;
$y=e^{a+b/x}\varepsilon$; s – образная кривая;

Экономический смысл коэффициента b: exp(b) определяет уровень насыщения.

Тест Бокса – Кокса – тест сравнения коэффициента R^2 для линейной и логарифмической вариантов модели.

Тест предполагает такое преобразование масштаба наблюдений y, при котором обеспечивалась бы возможность непосредственного сравнения RSS в линейной и логарифмических моделях.

1. Вычисляется среднее геометрическое значений y в выборке. Оно совпадает с экспонентой среднего арифметического lny, которое легко рассчитать: $e^{\frac{1}{n}\sum \ln y} = (y_1 * ... * y_n)^{\frac{1}{n}}$;
2. Пересчитываются наблюдения y, они делятся на это значение: $y_i^* = \frac{y_i}{среднеегеометрическое\_y}$;
3. оценивается регрессия для линейной модели с использованием y^* вместо y в качестве зависимой переменной и для логарифмической модели с использованием $\ln y^*$ вместо lny; во всех других отношениях модели должны оставаться неизменными. Теперь значения RSS для двух регрессий сравнимы, и, следовательно, модель с меньшей суммой квадратов отклонений обеспечивает лучшее соответствие.

4. Для того чтобы проверить, обеспечивает ли одна из моделей значимо лучшее соответствие, можно вычислить величину $-\frac{n}{2}\ln z$, где z – отношение RSS в пересчитанных регрессиях, а n – число наблюдений, и взять ее абсолютное значение. Эта статистика имеет распределение χ^2 (хи - квадрат) с одной степенью свободы. Если она превышает критическое значение χ^2 при выбранном уровне значимости, то делается вывод о наличии значимой разницы в качестве оценивания.

Основные типы нелинейных, используемых при количественной оценке связей во множественной регрессии:

$y = a\, x_1^{b1}\, x_2^{b2}, \ldots, x_m^{bm}$ – степенная функция;

Экономический смысл коэффициентов b_1, b_2, b_m: b_1, b_2, b_m - коэффициенты эластичности, показывают на сколько % изменится в среднем результат при изменении соответствующего фактора на 1% при неизменности действия других факторов

$y = 1/(a_0 + a_1 x_1 + a_2 x_2 +, \ldots, +a_m x_m)$ – гипербола;

Эта функция используется для аппроксимации прогнозируемой переменной y, имеющей резкие выбросы в некоторых точках данных.

$y = ab_1^{x1} b_2^{x2} \ldots b_m^{xm} \varepsilon$ - функция постоянного роста; коэффициенты $b_1^{x1}, b_2^{x2}, b_m^{xm}$ - коэффициенты роста по фактору x_i.

$y = e^{a0 + a1\, x1 + a2\, x2 +, \ldots, + am\, xm}$ – экспонента.

F-критерий Фишера для нелинейной регрессии:

$$F = \frac{R^2}{1 - R^2} * \frac{n - m - 1}{m};$$

где: n – число наблюдений, а m – число параметров при переменных x.

Вопросы для обсуждения:

1. Объясните необходимость построения нелинейных моделей парной и множественной регрессии в экономике.
2. В чем принципиальное отличие нелинейных, внутренне линейных, функций и нелинейных, внутренне нелинейных, функций?
3. Приведите области в экономике, в которых можно использовать: Гиперболу, степенную функцию, показательную, параболу, кубический многочлен.
4. Какие методы используются для сведения нелинейных функций к линейному виду?
5. Почему не сравнимы между собой коэффициенты детерминации линейной модели и модели, построенной при использовании логарифмов этих же данных?
6. В чем особенность оценки статистической значимости нелинейных моделей парной регрессии?

Задания для самостоятельной работы:

Задание

Даны статистические данные, описывающие зависимость y от $x_1, \dots x_m$.

1. Постройте уравнение множественной регрессии - дайте интерпретацию модели.
2. Оцените значимость коэффициентов регрессии. Постройте доверительные интервалы. Если коэффициенты окажутся статистически незначимыми, какова причина данного результата?
3. Рассчитайте F – критерий Фишера для проверки качества оценивания.
4. Спрогнозируйте значение y для какого – либо набора x_j..

№1.

Кривая Филипса описывает связь темпа роста зарплаты и уровня безработицы. А именно: $\delta\omega_t = \beta_1 + \beta_2 * \frac{1}{u_t} + \varepsilon_t$, где ω_t –уровень заработной платы, $\delta\omega_t = 100(\omega_t - \omega_{t-1})/\ \omega_{t-1}$ - темп роста зарплаты (в процентах) и u_t – процент безработных в год t. Используя данные для некоторой страны постройте уравнение парной регрессии и проверьте наличие значимой связи между $\delta\omega$ и u. Найдите «естественный уровень безработицы», т. е. такой уровень безработицы, при котором $\delta\omega$=0.

Год t	ω_t	u_t
1	1.62	1
2	1.65	1.4
3	1.79	1.1
4	1.94	1.5
5	2.03	1.5
6	2.12	1.2
7	2.26	1.0
8	2.44	1.1
9	2.57	1.3
10	2.66	1.8
11	2.73	1.9
12	2.8	1.5
13	2.92	1.4

№2.

Менеджер новой чебуречной не уверен в правильности выбранной цены на чебуреки, поэтому в течение 12 недель он варьирует цену и записывает количество проданных чебуреков. Получены следующие данные.

1. Постройте линейную модель парной регрессии. Найдите оптимальную, в смысле максимума выручки от продаж цену чебурека. Какую ценовую политику следует предпринять менеджеру чебуречной?
2. Постройте модель степенной функции и оцените, эластичен ли спрос на чебуреки? Совпадает ли ваш вывод о ценовой политике с предыдущим?
3. Проведите тест Бокса – Кокса.

Неделя	1	2	3	4	5	6	7	8	9	10	11	12
Количество проданных чебуреков q, шт.	795	915	965	892	585	644	714	1180	851	779	625	1001
Цена чебуреков, p, руб.	12,3	11,5	11	12	13,5	12,5	12,8	9,9	12,2	12,5	13	10,5

№3.

После финансового кризиса спрос на чебуреки упал, и менеджер был вынужден тратить часть средств на рекламу. Для изучения зависимости объема продаж от цены и расходов на рекламу менеджер использует следующую модель: $q_t=a_0+ a_1p_t+a_2c_t+a_3c_t^2+\varepsilon_t$.

t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
q_t	525	567	396	726	265	615	370	789	513	661	407	608	399	631	545	512	845	571
p_t	5,9	6,5	6,5	6,1	6,6	5,2	5,1	5,1	6,7	5,5	6,6	6,9	6,9	6,5	6,5	6,8	5,1	6,1
c_t	479	361	549	278	574	134	581	339	374	359	519	327	469	379	429	271	221	309

1. Пусть себестоимость производства одного чебурека = 2 руб. Найдите формулу выручки и прибыли
2. Найдите оптимальную цену, при которой прибыль достигает максимального значения, при расходах на рекламу = 280 руб.
3. Найдите оптимальный уровень расходов на рекламу, при котором прибыль достигает максимального значения, при цене чебурека =6 руб.
4. Помогите менеджеру найти оптимальное решение для цены и объема расходов на рекламу, при которых прибыль оптимальна (достигает максимального значения).

№4.

Дана зависимость выпуска Q от трудозатрат L и капиталовложений K 15 фирм некоторой отрасли. Оцените по этим данным производственную функцию Кобба-Дугласа $Q=\alpha L^{\beta 1}K^{\beta 2}$.

Фирма №	Q	L	K
1	2350	2334	1570
2	2470	2425	1850
3	2110	2230	1150
4	2560	2463	1940
5	2650	2565	2450
6	2240	2278	1340
7	2430	2380	1700
8	2530	2437	1860
9	2550	2446	1880
10	2450	2403	1790
11	2290	2301	1480
12	2160	2253	1240
13	2400	2367	1660
14	2490	2430	1850
15	2590	2470	2000

Рассчитайте объем выпуска при L = 2500 и K = 1800

№5.

Выберите модель нелинейной парной регрессии, описывающей зависимость между ежегодным потреблением бананов y, кг. и годовым доходом, x, тыс. руб.

Семья	1	2	3	4	5	6	7	8	9	10
Y	1,93	7,14	8,78	9,69	10,09	10,42	10,62	10,71	10.79	11,13
x	10	20	30	40	50	60	70	80	90	100

№6.

Выберите модель нелинейной парной регрессии, описывающей зависимость между ежегодным потреблением апельсинов y, кг. и годовым доходом, x, тыс. руб. для 10 семей.

Семья	1	2	3	4	5	6	7	8	9	10
Y	1,15	4,71	5,59	7,43	7,47	7,33	7,7	8,15	8,35	8,59
x	10	20	30	40	50	60	70	80	90	100

№7.

Динамика выпуска продукции y, млн. долл. некоторой страны за 30 лет характеризуется данными. Постройте экспоненциальную модель.

Год, t	1	2	3	4	5	6	7	8	9	10
y	1089	1006	1450	1273	1983	2076	2017	2193	2170	2378
Год, t	11	12	13	14	15	16	17	18	19	20
y	2914	3100	3741	3286	3706	5608	4326	6400	6373	9995
Год, t	21	22	23	24	25	26	27	28	29	30
y	7709	7977	9634	9818	9453	12553	9445	18299	12888	18581

№8.

Выберите модель нелинейной парной регрессии, описывающей зависимость между рентабельностью продукции y, %, от ее трудоемкости x, чел. на ед. продукции для 15 предприятий.

№	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Y	58,8	65,9	23,5	15,4	13,9	12,3	19,9	13,6	9,7	11,2	10,3	10,5	9,13	7,5	6,4
x	1	1,2	1,5	1,8	2	2,1	2,3	2,4	2,8	3	3,1	3,3	3,7	4,1	4,6

№9.

Имеются данные об индексах реального объема производства, реальных капитальных затрат и реальных затрат труда в промышленности некоторой страны за 20 лет. Постройте функцию Кобба – Дугласа.

Год	Y	K	L	Год	Y	K	L
1	100	100	100	11	169	197	143
2	110	109	109	12	175	210	147
3	109	119	108	13	180	217	149
4	115	120	116	14	191	225	153
5	126	135	120	15	195	238	155
6	133	139	122	16	192	247	151
7	142	152	127	17	199	268	153
8	152	166	135	18	244	299	176
9	154	178	139	19	268	337	189
10	161	189	138	20	282	370	195

Тема 5. Моделирование одномерных временных рядов.

1. Понятие временного ряда. Компоненты временного ряда.
2. Автокорреляция временного ряда и выявление его структуры.
3. Моделирование сезонных и циклических колебаний: метод скользящей средней.
4. Моделирование сезонных и циклических колебаний: применение фиктивных переменных.
5. Моделирование тенденции временного ряда и случайной компоненты.
6. Моделирование временного ряда при наличии структурных изменений.

Ключевые слова.

Временной ряд. Трендовая, циклическая и случайные компоненты. Аддитивная модель. Мультипликативная модель. Автокорреляция. Лаг. Автокорреляционная функция. Коррелограмма. Структура временного ряда. Аналитическое выравнивание временного ряда. Метод скользящей средней. Структурные изменения. Кусочно-линейная модель регрессии. Тест Чоу.

Основные теоретические аспекты темы:

Временной ряд – совокупность значений какого-либо показателя за несколько последовательных моментов или периодов времени.

Аддитивная модель – модель вида: Y=T+S+E,
где T - трендовая компонента;
S – циклическая компонента;
E – случайная компонента.

Мультипликативная модель – модель вида: Y=T*S*E.

Автокорреляция временного ряда – корреляционная зависимость между последовательными уровнями временного ряда.

Лаг – число периодов, по которым рассчитывается коэффициент автокорреляции.

Коэффициент автокорреляции уровней ряда первого порядка:

$$r_1 = \frac{\sum_{t=2}^{n}(y_t - \overline{y_1}) * (y_{t-1} - \overline{y_2})}{\sqrt{\sum_{t=2}^{n}(y_t - \overline{y_1})^2 * \sum_{t=2}^{n}(y_{t-1} - \overline{y_2})^2}}; \quad \overline{y_1} = \frac{\sum_{t=2}^{n} y_t}{n-1}; \quad \overline{y_2} = \frac{\sum_{t=2}^{n} y_{t-1}}{n-1}.$$

Автокорреляционная функция временного ряда – последовательность коэффициентов автокорреляции уровней временного ряда.

Коррелограмма – график зависимости значений автокорреляционной функции от величины лага.

Свойства коэффициентов автокорреляции:

1. Он строится по аналогии с линейным коэффициентом корреляции и, таким образом, характеризует тесноту только линейной связи текущего и предыдущего уровней ряда. Поэтому по коэффициенту автокорреляции можно судить о наличии линейной (или близкой к линейной) тенденции. Для временных рядов, имеющих сильную нелинейную тенденцию, коэффициент автокорреляции уровней исходного ряда может приближаться к нулю.
2. По знаку коэффициента автокорреляции нельзя делать вывод о возрастающей или убывающей тенденции в уровнях ряда.

Анализ структуры ряда:

- Если r_1 наиболее высокий, то исследуемый ряд содержит только тенденцию;
- Если r_t наиболее высокий, то ряд содержит циклические колебания с периодичностью в t моментов времени;
- Если ни один из коэффициентов автокорреляции не является значимым, то либо ряд не содержит тенденции и циклических колебаний, либо ряд содержит сильную нелинейную тенденцию.

Аналитическое выравнивание временного ряда – способ моделирования тенденции временного ряда: построение аналитической функции, характеризующей зависимость уровней ряда от времени, или тренда.

Способы определения типа тенденции:

- Качественный анализ
- Визуальный анализ графика
- Коэффициенты автокорреляции: если временной ряд содержит нелинейную тенденцию, например, в форме экспоненты, то коэффициент автокорреляции первого порядка по логарифмам уровней исходного ряда будет выше, чем соответствующий коэффициент, рассчитанный по уровням ряда. Чем сильнее выражена нелинейная тенденция в изучаемом временном ряде, тем в большей степени будут различаться значения указанных коэффициентов.
- Перебор основных форм тренда, расчет скорректированного R^2.

Функции, используемые для построения трендов:

Линейный тренд: $y_t=a+b*t$;

Нелинейные функции: $y_t=a+b/t$ - гипербола;

$y=a*t^b$ – степенная функция;

$y_t=a*b^t$ - экспоненциальная функция;

$y= a_0+a_1 t +a_2 t^2 +, . . . , +a_m t^m$ – параболы разных порядков.

Алгоритм построения аддитивной и мультипликативной модели: метод скользящей средней.

Шаг1. Выравнивание исходных уровней ряда методом скользящей средней:

1. Суммируем уровни ряда последовательно за каждый промежуток времени, в котором наблюдаются колебания со сдвигом на один момент времени и определяем условные величины показателя Y.
2. Делим полученные величины на число моментов времени в промежутке и находим скользящие средние.
3. Находим средние значения из двух последовательных скользящих

Шаг 2. Оценка сезонной компоненты:

1. Находим оценку сезонной компоненты, как разность между фактическими уровнями ряда и центрированными скользящими средними. В мультипликативной модели – как частное от деления фактических уровней ряда на центрированные скользящие средние.
2. Находим средние оценки сезонной компоненты за каждый промежуток времени, в котором наблюдаются колебания $\overline{S_i}$.
3. Исходя из условия взаимопогашения сезонных воздействий определяем корректирующий коэффициент k.
   - В аддитивной модели сумма значений сезонной компоненты по всем кварталам должна быть равна нулю, тогда $k = \frac{\sum S_i}{n}$.
   - В мультипликативной модели сумма значений сезонной компоненты по всем кварталам должна быть равна числу периодов в цикле, т.е. четырем в случае четырех кварталов: $k = \frac{n}{\sum S_i}$; где n – период колебаний.
4. Рассчитываем скорректированные значения сезонных компонент: в аддитивной модели: $S_i = \overline{S_i} - k$; в мультипликативной модели: $S_i = \overline{S_i} * k$;

Шаг 3. Элиминирование влияния сезонной компоненты:

- Находим значения T+E как Y-S – в аддитивной модели.
- Находим значения T*E как Y/S – в мультипликативной модели.

Шаг 4. Определение трендовой компоненты ряда.

1. Трендовая компонента ряда определяется с помощью построения регрессионной модели, параметры которой находятся методом наименьших квадратов.
2. С помощью уравнения регрессии находим теоретические уровни трендовой компоненты T для каждого момента времени t.

Шаг 5. Находим значения T+S в аддитивной модели или T*S в мультипликативной модели.

Шаг 6. Находим случайную компоненту E= Y-(T+S) в аддитивной модели и E= Y/(T*S) в мультипликативной модели

Шаг 7. Оценка качества модели.

1. Находим сумму квадратов случайной компоненты.

2. Находим отношение суммы квадратов случайной компоненты к общей сумме квадратов отклонений уровней ряда от его среднего значения: $\frac{\sum E^2}{\sum (y_t - \overline{y})^2} * 100\%$

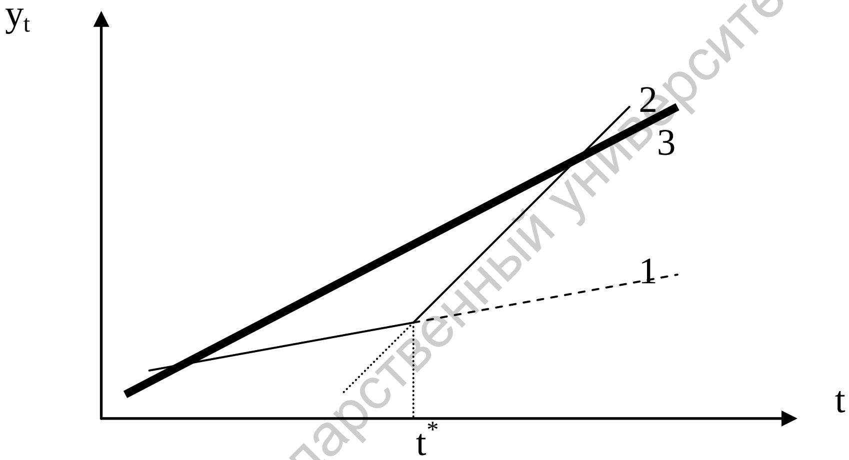
Применение фиктивных переменных в оценке сезонных компонент:
Строится модель множественной регрессии вида:
$y = a + bt + c_1 d_1 + c_2 d_2 + c_3 d_3 + \varepsilon$ **;**
a, b, c_1, c_2, c_3 - параметры модели**;**
d_1, d_2, d_3 - фиктивные переменные**;**

$$d_1 = \begin{cases} 1, для\_первого\_квартала \\ 0, в\_остальных\_случаях \end{cases};$$

$$d_2 = \begin{cases} 1, для\_второго\_квартала \\ 0, в\_остальных\_случаях \end{cases}$$

$$d_3 = \begin{cases} 1, для\_третьего\_квартала \\ 0, в\_остальных\_случаях \end{cases}.$$

Анализ значимости структурных изменений:



Если структурное изменение значимо, то для моделирования тенденции данного временного ряда следует использовать кусочно – линейные модели регрессии, то есть разделить исходную совокупность на две подсовокупности (до момента времени t^* и после момента времени t^*) и построить отдельно по каждой подсовокупности уравнения линейной регрессии.

Последствия:

«+» - снижение остаточной суммы квадратов по сравнению с единым для всей совокупности уравнением тренда;

«-» - потеря числа наблюдений и, следовательно, снижение числа степеней свободы в каждом уравнении кусочно – линейной модели.

Если структурные изменения незначительно повлияли на характер тенденции ряда, то ее можно описать с помощью единого для всей совокупности данных уравнения тренда

Последствия:

«+» - сохранение числа наблюдений исходной совокупности.

«-» - остаточная сумма квадратов будет выше по сравнению с кусочно – линейной моделью.

Выбор одной из двух моделей (кусочно – линейной или единого уравнения тренда) будет зависеть от соотношения между снижением остаточной дисперсии и потерей числа степеней свободы при переходе от единого уравнения регрессии к кусочно – линейной модели.

Тест Чоу:

Гипотеза H_0: о структурной стабильности тенденции изучаемого временного ряда.

Номер уравнения	Вид уравнения	Число наблюдений в совокупности	Остаточная сумма квадратов	Число параметров в уравнении	Число степеней свободы остаточной дисперсии
Кусочно – линейная модель					
1	$Y_1=a_1+b_1t$	n_1	RSS_1	k_1 (2)	n_1-k_1
2	$Y_2=a_2+b_2t$	n_2	RSS_2	k_2 (2)	n_2-k_2
Уравнение тренда по всей совокупности					
3	$Y_3=a_3+b_3t$	n	RSS_3	k_3(2)	$n-k_3$

Остаточная сумма квадратов по кусочно – линейной модели:

$RSS_{кл} = RSS_1 + RSS_2$

Соответствующее ей число степеней свободы составит:

$(n- k_1 - k_2) = (n_1-k_1) + (n_2-k_2)$.

Сокращение остаточной дисперсии при переходе от единого уравнения тренда к кусочно – линейной модели:

$\Delta RSS = RSS_3 - RSSкл$;

Число степеней свободы, соответствующее ΔRSS:

$(n-k_3) - (n- k_1 - k_2) = (k_1 + k_2 - k_3)$

F – критерий $= \dfrac{\Delta RSS/(k_1+k_2+k_3)}{RSSкл/(n-k_1-k_2)}$;

Найденное значение F – критерия сравнивают с табличным, полученным по таблицам распределения Фишера для определенного уровня значимости и числа степеней свободы $(k_1 + k_2 - k_3)$ и $(n- k_1 - k_2)$.

Если Fфакт > Fтабл, то гипотеза о структурной стабильности тенденции отклоняется, а влияние структурных изменений на динамику изучаемого показателя признают значимым. В этом случае моделирование тенденции временного ряда следует осуществлять с помощью кусочно – линейной модели.

Причины структурных изменений обуславливают различия в оценках параметров уравнений (1) и (2):

- Изменение численной оценки a_1 и a_2 при условии, что различия между b_1 и b_2 статистически незначимы. Геометрически это означает, что прямые параллельны. В данной ситуации можно говорить о скачкообразном изменении уровней ряда в момент времени t^* при неизменном среднем абсолютном приросте за период;
- Изменение численной оценки b_1 и b_2 при условии, что различия между a_1 и a_2 статистически незначимы. Геометрически это означает, что прямые пересекают ось ординат в одной точке. В данной ситуации изменение тенденции связано с изменением среднего абсолютного прироста временного ряда t^* .
- Изменение численной оценки a_1 и a_2 и b_1 и b_2 . Геометрически эта ситуация означает, что изменение характера тенденции сопровождается изменением как начального уровня ряда, так и среднего за период абсолютного прироста.

<u>Вопросы для обсуждения:</u>

1. Объясните, почему временной ряд представляет собой совокупность трендовой, циклической и случайной компоненты?
2. Какой вид связи между соседними уровнями ряда характеризует коэффициент автокорреляции?
3. В чем сходство и различие коэффициента корреляции в регрессионном анализе и коэффициента автокорреляции?
4. Объясните, что представляет собой структура временного ряда? Какой анализ позволяет ее определять?
5. Как регрессионный анализ применяется в моделировании одномерных временных рядов?
6. Какой критерий лежит при выборе построения аддитивной или мультипликативной модели временного ряда?
7. Назовите положительные и отрицательные моменты в построении кусочно-линейных и единого уравнения тренда при наличии структурных изменений в динамике переменных.
8. Каков критерий выбора построения модели временного ряда при наличии структурных изменений в динамике переменных?

<u>Пример построения аддитивной модели временного ряда с помощью пакета Excel и оценка ее значимости.</u>

Задание.

Собраны статистические данные потребления электроэнергии за 4 года.

1. Построить коррелограмму временного ряда.
2. Методом скользящей средней вычислить сезонную, трендовую и случайную компоненту аддитивной модели.

3. С помощью фиктивных переменных вычислить сезонную, трендовую и случайную компоненту аддитивной модели. (Приложение В3)

№ квартала t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Y_t	6	4,4	5	9	7,2	4,8	6	10	8	5,6	6,4	11	9	6,6	7	10,8

1. Коррелограмма временного ряда потребления электроэнергии:

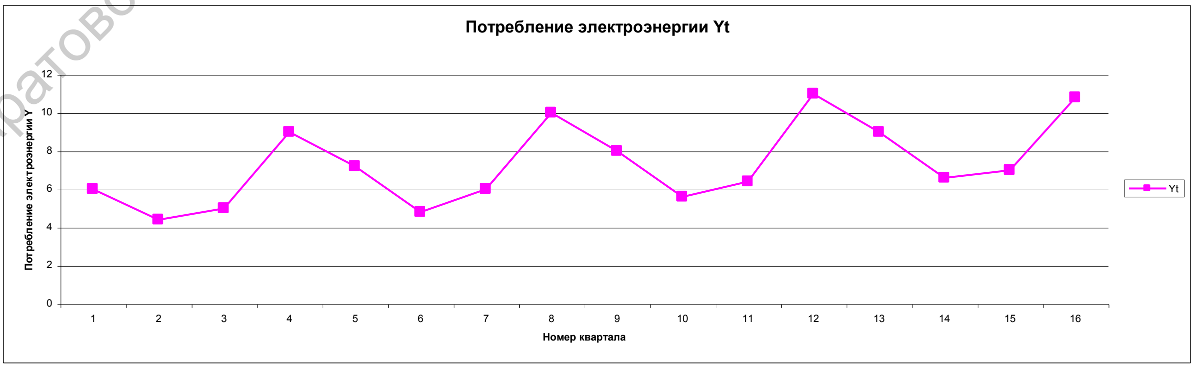
Лаг	Коэффициент автокорреляции уровней
1	0,165
2	0,5666
3	0,113
4	0,983
5	0,1187
6	0,722
7	0,003
8	0,97

2. Метод скользящей средней

Шаг 1. Выравнивание исходных уровней ряда методом скользящей средней

Таблица 1.

№ квартала	Yt	Итого за 4 квартала	Скользя-щая средняя за 4 квартала	Центриро-ванная скользя-щая средняя	Оценка сезонной компонен-ты
1	6				
2	4,4	24,4	6,1		
3	5	25,6	6,4	6,25	-1,25
4	9	26	6,5	6,45	2,55
5	7,2	27	6,75	6,625	0,575
6	4,8	28	7	6,875	-2,075
7	6	28,8	7,2	7,1	-1,1
8	10	29,6	7,4	7,3	2,7
9	8	30	7,5	7,45	0,55
10	5,6	31	7,75	7,625	-2,025
11	6,4	32	8	7,875	-1,475
12	11	33	8,25	8,125	2,875
13	9	33,6	8,4	8,325	0,675
14	6,6	33,4	8,35	8,375	-1,775
15	7	24,4			
16	10,8	17,8			



Шаг 2. Оценка сезонной компоненты:

Таблица 2.

Показатели	Год	№ квартала i			
		1	2	3	4
	1	-	-	-1,25	2,55
	2	0,575	-2,075	-1,1	2,7
	3	0,55	-2,025	-1,475	2,875
	4	0,675	-1,775	-	-
Итого за i квартал		1,8	-5,875	-3,825	8,125
Средняя оценка сезонной компоненты для i – го квартала		0,6	-1,958	-1,275	2,708
Скорректированная сезонная компонента Si		0,58125	-1,978	-1,294	2,69

Корректирующий коэффициент $k = \dfrac{0{,}6 - 1{,}958 - 1{,}275 + 2{,}71}{4} = 0{,}0185$

Шаг 3. Элиминирование влияния сезонной компоненты: (Y-S) (гр.4.)

Шаг 4. Определение трендовой компоненты ряда. (гр. 5.)

Шаг 6. Находим значения T+S (гр.6.)

Шаг 7. Находим случайную компоненту E= Y-(T+S) (гр. 7.)

Таблица 3.

№ t	Y_t	S_i	T+E=Y-S	T	T+S	E=Y-(T+S)	E^2
1	6	0,581	5,419	5,9019	6,4829	-0,4829	0,2332
2	4,4	-1,977	6,377	6,0883	4,1113	0,2887	0,0833
3	5	-1,294	6,294	6,2747	4,9807	0,0193	0,0004
4	9	2,69	6,31	6,4611	9,1511	-0,1511	0,0228
5	7,2	0,581	6,619	6,6476	7,2286	-0,0286	0,0008
6	4,8	-1,977	6,777	6,834	4,857	-0,057	0,0032
7	6	-1,294	7,294	7,0204	5,7264	0,2736	0,0749
8	10	2,69	7,31	7,2068	9,8968	0,1032	0,0107
9	8	0,581	7,419	7,3932	7,9742	0,0258	0,0007
10	5,6	-1,977	7,577	7,5796	5,6026	-0,0026	7E-06
11	6,4	-1,294	7,694	7,766	6,472	-0,072	0,0052
12	11	2,69	8,31	7,9524	10,642	0,3576	0,1279
13	9	0,581	8,419	8,1388	8,7198	0,2802	0,0785
14	6,6	-1,977	8,577	8,3252	6,3482	0,2518	0,0634
15	7	-1,294	8,294	8,5117	7,2177	-0,2176	0,0474
16	10,8	2,69	8,11	8,6981	11,388	-0,5881	0,3458
							1,098126

Нахождение модели Тренда T=a+b*t

0,186412	5,7155
0,015189	0,146868
0,914959	0,280067
150,6265	14
11,81478	1,098126

T=5,7155+0,18641*t

Сумма квадратов абсолютных ошибок = 1,0981

Шаг 8. Оценка качества модели.

Сумма квадратов абсолютных ошибок: $\Sigma E^2 = 1{,}0981$

Отношение суммы квадратов случайной компоненты к общей сумме квадратов отклонений уровней ряда от его среднего значения:

$$\frac{\sum E^2}{\sum (y_t - \bar{y})^2} * 100\% = 1{,}5\%$$

Вывод: Построенная аддитивная модель объясняет 98,5% общей вариации уровней временного ряда потребления электроэнергии за 16 кварталов исследуемых 4 – х лет и ее можно использовать в прогнозах будущего потребления электроэнергии**.**

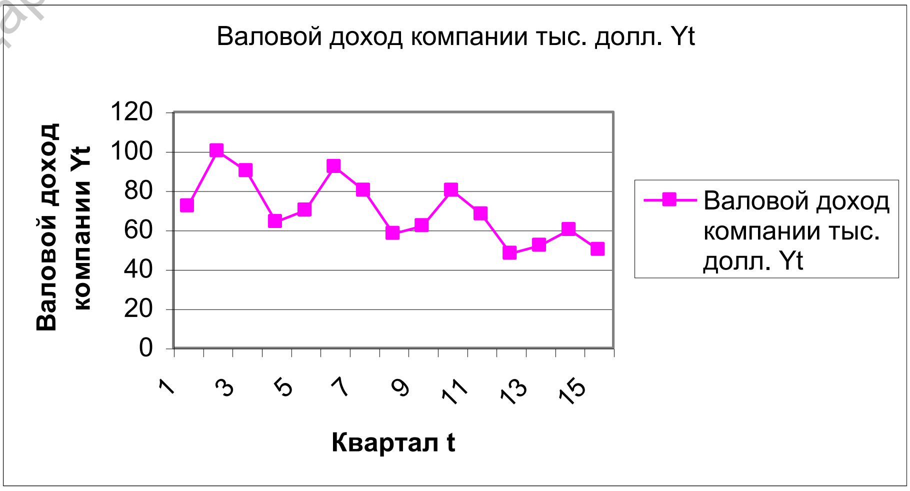
3. Вычисление сезонной и трендовой компонент с помощью фиктивных переменных. (Приложение)

Задания для самостоятельной работы**:**

№ 1 .

По статистическим данными постройте модель временного ряда. С помощью коэффициентов автокорреляции определите ее структуру и тип модели. Спрогнозируйте с помощью модели валовой доход в 5-м году работы компании. С помощью фиктивных переменных вычислить сезонную, трендовую и случайную компоненту аддитивной модели

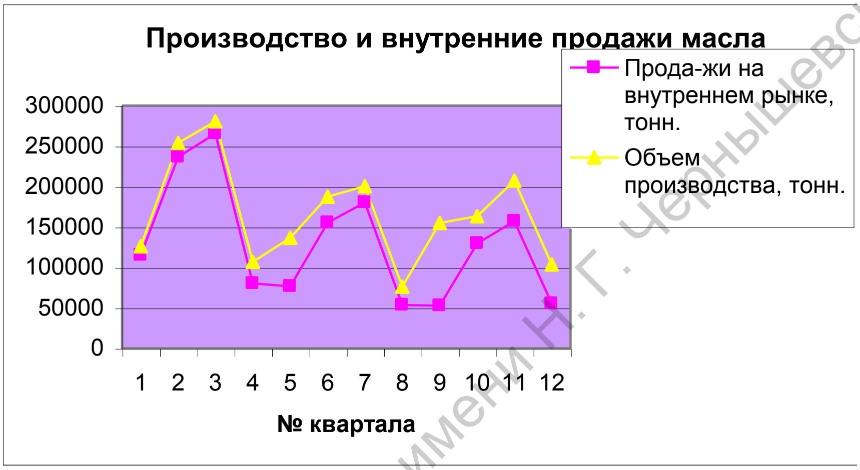
№ квартала t	Валовой доход компании тыс. долл. Yt
1	72
2	100
3	90
4	64
5	70
6	92
7	80
8	58
9	62
10	80
11	68
12	48
13	52
14	60
15	50
16	30



№2

По статистическим данным, описывающим поквартальное производство масла и объем его продаж на внутреннем рынке за 2 гола постройте модели временных рядов и спрогнозируйте по ним величины производства и объем продаж в следующие 3 года. С помощью фиктивных переменных вычислить сезонную, трендовую и случайную компоненту аддитивной модели

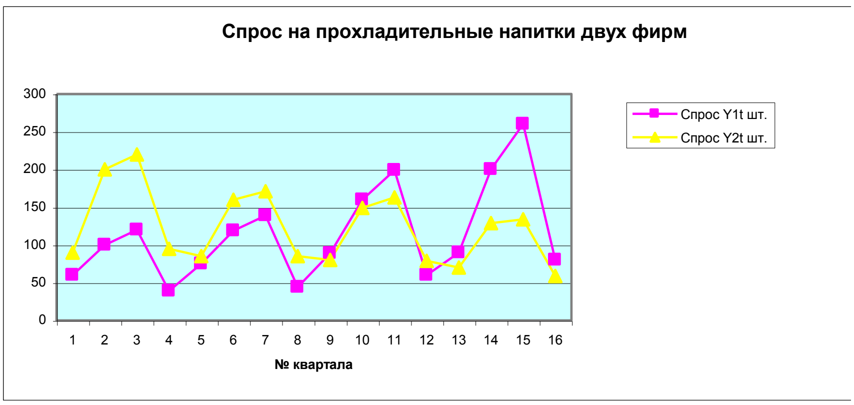
№ кварта-ла t	Прода-жи на внутреннем рынке, тонн.	Объем производства, тонн.
1	115007	126091
2	236426	253671
3	265466	280579
4	80249	106842
5	76681	136515
6	155130	187225
7	180000	200409
8	53461	76059
9	52845	154753
10	129759	163209
11	156879	207323
12	55145	103566



№3.

По статистическим данным, описывающим объем спроса на прохладительные напитки двух фирм в течение 4-х лет, постройте модели временных рядов, описывающих динамику спроса обеих фирм. С помощью фиктивных переменных вычислить сезонную, трендовую и случайную компоненту аддитивной модели. Спрогнозируйте квартал, когда одна из фирм покинет рынок. Каков будет объем спроса в этот момент у фирмы- конкурента?

№ кварта-ла t	Спрос Y1t шт.	Спрос Y2t шт.
1	60	90
2	100	200
3	120	220
4	39	95
5	75	85
6	119	160
7	139	171
8	44	85
9	89	80
10	160	149
11	199	163
12	60	79
13	90	70
14	200	129
15	260	134
16	80	59



№4.

По данными постройте модель временного ряда продаж компании, млн. долл. Определите ее структуру. Спрогнозируйте с помощью модели продажи в 2007 - м году.

Год	Квартал	Продажи	Год	Квартал	Продажи
1999	1	2292	2003	1	2643
	2	2450		2	2811
	3	2363		3	2679
	4	2477		4	2736
2000	1	2063	2004	1	2692
	2	2358		2	2871
	3	2316		3	2900
	4	2366		4	2811
2001	1	2268	2005	1	2497
	2	2533		2	2792
	3	2479		3	2838
	4	2625		4	2780
2002	1	2616	2006	1	2778
	2	2793		2	3066
	3	2656		3	3213
	4	2746		4	2928

№5

Даны поквартальные данные о прибыли компании за последние четыре года, тыс. долл. Постройте мультипликативную модель временного ряда. Спрогнозируйте прибыль компании во втором квартале шестого года работы компании.

	Год 1				Год 2				Год 3				Год 4			
Год	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Прибыль	72	70	62	52	100	92	80	60	90	80	68	50	64	58	48	30

№6.

Даны данные о квартальных продажах фирмы, тыс. долл. Постройте мультипликативную модель временного ряда. Спрогнозируйте объем продаж в 1996 г.

Год	Квартал	Продажи	Год	Квартал	Продажи
1990	1	232,7	1993	1	178,3
	2	309,2		2	274,5
	3	310,7		3	295,4
	4	293,0		4	286,4
1991	1	205	1994	1	190,8
	2	234,4		2	263,5
	3	285,4		3	318,8
	4	258,7		4	305,3
1992	1	193,2	1995	1	242,6
	2	263,7		2	318,8
	3	292,5		3	329,6
	4	315,2		4	338,2

Тема 6. Системы эконометрических уравнений.

1. Понятие системы эконометрических уравнений.
2. Структурная и приведенная форма модели.
3. Идентификация. Необходимое и достаточное условие идентификации.
4. Оценивание параметров структурной модели.

Ключевые слова.

Система независимых уравнений. Система рекурсивных уравнений. Система взаимозависимых уравнений. Структурная форма модели. Структурные коэффициенты модели. Структурные коэффициенты модели. Приведенная форма модели. Идентификация. Идентифицируемые, неидентифицируемые и сверхидентифицируемые структурные модели. Косвенный метод наименьших квадратов. Двухшаговый метод наименьших квадратов.

Основные теоретические аспекты темы:

Система независимых уравнений – система, в которой каждая зависимая переменная y рассматривается как функция одного и того же набора факторов x то есть система вида:

$Y_1=a_{11}x_1 + a_{12}x_2 +\ldots+ a_{1m}x_m +\varepsilon_1;$
$Y_2=a_{21}x_1 + a_{22}x_2 +\ldots+ a_{2m}x_m +\varepsilon_2;$
………………………………..
$Y_n=a_{n1}x_1 + a_{n2}x_2 +\ldots+ a_{nm}x_m +\varepsilon_n.$

Система рекурсивных уравнений – система, в которой зависимая переменная одного уравнения выступает в виде фактора x в другом уравнении, то есть система вида:

$Y_1=a_{11}x_1 + a_{12}x_2 +\ldots+ a_{1m}x_m +\varepsilon_1;$
$Y_2= b_{21}y_1 +a_{21}x_1 + a_{22}x_2 +\ldots+ a_{2m}x_m +\varepsilon_2;$
$Y_3= b_{31}y_1 + b_{32}y_2+a_{31}x_1 + a_{32}x_2 +\ldots+ a_{3m}x_m +\varepsilon_2;$
………………………………………………………………..
$Y_n= b_{n1}y_1 + b_{n2}y_2 +\ldots+ b_{nn-1}y_{n-1} + a_{n1}x_1 + a_{n2}x_2 +\ldots+ a_{nm}x_m +\varepsilon_n.$

Система взаимозависимых уравнений (система совместных одновременных уравнений) – система в которой одни и те же зависимые переменные в одних уравнениях входят в левую часть, а в других уравнениях – в правую, то есть система вида:

$Y_1= b_{12}y_2 + b_{13}y_3 +\ldots+ b_{1n}y_n + a_{11}x_1 + a_{12}x_2 +\ldots+ a_{1m}x_m +\varepsilon_1;$
$Y_2= b_{21}y_1 +b_{23}y_3 +\ldots+ b_{2n}y_n + a_{21}x_1 + a_{22}x_2 +\ldots+ a_{2m}x_m +\varepsilon_2;$
……………………………………………………………….
$Y_n= b_{n1}y_1 + b_{n2}y_2 +\ldots+ b_{nn-1}y_{n-1} + a_{n1}x_1 + a_{n2}x_2 +\ldots+ a_{nm}x_m +\varepsilon_n.$

Приведенная форма модели – система линейных функций эндогенных переменных от экзогенных:

$Y_1=\delta_{11}x_1 +\delta_{12}x_2+\ldots+ \delta_{1m}x_m$;
$Y_2=\delta_{21}x_1 +\delta_{22}x_2+\ldots+ \delta_{2m}x_m$;
………………………………..
$Y_n=\delta_{n1}x_1 + \delta_{n2}x_2+\ldots+ \delta_{nm}x_m$.

Где δ_{ij} – коэффициенты приведенной формы модели.

Идентификация – это единственность соответствия между приведенной и структурной формами модели.

Модель идентифицируема – если все ее структурные коэффициенты определяются однозначно, единственным образом по коэффициентам приведенной формы модели, то есть если число параметров структурной формы модели равно числу параметров приведенной формы модели.

Модель неидентифицируема – если число структурных коэффициентов больше числа приведенных коэффициентов и следовательно, структурные коэффициенты не могут быть оценены через коэффициенты приведенной формы модели

Модель сверхидентифицируема – если число структурных коэффициентов меньше числа приведенных коэффициентов и следовательно, на основе приведенных коэффициентов можно получить два или более значений одного структурного коэффициента.

Необходимое условие идентифицируемости модели:

Чтобы уравнение было идентифицируемо, необходимо, чтобы число экзогенных переменных (D), отсутствующих в данном уравнении, но присутствующих в системе, было равно числу эндогенных переменных в данном уравнении (H) без одного.
D+1=H – уравнение идентифицируемо;
D+1<H – уравнение неидентифицируемо;
D+1>H – уравнение сверхидентифицируемо.

Достаточное условие идентифицируемости модели:

Уравнение идентифицируемо, если по отсутствующим в нем экзогенным и эндогенным переменным можно из коэффициентов при них в других уравнениях системы получить матрицу, определитель которой не равен нулю, а ранг матрицы не меньше, чем число эндогенных переменных в системе без одного.

Алгоритм косвенного метода наименьших квадратов:

- Структурная модель преобразовывается в приведенную форму модели.
- Для каждого уравнения приведенной формы модели обычным МНК оцениваются приведенные коэффициенты.
- Коэффициенты приведенной формы модели трансформируются в параметры структурной формы модели.

Алгоритм двухшагового метода наименьших квадратов:

- Определяется приведенная форма модели, и находятся на ее основе оценки теоретических значений эндогенных переменных.
- Определяются структурные коэффициенты модели по данным теоретических (расчетных) значений эндогенных переменных.

Вопросы для обсуждения:

1. Объясните, почему построение систем эконометрических уравнений важно в экономических исследованиях?
2. В чем сходство и различие моделей эконометрических уравнений с простыми моделями множественной регрессий?
3. Приведите примеры экономических процессов и явлений, которые могут быть описаны системами независимых, рекурсивных и взаимозависимых уравнений.
4. Почему необходимо преобразовывать структурную форму модели в приведенную?
5. В каком случае вся модель является идентифицируемой и сверхидентифицируемой?

Задания для самостоятельной работы:

№ 1.

Проверьте, идентифицируема ли эконометрическая модель**:**

$Y_1 = b_{12}y_2 + b_{13}y_3 + a_{11}x_1 + a_{12}x_2$;

$Y_2 = b_{21}y_1 + a_{21}x_1 + a_{22}x_2 + a_{23}x_3$;

$Y_3 = b_{31}y_1 + b_{32}y_2 + a_{31}x_1 + a_{33}x_3 + a_{34}x_4$.

№ 2.

Проверьте, идентифицируема ли эконометрическая модель**:**

$Y_1 = b_{12}y_2 + b_{13}y_3 + a_{11}x_1 + a_{12}x_2$;

$Y_2 = b_{21}y_1 + a_{22}x_2 + a_{23}x_3$;

$Y_3 = b_{31}y_1 + b_{32}y_2 + a_{31}x_1 + a_{33}x_3 + a_{34}x_4$.

№3.

Проверьте, каждое уравнение системы на необходимое и достаточное условие идентификации.

$Y_1 = b_{12}y_2 + b_{13}y_3 + a_{11}x_1 + a_{12}x_2$;

$Y_2 = b_{21}y_1 + a_{22}x_2 + a_{23}x_3 + a_{24}x_4$;
$Y_3 = b_{31}y_1 + b_{32}y_2 + a_{31}x_1 + a_{32}x_2$.

№4.

Постройте, используя статистику в таблице, эконометрическую модель косвенным методом наименьших квадратов**:**
$Y_1 = b_{12}y_2 + a_{11}x_1 + \varepsilon_1$;
$Y_2 = b_{21}y_1 + a_{22}x_2 + +\varepsilon_2.$

№ региона	Y_1	Y_2	X_1	X_2
1	2	5	1	3
2	3	6	2	1
3	4	7	3	2
4	5	8	2	5
5	6	5	4	6

№5.

Постройте, используя статистику в таблице, эконометрическую модель двухшаговым методом наименьших квадратов**:**
$Y_1 = b_{12}(y_2 + x_1) + \varepsilon_1$;
$Y_2 = b_{21}y_1 + a_{22}x_2 + +\varepsilon_2.$

№ региона	Y_1	Y_2	X_1	X_2
1	2	5	1	3
2	3	6	2	1
3	4	7	3	2
4	5	8	2	5
5	6	5	4	6

ТЕСТ ДЛЯ ОЦЕНКИ ОСТАТОЧНЫХ ЗНАНИЙ

1. Эконометрическая модель - это модель:
а) гипотетического экономического объекта;
б) конкретно-существующего экономического объекта, построенная на гипотетических данных;
в) конкретно-существующего экономического объекта, построенная на статистических данных.

2. Модель, отражающая положительную зависимость предложения денег от ставки процента, является:
а) мезомоделью;
б) макромоделью;
в) микромоделью.

3. Предопределенные переменные включают:
а) все экзогенные и эндогенные переменные;
б) только экзогенные переменные;
в) все экзогенные переменные и лаговые эндогенные переменные;
г) лаговые экзогенные и эндогенные переменные.

4. Чем точнее информация об исследуемом объекте, тем:
а) больше доля "черного ящика";
б) меньше доля "черного ящика";
в) качество информации не влияет на долю "черного ящика" в моделировании.

5. Степени свободы в наборе данных определяют число единиц данных:
а) независимых друг от друга, которые могут нести отдельные элементы информации;
б) зависимых друг от друга, которые могут нести отдельные элементы информации;
в) независимых друг от друга, которые могут нести общие элементы информации.

6. Оценочные значения характеристик рассчитываются по данным:
а) генеральной совокупности;
б) выборки;
в) как по выборочным данным, так и по данным генеральной совокупности.

7. Чем больше σ, тем выборка:
а) лучше;
б) хуже.

8. Стохастические модели характеризуются:

а) наличием случайной составляющей;
б) отсутствием случайной составляющей;
в) случайная составляющая не играет роли в определении типа модели.

9. Причинно – следственная связь между двумя переменными в экономике называется:
а) регрессионной;
б) корреляционной.

10. Коэффициенты регрессии, найденные методом наименьших квадратов являются:
а) истинными коэффициентами регрессии;
б) оценками истинных коэффициентов регрессии;
в) не являются ни теми ни другими.

11. Если модель парной регрессии описывает зависимость спроса товара от цены, то коэффициент а:
а) имеет экономический смысл;
б) не имеет экономического смысла;
в) определить невозможно.

12. Если коэффициент корреляции равен 0, 24, то наблюдается:
а) положительная сильная линейная связь;
б) отрицательная слабая линейная связь;
в) положительная слабая линейная связь.

13. Если коэффициент корреляции равен –9, 77, то наблюдается:
а) положительная сильная линейная связь;
б) отрицательная слабая линейная связь;
в) положительная слабая линейная связь;
г) отрицательная сильная линейная связь.

14. В эксперименте Монте – Карло:
а) неизвестны истинные значения коэффициентов регрессии;
б) известны заранее истинные коэффициенты регрессии;
в) истинные коэффициенты регрессии находятся методом наименьших квадратов.

15. В модели y=15+6x коэффициент 6 показывает:
а) среднее изменение фактора y при изменении x на единицу;
б) общее изменение фактора y при изменении x на единицу;
в) коэффициент 6 не имеет экономического смысла.

16. Одним из условий Гаусса – Маркова является условие:
а) зависимости случайных компонент для каждого наблюдения;

б) независимость случайных компонент для каждого наблюдения;
в) неравенство матожидания случайной компоненты нулю.

17.В соответствии теореме Гаусса – Маркова, коэффициенты регрессии, построенные МНК, являются:
а) смещенными оценками;
б) несмещенными оценками;
в) не являются оценками.

18.Оценивание каждого параметра в выборке поглощает:
а) три степени свободы в выборке;
б) две степени свободы;
в) одну степень свободы.

19.Ошибка первого рода имеет место когда:
а) отвергается ложная нулевая гипотеза;
б) принимается ложная нулевая гипотеза;
в) отвергается истинная нулевая гипотеза.

20.Доверительные интервалы в регрессионном анализе строятся для:
а) истинных значений параметров модели;
б) оценок истинных значений параметров модели;
в) для статистических данных y и x.

21.Если коэффициент детерминации равен 0,65, то модель описывает:
а) 65% вариации признака y;
б) 35% вариации признака y;
в) по значению коэффициента детерминации невозможно ответить на данный вопрос.

22. Факторная сумма квадратов отклонения вычисляется по формуле:
а) $\sum_i (y_i - \bar{y})^2 = \sum_i y^2 - n * \bar{y}^2$;
б) $\sum_i (y_x - \bar{y})^2 = b^2 * (\sum_i x^2 - n * \bar{x}^2)$.
в) $\sum_{i=1}^{n} (y_i - y_x)^2$.

22.Ошибка аппроксимации для хорошего подбора данных должна быть:
а) больше 7%;
б) меньше 7%;
в) в пределах 5-7%.

23.Интервальная оценка прогнозного значения принимает наименьшее значение:

а) в средней точке выборки;
б) в дальних точках от среднего значения;
в) в ближних точках к среднему значению.

24.Мультиколлинеарность факторов в модели множественной регрессии:
а) не всегда ведет к несостоятельным оценкам;
б) всегда ведет к несостоятельным оценкам;
в) не позволяет применять МНК вообще.

25.Две переменные коллинеарны, если:
а) $r_{xixj} \geq 0,7$;
б) $r_{xixj} < 0,7$;

26.Если определитель матрицы межфактороной корреляции равен 1, то это свидетельствует о:
а) полной мультиколлинеарности факторов;
б) полном отсутствии мультиколлинеарности факторов;
в) о высокой мультиколлинеарности факторов.

26.К методам преодоления межфакторной корреляции относится:
а) увеличение дисперсии случайного параметра;
б) уменьшение числа наблюдений;
в) увеличение дисперсии объясняющих переменных.

27.Особенностью коэффициентов «чистой» регрессии является:
а) сравнимость между собой;
б) несравнимость между собой;
в) отсутствие экономической интерпретации.

28.Особенностью коэффициентов регрессии стандартизованного уравнения является:
а) сравнимость между собой;
б) несравнимость между собой;
в) отсутствие экономической интерпретации.

29. Скорректированный индекс множественной регрессии:
а) завышает обычный индекс множественной регрессии;
б) занижает обычный индекс множественной регрессии;
в) равен обычному индексу множественной регрессии.

29.Фиктивные переменные отражают в модели:
а) количественные показатели;
б) качественные показатели;
в) как те, так и другие.

30.МНК применим к моделям:
а) линейным;
б) нелинейным, внутренне нелинейным;
в) нелинейным, внутренне линейным.

31.МНК применим к моделям:
а) нелинейным, внутренне нелинейным;
б) нелинейным, внутренне линейным
в) нелинейным, внутренне линейным, сведенным к линейному виду.

32.Показатель степени в степенной функции является:
а) показателем чистой регрессии;
б) показателем эластичности;
в) показателем постоянного роста.

33. Основание в показательной функции является:
а) показателем чистой регрессии;
б) показателем эластичности;
в) показателем постоянного роста.

34.Коэффициенты детерминации для линейного и логарифмического уравнения:
а) не сравнимы между собой;
б) сравнимы между собой;
в) не возникает необходимости их сравнивать.

35.Если сумма показателей степени в производственной функции Кобба – Дугласа равны 1, то наблюдается:
а) положительный эффект масштаба;
б) отрицательный эффект масштаба;
в) постоянный эффект масштаба.

36.Для построения степенной модели МНК необходимо статистические данные:
а) потенциировать;
б) логарифмировать;
в) не изменять, использовать в единицах измерения.

37.В случае неравномерной амплитуды колебаний необходимо строить:
а) аддитивную модель временного ряда;
б) мультипликативную модель временного ряда;
в) не имеет значения.

38.Высокий коэффициент автокорреляции первого порядка свидетельствует о наличии:

а) сезонности;
б) линейной связи;
в) нелинейной связи.

39.В аддитивной модели:
а) сумма значений сезонной компоненты по всем кварталам должна быть равна нулю;
б) сумма значений сезонной компоненты по всем кварталам должна быть равна числу периодов в цикле.

40.Для отражения сезонности для четырех времен года необходимо использовать:
а) две фиктивных переменных;
б) три фиктивных переменных;
в) четыре фиктивных переменных.

41.Положительным моментом построения кусочно – линейной модели является:
а) снижение остаточной суммы квадратов по сравнению с единым для всей совокупности уравнением тренда;
б) потеря числа наблюдений и, следовательно, снижение числа степеней свободы в каждом уравнении кусочно – линейной модели;
в) сохранение числа наблюдений исходной совокупности.
г) остаточная сумма квадратов будет выше по сравнению с единой моделью.

42.В мультипликативной модели оценка сезонной компоненты, находится как:
а) как разность между фактическими уровнями ряда и центрированными скользящими средними.
б) как частное от деления фактических уровней ряда на центрированные скользящие средние.

43.В случае структурных изменений изменение численной оценки a_1 и a_2 при условии, что различия между b_1 и b_2 статистически незначимы геометрически означает:
а) что прямые параллельны;
б) что прямые пересекают ось ординат в одной точке
в) что изменение характера тенденции сопровождается изменением как начального уровня ряда, так и среднего за период абсолютного прироста.

ОБЩАЯ ЛИТЕРАТУРА ПО КУРСУ:

1. Айвазян С.А. Мхитарян В.С. Прикладная статистика и основы эконометрики. Москва, 1998.
2. Аистов А.В., Максимов А.Г. Эконометрика шаг за шагом. – М.: Из – во ГУ ВШЭ, 2006.
3. Валентинов В.А. Эконометрика: Учебник – М.: Издательско – торговая корпорация «Дашков и К», 2006.
4. Гладилин А.В., Герасимов А.Н.. Громов Е.И. Эконометрика. Учебное пособие. – М.: Конкурс, 2006.
5. Доугерти Кристофер Введение в эконометрику. Перевод с английского. Москва, Экономический факультет МГУ, 2001.
6. Доугерти Кристофер Введение в эконометрику. Учебник. 2-е издание. / Перевод с англ. –М.: ИНФРА - М, 2007. – 432 с.
7. Кулинич Е.И. Эконометрия. Москва, «Финансы и статистика», 1999.
8. Катышев П.К., Магнус Я.Р., Пересецкий А.А. Сборник задач к начальному курсу эконометрики. Москва, «Дело», 2002.
9. Колемаев В.А. Эконометрика: Учебник. – М.: ИНФРА – М, 2007. – 160 с.
10. Магнус Я.Р., Катышев П.К., Пересецкий А.А. Эконометрика. Начальный курс. Москва, «Дело», 2000.
11. Мардас А.Н. Эконометрика. Краткий курс. Санкт – Петербург, «Питер», 2001.
12. Минько А.А. Прогнозирование в бизнесе с помощью Excel. Просто как дважды два. – М.: Эксмо, 2007. – 208 с.
13. Приходько А.И. Практикум по эконометрике: егрессионный анализ средствами Excel.- Ростов н/Д.: Феникс, 2007. – 256 с.
14. Ханк Джон Э, Дин У. Уичерн, Артур Дж. Райтс Бизнес – прогнозирование, 7 – е издание: Пер. с англ. – М.: Издательский дом «Вильямс», 2003. – 656 с.
15. Эконометрика. Под ред. Член. – корр. И.И. Елисеевой. Москва, «Финансы и статистика», 2001.
16. Эконометрика: Учебник. / И.И. Елисеева, С.В. Курышева, Т.В. Костеева и др.; Под ред. И.И. Елисеевой. – 2 –е изд., перераб. и доп. - М.: «Финансы и статистика», 2006. – 576 с.
17. Эконометрика: Учебник / Н.П. Тихомиров. Е.Ю. Дорохина – 2 –е изд. – М.: Из – во «Экзамен», 2007.