Answer on Question #65395 - Math - Statistics and Probability

Question: The regression equation of y on x and that of x on y are 8x-10y+66=0 and 40x-18 y = 214 respectively, and the variance of x is 9.

i) What are the mean values of x and y?
ii) Find σy.
iii) Find the coefficient of correlation between x and y.

Solution:

i) Regression equation of y on x:

$$y - \mu_y = r_{x,y} \cdot \frac{\sigma_y}{\sigma_x}(x - \mu_x)$$

and regression equation of x on y:

$$x - \mu_x = r_{x,y} \cdot \frac{\sigma_x}{\sigma_y}(y - \mu_y),$$

where μ_x and μ_y are mean values of x and y, σ_x and σ_y are standard deviation of x and y respectively and $r_{x,y}$ is the Pearson's correlation coefficient.

The Pearson's correlation coefficient is defined as $r_{x,y} = \frac{cov(x,y)}{\sigma_x \cdot \sigma_y}$, where $cov(x, y)$ is a covariance between two random variables x and y.

Let us solve the given regression equation of y on x with respect to y and the given regression equation of x on y with respect to x. We get $y = 0.8\,x + 6.6$ and $x = 0.45y + 5.35$ respectively.

The intercept of the first line $a1$ is equal to $\mu_y - b1 \cdot \mu_x$, where $b1$ is a slope of the first line. The intercept of the second line $a2$ is equal to $\mu_x - b2 \cdot \mu_y$, where $b2$ is a slope of the second line. Thus, we come to the system of two equations,

$$\mu_y - 0.8 \cdot \mu_x = 6.6,$$

$$\mu_x - 0.45 \cdot \mu_y = 5.35,$$

which determines μ_x and μ_y.

We solve the first equation with respect to μ_y: $\mu_y = 0.8 \cdot \mu_x + 6.6$ and put it in the second equation $\mu_x - 0.45 \cdot (0.8 \cdot \mu_x + 6.6) = 5.35$. We solve the last equation:

$$0.64 \cdot \mu_x = 8.32;$$

$$\mu_x = 13.$$

Hence, $\mu_y = 0.8 \cdot 13 + 6.6 = 17$.

Answer: $\mu_x = 13$, $\mu_y = 17$.

ii) The slope $b1$ of the line which corresponds to the solved regression equation of y on x is equal to $\frac{cov(x,y)}{\sigma_x^2}$. It is given, that $\sigma_x^2 = Var\ x\ = 9$. So, we obtain the equation

$$0.8 = \frac{cov(x,y)}{9}.$$

Hence, $cov(x, y) = 9 \cdot 0.8 = 7.2$.

From the solved regression equation of x on y we get

$0.45 = \frac{cov(x,y)}{\sigma_y^2}$.

Hence $\sigma_y^2 = \frac{cov(x,y)}{0.45} = \frac{7.2}{0.45} = 16$ and $\sigma_y = \sqrt{16} = 4$.

Answer: $\sigma_y = 4$

iii) The Pearson's correlation coefficient is equal to $r_{x,y} = \frac{cov(x,y)}{\sigma_x \cdot \sigma_y} = \frac{7.2}{3 \cdot 4} = 0.6$.

Answer: $r_{x,y} = 0.6$.

For linear regression see for example
http://onlinestatbook.com/Online_Statistics_Education.pdf (P.462-467)
http://faculty.cas.usf.edu/mbrannick/regression/regbas.html
http://onlinestatbook.com/2/regression/intro.html.