

# **ОСНОВЫ ЭКОНОМЕТРИКИ**

РОССИЙСКИЙ ГОСУДАРСТВЕННЫЙ ОТКРЫТЫЙ  
ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ ПУТЕЙ СООБЩЕНИЯ

Кафедра экономической теории

Учебно-методическое пособие по курсу

ОСНОВЫ ЭКОНОМЕТРИКИ

Саратов 2003

**УДК**  
**ББК**

Эконометрика./ Сост. Леванова Л.Н. – Саратов., 2003

Учебно-методическое пособие подготовлено в соответствии с положениями и требованиями Государственного образовательного стандарта высшего профессионального образования, включает основные вопросы лекций и семинарских занятий, краткие теоретические положения эконометрики, основные формулы для решения задач и построения эконометрических моделей, вопросы и задачи для практических занятий.

Для студентов и преподавателей экономических специальностей.

Составитель

Кандидат экономических наук, доцент Л.Н. Леванова

Рекомендуют к печати:

Кафедра экономической теории Российского государственного открытого  
технического университета путей сообщения

Кандидат экономических наук, доцент В.А. Узбеков

**ОБЩАЯ ЛИТЕРАТУРА ПО КУРСУ:**

1. Айвазян С.А. Мхитарян В.С. Прикладная статистика и основы эконометрики. Москва, 1998.
2. Доугерти Кристофер Введение в эконометрику. Перевод с английского. Москва, Экономический факультет МГУ, 2001.
3. Кулинич Е.И. Эконометрия. Москва, «Финансы и статистика», 1999.
4. Катышев П.К., Магнус Я.Р., Пересецкий А.А. Сборник задач к начальному курсу эконометрики. Москва, «Дело», 2002.
5. Магнус Я.Р., Катышев П.К., Пересецкий А.А. Эконометрика. Начальный курс. Москва, «Дело», 2000.
6. Мардас А.Н. Эконометрика. Краткий курс. Санкт – Петербург, «Питер», 2001.
7. Эконометрика. Под ред. Член. – корр. И.И. Елисейевой. Москва, «Финансы и статистика», 2001.

## Тема 1. Введение в эконометрику.

1. Предмет и методы эконометрики.
2. Модель. и моделирование. Экономические и эконометрические модели: свойства и классификация. Этапы моделирования.
3. Переменные в моделях и их типы.

### Ключевые слова:

Модель. Моделирование. Адекватность модели. Математическая модель. Экономические и эконометрические модели. Модели прогноза и имитации. Модели микроэкономики, мезоэкономики и макроэкономики. Теоретико-аналитические и прикладные модели. Дискриптивные и нормативные модели. Детерминистские и стохастические модели. Статические и динамические модели. Экзогенные переменные. Эндогенные переменные. Предопределенные переменные. Лаговые эндогенные переменные.

### 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.

**Эконометрическая модель** – вероятно – статистическая модель, описывающая механизм функционирования экономической или социально – экономической системы.

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

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

**Эндогенные переменные 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.

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, какая из моделей: детерминистская или стохастическая является более адекватной?
4. Как Вы считаете, если результаты (эндогенные переменные, выходные параметры) модели явно неверные, в чем может быть причина неудачного моделирования?
5. Как Вы считаете, каковы минусы агрегирования при макроэкономическом моделировании?
6. В чем заключается специфичность определения точности измерений социально-экономических явлений?
7. Как Вы считаете, почему изучаемая дисциплина появилась в российском образовании сравнительно недавно?

Тест:

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.

## Тема 2. Модель парной регрессии.

1. Уравнение и вид функции парной регрессии.
2. Оценка параметров уравнения линейной регрессии. Метод наименьших квадратов. Экономический смысл параметров.
3. Оценка адекватности модели и существенности параметров линейной регрессии.
4. Нелинейная регрессия. Применение модели в экономике. Тесты Бокса-Кокса.

### Ключевые слова:

Спецификация модели. Результативный признак, признак-фактор и стохастическая переменная в модели. Параметры регрессии. Метод наименьших квадратов. Стандартное отклонение случайной величины. Коэффициент вариации случайной величины. Коэффициент корреляции. Коэффициент детерминации. Общая сумма квадратов отклонений. Факторная сумма квадратов отклонений. Остаточная сумма квадратов отклонений. Число степеней свободы. Дисперсия на одну степень свободы. F- критерий Фишера. Стандартные ошибки коэффициентов уравнения регрессии. t – критерий Стьюдента. Доверительные интервалы. Интервалы прогноза. Средняя ошибка аппроксимации.

### 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 (результативный признак),

$x$  – независимая, объясняющая, переменная (признак- фактор),

$\varepsilon$  - возмущение или стохастическая переменная, включающая влияние неучтенных в модели факторов.

### **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;

$y=a+blg x$ -логарифмическая функция;

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

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

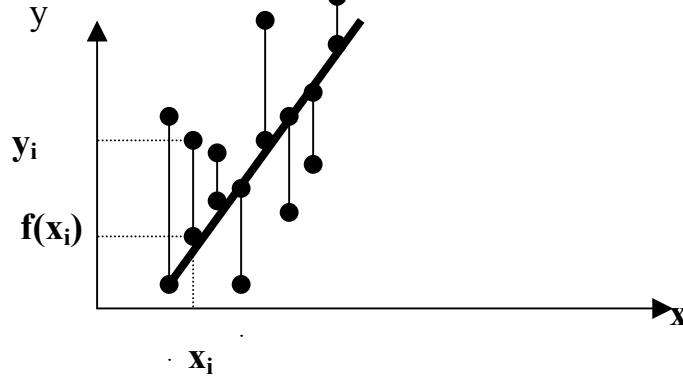
$y=1/(a+bx+cx^2)$ ;

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

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

$$Q = \sum_i e_i^2 = \sum_i (y_i - f(x_i))^2 \rightarrow \min$$

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



### Расчет параметров уравнения линейной регрессии $y=a+bx+\varepsilon$ :

Пусть собраны  $n$  статистических значений независимой, объясняющей переменной  $x$ :  $x_1, \dots, x_n$  и  $n$  статистических значений зависимой переменной  $y$ :  $y_1, \dots, y_n$ . Тогда, используя метод наименьших квадратов, параметры уравнения линейной регрессии можно вычислить следующим образом:

$$b = \frac{\overline{yx} - \bar{y} * \bar{x}}{x^2 - \bar{x}^2}; \quad a = \bar{y} - b * \bar{x}.$$

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.

**Стандартное отклонение случайной величины  $x$  ( $\sigma_x$ )** – мера разброса случайной величины вокруг среднего значения.

$$\sigma_x = \sqrt{x^2 - \bar{x}^2}.$$

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

$$V_x = \frac{\sigma_x}{\bar{x}};$$



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

$$r_{xy} = \frac{\text{cov}(x, y)}{\sigma_x \sigma_y} = \frac{\overline{yx} - \bar{y} * \bar{x}}{\sigma_x \sigma_y}. \quad r_{xy} \in [-1; 1].$$

Если:  $r_{xy} = -1$ , то наблюдается строгая отрицательная связь;  
 $r_{xy} = 1$ , то наблюдается строгая положительная связь;  
 $r_{xy} = 0$ , то линейная связь отсутствует.

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

**Общая сумма квадратов отклонений ( $S_{\text{общ}}$ ):**

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

**Факторная сумма квадратов отклонений ( $S_{\text{факт}}$ ):**

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

**Остаточная сумма квадратов отклонений ( $S_{\text{ост}}$ ):**

$$S_{\text{ост}} = S_{\text{общ}} - S_{\text{факт}}.$$

**Дисперсия на одну степень свободы:**

$$D_{\text{общ}} = S_{\text{общ}} / (n-1); \quad D_{\text{факт}} = S_{\text{факт}} / 1; \quad D_{\text{ост}} = S_{\text{ост}} / (n-2).$$

**F – критерий Фишера:**

$$F_{\text{выч}} = D_{\text{факт}} / D_{\text{ост}}.$$

Если  $F_{\text{выч}} > F_{\text{таб}}$ , то уравнение регрессии значимо.

**Стандартные ошибки коэффициентов регрессии:**

$$m_b = \sqrt{\frac{\sum_{i=1}^n (y_i - y_x)^2 / (n-2)}{\sum_i (x_i - \bar{x})^2}}; \quad m_a = \sqrt{\frac{\sum_{i=1}^n (y_i - y_x)^2 * \sum_i x^2}{(n-2) * n * \sum_i (x_i - \bar{x})^2}};$$

**Фактическое значение t-критерия Стьюдента:**

$$t_b = b/m_b; \quad t_a = a/m_a.$$

**Доверительные интервалы :**  $b+t* m_b$  и  $b-t* m_b$ ;  $a+t* m_a$  и  $a-t* m_a$ .

**Корреляция для нелинейной регрессии:**  $R = \sqrt{1 - \frac{\sum_{i=1}^n (y_i - y_x)^2}{\sum_i (y_i - \bar{y})^2}}$ ;  $R \in [0; 1]$

Чем ближе  $R$  к 1, тем теснее связь рассматриваемых признаков.

### **F-критерий Фишера для нелинейной регрессии:**

$$F = \frac{R^2}{1-R^2} * \frac{n-m-1}{m}; \text{ где: } n - \text{ число наблюдений, а } m - \text{ число}$$

параметров при переменных  $x$ .

### **Средняя ошибка аппроксимации:**

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

### Вопросы для обсуждения:

1. Объясните, чем вызвано появление в модели парной регрессии стохастической переменной  $\varepsilon$  ?
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?

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

#### **Задание**

По статистическим данным, описывающим зависимость удельного веса бракованной продукции от удельного веса рабочих со специальной подготовкой на предприятиях построить уравнение парной регрессии с помощью программы Excel и определить его значимость.

№	Удельный вес рабочих со специальной подготовкой, %, x	Удельный вес бракованной продукции, %, y	yx	x <sup>2</sup>	y <sup>2</sup>	y <sub>x</sub>
1	15	18	270	225	324	15,81
2	25	12	300	625	144	13,41
3	35	10	350	1225	100	11,01
4	45	8	360	2025	64	8,61
5	55	6	330	3025	36	6,21
6	65	5	325	4225	25	3,81
7	70	3	210	4900	9	2,61

Расчет параметров уравнения с помощью встроенной функции ЛИНЕЙН:

-0,238244	19,40793201	Параметр a = 19,41
0,027796	1,339264901	
0,936275	1,395764536	Параметр b = -0,24
73,46237	5	
143,1163	9,740793201	Уравнение парной регрессии y=19,41-0,24*x.

Стандартная ошибка параметра a = 1,34

t- критерий Стьюдента  $t_a = 19,41/1,34=14,5$

Табличное значение t - критерия Стьюдента  $t_{табл} = 2,57$

Стандартная ошибка параметра b = 0,028

t - критерий Стьюдента  $t_b = -0,24 / 0,028=8,5$

**Вывод: параметры регрессии значимы.**

Коэффициент корреляции  $r_{xy}=0,9676156$

Коэффициент детерминации  $r_{xy}^2=0,93628$

Факторная сумма квадратов S факт = 143,116

Остаточная сумма квадратов S ост = 9,74079

Общая сумма квадратов S общ = 152,857

Факторная дисперсия на одну степень свободы D факт = 143,116

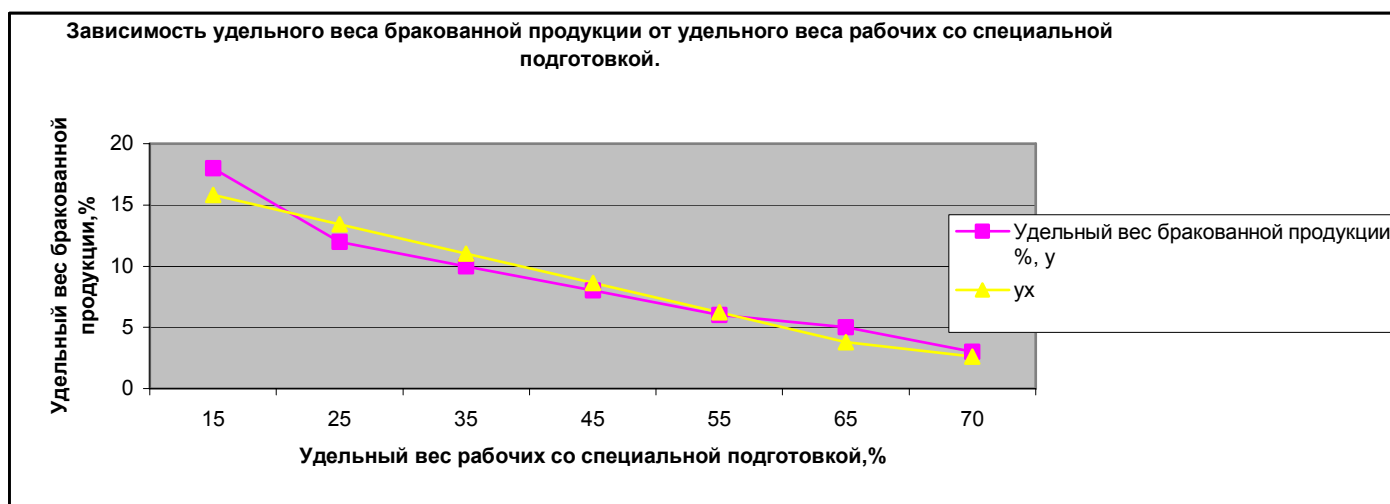
Остаточная дисперсия на одну степень свободы D ост = 1,948

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 адекватна и ее можно использовать для прогнозов.**



**Задание 1.**

По статистическим данным, описывающим зависимость уровня рентабельности на предприятии от скорости товарооборота построить уравнение парной регрессии с помощью программы Excel и определить его значимость.

№	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, %	0,78	0,38	0,21	0,51	0,95	1,05	0,83	0,98

**Задание 2**

По статистическим данным, описывающим зависимость индекса Лернера от рыночной доли фирмы построить уравнение парной регрессии с помощью программы Excel и определить его значимость.

№	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

**Задание 3**

По статистическим данным, описывающим зависимость уровня рентабельности на предприятии от удельного веса продовольственных товаров в товарообороте построить уравнение парной регрессии с помощью программы Excel и определить его значимость.

№	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

**Задание 4**

По статистическим данным, описывающим зависимость объема спроса на товар от его цены построить уравнение парной регрессии с помощью программы Excel и определить его значимость.

№	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

**Задание 5.**

В таблице приведены значения индекса реализации к 1992 г. в неизменных ценах в промышленности в целом и индекс избыточной занятости

к 1992 г. Постройте уравнение регрессии с помощью программы Excel и определите его значимость.

Год	1992	1993	1994	1995	1996
Индекс реализации к 1992 г.	1	0,87	0,68	0,59	0,5
Индекс избыточной занятости	1	1,023	1,206	1,240	1,310

### Задание 6.

В таблице приведены значения выручки от экспорта 1 тонны синтетического каучука за 10 кварталов и цены его на внутреннем рынке. Постройте уравнение регрессии с помощью программы Excel и определите его значимость. Спрогнозируйте значение экспорта каучука при цене 3000 долл. за тонну.

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

### Задание 7.

По статистическим данным, описывающим зависимость значения рентабельности производства синтетического каучука от индекса Лернера построить уравнение парной регрессии с помощью программы Excel и определить его значимость. Спрогнозировать значение рентабельности в 2004 г., если ожидается, что индексе Лернера составит 0,4.

Год	1991	1992	1993	1994	1995	1996	1997
Индекс Лернера L	0,14	0,33	0,21	0,14	0,22	0,25	0,28
Рентабельность, %	15,8	49	26,2	15,7	27,4	30	35

### Задание 8.

В таблице представлены расходы на агрегированное потребление Y и агрегированный располагаемый доход X в некоторой национальной экономике в течение 12 лет – с 1986 по 1987 г. Существует ли линейная зависимость данных показателей? Рассчитайте модель парной регрессии и оцените ее значимость.

Год	t	$Y_t$	$X_t$
1986	1	152	170
1987	2	159	179
1988	3	162	187
1989	4	165	189
1990	5	170	193
1991	6	172	199
1992	7	177	200
1993	8	179	207
1994	9	184	215
1995	10	186	216
1996	11	190	220
1997	12	191	225

### Задание 9.

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  $\omega_t$  – 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	$\omega_t$	$u_t$
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

### Тема 3. Модель множественной регрессии.

1. Уравнение и вид функции множественной регрессии.
2. Отбор факторов при построении модели множественной регрессии.
3. Оценка параметров уравнения множественной линейной регрессии. Метод наименьших квадратов.
4. Частные уравнения регрессии и частные коэффициенты эластичности.
5. Оценка адекватности модели и существенности параметров модели множественной регрессии.
6. Фиктивные переменные в модели множественной регрессии.

#### Ключевые слова:

Спецификация модели. Результативный признак, признак-факторы и стохастическая переменная в модели. Параметры регрессии. Интеркорреляция факторов модели. Матрица показателей корреляции. Мультиколлинеарность факторов. Определитель матрицы межфакторной корреляции. Метод наименьших квадратов. Частные уравнения регрессии. Частные коэффициенты эластичности. Индекс множественной корреляции. Скорректированный индекс множественной корреляции. Коэффициент частной корреляции Коэффициент частной корреляции. F- критерий Фишера модели множественной регрессии. Частный F – критерий Фишера. t – критерий Стьюдента. Последовательный и частный F – критерий. Средняя квадратическая ошибка коэффициента регрессии. Фиктивные переменные.

#### Основные теоретические аспекты темы.

**Множественная регрессия** – регрессия между переменными  $y$  и  $x_1, x_2, \dots, x_m$  то есть модель вида:  $y=f(x_1, x_2, \dots, x_m)+\varepsilon$ , где:

$y$  - зависимая переменная (результативный признак),

$x_1, x_2, \dots, x_m$  – независимые, объясняющие, переменные ( признак-факторы),

$\varepsilon$  - возмущение или стохастическая переменная, включающая влияние неучтенных в модели факторов.

#### **Основные типы функций, используемые при количественной оценке связей:**

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

Параметры  $a_1, a_2, a_m$  называются **коэффициентами «чистой» регрессии** и характеризуют среднее изменение результата с изменением соответствующего фактора на единицу при неизменном значении других факторов, закрепленных на среднем уровне.

Нелинейные функции:  $y=a x_1^{b_1} x_2^{b_2} \dots, x_m^{b_m}$  –степенная функция;

$b_1, b_2, b_m$  - коэффициенты эластичности; показывают на сколько % изменится в среднем результат при изменении соответствующего фактора на 1% и при неизменности действия других факторов.

$$y = 1/(a_0 + a_1 x_1 + a_2 x_2 + \dots + a_m x_m) - \text{гипербола};$$

$$y = e^{a_0 + a_1 x_1 + a_2 x_2 + \dots + a_m x_m} - \text{экспонента}.$$

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

**Определитель матрицы межфакторной корреляции** (на примере модели  $y = a_0 + a_1 x_1 + a_2 x_2 + a_3 x_3$ )

$$\text{Det}|R| = \begin{vmatrix} r_{x_1 x_1} & r_{x_2 x_1} & r_{x_3 x_1} \\ r_{x_1 x_2} & r_{x_2 x_2} & r_{x_3 x_2} \\ r_{x_1 x_3} & r_{x_2 x_3} & r_{x_3 x_3} \end{vmatrix}.$$

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.

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

$$Q = \sum_i e_i^2 = \sum_i (y_i - f(x_{ij}))^2 \rightarrow \min.$$

где:  $y_i$  – статистические значения зависимой переменной;

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

**Расчет параметров уравнения множественной линейной регрессии**

$y = a_0 + a_1 x_1 + a_2 x_2 + \dots + a_m x_m + \varepsilon$  с помощью метода наименьших квадратов:

Пусть собраны  $n$  статистических значений каждой из независимых, объясняющих переменных:  $x_{1j}, \dots, x_{nj}$  и  $n$  статистических значений зависимой переменной  $y$ :  $y_1, \dots, y_n$ . Тогда, можно сформировать матрицу  $X$  и вектор столбец  $Y$ :

$$X = \begin{pmatrix} 1 & x_{11} & x_{12} & x_{1m} \\ 1 & x_{21} & x_{22} & x_{2m} \\ \cdot & \cdot & \cdot & \cdot \\ 1 & x_{n1} & x_{n2} & x_{nm} \end{pmatrix}; Y = \begin{pmatrix} y_1 \\ y_2 \\ \cdot \\ y_n \end{pmatrix}.$$

Нахождение параметров  $a_0, a_1, \dots, a_m$  производится по следующей формуле:

$$A = (a_0, a_1, \dots, a_m) = (X^T X)^{-1} X^T Y.$$

**Частным уравнением регрессии модели  $y = a_0 + a_1 x_1 + a_2 x_2 + \dots + a_m x_m + \varepsilon$**  называется уравнение вида  $y_{x_i} = f(x_1, \dots, x_{(i-1)}, x_{(i+1)}, \dots, x_m) = f(\overline{x_1}, \dots, \overline{x_{i-1}}, x_i, \overline{x_{i+1}}, \dots, \overline{x_m})$ , то есть



уравнения регрессии, которые связывают результативный признак с соответствующими факторами  $x$  при закреплении других учитываемых в уравнении регрессии факторов на среднем уровне.

**Индекс множественной корреляции:**  $R = \sqrt{1 - \frac{\sum_{i=1}^n (y_i - y_x)^2}{\sum_i (y_i - y)^2}}$ ;  $R \in [0; 1]$

Чем ближе  $R$  к 1, тем теснее связь рассматриваемых признаков.

**Скорректированный индекс множественной корреляции:**

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

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

$$r_{yx_1x_2\dots x_{(i-1)}x_{(i+1)}\dots x_m} = \sqrt{1 - \frac{1 - R_{yx_1x_2\dots x_i\dots x_m}^2}{1 - R_{yx_1x_2\dots x_{(i-1)}x_{(i+1)}\dots x_m}^2}};$$

где:  $R_{yx_1x_2\dots x_i\dots x_m}^2$  – множественный коэффициент детерминации всего комплекса  $m$  факторов с результатом;

$R_{yx_1x_2\dots x_{(i-1)}x_{(i+1)}\dots x_m}^2$  – тот же показатель детерминации, но без введения в модель фактора  $x_i$ .

**F-критерий Фишера модели множественной регрессии:**

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

где:  $n$  – число наблюдений, а  $m$  – число параметров при переменных  $x$  (в линейной регрессии совпадает с числом включенных в модель факторов).

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

$$F_{xi} = \frac{R_{yx_1\dots x_i\dots x_m}^2 - R_{yx_1\dots x_{(i-1)}x_{(i+1)}\dots x_m}^2}{1 - R_{yx_1\dots x_i\dots x_m}^2} * \frac{n - m - 1}{1};$$

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

**t-критерий Стьюдента для коэффициента регрессии при  $i$  – м факторе:**

$$t_{bi} = \sqrt{F_{xi}} = \frac{b_i}{m_{bi}};$$

### Средняя квадратическая ошибка коэффициента регрессии:

$$m_{bi} = \frac{\sigma_y \sqrt{1 - R_{yx1...xm}^2}}{\sigma_{xi} \sqrt{1 - R_{xix1...xm}^2}} * \frac{1}{\sqrt{n - m - 1}};$$

где :  $\sigma_y$  – среднее квадратическое отклонение для признака  $y$ ;  
 $\sigma_x$  – среднее квадратическое отклонение для признака  $x_i$ ;  
 $R_{yx1...xm}^2$  – коэффициент детерминации для уравнения множественной регрессии;  
 $R_{xix1...xm}^2$  – коэффициент детерминации для зависимости фактора  $x_i$  со всеми другими факторами уравнения множественной регрессии;  
 $n-m-1$  – число степеней свободы для остаточной суммы квадратов отклонений.

#### 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. В чем Вы видите специфику спецификации модели множественной регрессии?
4. Каким требованиям должны отвечать факторы модели множественной регрессии и почему?
5. Как должны соотноситься коэффициенты детерминации для  $m$  и  $m+1$  факторов модели?
6. Объясните практическое применение в экономике частных коэффициентов эластичности.
7. В чем заключается смысл расчета скорректированного индекса корреляции и какова связь его с индексом корреляции при различных количествах вводимых в модель факторов?

### **Пример построения модели множественной регрессии и оценка ее значимости.**

#### **Задание**

Модель Ланкастера. По статистическим данным, описывающим зависимость цены от потребительских свойств станка ( $P$  – основной размер станка;  $N$  – мощность главного привода;  $n$  – максимальная частота вращения шпинделя;  $УА$  – уровень автоматизации;  $T$  – класс точности), построить модель множественной регрессии с помощью программы Excel и определить ее значимость. Определите влияние дифференциации продукции на цену.

№	Р, мм	N, кВт	п, об/мин	УА	Точность	Цена станка Y	Y <sub>x</sub>	Y-Y <sub>x</sub>	(Y-Y <sub>x</sub> ) <sup>2</sup>	Y-Y <sub>ср</sub>	(Y-Y <sub>ср</sub> ) <sup>2</sup>
1	400	11	2000	3	1	53,6	53,41	0,19	0,0361	19,375	375,3906
2	400	10	1500	1	1	43,6	41,012	2,588	6,69774	9,375	87,89063
3	400	8	2000	1	1	35	36,784	-1,784	3,18266	0,775	0,600625
4	400	8	2500	1	1,6	39	38	1	1	4,775	22,80063
5	320	8	2500	1	1	27	28,84	-1,84	3,3856	-7,225	52,20063
6	320	8	3000	3	1,6	44	41,54	2,46	6,0516	9,775	95,55063
7	320	6,3	2000	1	1	26	26,7262	-0,7262	0,52737	-8,225	67,65063
8	250	8	3000	3	1,6	32,7	35,289	-2,589	6,70292	-1,525	2,325625
9	250	6,3	3000	1	1,6	22,5	20,8912	1,6088	2,58824	-11,725	137,4756
10	250	5,5	2500	1	1	20	18,304	1,696	2,87642	-14,225	202,3506
11	320	8	2500	1	1,6	29,8	30,856	-1,056	1,11514	-4,425	19,58063
12	400	8	2000	1	1,6	37,5	38,8	-1,3	1,69	3,275	10,72563
									35,8538		1074,543

Матрица X =

1	400	11	2000	3	1
1	400	10	1500	1	1
1	400	8	2000	1	1
1	400	8	2500	1	1,6
1	320	8	2500	1	1
1	320	8	3000	3	1,6
1	320	6,3	2000	1	1
1	250	8	3000	3	1,6
1	250	6,3	3000	1	1,6
1	250	5,5	2500	1	1
1	320	8	2500	1	1,6
1	400	8	2000	1	1,6

**Вектор A =  $(X^T * X)^{-1} * X^T Y = (-18, 55; 0,0893; 1,714; -0,0016; 5,742; 3,36)$**

Y среднее = 34,23

Остаточная сумма квадратов S остат = 35,854

Общая сумма квадратов S общ = 1074,5

Факторная сумма квадратов S факт = 1074,5-35,854=1038,646

Индекс множественной корреляции R= 0,983

Индекс множественной детерминации R<sup>2</sup> = 0,966

F - критерий Фишера вычисляемый F<sub>выч</sub> = 34,1

Табличное значение F - критерия Фишера при уровне значимости 5% F<sub>табл</sub> = 4,39

Табличное значение F - критерия Фишера при уровне значимости 1% F<sub>табл</sub> = 8,75

**Вывд: F выч > F табл при обоих уровнях значимости, следовательно модель множественной регрессии адекватна, и ее можно использовать для суждения о стоимости аналогичных станков в исследованном диапазоне факторов.**

**Задание 1.**

По статистическим данным, описывающим зависимость уровня рентабельности торговой деятельности от удельного веса продовольственных товаров и оплаты труда, построить модель множественной регрессии с помощью программы Excel и определить ее значимость.

№ 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

**Задание 2.**

По статистическим данным, описывающим зависимость уровня рентабельности торговой деятельности от среднемесячного товарооборота в расчете на душу населения, удельного веса продовольственных товаров в товарообороте, времени обращения товаров, среднемесячной оплаты труда и трудоемкости товарооборота (численности работников на 100000 ед. товарооборота) построить модель множественной регрессии с помощью программы Excel и определить ее значимость.

№	Факторы					Уровень рентабельности, %
	Среднемесячный товарооборот в расчете на душу населения	Удельный вес продовольственных товаров в товарообороте, %	Время обращения товаров, дней	Среднемесячная оплата труда	Трудоемкость 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

**Задание 3.**

По статистическим данным, описывающим зависимость производительности труда по плодоконсервным заводам области за год от удельного веса рабочих с технической подготовкой и удельного веса механизированных работ, построить модель множественной регрессии с помощью программы Excel и определить ее значимость.

№ 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

#### Задание 4.

По статистическим данным, описывающим зависимость накопления пяти случайно выбранных семей от дохода и размера, построить модель множественной регрессии с помощью программы Excel и определить ее значимость. 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

#### Задание 5.

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:  $q_t = a_0 + a_1 p_t + a_2 c_t + a_3 c_t^2 + \varepsilon_t$ . В таблице приведены данные наблюдений за 18 недель. Построить модель множественной регрессии. Оцените ее значимость. Найдите оптимальную цену при расходах на рекламу = 280 руб. Найдите оптимальный уровень расходов на рекламу при цене чебурека = 6 руб. Помогите менеджеру найти оптимальное решение (максимизирующее чистый доход).

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
$c_t$	4,8	3,6	5,5	2,7	5,7	1,3	5,8	3,4	3,7	3,6	5,1	3,3	4,7	3,8	4,3	2,7	2,2	3,1

**Задание 6.**

В кейнсианской теории спрос на деньги зависит от доходов и процентных ставок. Рассмотрим модель:  $m_t = a_0 + a_1 y_t + a_2 i_t + \varepsilon_t$ ; где  $m_t$  – агрегат денежной массы М1 (млрд. долл.),  $y_t$  – валовой национальный продукт (млрд. долл.);  $i_t$  – процентные ставки по государственным облигациям (%).

Рассчитайте модель множественной регрессии и оцените спрос на деньги при 1) ВВП = 1000 и  $i = 10$ ; 2) ВВП = 2500 и  $i = 5$ . Оцените значимость модели.

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

**Задание 7.**

В таблице приведены реальный доход на душу населения  $y$  (тыс. долл.), процент рабочей силы, занятой в сельском хозяйстве –  $x_1$ , и средний уровень образования населения в возрасте после 25 лет  $x_2$  (число лет, проведенных в учебных заведениях) для 15 развитых стран с 1983 г. Постройте модель множественной линейной регрессии и оцените ее значимость.

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

**Задание 8.**

В таблице приведена зависимость выпуска  $Q$  от трудозатрат  $L$  и капиталовложений  $K$ . Оцените по этим данным производственную функцию Кобба-Дугласа  $Q = \alpha L^{\beta_1} K^{\beta_2}$ . Оцените значимость модели. Рассчитайте объем выпуска при  $L = 2500$  и  $K = 1800$ .

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

## Тема 4. Моделирование одномерных временных рядов.

1. Понятие временного ряда. Компоненты временного ряда.
2. Автокорреляция временного ряда и выявление его структуры.
3. Моделирование тенденции временного ряда, сезонных, циклических колебаний и случайной компоненты.
4. 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. 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:**

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

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

**Коррелограмма** – график зависимости значений автокорреляционной функции от величины лага.

**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.

**Функции, используемые для построения трендов:**

Линейный тренд:  $y_t = a + b \cdot t$ ;

Нелинейные функции:

$y_t = a + b/t$  - гипербола;

$y_t = a \cdot t^b$  – степенная функция;

$y_t = a \cdot b^t$  - экспоненциальная функция;

$y_t = a_0 + a_1 t + a_2 t^2 + \dots + a_m t^m$  – параболы разных порядков.

**Алгоритм построения аддитивной или мультипликативной модели.****Шаг 1. 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 средних – центрированные скользящие средние.

**Шаг 2. Оценка сезонной компоненты:**

1. Находим оценку сезонной компоненты, как разность между фактическими уровнями ряда и центрированными скользящими средними.
2. Находим средние оценки сезонной компоненты за каждый промежуток времени, в котором наблюдаются колебания  $\bar{S}_i$ .
3. Исходя из условия взаимопогашения сезонных воздействий определяем корректирующий коэффициент  $k$ : в аддитивной модели  $k = \frac{\sum S_i}{n}$ ; в мультипликативной модели -  $k = \frac{n}{\sum S_i}$ ; где  $n$  – период колебаний.
4. Рассчитываем скорректированные значения сезонных компонент: в аддитивной модели:  $S_i = \bar{S}_i - k$ ; в мультипликативной модели:  $S_i = \bar{S}_i \cdot k$ ;



Шаг 3. Элиминирование влияния сезонной компоненты:

Находим значения  $T+E$  как  $Y-S$  – в аддитивной модели.  $T \cdot E$  как  $Y/S$

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 уровни трендовой компоненты  $T$  для каждого момента времени  $t$ .

Шаг 6. Находим значения  $T+S$  или  $T \cdot S$ .Шаг 7. Находим случайную компоненту  $E = Y - (T+S)$  или  $E = Y / (T \cdot S)$ Шаг 8. 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:

$$\text{значения: } \frac{\sum E^2}{\sum (y_t - \bar{y})^2} * 100\%$$

Вопросы для обсуждения:

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?

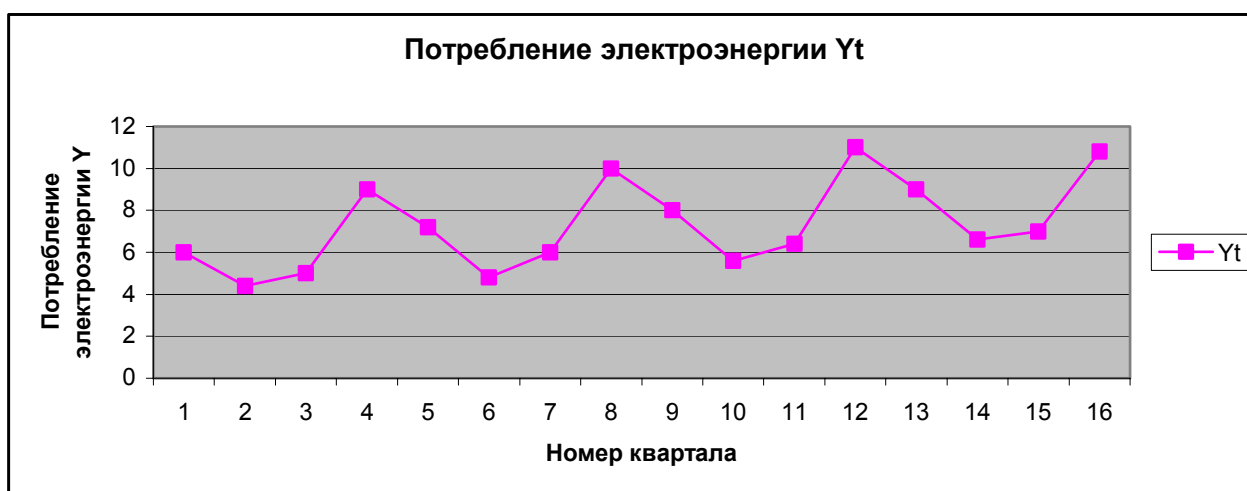
**Пример построения аддитивной модели временного ряда с помощью пакета Excel.****Задание.**

Построить аддитивную модель временного ряда, описывающего потребление продукции естественной монополии (электроэнергии) за 4 года:

№ 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

Шаг1. Выравнивание исходных уровней ряда методом скользящей средней  
Таблица 1.

№ квартала	$Y_t$	Итого за 4 квартала	Скользящая средняя за 4 квартала	Центрированная скользящая средняя	Оценка сезонной компоненты
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
	<b>1</b>	-	-	-1,25	2,55
	<b>2</b>	0,575	-2,075	-1,1	2,7
	<b>3</b>	0,55	-2,025	-1,475	2,875
	<b>4</b>	0,675	-1,775	-	-
<b>Итого за i квартал</b>		1,8	-5,875	-3,825	8,125
<b>Средняя оценка сезонной компоненты для i – го квартала</b>		0,6	-1,958	-1,275	2,708
<b>Скорректированная сезонная компонента Si</b>		0,58125	-1,978	-1,294	2,69

Корректирующий коэффициент  $k = \frac{0,6 - 1,958 - 1,275 + 2,71}{4} = 0,185$

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 <sub>t</sub>	S <sub>i</sub>	T+E=Y-S	T	T+S	E=Y-(T+S)	E <sup>2</sup>
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

Нахождение модели Тренда T=a+b\*t

1,098126

0,186412	5,7155	T=5,7155+0,18641*t
0,015189	0,146868	
0,914959	0,280067	
150,6265	14	
11,81478	1,098126	

Сумма квадратов абсолютных ошибок = 1,0981

**Шаг 8. Оценка качества модели.**

Сумма квадратов абсолютных ошибок:  $\sum E^2 = 1,0981$

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

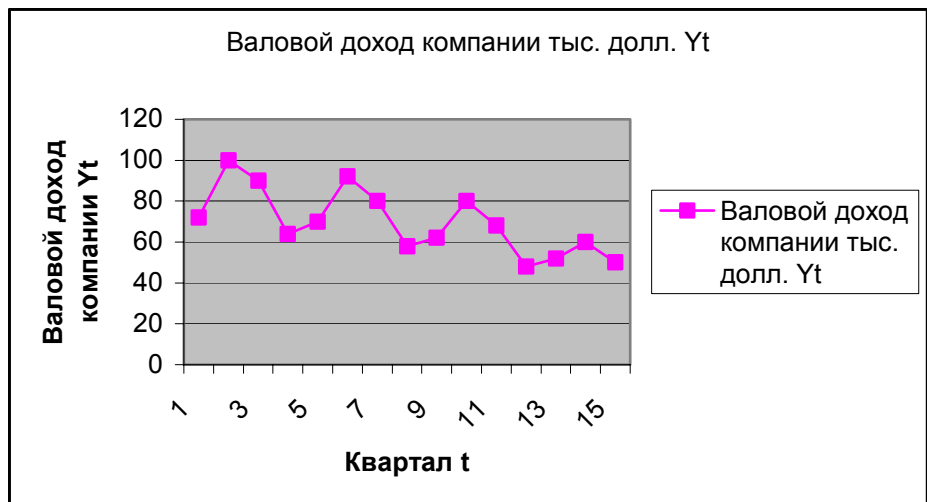
$$\frac{\sum E^2}{\sum (y_t - \bar{y})^2} * 100\% = 1,5\%$$

**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.

**Задание 1.**

По статистическим данным описывающим поквартальные данные валового дохода компании за последние 4 года 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.

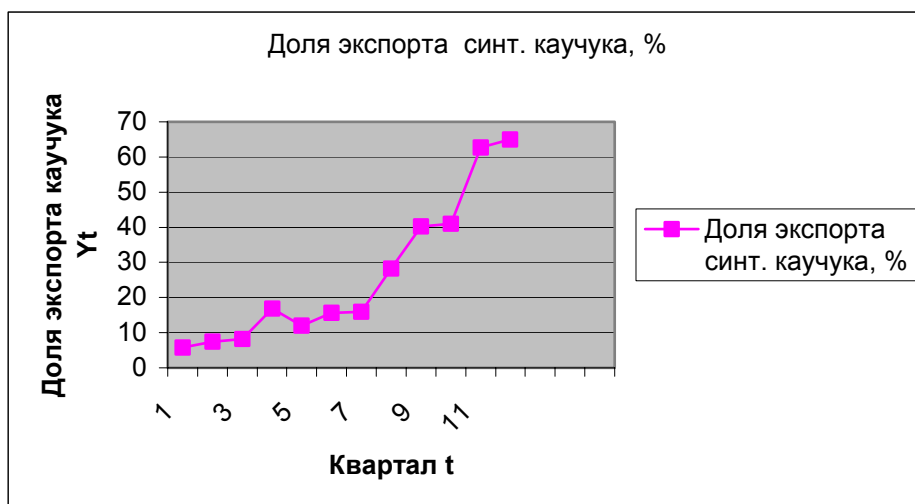
№ квартал a t	Валовой доход компании тыс. долл. Yt
1	72
2	100
3	90
4	64
5	70
6	92
7	80
8	58
9	62
10	80
11	68
12	48
13	52
14	60
15	50



## Задание 2

По статистическим данным описывающим поквартальные данные доли экспорта в отечественном производстве синтетического каучука (1992-1996 гг.). постройте модель временного ряда. Можно ли, используя построенную модель, спрогнозировать долю экспорта синтетического каучука в следующих годах.

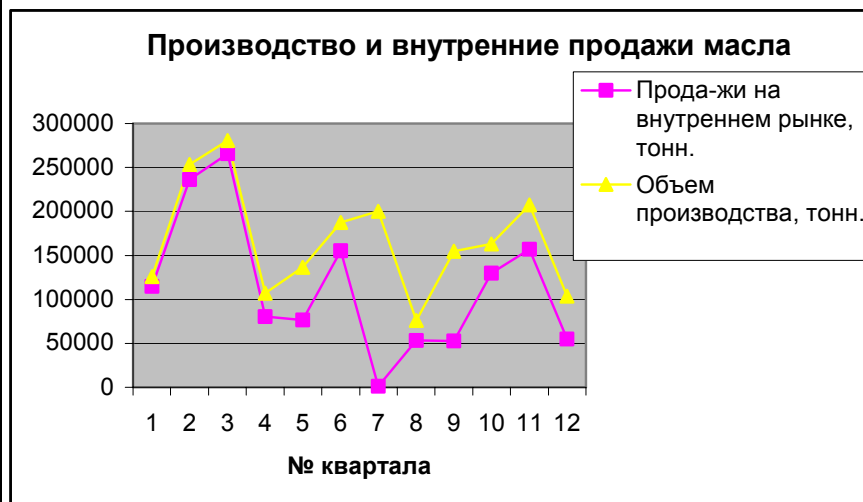
№ квартала $t$	Доля экспорта синт. каучука, %
1	5,7
2	7,4
3	8,2
4	16,9
5	12
6	15,6
7	16
8	28,3
9	40,3
10	41
11	62,7
12	65



## Задание 3

По статистическим данным, описывающим поквартальное производство масла и объем его продаж на внутреннем рынке за 2 года постройте модели временных рядов и спрогнозируйте по ним величины производства и объем продаж в следующие 3 года.

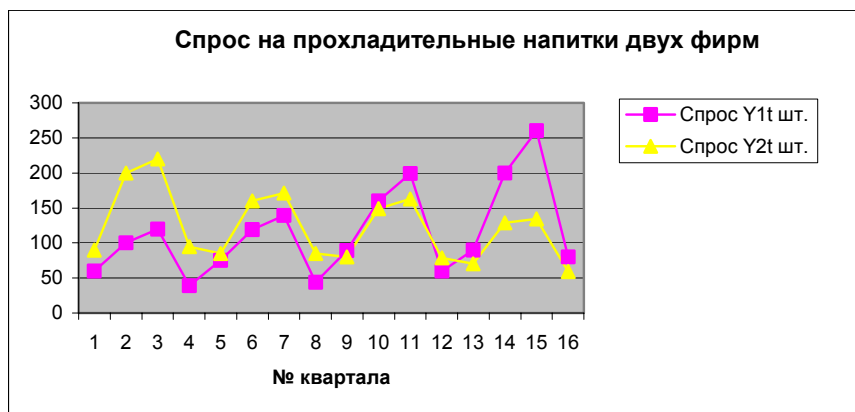
№ квартала $t$	Продажи на внутреннем рынке, тонн.	Объем производства, тонн.
1	115007	126091
2	236426	253671
3	265466	280579
4	80249	106842
5	76681	136515
6	155130	187225
7	1109	200409
8	53461	76059
9	52845	154753
10	129759	163209
11	156879	207323
12	55145	103566



### Задание 4

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. Спрогнозируйте квартал, когда одна из фирм покинет рынок. Каков будет объем спроса в этот момент у фирмы- конкурента?

№ квартала 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



## Тема 5. 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 у рассматривается как функция одного и того же набора факторов  $x$  то есть система вида:

$$Y_1 = a_{11}x_1 + a_{12}x_2 + \dots + a_{1m}x_m + \varepsilon_1;$$

$$Y_2 = a_{21}x_1 + a_{22}x_2 + \dots + a_{2m}x_m + \varepsilon_2;$$

.....

$$Y_n = a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nm}x_m + \varepsilon_n.$$

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

$$Y_1 = a_{11}x_1 + a_{12}x_2 + \dots + a_{1m}x_m + \varepsilon_1;$$

$$Y_2 = b_{21}y_1 + a_{21}x_1 + a_{22}x_2 + \dots + a_{2m}x_m + \varepsilon_2;$$

$$Y_3 = b_{31}y_1 + b_{32}y_2 + a_{31}x_1 + a_{32}x_2 + \dots + a_{3m}x_m + \varepsilon_3;$$

.....

$$Y_n = b_{n1}y_1 + b_{n2}y_2 + \dots + b_{nn-1}y_{n-1} + a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nm}x_m + \varepsilon_n.$$

**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:

$$Y_1 = b_{12}y_2 + b_{13}y_3 + \dots + b_{1n}y_n + a_{11}x_1 + a_{12}x_2 + \dots + a_{1m}x_m + \varepsilon_1;$$

$$Y_2 = b_{21}y_1 + b_{23}y_3 + \dots + b_{2n}y_n + a_{21}x_1 + a_{22}x_2 + \dots + a_{2m}x_m + \varepsilon_2;$$

.....

$$Y_n = b_{n1}y_1 + b_{n2}y_2 + \dots + b_{nn-1}y_{n-1} + a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nm}x_m + \varepsilon_n.$$

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

$$Y_1 = \delta_{11}X_1 + \delta_{12}X_2 + \dots + \delta_{1m}X_m ;$$

$$Y_2 = \delta_{21}X_1 + \delta_{22}X_2 + \dots + \delta_{2m}X_m ;$$

.....

$$Y_n = \delta_{n1}X_1 + \delta_{n2}X_2 + \dots + \delta_{nm}X_m .$$

Где  $\delta_{ij}$  – 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?

**Задание 1.**

Проверьте, идентифицируема ли эконометрическая модель:

$$\begin{aligned} Y_1 &= b_{12}y_2 + b_{13}y_3 + a_{11}x_1 + a_{12}x_2 ; \\ Y_2 &= b_{21}y_1 + a_{21}x_1 + a_{22}x_2 + a_{23}x_3 ; \\ Y_3 &= b_{31}y_1 + b_{32}y_2 + a_{31}x_1 + a_{33}x_3 + a_{34}x_4 . \end{aligned}$$

**Задание 2.**

Проверьте, идентифицируема ли эконометрическая модель:

$$\begin{aligned} Y_1 &= b_{12}y_2 + b_{13}y_3 + a_{11}x_1 + a_{12}x_2 ; \\ Y_2 &= b_{21}y_1 + a_{22}x_2 + a_{23}x_3 ; \\ Y_3 &= b_{31}y_1 + b_{32}y_2 + a_{31}x_1 + a_{33}x_3 + a_{34}x_4 . \end{aligned}$$

**Задание 3.**

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

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

$$Y_2 = b_{21}y_1 + a_{22}x_2 + a_{23}x_3 + a_{24}x_4 ;$$

$$Y_3 = b_{31}y_1 + b_{32}y_2 + a_{31}x_1 + a_{32}x_2 .$$

#### Задание 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

#### Задание 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