

Практическое занятие Регрессионный анализ

Задание 1.

Вариант 2. Рассчитать и построить график уравнения прямолинейной регрессии и сделать вывод о точности расчета уравнений, если данные выборки таковы:

x_i , уд/мин ~ 162; 174; 168; 174; 168; 156; 162; 162; 144; 174; 162; 144.

y_i , уд/мин ~ 72; 60; 60; 78; 72; 60; 84; 78; 66; 66; 60; 66.

Методические рекомендации к расчету.

1. Занести данные тестирования в рабочую таблицу и сделать соответствующие расчеты.

x_i	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$	y_i	$(y_i - \bar{y})$	$(y_i - \bar{y})^2$	$(x_i - \bar{x}) \cdot (y_i - \bar{y})$
1						
2						
\bar{x}	$\sum (x_i - \bar{x})$	$\sum (x_i - \bar{x})^2$	\bar{y}	$\sum (y_i - \bar{y})$	$\sum (y_i - \bar{y})^2$	$\sum (x_i - \bar{x})(y_i - \bar{y})$

2. Рассчитать значение нормированного коэффициента корреляции по формуле:

$$r_{xy} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \cdot \sum (y_i - \bar{y})^2}}$$

3. Рассчитать конечный вид уравнений прямолинейной регрессии по формулам

$$y_x = \bar{y} + \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} \cdot (x - \bar{x})$$

$$x_y = \bar{x} + \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (y_i - \bar{y})^2} \cdot (y - \bar{y})$$

4. Рассчитать абсолютные погрешности уравнений регрессии по формулам

$$\sigma_{y/x} = \sigma_y \cdot \sqrt{1 - r^2} = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}} \cdot \sqrt{1 - r^2}$$

$$\sigma_{x/y} = \sigma_x \cdot \sqrt{1 - r^2} = \sqrt{\frac{\sum (y_i - \bar{y})^2}{n-1}} \cdot \sqrt{1 - r^2}$$

5. Рассчитать относительные погрешности уравнений регрессии по формулам

$$\sigma_{y/x}^I = \frac{\sigma_{y/x}}{\bar{y}} \cdot 100\%$$

$$\sigma_{x/y}^I = \frac{\sigma_{x/y}}{\bar{x}} \cdot 100\%$$

6. Сделать вывод.

Решение:**Решение:**

1. Вычислим среднее арифметическое значение для выборок:

$$(\bar{x}) = (162 + 174 + 168 + 174 + 168 + 156 + 162 + 162 + 144 + 174 + 162 + 144) / 12 = 162,5$$

$$(\bar{y}) = (72 + 60 + 60 + 78 + 72 + 60 + 84 + 78 + 66 + 66 + 60 + 66) / 12 = 68,5$$

Используя полученные значения вычисляем параметры, необходимые для расчёта величин, перечисленных в пунктах со 2-го по 5й:

x_i	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$	y_i	$(y_i - \bar{y})$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})(y_i - \bar{y})$
162	-0,5	0,25	72	3,5	12,25	-1,75
174	11,5	132,25	60	-8,5	72,25	-97,75
168	5,5	30,25	60	-8,5	72,25	-46,75
174	11,5	132,25	78	9,5	90,25	109,25
168	5,5	30,25	72	3,5	12,25	19,25
156	-6,5	42,25	60	-8,5	72,25	55,25
162	-0,5	0,25	84	15,5	240,25	-7,75
162	-0,5	0,25	78	9,5	90,25	-4,75
144	-18,5	342,25	66	-2,5	6,25	46,25
174	11,5	132,25	66	-2,5	6,25	-28,75
162	-0,5	0,25	60	-8,5	72,25	4,25
144	-18,5	342,25	66	-2,5	6,25	46,25
\bar{x}	$\sum (x_i - \bar{x})$	$\sum (x_i - \bar{x})^2$	\bar{y}	$\sum (y_i - \bar{y})$	$\sum (y_i - \bar{y})^2$	$\sum (x_i - \bar{x})(y_i - \bar{y})$
162,5	0	1185	68,5	0	753	93

2. Нормированный коэффициент корреляции Пирсона, рассчитанный по вышеприведённым данным:

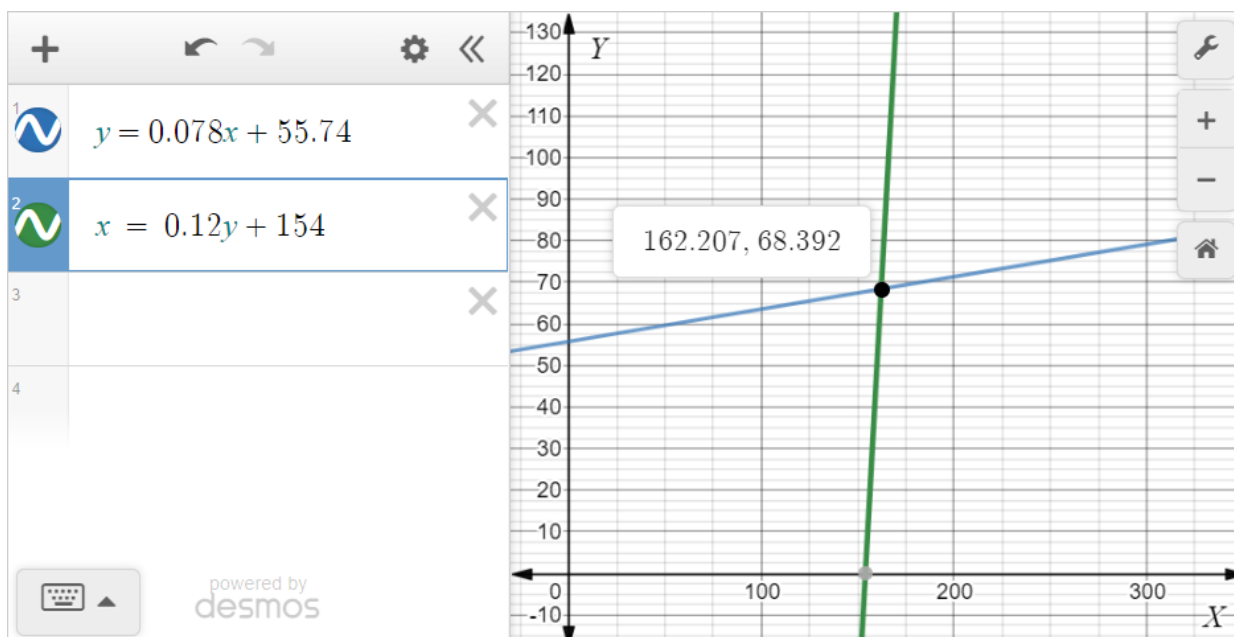
$$r_{xy}^P = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \cdot \sum (y_i - \bar{y})^2}} = 0,09$$

3. Конечный вид уравнений прямолинейной регрессии:

$$y_x = \bar{y} + \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} \cdot (x - \bar{x}) = 0.078x + 55,74$$

$$x_y = \bar{x} + \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (y_i - \bar{y})^2} \cdot (y - \bar{y}) = 0.12y + 154$$

График уравнения прямолинейной регрессии:



4. Абсолютные погрешности уравнений регрессии:

$$\sigma_{y/x} = \sigma_y \cdot \sqrt{1 - r^2} = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}} \cdot \sqrt{1 - r^2} = 10,3$$

$$\sigma_{x/y} = \sigma_x \cdot \sqrt{1 - r^2} = \sqrt{\frac{\sum (y_i - \bar{y})^2}{n-1}} \cdot \sqrt{1 - r^2} = 8,2$$

5. Относительные погрешности уравнений регрессии:

$$\sigma_{y/x}^I = \frac{\sigma_{y/x}}{\bar{y}} \cdot 100\% = 15\%$$

$$\sigma_{x/y}^I = \frac{\sigma_{x/y}}{\bar{x}} \cdot 100\% = 5\%$$

6. Вывод: построенный график регрессии свидетельствует о крайне слабой зависимости исследуемых пар значений, поскольку его линии пересекаются под углом, близким к прямому.

Рассчитанный коэффициент корреляции Пирсона = 0,09 также свидетельствует о крайне слабой корреляции пар значений.

Полученные значения относительных погрешностей уравнений регрессии свидетельствуют о приемлемой точности прогноза (не хуже 15%) значений одного показателя по заранее известным значениям другого.