

ЭКОНОМЕТРИКА

БРС эконометрика

- 1. Модель парной линейной регрессии
- 2. Модель множественной линейной регрессии
- 3. Итоговый коллоквиум
- 4. Итоговая контрольная работа
- 5. Модель временного ряда
- 6. Презентация

5

10

10

5

10

10

ИТОГ

Эконометрика — это наука, в которой на базе реальных статистических данных

- строятся,
- анализируются и
- совершенствуются

математические модели реальных экономических явлений.

*Эконометрическая модель –
главный инструмент
эконометрических исследований*

- В эконометрике модель относится к классу математических моделей.
- Модели экономических объектов, создаваемые именно в эконометрике, являются *дескриптивными* (описывающими реальность такой, как она есть, в отличие от оптимизационных моделей).

Этапы построения эконометрических моделей

- постановка задачи; набора показателей, взаимосвязи между которыми нас интересуют;
- априорный, предмодельный анализ содержательной сущности моделируемого явления;
- информационно-статистический этап – получение данных, анализ их качества;
- этап спецификации модели;
- *этап параметризации модели – определяем вид эконометрической модели, выражаем в математической форме взаимосвязь между её переменными, формулируем исходные предпосылки и ограничения модели.*
- этап идентификации модели – статистический анализ модели, оценка качества её параметров по имеющимся статистическим данным;
- этап верификации модели, анализа её точности и адекватности;
- интерпретация полученных результатов.

Три основных класса моделей

Регрессионные модели с одним уравнением

Модели временных рядов

Системы одновременных уравнений

Виды зависимости

- **Функциональной** называется взаимосвязь, при которой каждому значению одного показателя соответствует строго определенное значение другого.
- **Статистической** (стохастической, вероятностной) называется взаимосвязь, при которой одному значению первого показателя может соответствовать несколько значений второго показателя.
- Среди статистических зависимостей наибольший интерес представляют корреляционные зависимости.
- **Корреляционная** зависимость заключается в том, что средняя величина одного показателя изменяется в зависимости от значения другого.
- **Корреляционной зависимостью** между двумя переменными величинами называется функциональная зависимость между значениями одной из них и условным математическим ожиданием (средним значением) другой.

Модель парной линейной регрессии

- 1. Постройте поле корреляции и сформулируйте гипотезу о форме связи.
- 2. Рассчитайте параметры выборочного уравнения линейной регрессии с помощью МНК.
- 3. Оцените тесноту связи с помощью показателей корреляции (выборочный коэффициент корреляции) и детерминации.
- 4. Используя критерий Стьюдента оцените статистическую значимость коэффициентов регрессии и корреляции.
- 5. Постройте интервальные оценки параметров регрессии. Проверьте, согласуются ли полученные результаты с выводами, полученными в предыдущем пункте.

Модель парной линейной регрессии

- 6. Постройте таблицу дисперсионного анализа для оценки значимости уравнения в целом.
- 7. С помощью теста Гольдфельда – Кванта исследуйте гетероскедастичность остатков. Сделайте выводы.
- 8. В случае пригодности линейной модели рассчитайте прогнозное значение результата, если значение фактора увеличится на 5% от его среднего уровня. Определите доверительный интервал прогноза для уровня значимости $=0,05$.
- 9. Оцените полученные результаты, проинтерпретируйте полученное уравнение регрессии.

Количество пропусков	Результаты теста
3	10
4	8
5	7
5	6
6	4
6	5
7	4
7	2
8	3
9	1

Оценка параметров регрессии МНК

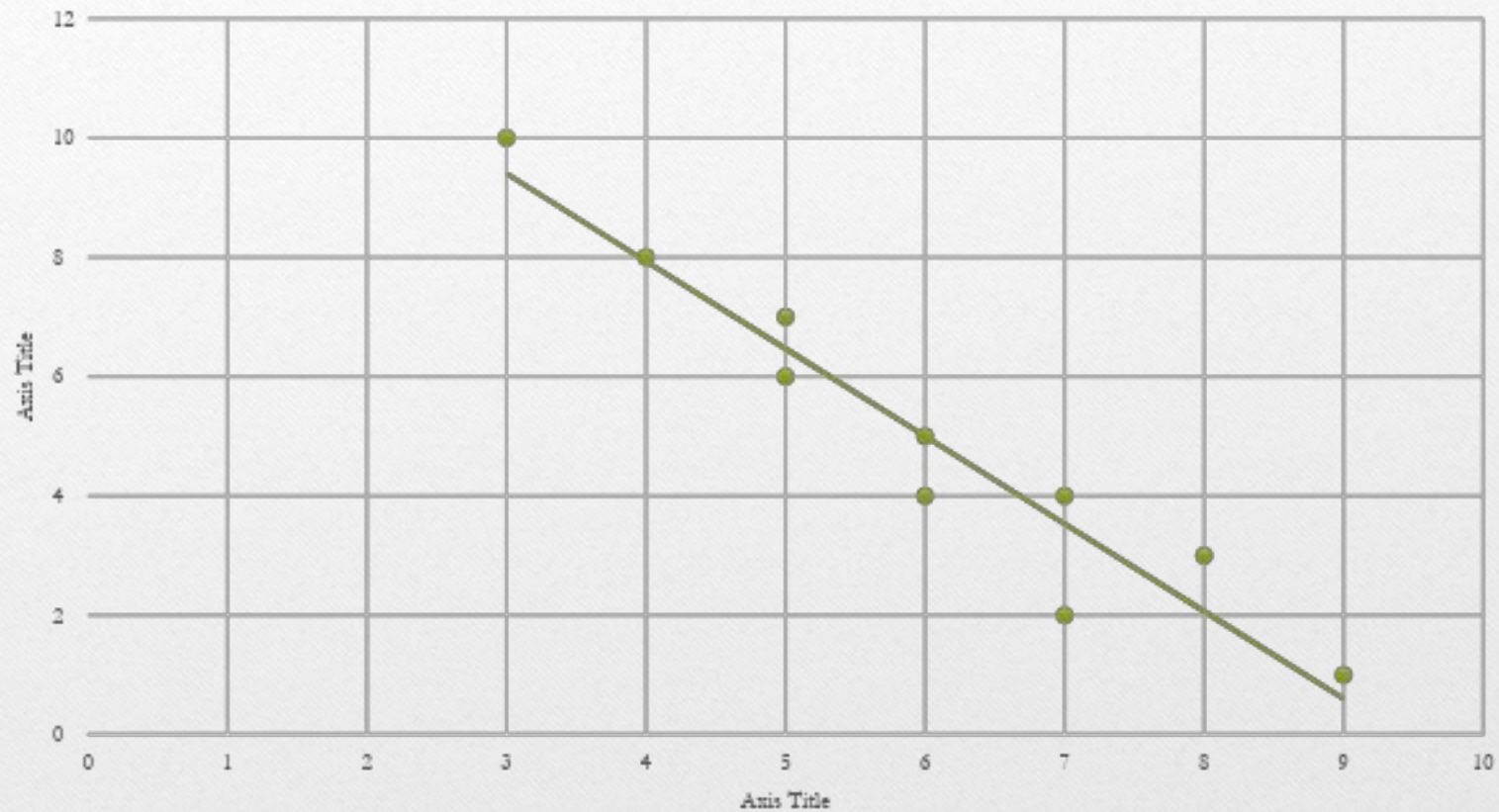
$$\hat{y}_i = b_0 + b_1 x_i$$

$$b_0 = \bar{y} - b_1 \bar{x}$$

$$b_1 = \frac{\text{cov}(x, y)}{\sigma_x^2} = \frac{\overline{x \cdot y} - \bar{y} \cdot \bar{x}}{\overline{x^2} - \bar{x}^2}$$

$$\hat{y} = -1,47x + 13,8$$

результаты теста



Оцените тесноту связи с помощью показателей корреляции (выборочный коэффициент корреляции) и детерминации

- Коэффициент корреляции

$$r_{xy} = b_1 \frac{\sigma_x}{\sigma_y} = \frac{\text{cov}(x, y)}{\sigma_x \cdot \sigma_y} = \frac{\overline{x \cdot y} - \bar{y} \cdot \bar{x}}{\sigma_x \cdot \sigma_y}$$

$$r_{xy} = \frac{n \cdot \sum_{i=1}^n y_i x_i - \sum_{i=1}^n y_i \cdot \sum_{i=1}^n x_i}{\sqrt{(n \cdot \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2) \cdot (n \cdot \sum_{i=1}^n y_i^2 - (\sum_{i=1}^n y_i)^2)}}$$

Коэффициент детерминации R^2

- Коэффициент детерминации характеризует долю дисперсии результативного признака Y , объясняемую регрессией, в общей дисперсии результативного признака. Соответственно величина $1 - R^2$ характеризует долю дисперсии Y , вызванную влиянием остальных, не учтенных в модели факторов.

- Коэффициент детерминации
$$R^2 = 1 - \frac{SS_{ост}}{SS_{общ}} = \frac{SS_R}{SS_{общ}}$$

- $R^2 = r_{xy}^2$

Используя критерий Стьюдента оцените статистическую значимость коэффициентов регрессии и корреляции

- $H_0: b_1 = 0$
- $H_1: b_1 \neq 0$

$$t = \frac{b_1}{S_{b_1}}$$

Статистика критерия при нулевой гипотезе имеет распределение Стьюдента с $(n-2)$ степенями свободы

$$S_{b_1} = \frac{S}{\sqrt{\sum (x_i - \bar{x})^2}}$$

Стандартная ошибка коэффициента регрессии

$$S^2 = \frac{\sum e_i^2}{n - 2}$$

Несмещенная оценка дисперсии случайных отклонений

Если вычисленное значение t – *статистики* - **|tфакт|** при заданном уровне значимости α больше критического (табличного) **t табл** , т.е.

$$|\mathbf{tфакт}| > \mathbf{t табл} = t(\alpha ; n-2),$$

то гипотеза $H_0 : b_1 = 0$, отвергается в пользу альтернативной при выбранном уровне значимости. Это подтверждает статистическую значимость коэффициента регрессии b_1 .

МНК

$$\hat{y} = -1,47x + 13,8$$

Уравнение парной
линейной регрессии

-1,46667

13,8

Стандартная
ошибка
коэффициента b_1

0,150923

0,942514

Стандартная ошибка
коэффициента b_0

Коэффициент
детерминации

0,921905

0,82664

Стандартная
ошибка
регрессии

94,43902

8

64,53333

5,466667

Регрессионная статистика	
Множественный R	0,960159
R-квадрат	0,921905
Нормированный R-квадрат	0,912143
Стандартная ошибка	0,82664
Наблюдения	10

ЗНАК!!!

Коэффициент корреляции

Коэффициент детерминации

Дисперсионный анализ	Число степеней свободы	Сумма квадратов отклонений			
	df	SS	MS	F	Значимость F
Регрессия	1	64,53333	64,53333	94,43902	1,05E-05
Остаток	8	5,466667	0,683333		
Итого	9	70			

	Коэффициенты	Стандартная ошибка	t-статистика	P-Значение	Нижние 95%	Верхние 95%
У-пересечение	13,8	0,942514	14,64169	4,65E-07	11,62656	15,97344
Количество пропусков	-1,46667	0,150923	-9,71797	1,05E-05	-1,8147	-1,11864

ВЫВОД ОСТАТКА		
Наблюдение	Предска занное результ аты теста	Остатки
1	9,4	0,6
2	7,933333	0,066667
3	6,466667	0,533333
4	6,466667	-0,46667
5	5	-1
6	5	0
7	3,533333	0,466667
8	3,533333	-1,53333
9	2,066667	0,933333
10	0,6	0,4

Количество пропусков График остатков



fx
Вставить функцию

Σ
Автосумма

Последние

Финансовые

Логические

Текстовые

Дата и время

Ссылки и массивы

Математические

Другие функции

Диспетчер имен

Присвоить имя

Использовать в формуле

Создать из выделенного

Влияющие ячейки

Зависимые ячейки

Показать формулы

Проверка наличия ошибок

Вычислить формулу

Убрать стрелки

Вычислить формулу

Библиотека функций

Определенные имена

Зависимости формул

D2 =ТЕНДЕНЦИЯ(B2:B11;A2:A11;F2;1)+ТЕНДЕНЦИЯ(B2:B11;A2:A11;F2;1)

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
1	Количество пропусков	результаты теста																			
2	3	10		ТЕНДЕНЦИЯ(B2:B11;A2:A11;F2;1)		6,3															
3	4	8																			
4	5	7																			
5	5	6																			
6	6	4																			
7	6	5																			
8	7	4																			
9	7	2																			
10	8	3																			
11	9	1																			
12	6	5																			
13	6,3																				
14																					
15																					
16																					
17																					
18																					
19																					
20																					
21																					
22																					
23																					
24																					
25																					
26																					
27																					
28																					
29																					
30																					
31																					
32																					
33																					
34																					

Аргументы функции

ТЕНДЕНЦИЯ

Известные_значения_y B2:B11 = {10;8;7;6;4;5;4;2;3;1}

Известные_значения_x A2:A11 = {3;4;5;5;6;6;7;7;8;9}

Новые_значения_x F2 = {6,3}

Конст 1 = ИСТИНА

= {4,56}

Возвращает значения в соответствии с линейной аппроксимацией по методу наименьших квадратов.

Конст логическое значение: константа в вычисляется обычным образом при значении ИСТИНА или отсутствии значения и равна 0 при значении ЛОЖЬ.

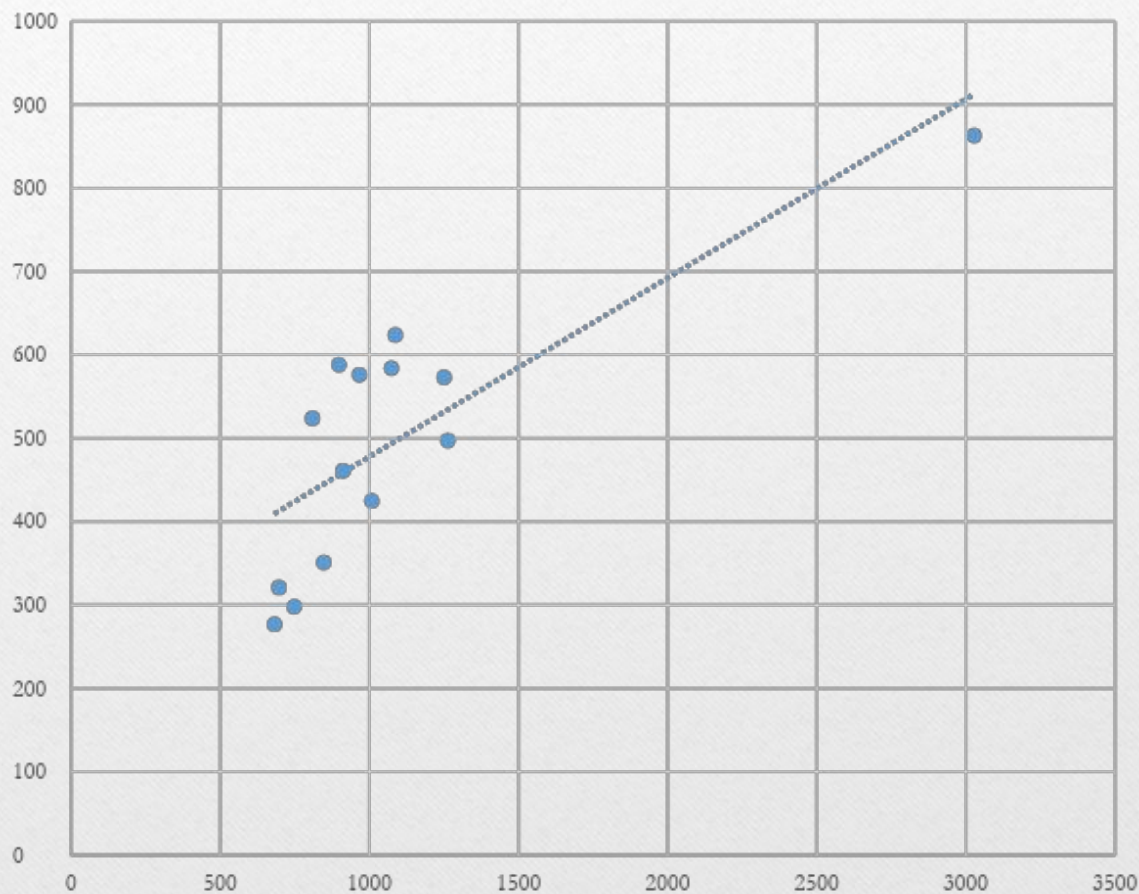
Значение: 9,12

[Справка по этой функции](#)

OK Отмена

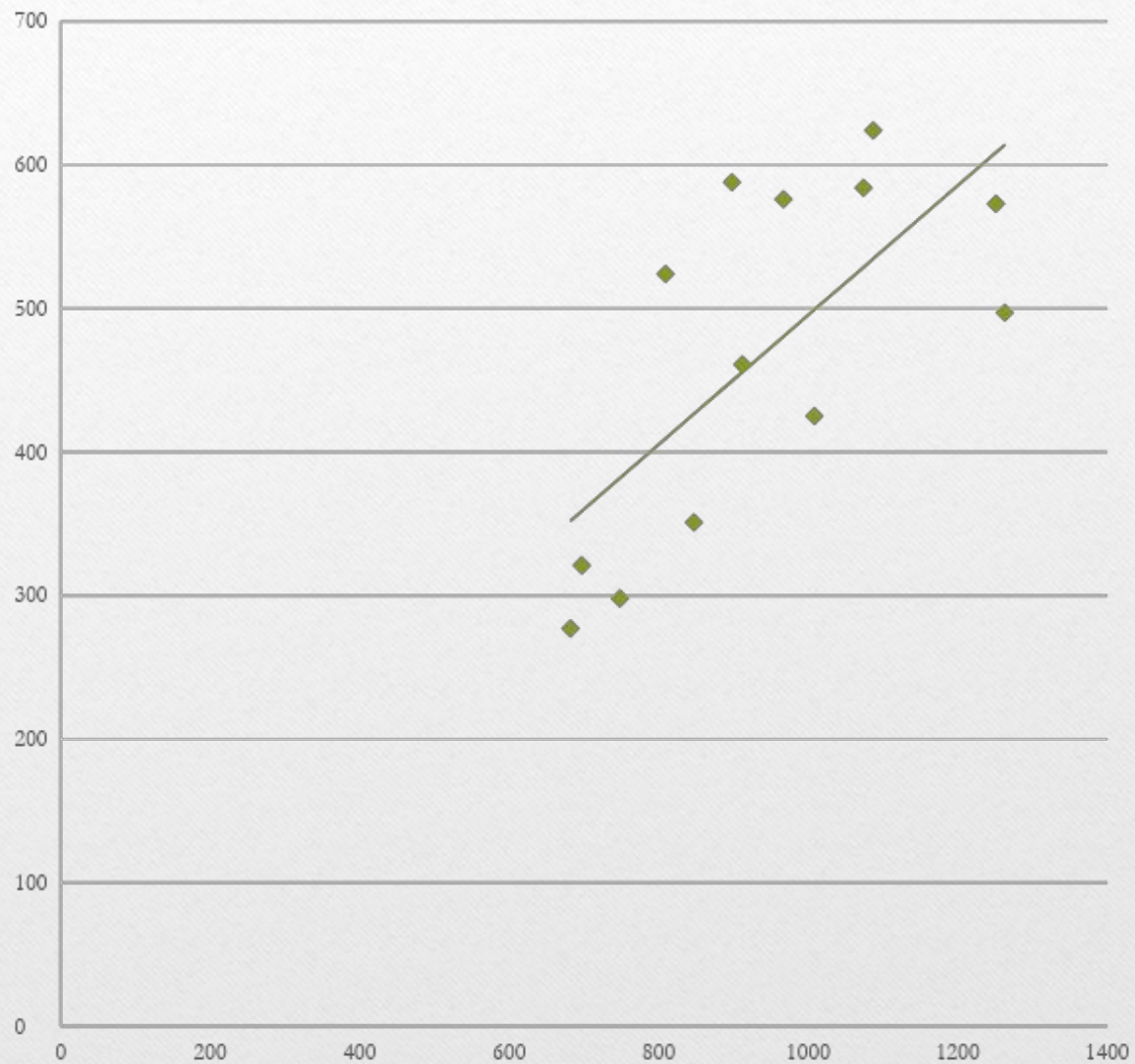
средняя зарботная плата	потребите льские расходы
912	461
809	524
748	298
847	351
1087	624
1074	584
1008	425
682	277
697	321
1251	573
967	576
898	588
1263	497
3027	863

потребительские расходы

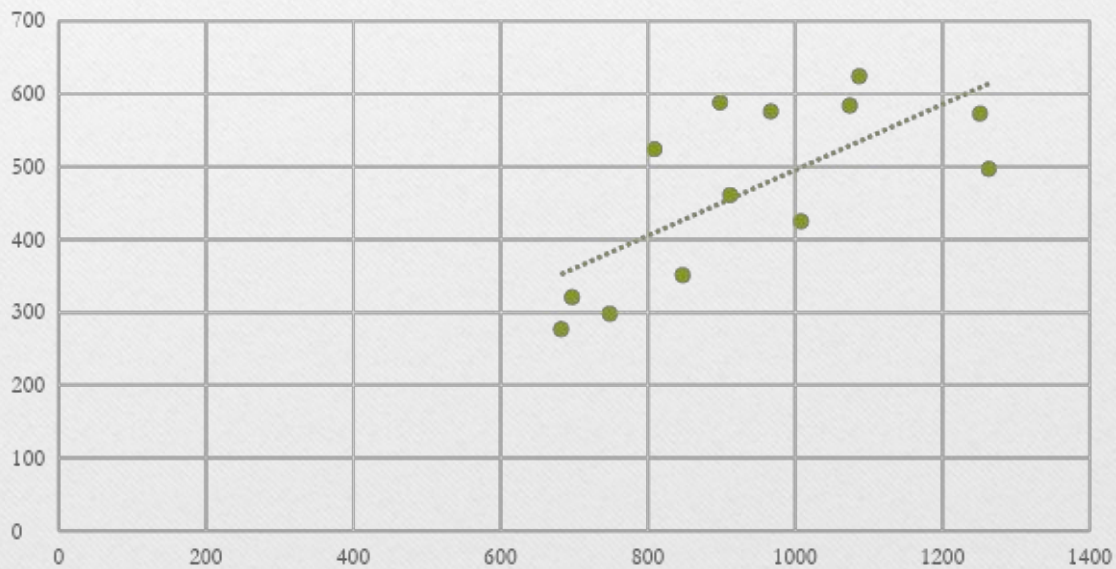
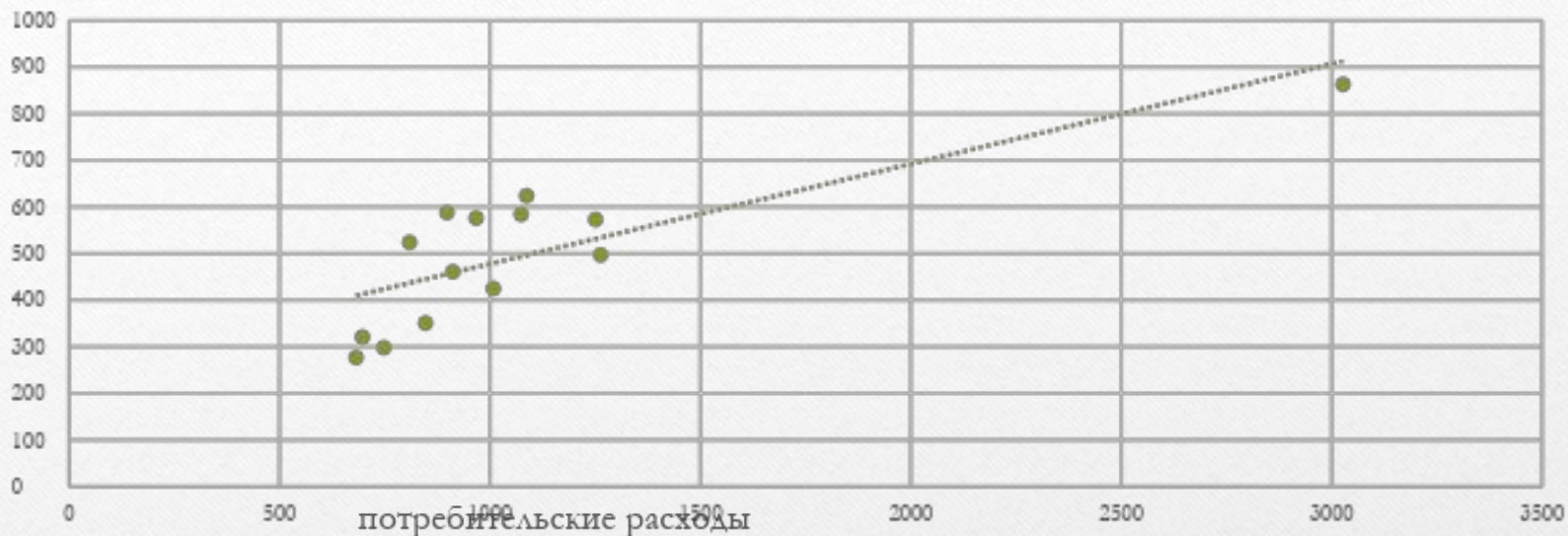


средняя зарботная плата	потребител ьские расходы
912	461
809	524
748	298
847	351
1087	624
1074	584
1008	425
682	277
697	321
1251	573
967	576
898	588
1263	497
3027	863

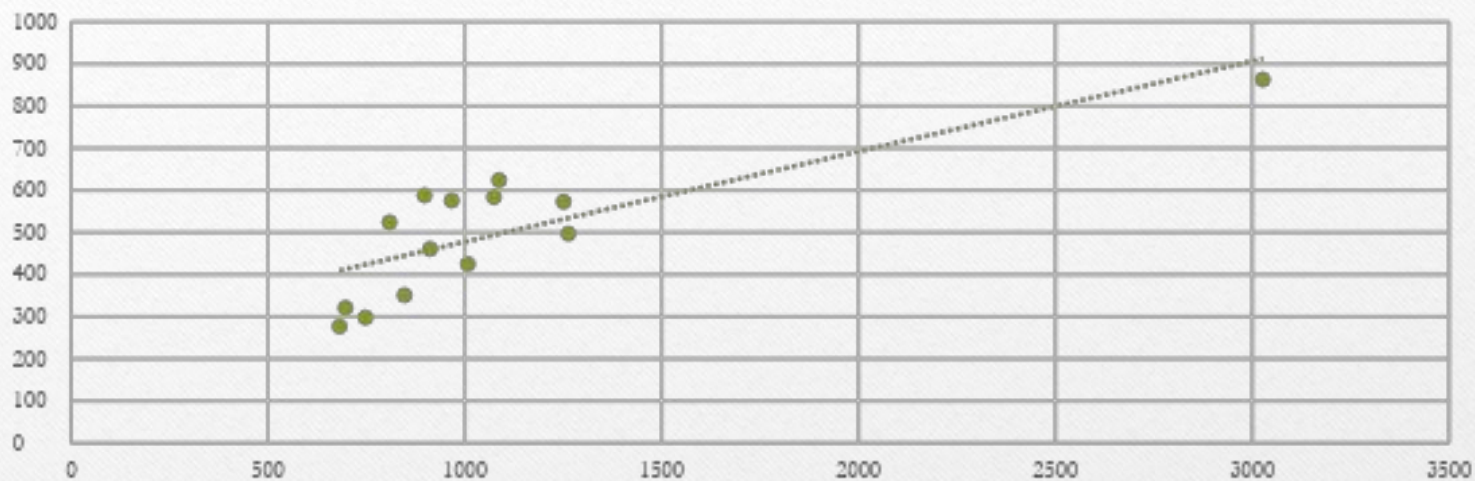
потребительские расходы



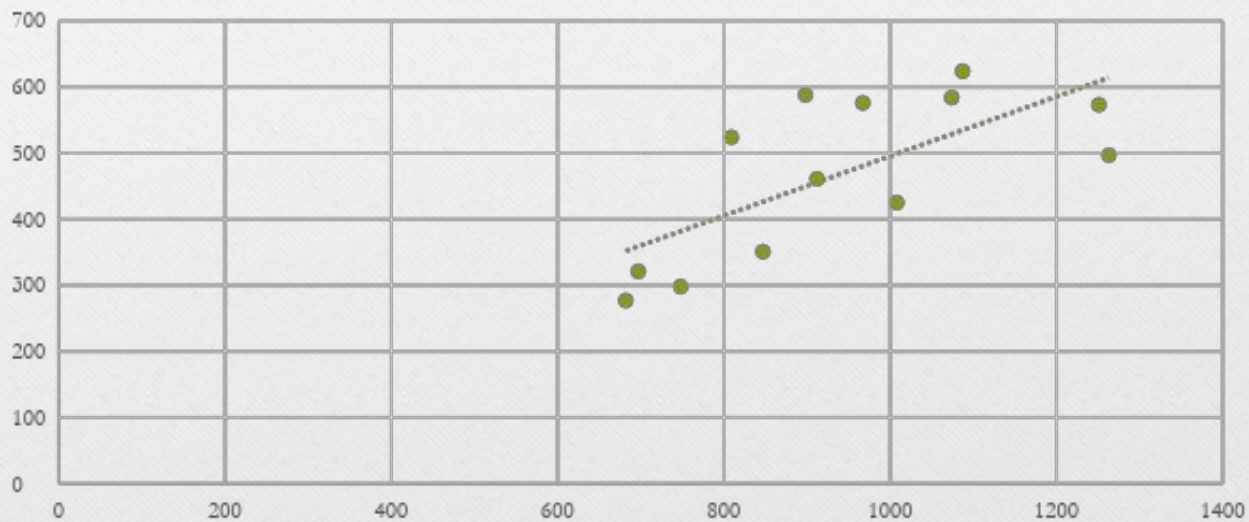
потребительские расходы



потребительские расходы



потребительские расходы

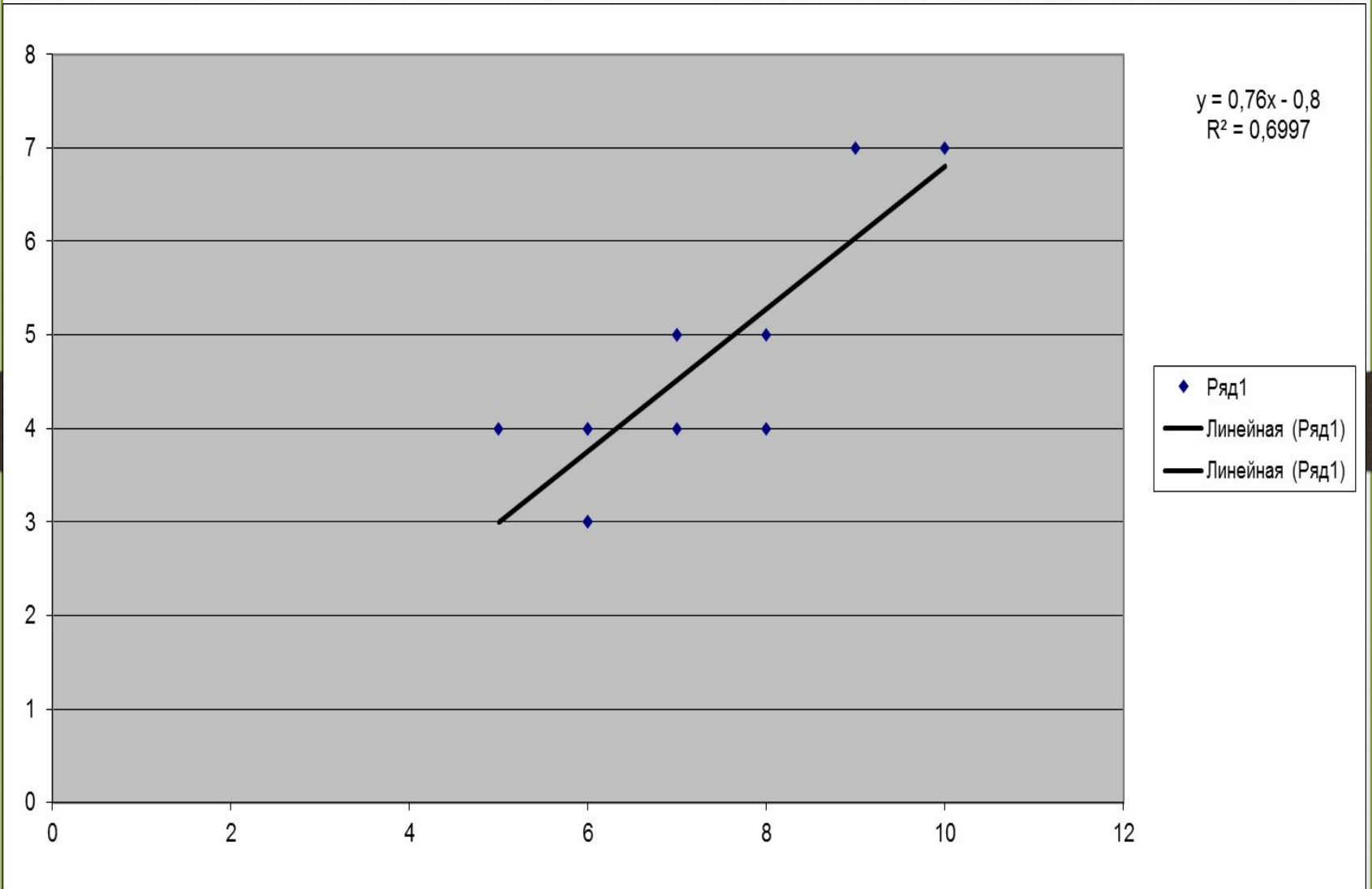


ПРИМЕР



10	7
6	4
8	4
8	5
6	4
7	5
6	3
7	4
9	7
6	3
5	4
7	5

x	y
5	4
6	4
6	4
6	3
6	3
7	5
7	4
7	5
8	4
8	5
9	7
10	7



$$y = 0,76x - 0,8$$
$$R^2 = 0,6997$$

- ◆ Ряд1
- Линейная (Ряд1)
- Линейная (Ряд1)

	y	x					<i>Остатки</i>
1	4	5	20	16	25	6,8	0,2
2	4	6	24	16	36	3,76	0,24
3	4	6	24	16	36	5,28	-1,28
4	3	6	18	9	36	5,28	-0,28
5	3	6	18	9	36	3,76	0,24
6	5	7	35	25	49	4,52	0,48
7	4	7	28	16	49	3,76	-0,76
8	5	7	35	25	49	4,52	-0,52
9	4	8	32	16	64	6,04	0,96
10	5	8	40	25	64	3,76	-0,76
11	7	9	63	49	81	3	1
12	7	10	70	49	100	4,52	0,48
Σ	55	85	407	271	625	55	0

0,76	-0,8
0,157434	1,136182
0,699736	0,753658
23,30399	10
13,23667	5,68

Значение коэффициента b_1	Значение коэффициента b_0
<p data-bbox="440 315 807 425">Стандартная ошибка коэффициента b_1 –</p> $S_{b_1} = \frac{S}{\sqrt{\sum (x_i - \bar{x})^2}}$	<p data-bbox="1136 315 1503 425">Стандартная ошибка коэффициента b_0 –</p> $S_{b_0} = \frac{S \sqrt{\sum x_i^2}}{\sqrt{n \sum (x_i - \bar{x})^2}}$
<p data-bbox="343 618 904 661">Коэффициент детерминации R^2</p> $R^2 = 1 - \frac{SS_{ост}}{SS_{общ}} = \frac{SS_R}{SS_{общ}}$	<p data-bbox="1029 618 1609 661">Стандартная ошибка регрессии –</p> $S^2 = \frac{\sum e_i^2}{n-2}$
<p data-bbox="498 853 749 889">F - статистика</p> $F = \frac{MS_R}{MS_{ост}} = \frac{\sum (\hat{y}_i - \bar{y})^2 / 1}{\sum (y_i - \hat{y}_i)^2 / (n-2)}$	<p data-bbox="1107 853 1532 889">Число степеней свободы</p> <p data-bbox="1290 925 1348 961">n-2</p>
<p data-bbox="334 1061 913 1103">Регрессионная сумма квадратов –</p> $SS_R = \sum_i (\hat{y}_i - \bar{y})^2$	<p data-bbox="1049 1061 1590 1103">Остаточная сумма квадратов –</p> $SS_{ост} = \sum_i (y_i - \hat{y}_i)^2$

ВЫВОД ИТОГОВ	
<i>Регрессионная статистика</i>	
Множественный R	0,836502052
R-квадрат	0,699735683
Нормированный R-квадрат	0,669709251
Стандартная ошибка	0,753657747
Наблюдения	12

Дисперсионный анализ					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Значимость F</i>
Регрессия	1	13,23666667	13,23666667	23,30399061	0,000694458
Остаток	10	5,68	0,568		
Итого	11	18,91666667			

	Коэффициент β_1	Стандартная ошибка	t-статистика	P-Значение	Нижние 95%	Верхние 95%
Y-пересечение	-0,8	1,136181804	-0,704112667	0,497431151	-3,331570809	1,731571
X	0,76	0,157433969	4,8274207	0,000694458	0,409215259	1,110785

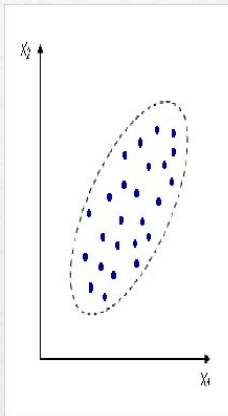
ВЫВОД ОСТАТКА		
Наблюдение	Предсказанное Y	Остатки
1	6,8	0,2
2	3,76	0,24
3	5,28	-1,28
4	5,28	-0,28
5	3,76	0,24
6	4,52	0,48
7	3,76	-0,76
8	4,52	-0,52
9	6,04	0,96
10	3,76	-0,76
11	3	1
12	4,52	0,48

Постройте поле корреляции и сформулируйте гипотезу о форме связи

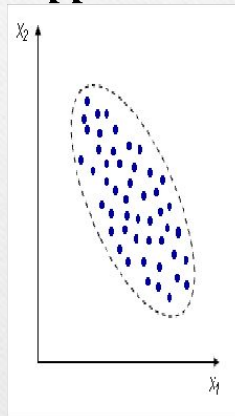
- Пусть имеется два ряда эмпирических данных
- $X (x_1, x_2, \dots, x_n)$ и
- $Y (y_1, y_2, \dots, y_n)$, соответствующие им точки с координатами (x_i, y_i) , где $i=1, 2, \dots, n$, отобразим на координатной плоскости. Такое изображение называется *полем корреляции*.
- Визуальный анализ корреляционного поля позволяет сделать предположение о форме взаимосвязи двух исследуемых показателей.
- По форме взаимосвязи корреляционные зависимости принято разделять на линейные и нелинейные.

Парная корреляция

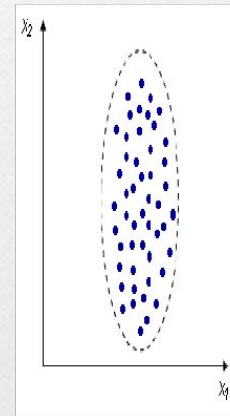
Форма корреляционной зависимости



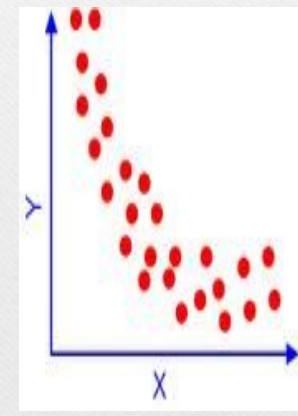
линейная положительная



линейная отрицательная



отсутствует



нелинейная