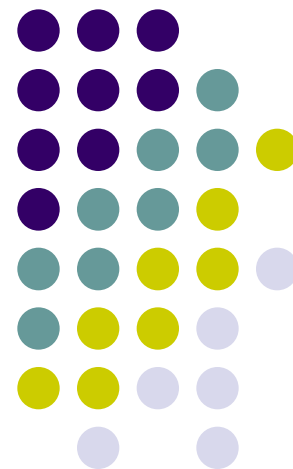


# КОРРЕЛЯЦИОННО-РЕГРЕССИОННЫЙ АНАЛИЗ

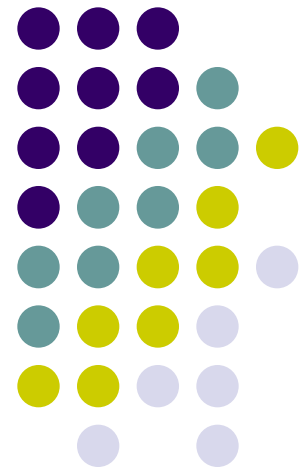
**Вопрос 1. Понятия и задачи корреляционно-регрессионного анализа.**

**Вопрос 2. Этапы проведения корреляционно-регрессионного анализа.**



---

## Вопрос 1. Понятия и задачи корреляционно- регрессионного анализа



## Виды зависимостей

### Функциональная

- зависимость, при которой значению одной переменной соответствует *точно заданное значение* другой переменной.

### Статистическая

- зависимость, при которой значению одной переменной соответствует *множество возможных значений* другой переменной.



Статистические зависимости проявляются только в массовом процессе, при большом числе единиц совокупности.

## Корреляционная зависимость

Статистическая зависимость называется **корреляционной**, если при изменении одной переменной изменяется математическое ожидание (среднее значение) другой.



## Понятие регрессии

Односторонняя статистическая зависимость выражается с помощью функции, которая называется **регрессий**.

## Виды регрессий

### 1. Количество переменных

- простая (парная)
- множественная

### 2. Аналитическая форма

- линейная
- нелинейная

### 3. Характер регрессии

- положительная
- отрицательная

## Корреляционно-регрессионный анализ

**Корреляционно-регрессионный анализ** – метод статистического моделирования для исследования формы связи между случайными величинами.

### *Корреляционный анализ*

- *исследование корреляционных связей*

### *Регрессионный анализ*

- *исследование односторонних статистических зависимостей*

### Основные вопросы исследования статистической зависимости

■ Существует ли зависимость между величинами?	Корреляционный анализ
■ Насколько сильной она является?	
■ Каков характер этой зависимости – прямой или обратный, линейный или нелинейный?	Регрессионный анализ
■ Каково наиболее вероятное значение $Y$ при заданном $X$ (прогнозирование)?	



## Задачи корреляционного анализа

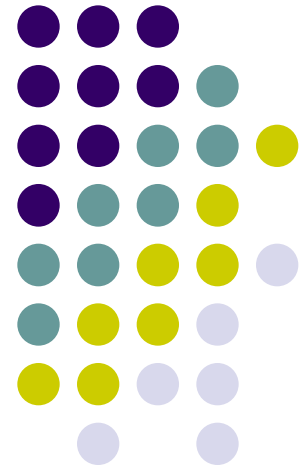
1. Измерение степени тесноты (силы) двух и более признаков.
2. Отбор факторов, оказывающих наиболее существенное влияние на результативный признак, на основании измерения тесноты связи между явлениями.
3. Обнаружение неизвестных причинных связей.

## Задачи регрессионного анализа

1. Установление формы зависимости (линейная или нелинейная; положительная или отрицательная и т. д.).
2. Определение функции регрессии в виде математического уравнения и установление влияния факторов на зависимую переменную.
3. Оценка неизвестных значений зависимой переменной.

---

## Вопрос 2. Этапы проведения корреляционно- регрессионного анализа



### Модели корреляционно-регрессионного анализа

Корреляционно-регрессионный анализ – основа для построения *моделей*, которые в математической форме описывают основные количественные связи между изучаемыми показателями и влияющими на них факторами.

Модели выступают в качестве средства анализа и прогнозирования конкретных экономических процессов на основе реальной статистической информации.

# Основные классы моделей

1. Регрессионные модели с одним уравнением, в которых зависимая переменная (результативный признак)  $Y$  представляется в виде функции

$$Y = f(X_1, \dots, X_k, \beta_1, \dots, \beta_\varphi),$$

где  $X_1, \dots, X_k$  – независимые переменные (факторы);

$\beta_1, \dots, \beta_\varphi$  – параметры.

# Основные классы моделей

**2. Системы одновременных уравнений** – система уравнений, в которой одни и те же переменные одновременно рассматриваются как зависимые в одних уравнениях и как независимые в других.

Общий вид системы одновременных уравнений:

$$\left\{ \begin{array}{l} y_1 = b_{12}y_2 + b_{13}y_3 + \dots + b_{1n}y_n + a_{11}x_1 + a_{12}x_2 + \dots + a_{1m}x_m \\ y_2 = b_{21}y_1 + b_{23}y_3 + \dots + b_{2n}y_n + a_{21}x_1 + a_{22}x_2 + \dots + a_{2m}x_m \\ \vdots \\ y_n = b_{n1}y_1 + b_{n2}y_2 + \dots + b_{nn-1}y_{n-1} + a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nm}x_m \end{array} \right.$$

## Этапы корреляционно-регрессионного анализа

(на примере построения многофакторной модели)

1. Определение цели моделирования и априорное исследование экономической проблемы.
2. Формирование перечня факторов и их логический анализ.
3. Сбор исходных данных и анализ информации.
4. Выбор формы связи изучаемых показателей.
5. Построение модели.
6. Проверка качества построенной модели.
7. Использование модели для экономического анализа и прогнозирования.

### **Этап 1. Определение цели моделирования и априорное исследование экономической проблемы**

**Целью моделирования** может быть:

- *анализ* изучаемого экономического объекта (процесса);
- *прогноз* экономических показателей изучаемого объекта;
- *имитация* развития объекта при различных значениях переменных;
- *выработка* управленческих решений.



### ***Этап 1. Определение цели моделирования и априорное исследование экономической проблемы***

На этом этапе проводится предварительный *анализ сущности* изучаемого явления или процесса.

На основе знаний макро- и микроэкономики конкретизируются явления и процессы.

В результате должны быть сформулированы экономически осмысленные и приемлемые гипотезы о зависимости экономических явлений.

### ***Этап 2. Формирование перечня факторов и их логический анализ***

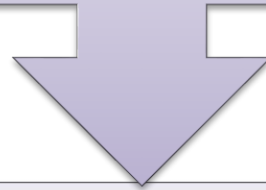
#### ***Требования, предъявляемые к включаемым в модель факторам:***

- каждый фактор должен быть обоснован теоретически;
- целесообразно включать факторы, оказывающие существенное воздействие на изучаемые показатели;
- факторы не должны быть линейно зависимы;
- факторы должны быть количественно измерены;
- в одну модель нельзя включать совокупный фактор и образующие его частные факторы.

### **Этап 2. Формирование перечня факторов и их логический анализ**

Отбор факторов осуществляется в две стадии:

На первой стадии отбирают факторы, исходя из сущности проблемы и ориентируясь на соображения профессионально-теоретического характера.



На второй стадии проводят анализ факторов на мультиколлинеарность.

### **Этап 2. Формирование перечня факторов и их логический анализ**

**Мультиколлинеарность** – наличие сильной корреляции между независимыми переменными.

Для определения мультиколлинеарности анализируют *корреляционную матрицу* между факторами и выявляют пары переменных, имеющих высокие коэффициенты корреляции (обычно более 0,7).

Если такие переменные существуют, то они дублируют друг друга и одну из них следует исключить из модели.

**Этап 2. Формирование перечня факторов  
и их логический анализ**

*Например, при изучении зависимости  $y=f(x_1, x_2, x_3)$  матрица парных коэффициентов корреляции оказалась следующей:*

	$y$	$x_1$	$x_2$	$x_3$
$y$	1			
$x_1$	0,9	1		
$x_2$	0,7	0,8	1	
$x_3$	0,8	0,2	0,5	1

*Переменные  $x_1$  и  $x_2$  коллинеарны. В уравнение регрессии целесообразно включить  $x_1$ , а не  $x_2$ , т.к. корреляция  $x_2$  с  $y$  слабее (0,7), чем корреляция  $x_1$  с  $y$  (0,9), а также слабее межфакторная корреляция с  $x_3$  ( $0,2 < 0,5$ ).*

### *Этап 3. Сбор исходных данных и анализ информации*

Статистическая информация может быть собрана в трёх видах: временные ряды, пространственные данные, панельные данные.



Условие 1. *Достаточное число единиц совокупности* (рекомендуется, чтобы число наблюдений было в 5-8 раз больше числа факторов).

Условие 2. *Достаточная однородность совокупности данных.*

### **Этап 3. Сбор исходных данных и анализ информации**

#### **Источники данных:**

- сайт *Федеральной службы государственной статистики* (Росстата) – URL: **<https://rosstat.gov.ru/>**
- база данных «Российский мониторинг экономического положения и здоровья населения» – URL: **<http://www.hse.ru/rlms/project>**
- сайт *Единого архива экономических и социологических данных* – URL: **<http://sophist.hse.ru>**

**и другие.**

### ***Этап 4. Выбор формы связи изучаемых показателей***

**Основная задача** – выбор вида функции регрессии.

Формулируются:

- гипотеза о форме связи (линейная или нелинейная, простая или множественная),
- исходные предпосылки и ограничения модели.



### *Этап 5. Построение модели*

На данном этапе:

- осуществляется оценка параметров (коэффициентов) модели;
- вычисляется ряд показателей, характеризующих точность регрессионного анализа.

### Этап 6. Проверка качества построенной модели

Процедуры данного этапа анализа :

- **проверка адекватности модели** (соответствие модели моделируемому реальному экономическому объекту или процессу);
- **оценка значимости уравнения регрессии** (с использованием  $F$ -критерия Фишера);
- **оценка качества уравнения регрессии** (с использованием коэффициента детерминации);
- **оценка значимости коэффициентов регрессии** (с использованием  $t$ -статистики);
- **оценка качества подбора уравнения** (с использованием ошибки аппроксимации).

### ***Этап 7. Использование модели для экономического анализа и прогнозирования***

Модель, удовлетворяющая всем необходимым требованиям, может быть использована для прогнозирования или объяснения исследуемых экономических процессов.

Модель позволяет предсказать среднее значение исследуемого экономического показателя на основе прогнозируемых значений факторов.