

РЕШЕНИЕ ЗАДАЧ НА ПОСТРОЕНИЕ ЛИНЕЙНОЙ РЕГРЕССИИ

ЗАДАНИЕ.

Распределение 100 новых видов тарифов на сотовую связь всех известных мобильных систем X (ден. ед.) и выручка от них Y (ден.ед.) приводится в таблице:

$x \backslash y$	10	15	20	25	30	35	Итого
20	4	2					6
30		5	3				8
40			5	45	5		55
50			2	8	7		17
60				4	7	3	14
Итого	4	7	10	57	19	3	100

Необходимо:

- 1) Вычислить групповые средние \bar{x}_i и \bar{y}_j и построить эмпирические линии регрессии;
- 2) Предполагая, что между переменными X и Y существует линейная корреляционная зависимость:
 - А) найти уравнения прямых регрессии, построить их графики на одном чертеже с эмпирическими линиями регрессии и дать экономическую интерпретацию полученных уравнений;
 - Б) вычислить коэффициент корреляции, на уровне значимости $\alpha = 0,05$ оценить его значимость и сделать вывод о тесноте и направлении связи между переменными X и Y ;
 - В) используя соответствующее уравнение регрессии, оценить среднюю выручку от мобильных систем с 20 новыми видами тарифов.

РЕШЕНИЕ.

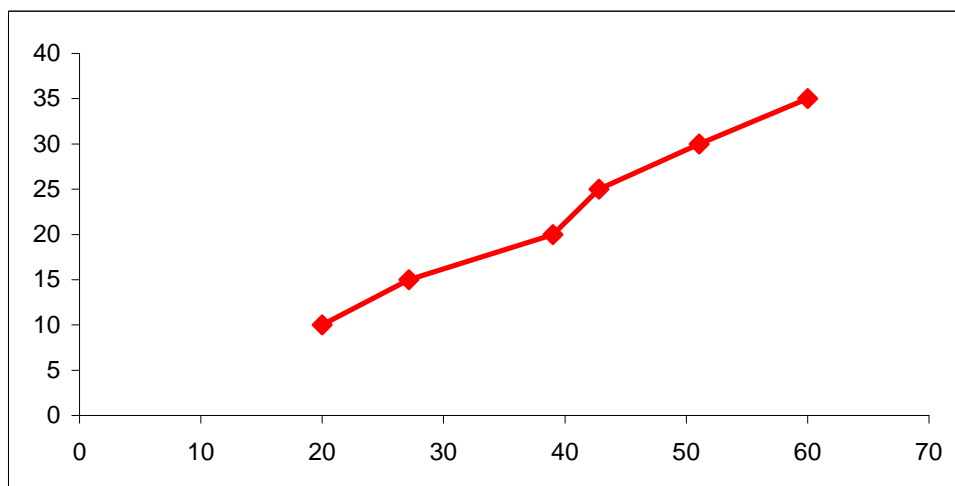
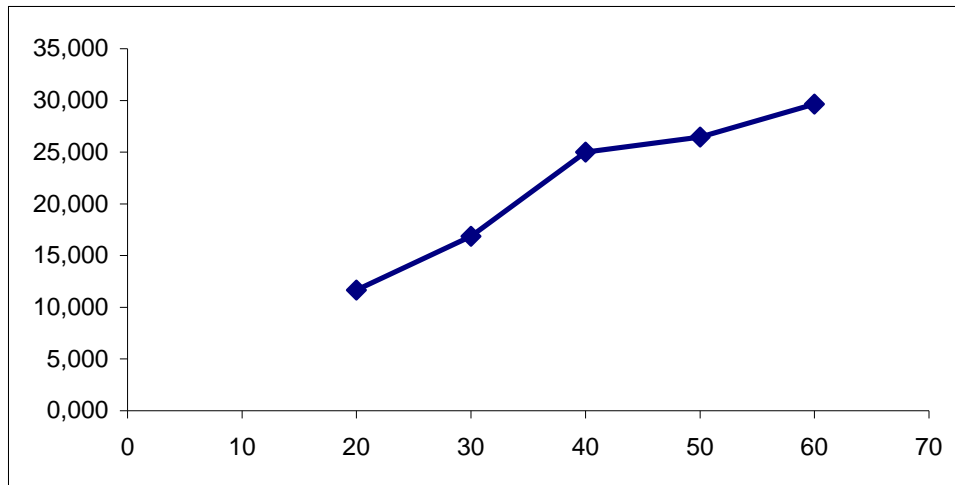
- 1) Найдем групповые средние по формулам: $\bar{x}_j = \frac{\sum_{i=1}^5 x_i n_{ij}}{n_j}$; $\bar{y}_i = \frac{\sum_{j=1}^6 y_j n_{ij}}{n_i}$.

Вычисления проведем в Excel, получаем:

\bar{x}_j	20	27,14	39	42,81	51,05	60

\bar{y}_i
11,667
16,875
25,000
26,471
29,643

Построим эмпирические линии регрессии (Y на X , X на Y).



Из вида эмпирических линий регрессии можно заключить, что между переменными наблюдается линейная зависимость.

Найдем уравнения прямых линий регрессии. Вычислим необходимые величины (расчеты в таблицах ниже):

x_i	20	30	40	50	60	Сумма
n_i	6	8	55	17	14	100
$x_i \cdot n_i$	120	240	2200	850	840	4250
$x_i^2 \cdot n_i$	2400	7200	88000	42500	50400	190500

y_j	10	15	20	25	30	35	Сумма
-------	----	----	----	----	----	----	--------------

n_j	4	7	10	57	19	3	100
$y_j \cdot n_j$	40	105	200	1425	570	105	2445
$y_j^2 \cdot n_j$	400	1575	4000	35625	17100	3675	62375

$$\sum_{i=1}^5 x_i n_i = 4250, \quad \sum_{i=1}^5 x_i^2 n_i = 190500,$$

$$\bar{x} = \frac{\sum_{i=1}^5 x_i n_i}{n} = \frac{4250}{100} = 42,5,$$

$$s_x^2 = \frac{\sum_{i=1}^5 x_i^2 n_i}{n} - \bar{x}^2 = \frac{190500}{100} - 42,5^2 = 98,75$$

$$\sum_{j=1}^6 y_j n_j = 2445, \quad \sum_{j=1}^6 y_j^2 n_j = 62375,$$

$$\bar{y} = \frac{\sum_{j=1}^6 y_j n_j}{n} = \frac{2445}{100} = 24,45,$$

$$s_y^2 = \frac{\sum_{j=1}^6 y_j^2 n_j}{n} - \bar{y}^2 = \frac{62375}{100} - 24,45^2 = 25,9475.$$

$$\sum_{i=1}^5 \sum_{j=1}^6 x_i y_j n_{ij} = 107850$$

$$\mu = \bar{xy} - \bar{x} \cdot \bar{y} = \frac{\sum_{i=1}^5 \sum_{j=1}^6 x_i y_j n_{ij}}{n} - \bar{x} \cdot \bar{y} = \frac{107850}{100} - 42,5 \cdot 24,45 = 39,375$$

$$b_{yx} = \frac{\mu}{s_x^2} = \frac{39,375}{98,75} = 0,3987$$

$$b_{xy} = \frac{\mu}{s_y^2} = \frac{39,375}{24,9475} = 1,5175$$

Уравнения прямых регрессии:

$$y_x - \bar{y} = b_{yx} (x - \bar{x}),$$

$$y_x - 24,45 = 0,3987(x - 42,5),$$

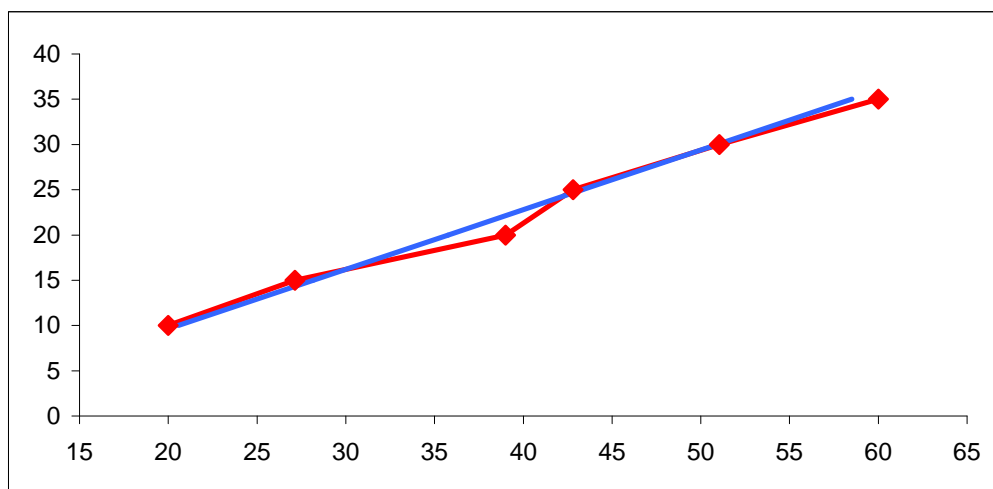
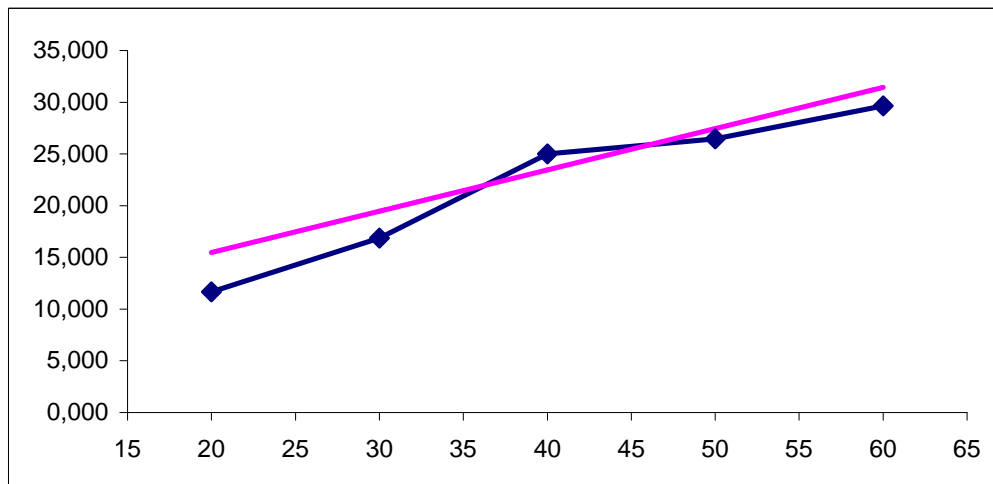
$$y_x = 0,3987x + 7,5053.$$

$$x_y - \bar{x} = b_{xy}(y - \bar{y}),$$

$$x_y - 42,5 = 1,5175(y - 24,45),$$

$$x_y = 1,5175y + 5,3971.$$

Построим графики линий регрессии на одном чертеже с эмпирическими линиями регрессии.



Экономическая интерпретация полученных уравнений:

$y_x = 0,3987x + 7,5053$ - при увеличении тарифа на одну денежную единицу, выручка растет в среднем на 0,3987 ден. ед.

$x_y = 1,5175y + 5,3971$ - при увеличении выручки на одну денежную единицу, тариф растет в среднем на 1,5775 ден. ед.

Вычислим коэффициент корреляции $r = \sqrt{b_{yx} b_{xy}} = \sqrt{0,3987 \cdot 1,5175} \approx 0,778$

На уровне значимости $\alpha = 0,05$ оценим значимость коэффициента корреляции. Вычислим значение критерия $t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} = \frac{0,778\sqrt{100-2}}{\sqrt{1-0,778^2}} \approx 12,259$

По таблице критерия Стьюдента для уровня значимости 0,05 находим $t_{0,95;98} = 1,985$. Так как наблюдаемое значение 12,259 больше критического, коэффициент корреляции значим.

Связь между переменными X и Y тесная, прямая.

Используя соответствующее уравнение регрессии, оценим среднюю выручку от мобильных систем с 20 новыми видами тарифов:

$$y_x(20) = 0,3987 \cdot 20 + 7,5053 \approx 15,479 \text{ ден. ед.}$$