

# Эконометрика



---

Введение.

Основные понятия

Чекменева Т.Д.

# ЭКОНОМЕТРИКА

**Эконометрика** – устанавливает и исследует количественные закономерности в экономике на основе методов математической статистики.

**Эконометрическая модель** служит основой для экономического анализа и прогнозирования, дает возможность для принятия обоснованных экономических решений.

# Литература

Сведения об учебниках			Количество экземпляров в библиотеке
Наименование, издательство	Автор	Год издания	
Эконометрика: Учебник для вузов. – М.: ЮНИТИ-ДАНА	Кремер Н. Ш., Путко Б. А.	2003,311с	10
Эконометрика: Учебник. – М.: Финансы и статистика	Домбровский В. В.	2004,342с	35
Эконометрика: Учебник для вузов. – М.: Финансы и статистика	Елисеева И. И.	2003,342с.	40
		2005,576с.	10
Практикум по эконометрике: Учеб. пособие. – М.: Финансы и статистика	Елисеева И. И.	2003	1
Эконометрические методы в макроэкономическом анализе	Замков О. О.	2001	10      3

# В эконометрике исследуются три основных класса моделей:

1) модели временных рядов

$$y_t = f(t) + \varepsilon_t$$

2) регрессионные модели с одним уравнением

где 
$$y = f(x_1, \dots, x_k, \beta_1, \dots, \beta_p) + \varepsilon$$

$y$  – зависимая (объясняемая) переменная,

$x_1, \dots, x_k$  – независимые (объясняющие) переменные,

$\beta_1, \dots, \beta_p$  – параметры;

3) **системы одновременных уравнений**, где каждое уравнение может, кроме объясняющих переменных, включать в себя объясняемые переменные из других уравнений.

Переменные в регрессионной модели делятся на **два класса** в зависимости от того, являются они или нет объектом объяснения.

1. **Эндогенные** переменные (зависимые) определяется моделируемым явлением.
2. **Экзогенные** переменные (независимые) входят в модель, но определяются независимо от нее.

# Эконометрические данные делят на два типа:

1. **Перекрёстные** (пространственные) данные (**cross-sectional data**) – это данные, полученные для разных однотипных объектов и относящиеся к одному периоду времени (например, данные о курсах валют в определенный день по всем обменным пунктам);
2. **Временные** ряды (**time series data**) – это данные, характеризующие один и тот же объект в различные моменты времени.

# Основные понятия теории вероятностей

- случайное событие ( $A$ )
- случайная величина ( $X$ )
- функция распределения  $F(x)$
- числовые характеристики с/в
- случайный вектор ( $X_1, \dots, X_n$ )

# Основные числовые характеристики случайной величины $X$

## 1. Математическое ожидание

для дискретной случайной величины:

$$E(x) = \sum_k x_k \cdot p_k$$

## 2. Дисперсия случайной величины

$$V(X) = E(X - EX)^2 = E(X^2) - (EX)^2$$

$\sigma = \sqrt{V(X)}$  – стандартное отклонение



### 3. Квантили и процентные точки распределения

Квантиль уровня  $q$  ( $q$ -квантиль)

непрерывного распределения  $F(x)$  – это число  $U_q$  (число  $q$  задано,  $0 < q < 1$ ) такое, что

$$P(X \leq U_q) = q$$

$Q$ -процентная точка непрерывного  
распределения  $F(x)$  – это число  $W_Q$  такое, что

$$P(X > W_Q) = \frac{Q}{100}$$

Между квантилем и процентной точкой  
имеется связь:  $U_q = W_{100(1-q)}$

# Числовые характеристики случайного вектора:

- Математическое ожидание

$$E(X) = (EX_1, \dots, EX_n)^T$$

- Дисперсия случайного вектора – квадратная матрица  $V(X)$ , составленная из коэффициентов ковариации компонент этого вектора

$$\text{Cov}(X_i, X_j) = E[(X_i - EX_i)(X_j - EX_j)]$$

# Парный коэффициент корреляции

$$r_{XY} = \frac{\text{cov}(X, Y)}{\sigma(x)\sigma(y)}$$

Если  $X, Y$  – независимы, то  $r = 0$ :

$$\text{cov}(X, Y) = E\left(X - \underset{=0}{EX}\right) \cdot E\left(Y - \underset{=0}{EY}\right) = 0$$

# Основные понятия и задачи математической статистики

- генеральная совокупность;
- выборка (выборочная совокупность);
- оценивание параметров;
- проверка гипотез;
- специальные распределения:  
 $\chi^2$ - распределение,  $t$ - распределение Стьюдента,  $F$ - распределение Фишера

# Нормальный закон распределения (Гаусса)

$X \sim N(m, \sigma^2)$  произвольная нормальная  
величина

$$m = E(X), \sigma^2 = V(X)$$

$X_0 \sim N(0, 1)$  стандартная нормальная  
величина

$$X_0 = \frac{X - m}{\sigma}$$

# Статистические оценки делятся на *точечные* и *интервальные*

## Свойства оценок (требования):

- **несмещенность:**  $E\hat{\theta} = \theta$
- **состоятельность:** оценка  $\hat{\theta}$  гарантирует приближение к истинному значению при  
$$n \rightarrow \infty$$
- **эффективность:**  $E(\hat{\theta} - \theta)^2$  - минимальная из всех возможных оценок параметра  $\theta$

# Точечные оценки (выборочные статистики)

- Оценка математического ожидания – среднее значение признака:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

- Оценка дисперсии признака:

- выборочная дисперсия:

$$\sigma^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

- исправленная дисперсия:

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

# Точечные оценки (выборочные статистики)

- Оценка коэффициента корреляции:

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \cdot \sum (y_i - \bar{y})^2}} = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\sigma_x \sigma_y}$$

$$|r| \leq 1$$

$$|r| > 0,7 \quad - \text{связь сильная}$$

$$|r| < 0,3 \quad - \text{связь слабая}$$



# Проверка статистических гипотез

$H_0 : \theta \in Z_0$       основная (нулевая) гипотеза

$H_1 : \theta \in Z_1$       альтернативная гипотеза

**Задача:** на основе наблюдений  $X_1, \dots, X_n$

проверить нулевую гипотезу, т.е. принять её  
либо отвергнуть в пользу альтернативной  
гипотезы

Процедура проверки гипотезы называется  
статистическим критерием (тестом).

# Статистический критерий

формулируется в виде неравенства:

$$t_n > C$$

$t_n$  – некоторая функция от выборки:

$$t_n = t_n(X_1, \dots, X_n) \text{ - проверочная статистика}$$

$C = t_{n,\alpha}$  - критическое значение (порог),  
определяющее критическую  
область ( $K_\alpha$ )

# Алгоритм проверки статистической гипотезы:

1. на основании наблюдений  $X_1, \dots, X_n$   
вычисляется значение статистики  $t_n$ ;
2. при заданном уровне значимости  $\alpha$   
находится критическая область  $K_\alpha$  (т.  
е. её критическая точка  $t_{n,\alpha}$ );
3. если  $t_n \in K_\alpha$  (т. е.  $t_n > t_{n,\alpha}$ ), то гипотеза  
 $H_0$  отвергается в пользу  $H_1$ ; в противном  
случае принимается гипотеза  $H_0$ .

# Ошибки проверки гипотез

1) *Ошибка 1-го рода*: отвергнуть  $H_0$ , когда она верна (ложная тревога).

Вероятность ошибки 1-го рода:  $P(H_1 | H_0) = \alpha$

или 
$$\alpha = P_{H_0}(X \in K_\alpha)$$

$\alpha$  – уровень значимости

2) *Ошибка 2-го рода*: принять гипотезу  $H_0$ , когда верна гипотеза  $H_1$ .

Вероятность:  $P(H_0 | H_1) = \beta$       или       $\beta = P_{H_1}(X \in \bar{K}_\alpha)$

# Односторонние и двусторонние критические области

1) *Односторонняя*:  $\alpha = P\{t_n > t_{n,\alpha}\}$

2) *Двусторонняя*:  
(для симметричных распределений)  $\alpha = P\{|t_n| > t_{n,\alpha}^*\}$

Если распределение симметрично, то двусторонние критические точки связаны с односторонними критическими точками соотношением:  $t_{n,\alpha}^* = t_{n,\alpha}/2$

# $p$ -value

Для рассмотренных распределений статистики  $t_n$  (нормальное,  $\chi^2$ , Стьюдента, Фишера) можно использовать еще одну критическую величину:

$p$ -значение (значимость),  $p$ -value.

Если *найдено значение* статистики, например,

$t_n = C$ , то  $p$ -значение вычисляется как **вероятность**:

- а)  $P(|t_n| > |C|)$  – для симметричных распределений;
- б)  $P(t_n > C)$  – для несимметричных распределений.

Очевидно, если  $p$ -значение  $< \alpha$ , то  $t_n > t_{n,\alpha}$ .