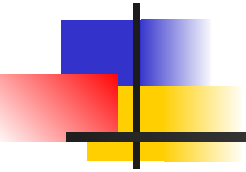


Парный регрессионный анализ






План:

1. Парный регрессионный анализ
2. Метод наименьших квадратов
3. Пример

Основные задачи статистического анализа

- 
- Выявление наличия или отсутствия взаимосвязи между изучаемыми факторами (корреляционный анализ)
 - Определение вида взаимосвязи между изучаемыми факторами (регрессионный анализ)
 - Проверка гипотезы о виде взаимосвязи между факторами



Парный регрессионный анализ

Корреляционный анализ показывает, что две переменные связаны друг с другом, однако он не дает представления о том, каким образом они связаны.

Для определения вида зависимости между переменными проводят регрессионный анализ

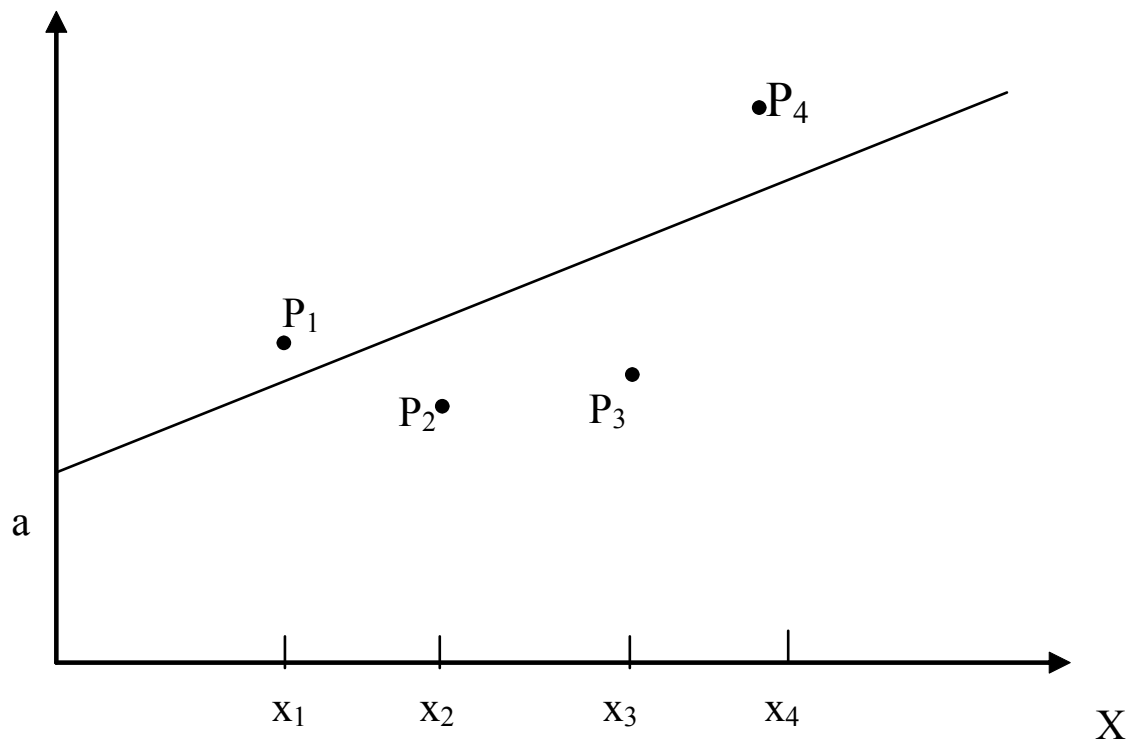


Регрессионный анализ

Пусть изучается зависимость между факторами **X** и **Y**. В результате **n** независимых опытов получены **n** пар чисел:

$$(x_1; y_1), (x_2; y_2), \dots, (x_n; y_n)$$

X	x_1	x_2	...	x_n
Y	y_1	y_2	...	y_n



Рассмотрим более подробно случай, когда одна переменная зависит от другой. Рассмотрим простейшую, т.е. линейную модель


$$y = \alpha + \beta x + u$$

где x – независимая переменная;

y – зависимая переменная;

α и β – истинные значения параметров регрессии;

u – случайная составляющая



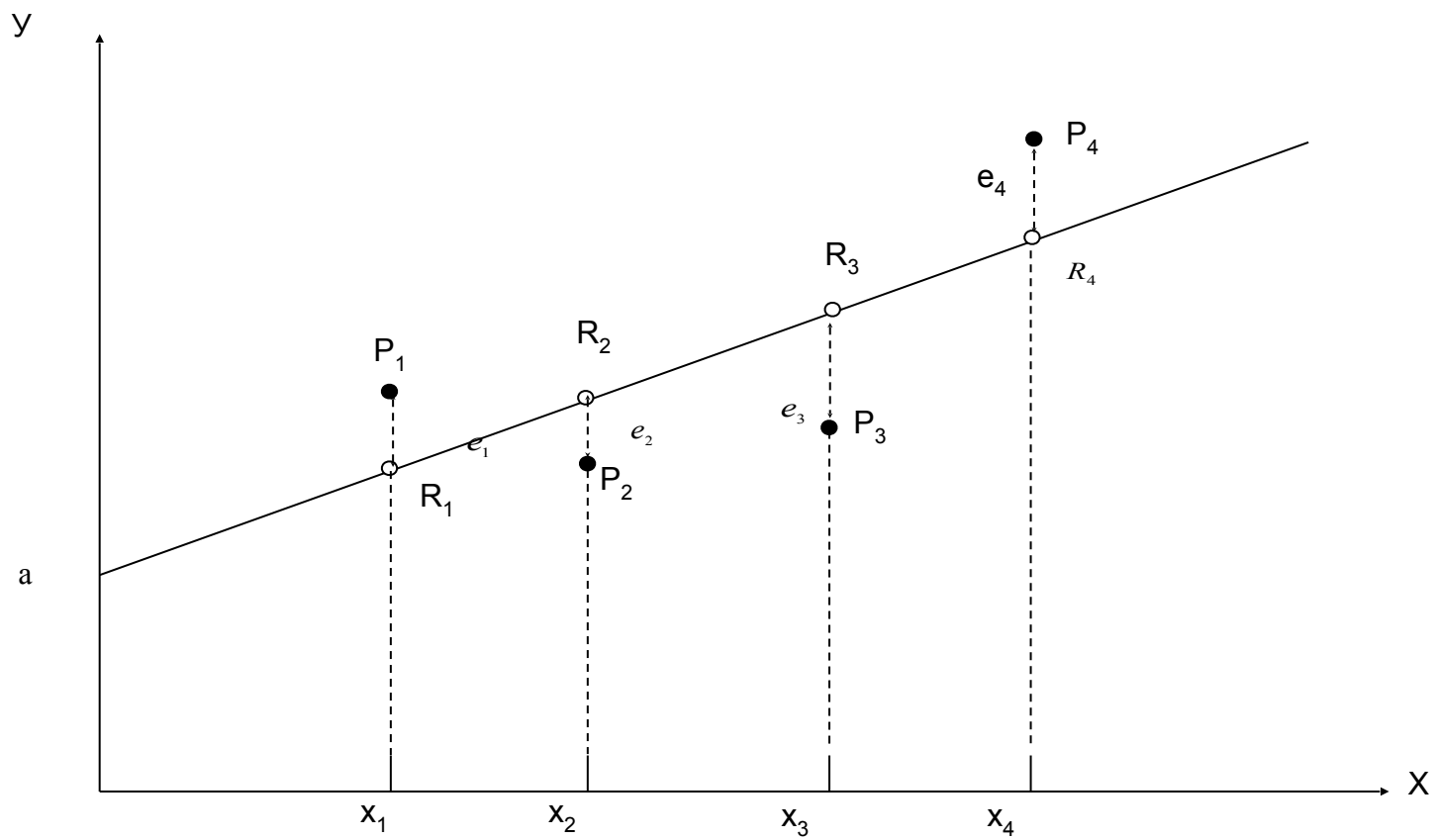
Оценка параметров по методу наименьших квадратов

Допустим, что мы имеем четыре наблюдения для X и Y .

Мы хотим построить линию регрессии таким образом, чтобы отклонения были минимальными

Уравнение этой прямой:

$$Y = a + bx$$



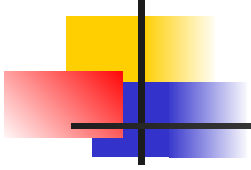
Один из способов решения поставленной проблемы состоит в минимизации суммы квадратов отклонений

$$S(a, b) = \sum_i (y_i - Y_i)^2 = \sum_{i=1}^n (y_i - a - bx_i)^2 \rightarrow \min$$



Система нормальных уравнений


$$\begin{cases} \frac{\partial S}{\partial a} = -2 & \sum_{i=1}^n (y_i - a - bx_i) = 0 \\ \frac{\partial S}{\partial b} = -2 & \sum_{i=1}^n (y_i - a - bx_i)x_i = 0 \end{cases}$$



Решение системы

$$a = \frac{\sum x^2 \sum y - \sum x \sum xy}{n \sum x^2 - (\sum x)^2}$$
$$b = \frac{n \sum xy - \sum x \sum y}{n \sum x^2 - (\sum x)^2}$$

Пример



№	Бонитет(X1)	Удобрения(X2)	Урожайность(Y)
1	31	75	22
2	34	72	23
3	40	79	23
4	44	81	24
5	51	83	25
6	56	81	25
7	62	90	27
8	64	95	29
9	69	100	31
10	75	95	33
11	81	110	34
12	83	115	36
13	88	110	35
14	95	120	36
15	98	130	37

Получено следующее уравнение
регрессии:

$$Y = 0,92 + 0,29x$$

Полученный результат можно истолковать следующим образом. Если количество вносимых удобрений увеличить на 1 единицу, то урожайность увеличится в среднем на 0,29 ц/га. Если $x = 0$, то прогнозируемый уровень урожайности равняется 0,92 ц/га



СПАСИБО ЗА ВНИМАНИЕ!



+ 998 71 237 1948



smirzaev@tiiame.uz