

Параболическое уравнение парной регрессии

Равноускоренное изменение результативного признака y при равномерном изменении фактора x говорит о параболической связи между признаками.

При этом в некоторой точке связь может поменять свое направление: прямая связь перейти в обратную и наоборот.

При *парной* параболической зависимости уравнение регрессии имеет вид:

$$\bar{y}_x = a_0 + a_1x + a_2x^2$$

Параметры уравнения находят из следующей системы нормальных уравнений:

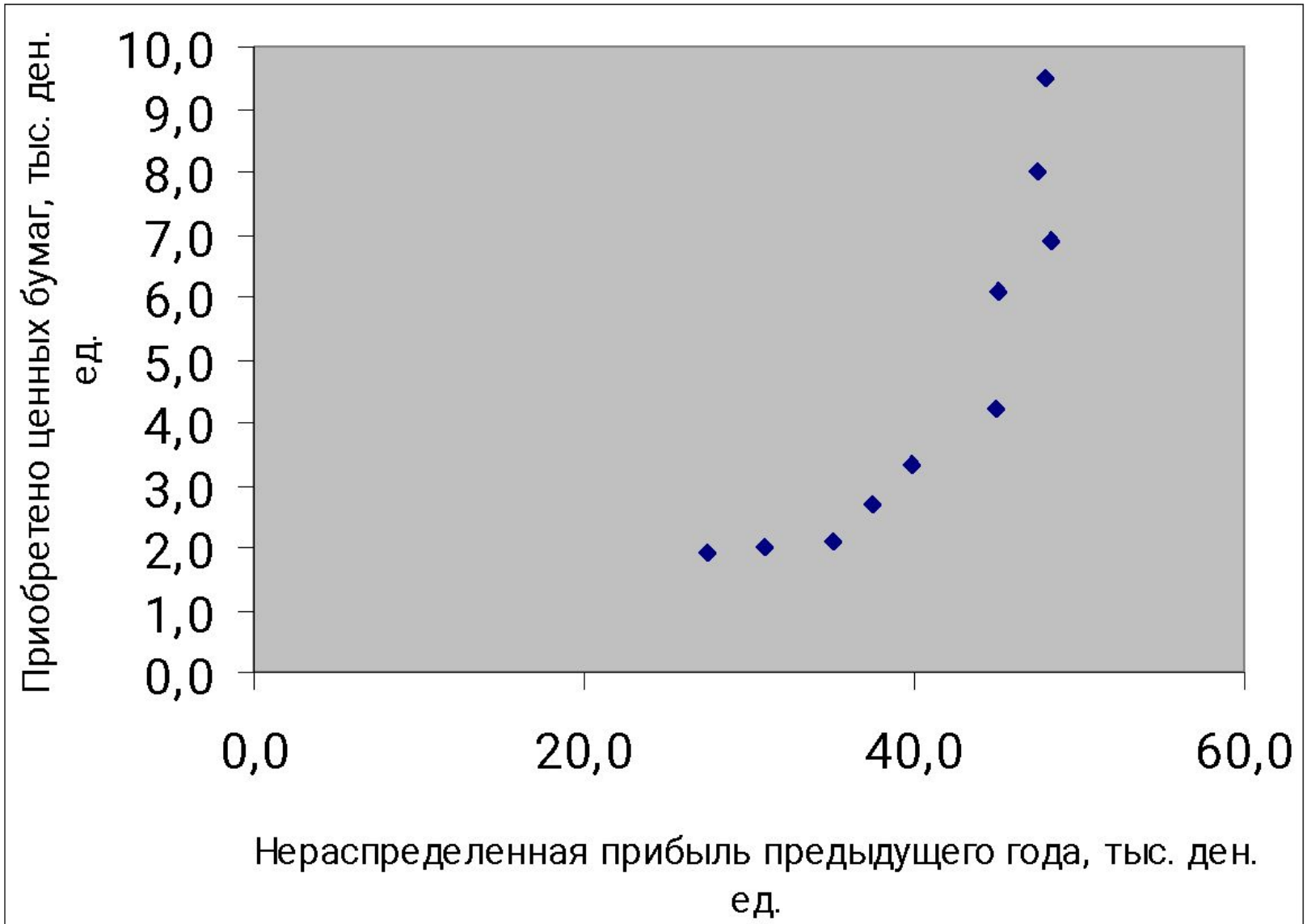
$$\begin{cases} na_0 + a_1 \sum x + a_2 \sum x^2 = \sum y, \\ a_0 \sum x + a_1 \sum x^2 + a_2 \sum x^3 = \sum yx, \\ a_0 \sum x^2 + a_1 \sum x^3 + a_2 \sum x^4 = \sum yx^2. \end{cases}$$

Пример.

Имеются следующие финансовые данные по 10 предприятиям промышленности (тыс. ден. ед):

Номер предприятия	<i>Нераспределенная прибыль предыдущего года</i>	<i>Приобретено ценных бумаг</i>
1	27,4	1,9
2	31,0	2,0
3	35,0	2,1
4	37,4	2,7
5	39,8	3,3
6	44,9	4,2
7	45,0	6,1
8	48,2	6,9
9	47,5	8,0
10	48,0	9,5

Анализ построенного корреляционного поля показывает целесообразность применения в качестве регрессионной модели уравнения параболы 2-ого порядка:



Номер предприятия	x	y	x^2	x^3	x^4	xy	x^2y	\bar{y}_x
1	27,4	1,9	751	20571	563641	52	1426	-0,61974
2	31,0	2,0	961	29791	923521	62	1922	1,2735
3	35,0	2,1	1225	42875	1500625	74	2573	3,1035
4	37,4	2,7	1399	52314	1956530	101	3777	4,06326
5	39,8	3,3	1584	63045	2509183	131	5227	4,91934
6	44,9	4,2	2016	90519	4064296	189	8467	6,39426
7	45,0	6,1	2025	91125	4100625	275	12353	6,4185
8	48,2	6,9	2323	111980	5397444	333	16030	7,09914
9	47,5	8,0	2256	107172	5090664	380	18050	6,966
10	48,0	9,5	2304	110592	5308416	456	21888	7,062
Сумма	404,2	46,7	16844	719983	31414944	2052	91713	46,7

Составим систему нормальных уравнений:

$$\begin{cases} 10a_0 + 404,2a_1 + 16844a_2 = 46,7 \\ 404,2a_0 + 16844a_1 + 719983a_2 = 2052 \\ 16844a_0 + 719983a_1 + 31414944a_2 = 91713 \end{cases}$$

Решив ее, получаем значения для оценок параметров регрессии:

$$a_0 = -22,674; a_1 = 1,052; a_2 = -0,009.$$

Искомое уравнение параболы имеет вид:

$$\bar{y}_x = -22,674 + 1,052x - 0,009x^2$$

Подставляя в полученное уравнение значения факторной переменной x , рассчитаем выравненные значения \bar{y}_x

Насколько «хорошо» данное уравнение описывает взаимосвязь признаков можно судить, рассчитав теоретическое корреляционное отношение:

$$\eta_{теор.} = \sqrt{\frac{\sum (\bar{y}_x - \bar{y})^2}{\sum (y - \bar{y})^2}} = \sqrt{\frac{65,3}{68,8}} = 0,97.$$

Как видим, связь между признаками сильная. Теоретический коэффициент детерминации, соответственно

$$\eta^2_{теор.} = \frac{65,3}{68,8} = 0,949,$$

говорит о том, что построенное параболическое уравнение описывает вариацию зависимой переменной на 94,9 %.

Гиперболическое уравнение парной регрессии

Гиперболическая зависимость имеет место, если связь между признаками обратная.

Обратная связь также может описываться линейным уравнением с отрицательным значением коэффициента регрессии. По расположению точек на корреляционном поле можно выбрать окончательный вид уравнения.

При выборе гиперболы факторный признак не может принимать нулевое значение.

Уравнение гиперболы при парной зависимости между признаками имеет вид:

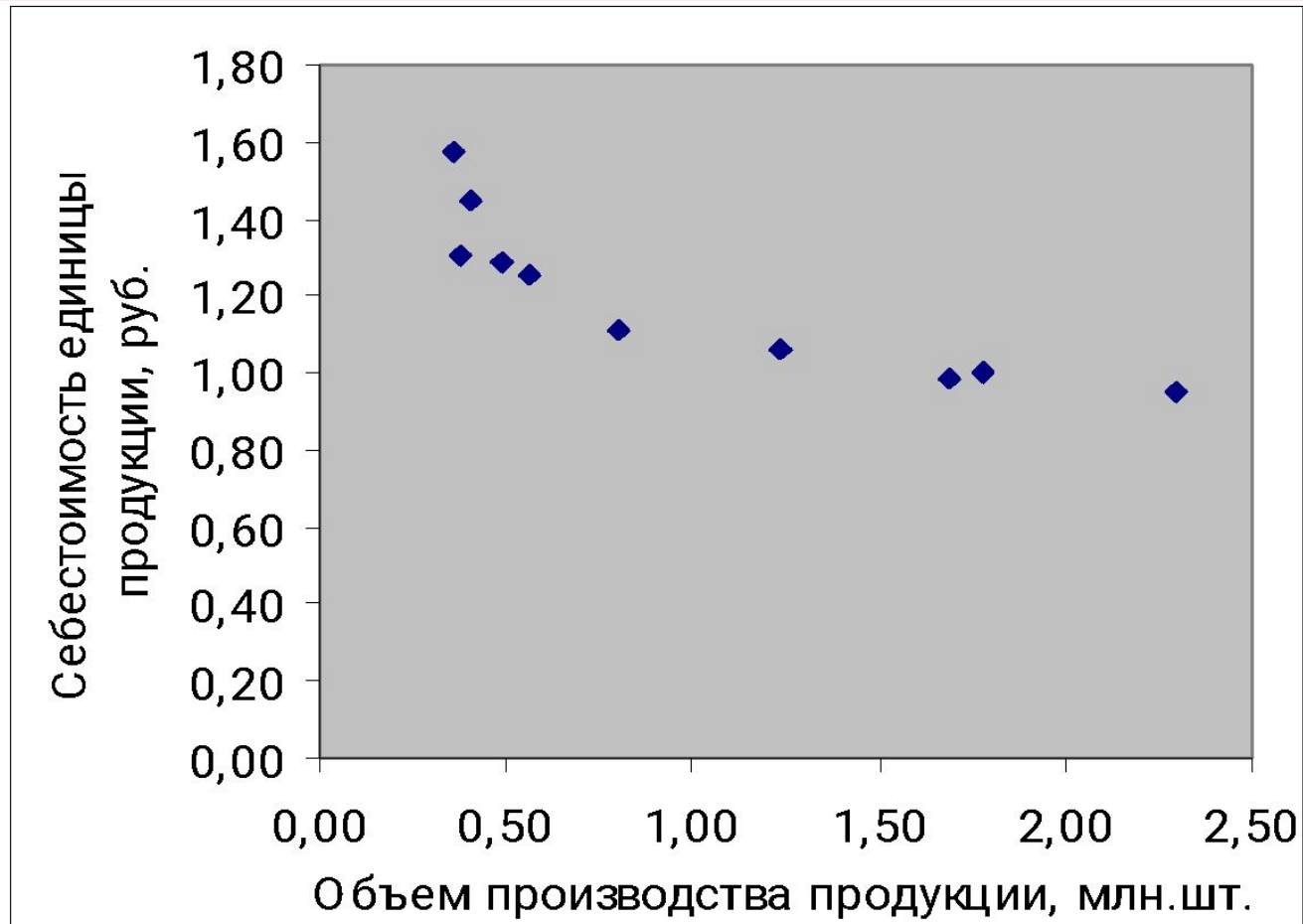
$$\bar{y}_x = a_0 + \frac{a_1}{x}$$

Для оценки параметров гиперболы в случае парной зависимости используется следующая система нормальных уравнений:

$$\begin{cases} na_0 + a_1 \sum \frac{1}{x} = \sum y, \\ a_0 \sum \frac{1}{x} + a_1 \sum \frac{1}{x^2} = \sum \frac{y}{x}. \end{cases}$$

Пример

Имеются данные по 10 предприятиям промышленности об объеме произведенной продукции определенного вида и ее себестоимости. Между признаками имеется обратная зависимость, что подтверждается специфическим расположением точек на корреляционном поле:



<i>№ пред.</i>	<i>Объем произ. продукции, млн.шт.</i>	<i>Себестои- мость ед. прод., руб.</i>	$\frac{1}{x}$	$\frac{1}{x^2}$	$\frac{y}{x}$	\bar{y}_x
1	0,36	1,57	2,778	7,7160	4,3694	1,47
2	0,41	1,45	2,439	5,9488	3,5366	1,40
3	0,38	1,30	2,632	6,9252	3,4211	1,44
4	0,49	1,29	2,041	4,1649	2,6327	1,31
5	0,56	1,25	1,786	3,1888	2,2321	1,25
6	0,80	1,11	1,250	1,5625	1,3875	1,13
7	1,24	1,06	0,806	0,6504	0,8548	1,04
8	1,78	1,00	0,562	0,3156	0,5618	0,98
9	1,69	0,98	0,592	0,3501	0,5799	0,99
10	2,30	0,95	0,435	0,1890	0,4130	0,95
Итого	10,01	11,96	15,32	31,011	19,989	11,96

$$\begin{cases} 10 \cdot a_0 + 15,32 \cdot a_1 = 11,96 \\ 15,32 \cdot a_0 + 31,0114 \cdot a_1 = 19,9889. \end{cases}$$

$a_0 = 0,857512$ и $a_1 = 0,220945$.

Таким образом, искомое уравнение гиперболы имеет вид:

$$\bar{y}_x = 0,857512 + \frac{0,220945}{x}$$

Теоретический коэффициент детерминации равен:

$$\eta_{теор.}^2 = \frac{\sum (\bar{y}_x - \bar{y})^2}{\sum (y - \bar{y})^2} = \frac{0,3682}{0,4007} = 0,9189$$

Полученное уравнение описывает вариацию зависимой переменной на 91,89 %.