

Кафедра Бизнес - статистики

Установочная лекция №2

Дисциплина «Статистика»

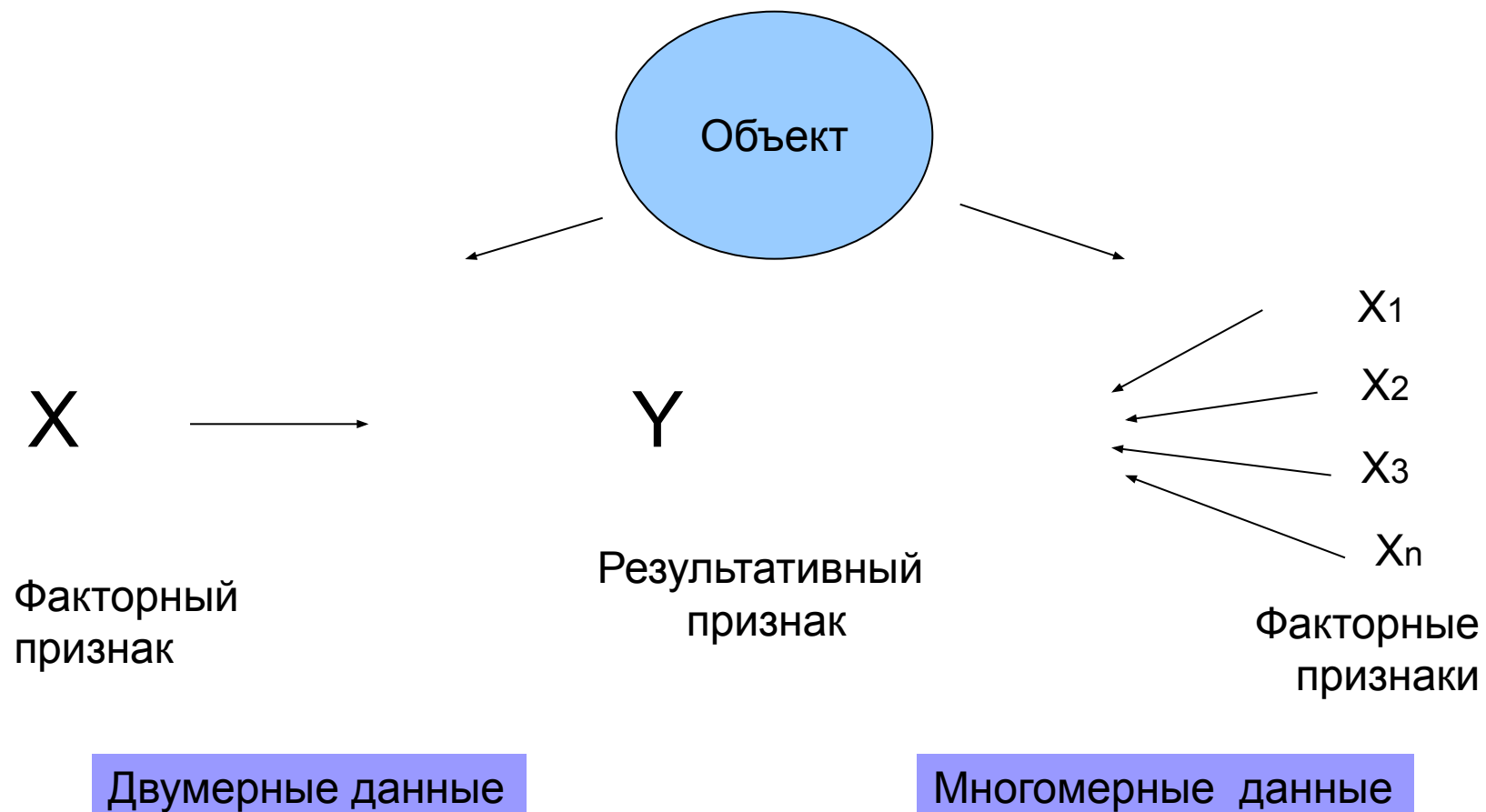
Общий план работы

- Обзор темы «Изучение взаимосвязи между социально-экономическими явлениями методами корреляционно-регрессионного анализа»
- Обзор темы «Анализ динамики социально-экономических явлений и процессов»

Тема: Изучение взаимосвязи между явлениями методами корреляционно-регрессионного анализа

- План:
- 1. Принципы изучения взаимосвязи.
- 2. Классификация видов взаимосвязи
- 3. Исследование взаимосвязи с помощью диаграммы рассеяния
- 4. Расчет линейного коэффициента корреляции
- 5. Ложная корреляция
- 6. Задачи применения регрессионного анализа
- 7. Вычисление и интерпретация параметров линейной парной регрессии

Принципы изучения взаимосвязи



Виды взаимосвязи

Взаимосвязь проявляется в изменении значений результативных показателей под влиянием факторных

- **Взаимосвязь**

Функциональная

Значение результативного показателя полностью определяется факторным

Для каждого X существует только одно значение Y

$$Y = X^2 + 10$$

$$X=2$$

$$Y = 14$$

Стохастическая

Значение результативного показателя меняется под воздействием множества факторов

Для каждого X могут существовать разные значения Y для разных объектов

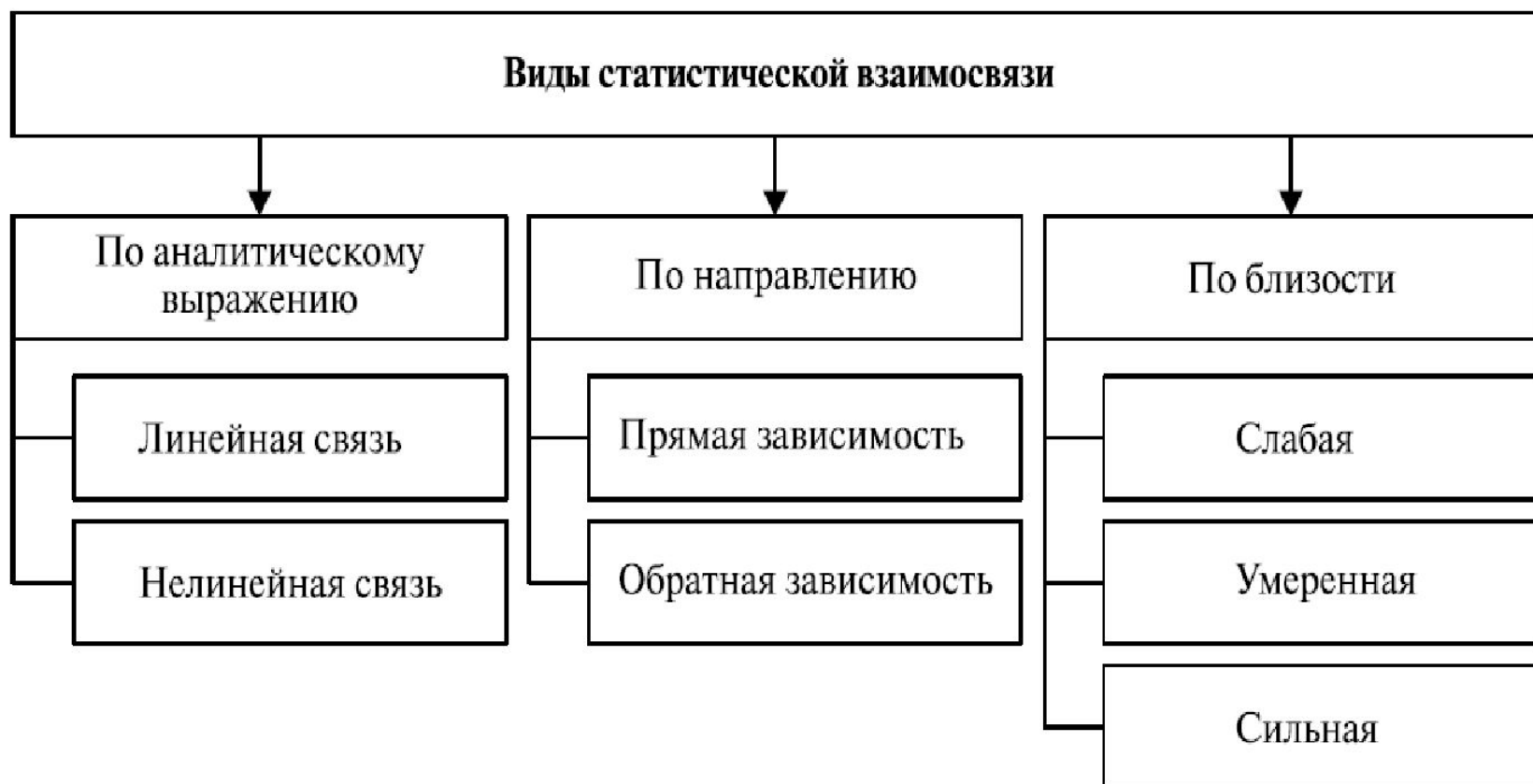
Пример: Результаты деятельности аудиторских фирм по итогам 2015 г.

№ фирмы	у Совокупная выручка, млн. руб.	X1 Среднесписочная численность аудиторов	X2 Кол-во заключенных договоров на оказание услуг, тысяч
1	144,0	109	3,1
2	123,0	56	3,3
3	108,0	80	2,6
4	80,0	26	1,7
5	70,0	50	2,4
6	67,0	44	1,4
7	53,0	15	2,6
8	52,0	28	1,5
9	49,0	23	1,8
10	46,0	24	1,2
11	41,0	20	0,9
12	40,0	21	1,2
13	36,0	43	1,4
14	34,0	10	1,4
15	32,0	21	1,1
Итого	975,0	570	27,60
Среднее	65,0	38	1,84

Сущность корреляционной связи

- **При корреляционной связи изменение среднего значения результативного признака обусловлено влиянием (= изменением) факторных признаков.**

Классификация видов взаимосвязи



Критерии оценки тесноты взаимосвязи

таблица 9.1.1

Количественные критерии оценки тесноты связи

Величина показателя связи	Теснота связи
до $\pm 0,3$	практически отсутствует
$\pm 0,3 - \pm 0,5$	слабая
$\pm 0,5 - \pm 0,7$	умеренная
$\pm 0,7 - \pm 1,0$	сильная
± 1	функциональная

Знак при коэффициенте связи указывает на направление связи:

> 0 – прямая взаимосвязь

< 0 – обратная взаимосвязь

Сущность корреляционного и регрессионного анализов

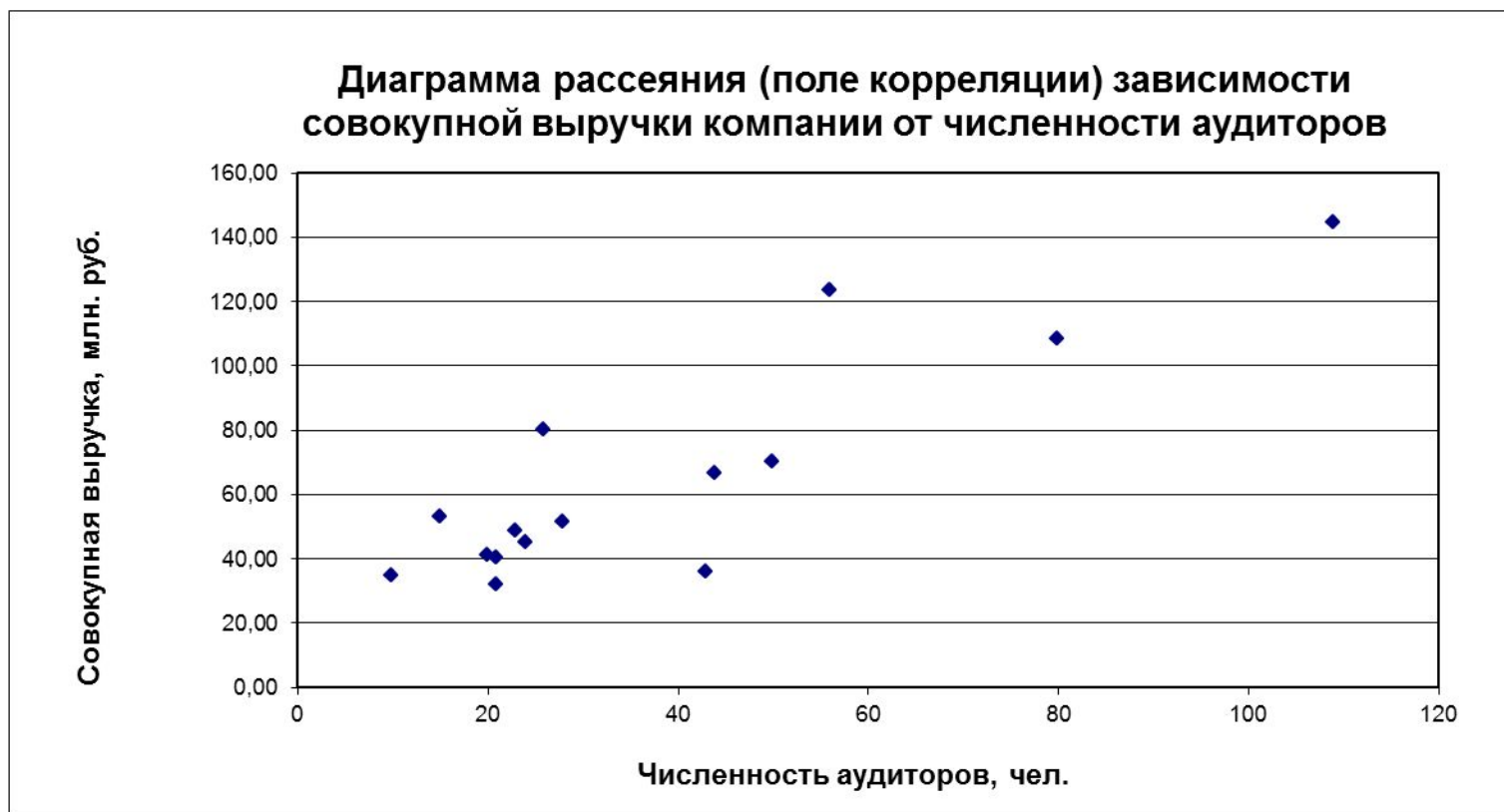
- **Корреляционный анализ** выявляет тесноту и направление взаимосвязи между показателями
- **Регрессионный анализ** дает аналитическое выражение связи в виде математической функции, которая позволяет вычислить (предсказать или спрогнозировать) значения одной переменной на основании другой.

Условия применения корреляционно-регрессионного анализа

1. Единицы исследуемой совокупности должны иметь одинаковую размерность и методологию расчета.
2. Переменные должны быть выражены количественно и являться случайно выбранными из единиц генеральной совокупности.
4. Единицы исследуемой совокупности должны быть независимыми друг от друга. Зависимость единиц совокупности друг от друга в статистике называется *автокорреляцией*.
5. Показатели должны быть однородными.
6. Совокупность исходных данных должна подчиняться нормальному закону распределения.
7. Количество единиц совокупности должно превышать количество факторных признаков минимум в 3–4 раза (лучше в 8–10 раз).
8. Факторные признаки не должны находиться между собой в функциональной зависимости. Существенная связь факторных признаков в статистике называется *мультиколлинеарностью*.

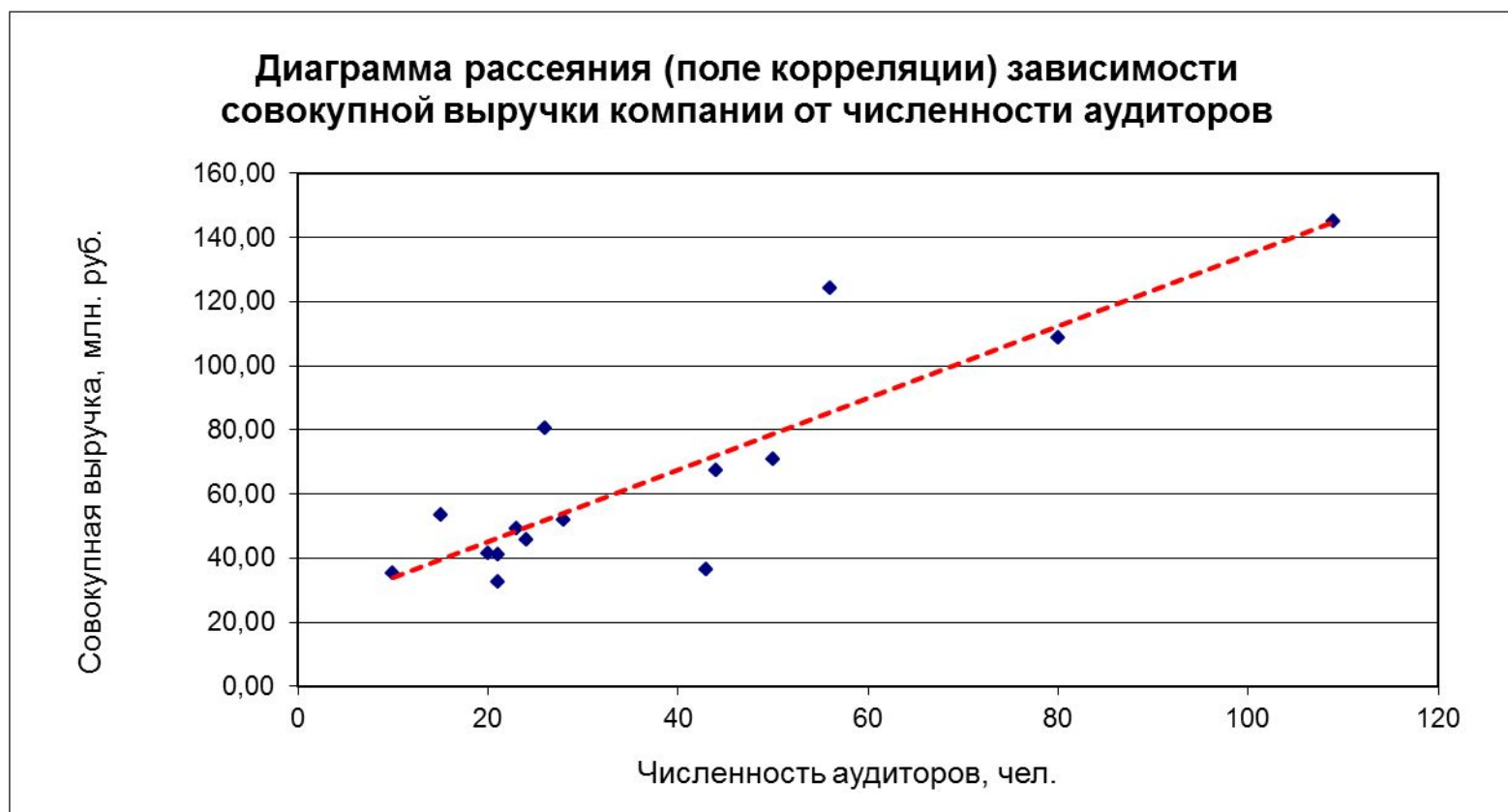
Графический анализ взаимосвязи

- Для графического анализа строится диаграмма рассеяния (поле корреляции)



Графический анализ взаимосвязи: ось концентрации облака точек

- Линия регрессии – ось концентрации облака точек

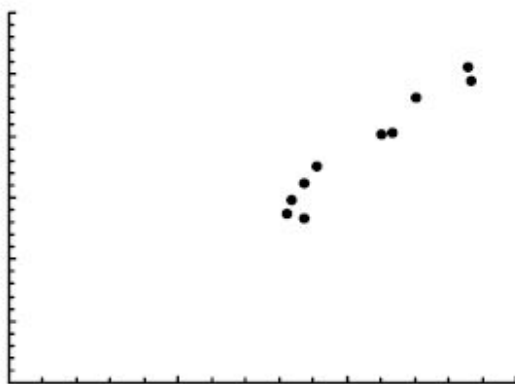


Графический анализ взаимосвязи:

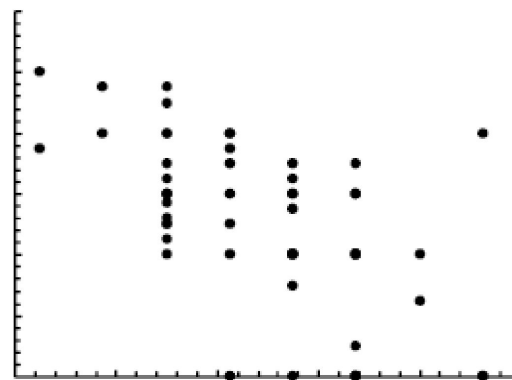
А) и Б) – линейная связь

В) и Г) – нелинейная связь

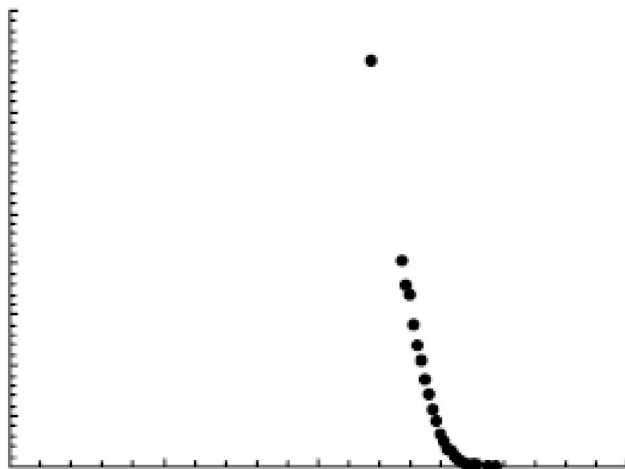
А)



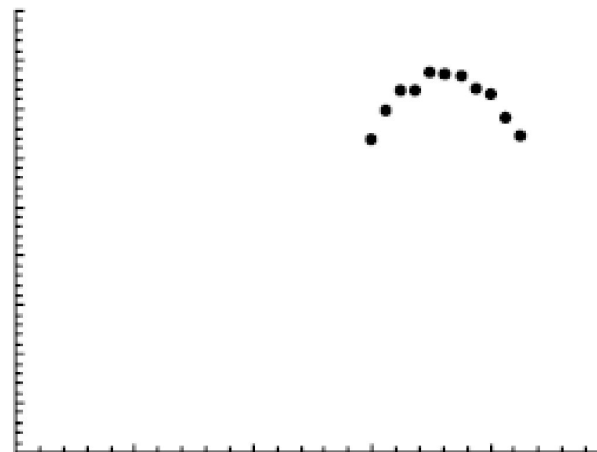
Б)



В)

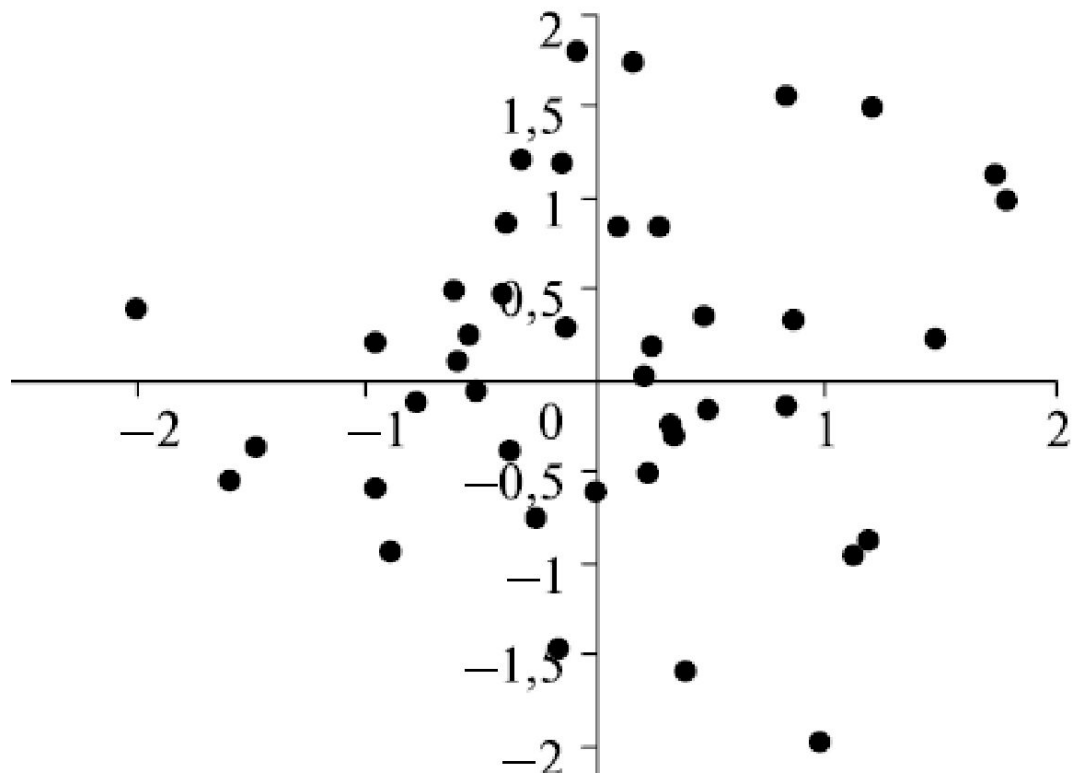


Г)



Графический анализ взаимосвязи: взаимосвязь отсутствует

Д)



Пример: Результаты деятельности аудиторских фирм по итогам 2011 г.

№ фирмы	у	X1	X2
	Совокупная выручка, млн. руб.	Среднесписочная численность аудиторов	Кол-во заключенных договоров на оказание услуг, тысяч
1	144,0	109	3,1
2	123,0	56	3,3
3	108,0	80	2,6
4	80,0	26	1,7
5	70,0	50	2,4
6	67,0	44	1,4
7	53,0	15	2,6
8	52,0	28	1,5
9	49,0	23	1,8
10	46,0	24	1,2
11	41,0	20	0,9
12	40,0	21	1,2
13	36,0	43	1,4
14	34,0	10	1,4
15	32,0	21	1,1
Итого	975,0	570	27,60
Среднее	65,0	38	1,84

Какой из показателей X1 или X2 оказывает более сильное влияние на размер совокупной выручки аудиторской фирмы?

Парный линейный коэффициент корреляции Пирсона

$$r = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\sigma_x \cdot \sigma_y},$$

$$\overline{xy} = \frac{\sum_{i=1}^n x_i y_i}{n}; \quad \bar{y} = \frac{\sum_{i=1}^n y_i}{n}; \quad \bar{x} = \frac{\sum_{i=1}^n x_i}{n}; \quad \sigma_x = \sqrt{\overline{x^2} - \bar{x}^2}; \quad \sigma_y = \sqrt{\overline{y^2} - \bar{y}^2}.$$

n — число наблюдений

Парный линейный коэффициент корреляции Пирсона

- Коэффициент корреляции, обозначаемый r , характеризует тесноту и направление линейной связи между двумя признаками.

$$r_{xy} = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\sigma_x \cdot \sigma_y}$$

Упрощенная формула расчета СКО

ВАЖНО:

применима только для несгруппированных данных

$$\sigma_x = \sqrt{\overline{x^2} - (\bar{x})^2}; \quad \sigma_y = \sqrt{\overline{y^2} - (\bar{y})^2};$$

$$\overline{X^2} = \frac{\sum_{i=1}^n x_i^2}{n};$$

$$\overline{Y^2} = \frac{\sum_{i=1}^n y_i^2}{n}$$

Парный линейный коэффициент корреляции Пирсона

Вспомогательная таблица для определения линейного коэффициента корреляции

Число наблюдений, n	Факторный признак, x	Результативный признак, y	y^2	x^2	xy
1	x_1	y_1	y_1^2	x_1^2	x_1y_1
2	x_2	y_2	y_2^2	x_2^2	x_2y_2
3	x_3	y_3	y_3^2	x_3^2	x_3y_3
Итого	$\sum_{i=1}^n x_i$	$\sum_{i=1}^n y_i$	$\sum_{i=1}^n y_i^2$	$\sum_{i=1}^n x_i^2$	$\sum_{i=1}^n x_i y_i$
Средняя	$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$	$\bar{y} = \frac{\sum_{i=1}^n y_i}{n}$	$\bar{y}^2 = \frac{\sum_{i=1}^n y_i^2}{n}$	$\bar{x}^2 = \frac{\sum_{i=1}^n x_i^2}{n}$	$\overline{xy} = \frac{\sum_{i=1}^n x_i y_i}{n}$

Парный линейный коэффициент корреляции Пирсона

Вспомогательная таблица для определения линейного коэффициента корреляции

Число наблюдений, n	Факторный признак, x	Результативный признак, y	y^2	x^2	xy
1	x_1	y_1	y_1^2	x_1^2	x_1y_1
2	x_2	y_2	y_2^2	x_2^2	x_2y_2
3	x_3	y_3	y_3^2	x_3^2	x_3y_3
Итого	$\sum_{i=1}^n x_i$	$\sum_{i=1}^n y_i$	$\sum_{i=1}^n y_i^2$	$\sum_{i=1}^n x_i^2$	$\sum_{i=1}^n x_i y_i$
Средняя	$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$	$\bar{y} = \frac{\sum_{i=1}^n y_i}{n}$	$\bar{y} = \frac{\sum_{i=1}^n y_i^2}{n}$	$\bar{x} = \frac{\sum_{i=1}^n x_i^2}{n}$	$\overline{xy} = \frac{\sum_{i=1}^n x_i y_i}{n}$

$$r = \frac{n \sum xy - \sum x \sum y}{\sqrt{\left[n \sum x^2 - (\sum x)^2 \right] \cdot \left[n \sum y^2 - (\sum y)^2 \right]}}$$

Пример: рассчитаем парный линейный коэффициент корреляции между численностью аудиторов (X1) и размером совокупной выручки аудиторской фирмы (Y)

№ фирмы	Y, млн. руб.	X1, чел	Y^2	X1^2	Y*X1
1	144,0	109	20736,0	11881	15696,00
2	123,0	56	15129,0	3136	6888,00
3	108,0	80	11664,0	6400	8640,00
4	80,0	26	6400,0	676	2080,00
5	70,0	50	4900,0	2500	3500,00
6	67,0	44	4489,0	1936	2948,00
7	53,0	15	2809,0	225	795,00
8	52,0	28	2704,0	784	1456,00
9	49,0	23	2401,0	529	1127,00
10	46,0	24	2116,0	576	1104,00
11	41,0	20	1681,0	400	820,00
12	40,0	21	1600,0	441	840,00
13	36,0	43	1296,0	1849	1548,00
14	34,0	10	1156,0	100	340,00
15	32,0	21	1024,0	441	672,00
Итого	975,0	570	80105,0	31874	48454,0
Среднее	65,0	38	5340,3	2124,9	3230,3

$$r = \frac{n\sum xy - \sum x \sum y}{\sqrt{[n\sum x^2 - (\sum x)^2] \cdot [n\sum y^2 - (\sum y)^2]}}$$

$$r = \frac{15 \cdot 48454 - 570 \cdot 975}{\sqrt{[15 \cdot 31874 - 570^2] \cdot [15 \cdot 80105 - 975^2]}}$$

$$r = 0,872$$

Парный линейный коэффициент корреляции Пирсона

Вспомогательная таблица для определения линейного коэффициента корреляции

массив1	Число наблюдений, n	Факторный признак, x	Результативный признак, y	y^2	x^2	xy
	1	x_1	y_1	y_1^2	x_1^2	x_1y_1
	2	x_2	y_2	y_2^2	x_2^2	x_2y_2
	3	x_3	y_3	y_3^2	x_3^2	x_3y_3
Итого		$\sum_{i=1}^n x_i$	$\sum_{i=1}^n y_i$	$\sum_{i=1}^n y_i^2$	$\sum_{i=1}^n x_i^2$	$\sum_{i=1}^n x_i y_i$
Средняя		$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$	$\bar{y} = \frac{\sum_{i=1}^n y_i}{n}$	$\bar{y} = \frac{\sum_{i=1}^n y_i^2}{n}$	$\bar{x} = \frac{\sum_{i=1}^n x_i^2}{n}$	$\overline{xy} = \frac{\sum_{i=1}^n x_i y_i}{n}$

массив2

$$r = \frac{n \sum xy - \sum x \sum y}{\sqrt{\left[n \sum x^2 - (\sum x)^2 \right] \cdot \left[n \sum y^2 - (\sum y)^2 \right]}}$$

Расчет при помощи MS Excel =ПИРСОН(массив1;массив2)
 =PEARSON(массив1;массив2)

Расчет r при помощи MS Excel

	У	X1	X2					
№ фирмы	Совокупная выручка, млн. руб.	Среднесписочная численность аудиторов	Кол-во заключенных договоров на оказание услуг, тысяч					
1	144,0	109	3,1	r Y X1 =	0,8724			
2	123,0	56	3,3					
3	108,0	80	2,6					
4	80,0	26	1,7	r Y X2 =	0,8506			
5	70,0	50	2,4					
6	67,0	44	1,4					
7	53,0	15	2,6					
8	52,0	28	1,5					
9	49,0	23	1,8					
10	46,0	24	1,2					
11	41,0	20	0,9					
12	40,0	21	1,2					
13	36,0	43	1,4					
14	34,0	10	1,4					
15	32,0	21	1,1					
Итого	975,0	570	27,60					
Среднее	65,0	38	1,84					

Проверка значимости r

t-критерий Стьюдента

$$t_{\text{набл}} = \frac{r}{\sqrt{1-r^2}} \sqrt{n-2} \quad t_{\text{кр}} = St^{-1}(\alpha; n-2)$$

$|t_{\text{набл}}| > t_{\text{кр}} \Rightarrow$ коэффициент значим

Пример: Результаты деятельности аудиторских фирм по итогам 2011 г.

№ фирмы	у	X1	X2
	Совокупная выручка, млн. руб.	Среднесписочная численность аудиторов	Кол-во заключенных договоров на оказание услуг, тысяч
1	144,0	109	3,1
2	123,0	56	3,3
3	108,0	80	2,6
4	80,0	26	1,7
5	70,0	50	2,4
6	67,0	44	1,4
7	53,0	15	2,6
8	52,0	28	1,5
9	49,0	23	1,8
10	46,0	24	1,2
11	41,0	20	0,9
12	40,0	21	1,2
13	36,0	43	1,4
14	34,0	10	1,4
15	32,0	21	1,1
Итого	975,0	570	27,60
Среднее	65,0	38	1,84

$$r_{yX1} = 0,87$$

$$r_{yX2} = 0,85$$

$$t_{кр} = \text{СТЪЮДРАСПРОБР} \\ (0,05; 13) = 2,1604$$

$$t_{рас\ y/x1} = \frac{0,87}{\sqrt{1 - 0,87^2}} \sqrt{15 - 2}$$

$$t_{рас\ y/x2} = \frac{0,85}{\sqrt{1 - 0,85^2}} \sqrt{15 - 2}$$

$$t_{рас\ r_{yX1}} = 6,326$$

$$t_{рас\ r_{yX2}} = 5,818$$

Ложная корреляция

Высокая корреляция между X и $Y \neq$ причинно-следственная взаимосвязь между X и Y !

Ложной корреляцией называют случай наличия умеренной или сильной взаимосвязи между показателями, вызванной действием третьего (скрытого) фактора.

Пример ложной корреляции

Количество пожарных

Ущерб от пожара



Причиной взаимосвязи является третий фактор: масштаб пожара

Парная линейная регрессия

Линейное уравнение регрессии:

Теоретическое значение Y при заданном X ← $\bar{y}_x = a_0 + a_1 x$ → Коэффициенты регрессии

- **a_0** — показывает усредненное влияние на результативный признак неучтенных в уравнении факторных признаков;
- **a_1** — показывает, насколько в среднем изменится значение результативного признака Y , при изменении факторного признака на единицу собственного измерения.

Парная линейная регрессия

$$\bar{y}_x = a_0 + a_1 x$$

Метод наименьших квадратов

$$S = \sum (y - \bar{y}_x)^2 \rightarrow \min.$$

$$\left\{ \begin{array}{l} na_0 + a_1 \sum_{i=1}^n x_i = \sum_{i=1}^n y_i \\ a_0 \sum_{i=1}^n x_i + a_1 \sum_{i=1}^n x_i^2 = \sum_{i=1}^n x_i y_i \end{array} \right. ,$$

Пример: рассчитаем параметры линейного уравнения регрессии между численностью аудиторов (X1) и размером совокупной выручки аудиторской фирмы (Y)

№ фирмы	Y, млн. руб.	X1, чел	Y^2	X1^2	Y*X1
1	144,0	109	20736,0	11881	15696,00
2	123,0	56	15129,0	3136	6888,00
3	108,0	80	11664,0	6400	8640,00
4	80,0	26	6400,0	676	2080,00
5	70,0	50	4900,0	2500	3500,00
6	67,0	44	4489,0	1936	2948,00
7	53,0	15	2809,0	225	795,00
8	52,0	28	2704,0	784	1456,00
9	49,0	23	2401,0	529	1127,00
10	46,0	24	2116,0	576	1104,00
11	41,0	20	1681,0	400	820,00
12	40,0	21	1600,0	441	840,00
13	36,0	43	1296,0	1849	1548,00
14	34,0	10	1156,0	100	340,00
15	32,0	21	1024,0	441	672,00
Итого	975,0	570	80105,0	31874	48454,0
Среднее	65,0	38	5340,3	2124,9	3230,3

$$\begin{cases} na_0 + a_1 \sum_{i=1}^n x_i = \sum_{i=1}^n y_i \\ a_0 \sum_{i=1}^n x_i + a_1 \sum_{i=1}^n x_i^2 = \sum_{i=1}^n x_i y_i \end{cases}$$

$$\begin{cases} 15a_0 + a_1 570 = 975 \\ a_0 570 + a_1 31874 = 48454 \end{cases}$$

$$\begin{cases} a_0 = 22,573 \\ a_1 = 1,1165 \end{cases}$$

$$Y_{x1} = 22,573 + 1,1165 \cdot X1$$

Интерпретация коэффициентов регрессии

$$Y_{x1} = 22,573 + 1,1165 * X1$$

$a_0 = 22,573$ показывает, в какой степени на размер совокупной выручки оказывают влияние другие факторы, не включенные в парную модель. Если исключить численность аудиторов, то совокупная выручка составит 22,573 млн. руб.

$a_1 = 1,1165$ показывает, что при увеличении численности аудиторов на 1 человека совокупная выручка аудиторской фирмы возрастает в среднем на 1,1165 млн. руб. в год.

Парная линейная регрессия

$$\bar{y}_x = a_0 + a_1 x$$

Расчет при помощи MS Excel:

**a1=НАКЛОН(известные значения_Y;
известные значения_X)**

$$a_0 = \bar{Y} - a_1 * \bar{X}$$

**a0=ОТРЕЗОК(известные значения_Y;
известные значения_X)**

Расчет коэффициентов регрессии при помощи MS Excel

№ фирмы	у Совокупная выручка, млн. руб.	X1 Среднесписочная численность аудиторов	X2 Кол-во заключенных договоров на оказание услуг, тысяч				
1	144,0	109	3,1	a1 =	1,11651		
2	123,0	56	3,3				
3	108,0	80	2,6				
4	80,0	26	1,7	a0 = 65 - 1,1165*38 =	22,5727		
5	70,0	50	2,4				
6	67,0	44	1,4				
7	53,0	15	2,6				
8	52,0	28	1,5				
9	49,0	23	1,8				
10	46,0	24	1,2				
11	41,0	20	0,9				
12	40,0	21	1,2				
13	36,0	43	1,4				
14	34,0	10	1,4				
15	32,0	21	1,1				
Итого	975,0	570	27,60				
Среднее	65,0	38	1,84				

Расчет коэффициентов регрессии при помощи MS Excel

	у	X1	X2				
№ фирмы	Совокупная выручка, млн. руб.	Среднесписочная численность аудиторов	Кол-во заключенных договоров на оказание услуг, тысяч				
1	144,0	109	3,1	a1 =	1,11651		
2	123,0	56	3,3				
3	108,0	80	2,6				
4	80,0	26	1,7	a0 = 65 - 1,1165*38 =	22,5727		
5	70,0	50	2,4				
6	67,0	44	1,4				
7	53,0	15	2,6				
8	52,0	28	1,5				
9	49,0	23	1,8				
10	46,0	24	1,2				
11	41,0	20	0,9				
12	40,0	21	1,2				
13	36,0	43	1,4				
14	34,0	10	1,4				
15	32,0	21	1,1				
Итого	975,0	570	27,60				
Среднее	65,0	38	1,84				

Расчет теоретических значений результативного показателя

- По уравнению регрессии получают теоретические значения Y путем подстановки в уравнение значений факторного признака X .
- Если фирма планирует увеличить число аудиторов до 48 человек, то она может получить
 $22,573 + 1,1165 * 48 = 76,2$ млн. руб.
совокупной выручки

Расчет теоретических значений Y по уравнению регрессии $Y_{x1}=22,573+1,1165*X1$

№ фирмы	Y Совокупная выручка, млн. руб.	X1 Среднесписочная численность аудиторов	Yx1
1	144,0	109	144,3 = 22,573 + 1,1165 * 109
2	123,0	56	85,1 = 22,573 + 1,1165 * 56
3	108,0	80	111,9 = 22,573 + 1,1165 * 80
4	80,0	26	51,6 = 22,573 + 1,1165 * 26
5	70,0	50	78,4 = 22,573 + 1,1165 * 50
6	67,0	44	71,7
7	53,0	15	39,3
8	52,0	28	53,8
9	49,0	23	48,3
10	46,0	24	49,4
11	41,0	20	44,9
12	40,0	21	46,0
13	36,0	43	70,6
14	34,0	10	33,7
15	32,0	21	46,0
Итого	975,0	570	975,0
Среднее	65,0	38	65,0

Модели множественной регрессии

Множественная регрессия

- линейная:

$$\bar{y}_x = a_0 + a_1 x_1 + a_2 x_2 + \dots + a_k x_k;$$

- степенная:

$$\bar{y}_x = a_0 x_1^{a_1} + x_2^{a_2} + x_3^{a_3} + \dots + x_k^{a_k};$$

- параболическая:

$$\bar{y}_x = a_0 + a_1 x_1^2 + a_2 x_2^2 + \dots + a_k x_k^2,$$

Пример множественной регрессии

$$Y_R = 0,314X_1 + 0,355X_2 + 0,230X_3. \quad (2)$$

Y - индекс развития банковской конкуренции

X_1 - индекс финансовой насыщенности региона банковскими услугами (по объему выданных кредитов)

X_2 - индекс развития сберегательного дела (депозиты на душу населения к доходам населения)

X_3 - индекс институциональной насыщенности региона банковскими услугами

Применение множественных моделей регрессии

Модели классификации клиентов в скоринге

Множественная линейная регрессия

$$p = w_0 + w_1 x_1 + w_2 x_2 + \dots + w_n x_n,$$

где p -- вероятность дефолта,

w -- весовые коэффициенты,

x -- характеристики клиента.

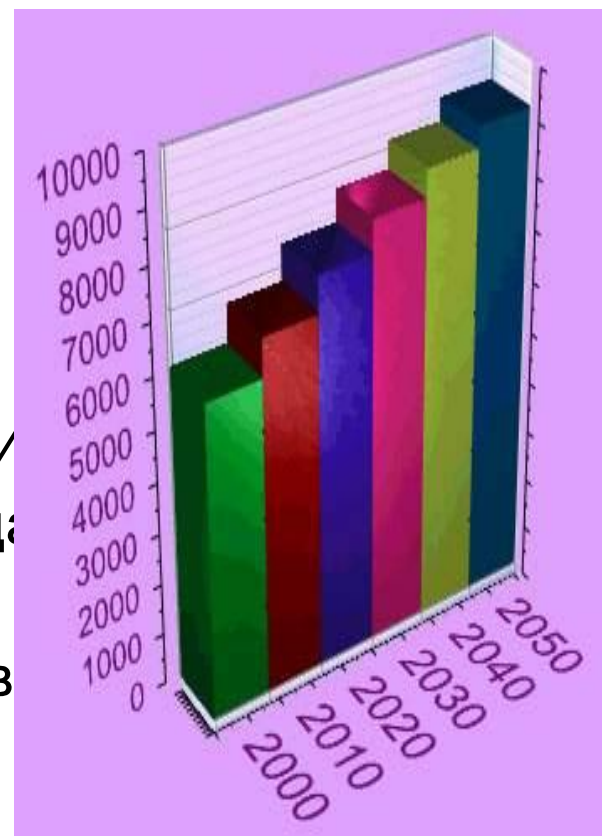
Логистическая регрессия

$$\log(p/(1-p)) = w_0 + w_1 x_1 + w_2 x_2 + \dots + w_n x_n.$$

ТЕМА: АНАЛИЗ ДИНАМИКИ СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ ЯВЛЕНИЙ

План

1. Понятие и классификация рядов динамики
2. Индивидуальные и средние аналитические показатели динамики
3. Простейшие методы прогнозирования
4. Методы выявления тенденции в рядах динамики
5. Аналитическое выравнивание рядов динамики



Понятие ряда динамики

- **Ряд динамики** (от англ. *time series* — временной ряд) — это последовательность изменяющихся во времени значений показателя, расположенных в хронологическом порядке.

Составляющими ряда динамики являются:

- **У** - значения показателя—**уровни ряда**
- **t** - **периоды** (годы, кварталы, месяцы, сутки, ...) или **моменты** времени (даты, часы, начало или конец года, квартала, месяца, дня, ...).

t Периоды или моменты	t_1	t_2	t_3	...	t_{n-1}	t_n
у Значения показателя	y_1	y_2	y_3	...	y_{n-1}	y_n

Классификация рядов динамики



Число построенных квартир и их средний размер в РФ

	2011	2012	2013	2014	2015
Число квартир, тыс.	786	838	929	1124	1195
Число однокомнатных квартир, % от общего количества	36	38	39	41	43
Средний размер квартир, м ² общей площади	79,3	78,4	75,8	74,9	71,4

**Численность персонала фирмы в I полугодии 2016 г.
(на 1-е число месяца)**

Дата	1.01	1.02	1.03	1.04	1.05	1.06
Численность персонала, чел.	780	810	880	930	940	970

Динамика объема розничного товарооборота в регионе

Месяц	Январь	Февраль	Март	Апрель	Май	Июнь
Товарооборот, млн руб.	3,4	3,2	3,6	3,8	3,5	3,7

Выполняем задание: Определите вид рядов динамики, характеризующих изменение следующих статистических показателей:

Статистический показатель	Вид ряда динамики
1. Численность населения за ряд лет (по состоянию на начало каждого года)	
2. Среднемесячная заработная плата работников по отраслям экономики (по годам)	
3. Число предприятий и организаций в РФ за ряд лет (по состоянию на 1 янв.)	
4. Вклады населения в учреждения Сбербанка РФ за ряд лет (на конец каждого года)	
5. Число родившихся (по годам)	
6. Денежные доходы и расходы населения (по годам)	
7. Индекс потребительских цен на товары и услуги населению (на первое число месяца за ряд лет)	
8. Распределение розничного товарооборота по всем каналам реализации по формам собственности (по годам)	
9. Удельный вес новой товарной продукции машиностроения в общем объеме продукции (по годам)	

Сопоставимость уровней

Основные причины несопоставимости:

- различие в единицах измерения и единицах счета;
- различие в методологии учета или расчета показателей;
- изменение круга охватываемых объектов вследствие перехода ряда объектов из одного подчинения в другое;
- изменение территориальных границ областей, районов, округов.

Смыкание рядов динамики

Динамика малых предприятий одного из городов Российской Федерации

	1999 г.	2000 г.	2001 г.	2002 г.	2003 г.	2004 г.	2005 г.	2006 г.
Количество малых предприятий: По старой методологии отнесения к кругу малых предприятий	54	89	125	—	—	—	—	—
По новой методологии отнесения к кругу малых предприятий	—	—	90	138	169	194	206	220
Сомкнутый ряд динамики абсолютных величин (по новой методологии)	39 $54 \cdot 0,72$	64 $89 \cdot 0,72$	90	138	169	194	206	220
Сомкнутый ряд динамики относительных величин, в % к 2001 г.	43	71	100	153	188	215	229	244

$$90 : 125 = 0,72$$

Аналитические показатели ряда динамики

Аналитические
показатели динамики

Индивидуальные

Средние

Δ_i

цепные
базисные

Абсолютный прирост

$\bar{\Delta}$

Tr_i

цепные
базисные

Темп роста

\bar{Tr}

Trp_i

цепные
базисные

Темп прироста

\bar{Trp}

Формулы для расчета индивидуальных аналитических показателей динамики

Абсолютный прирост

$$\Delta_{\text{ц}} = Y_i - Y_{i-1}$$

$$\Delta_{\text{б}} = Y_i - Y_1$$

Темп роста

$$T_{\text{рц}} = \frac{y_i}{y_{i-1}} 100\%$$

$$T_{\text{рб}} = \frac{y_i}{y_0} 100\%$$

Темп прироста

$$T_{\text{прц}} = \frac{\Delta_{\text{ц}}}{y_{i-1}} = \frac{y_i}{y_{i-1}} 100\% - 100\% = T_{\text{рц}} - 100\%$$

$$T_{\text{прб}} = \frac{\Delta_{\text{б}}}{y_0} 100\% - 100\% = T_{\text{рб}} - 100\%$$

ЦЕПНЫЕ ПОКАЗАТЕЛИ

БАЗИСНЫЕ ПОКАЗАТЕЛИ

Пример расчета ИАПД

**Объем экспорта РФ со странами дальнего зарубежья в 2007-2011 гг.
(млрд.долл.США)**

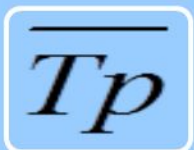
Год	Объем экспорта, млрд. долл. США	Абсолютный прирост, млрд.долл.США		Темп роста, %		Темп прироста, %		Абсолютное значение 1% прироста, млрд. долл.США
		по сравн. с предш. годом	по сравн. с 2007 г.	по сравн. с предш. годом	по сравн. с 2007 г.	по сравн. с предш. годом	по сравн. с 2007 г.	
	1	2	3	4	5	6	7	8
2007	300,0	-	-	-	-	-	-	-
2008	400,0	100,0	100,0	133,33	133,33	33,33	33,33	3,000
2009	255,0	-145,0	-45,0	63,75	85,00	-36,25	-15,00	4,000
2010	338,0	83,0	38,0	132,55	112,67	32,55	12,67	2,550
2011	438,0	100,0	138,0	129,59	146,00	29,59	46,00	3,380

Формулы для расчета средних аналитических показателей динамики



Средний абсолютный прирост

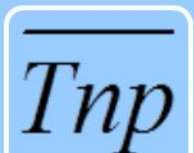
$$\bar{\Delta} = \frac{y_2 - y_1 + y_3 - y_2 + \dots + (y_n - y_{n-1})}{n-1} = \frac{y_n - y_1}{n-1}$$



Средний темп роста

$$\bar{T}_p = \sqrt[n-1]{\frac{y_2}{y_1} \cdot \frac{y_3}{y_2} \cdot \frac{y_4}{y_3} \cdot \dots \cdot \frac{y_n}{y_{n-1}}} \cdot 100\% = \sqrt[n-1]{\frac{y_n}{y_1}} \cdot 100\%$$

Для расчета \bar{T}_p применяется формула средней геометрической величины



Средний темп прироста

$$\bar{T}_{np} = \bar{T}_p - 100\%$$

Пример расчета САПД

Год	Объем экспорта, млрд. долл. США
2007	300
2008	400
2009	255
2010	338
2011	438

$$\bar{\Delta} = \frac{y_n - y_1}{n - 1} = \frac{438,0 - 300,0}{5 - 1} = \frac{138,0}{4} = 34,5$$

$$\bar{T}_p = \sqrt[4]{\frac{438,0}{300,0}} = \sqrt[4]{1,4600} = 1,0992 \quad \text{или } 109,92\%$$

$$\bar{T}_{np} = 109,92\% - 100\% = 9,92\%.$$

Простейшие методы прогнозирования

По среднему
абсолютному
приросту

$$\hat{y}_{n+k} = y_n + \bar{\Delta} \cdot k$$

По среднему
темпу роста

$$\hat{y}_{n+k} = y_n \cdot \overline{Tr}^k$$

Экстраполяция
тренда

$$y = f(t)$$

Продлить ряд
значений t и
подставить в
уравнение
тренда

Пример расчета прогнозных значений

Год	Объем экспорта, млрд. долл. США	$\hat{y}_{n+k} = y_n + \bar{\Delta} \cdot k$ $\bar{\Delta} = 34,5$	$\hat{y}_{n+k} = y_n \cdot T_p^k$ $T_p = 1,0992$
2007	300		
2008	400		
2009	255		
2010	338		
2011	438		
2012	$k=1$	472,5 = 438+34,5*1	481,45 =438*1,0992 ¹
2013	$k=2$	507,0 = 438+34,5*2	529,21 =438*1,0992 ²
2014	$k=3$	541,5 = 438+34,5*3	581,71 =438*1,0992 ³

Методы выявления тенденции в рядах динамики



Метод укрупнения интервалов

месяц	У	Квартальные уровни	Срнемесячные из покварт. уровней
1	13,3		-
2	13,4		13,40
3	13,5	40,2	
4	13,4		
5	13,6		13,60
6	13,8	40,8	
7	14,1		
8	14,2		14,13
9	14,1	42,4	
10	14,2		
11	14,6		14,57
12	14,9	43,7	-

Сглаживание ряда динамики нечетноуровневыми скользящими средними

месяц	Y	3-х уровн. скользящие суммы	3-х уровн. скользящие средние
1	13,3	-	-
2	13,4	-	13,40
3	13,5	40,2	13,43
4	13,4	40,3	13,50
5	13,6	40,5	13,60
6	13,8	40,8	13,83
7	14,1	41,5	14,03
8	14,2	42,1	14,13
9	14,1	42,4	14,17
10	14,2	42,5	14,30
11	14,6	42,9	14,57
12	14,9	43,7	-

Сглаживание ряда динамики четноуровневыми скользящими средними

месяц	Y	4-х уровн. скольз. суммы	4-х уровн. скользящие средние	4-х уровн. скольз. средние центр.
1	13,3			
2	13,4			
3	13,5			13,44
4	13,4	53,6	13,40	13,53
5	13,6	53,9	13,48	13,65
6	13,8	54,3	13,58	13,83
7	14,1	54,9	13,73	13,99
8	14,2	55,7	13,93	14,10
9	14,1	56,2	14,05	14,21
10	14,2	56,6	14,15	14,36
11	14,6	57,1	14,28	
12	14,9	57,8	14,45	

Сущность метода аналитического выравнивания

Аналитическое выравнивание предполагает представление уровней ряда динамики в виде функции времени:

$$Y = f(t).$$

$$\sum_{t=1}^n (y_t - \bar{y}_t)^2 \rightarrow \min,$$

где: y_t – фактическое значение уровня ряда динамики;
 \bar{y}_t – расчетное значение;
 n – количество уровней в ряду динамики.

Способы задания фактора времени t

Линейно

- y y_1 y_2 y_3 y_4 y_5 y_6 y_7
- t 1 2 3 4 5 6 7

Методом
условного
нуля

- n - нечетное
- y y_1 y_2 y_3 y_4 y_5 y_6 y_7
- t -3 -2 -1 0 1 2 3
- n - четное
- y y_1 y_2 y_3 y_4 y_5 y_6
- t -5 -3 -1 1 3 5

Расчет параметров линейного уравнения тренда

$$\bar{y}_t = a_0 + a_1 t$$

Фактор t задан линейно

$$\begin{cases} na_0 + a_1 \sum t = \sum y \\ a_0 \sum t + a_1 \sum t^2 = \sum yt \end{cases}$$

Фактор t задан методом условного нуля

$$\begin{cases} a_0 n = \sum y \\ a_1 \sum t^2 = \sum ty \end{cases}$$

$$\begin{cases} a_0 = \frac{\sum y}{n} \\ a_1 = \frac{\sum ty}{\sum t^2} \end{cases}$$

Расчет параметров параболического уравнения тренда

$$y_t = a_0 + a_1 t + a_2 t^2$$

Фактор t задан
линейно

Фактор t задан методом условного
нуля

$$\begin{cases} na_0 + a_1 \sum t + a_2 \sum t^2 = \sum y \\ a_0 \sum t + a_1 \sum t^2 + a_2 \sum t^3 = \sum yt \\ a_0 \sum t^2 + a_1 \sum t^3 + a_2 \sum t^4 = \sum yt^2 \end{cases}$$

$$\begin{cases} a_0 n + a_2 \sum t^2 = \sum y \\ a_1 \sum t^2 = \sum ty \\ a_0 \sum t^2 + a_2 \sum t^4 = \sum t^2 y \end{cases}$$

Критерий выбора модели

- **Лучше описывает тенденцию развития та модель, у которой расчетные значения максимально близки к исходным наблюдаемым значениям уровней.**

Соответственно, выбирается та модель, при которой

$$\sum_{t=1}^n (y_t - \bar{y}_t)^2 \rightarrow \min,$$

Эту модель можно использовать для целей прогнозирования. Как правило, при построении модели для целей прогнозирования фактор времени задается линейно.

Спасибо за Ваше внимание!

Желаю Вам успехов!

С уважением, Е.В.Улитина