UNIVERSITY OF BATNA
FACULTY OF MATHEMATICS
& COMPUTER SCIENCE
COMMON CORE DEPARTMENT MI

Subject: Probability and Statistics
MI+Engineering
Academic Year: 2023/2024

Solution TD No. 4
(bivariate statistics)

Exercise 1

1. Calculation of Covariance The covariance $\text{Cov}(X,Y)$ is calculated using the formula:

$$\text{Cov}(X,Y) = \frac{1}{n}\sum(x_i y_i) - \bar{X}\bar{Y}$$

Given the data, we compute:

$$\text{Cov}(X,Y) = 412.8 - 7 \times 65.2 = -43.6$$

where:

$$\bar{Y} = \frac{1}{n}\sum y_i = 65.2, \quad \bar{X} = \frac{1}{n}\sum x_i = 7$$

Thus, the covariance $\text{Cov}(X,Y)$ is -43.6.

2. Regression Line Equation The equation of the regression line $Y = aX + b$ is determined where:

$$a = \frac{\text{Cov}(X,Y)}{\text{Var}(X)}, \quad b = \bar{Y} - a\bar{X}$$

From the calculations:

$$\text{Var}(X) = \frac{1}{n}\sum x_i^2 - \bar{X}^2 = 61.8 - 49 = 12.8$$

$$a = \frac{-43.6}{12.8} = -3.4, \quad b = 65.2 + 3.4 \times 7 = 89$$

Thus, the regression line is:

$$Y = -3.4X + 89$$

3. Correlation Coefficient The linear correlation coefficient r is computed as:

$$r = \frac{\text{Cov}(X,Y)}{\sigma_X \sigma_Y}$$

Where:

$$\text{Var}(Y) = \frac{1}{n}\sum y_i^2 - \bar{Y}^2 = 150.16$$

$$\sigma_Y = \sqrt{150.16} = 12.25$$

Given σ_X and σ_Y, calculate r. $r = \dfrac{-46.6}{\sqrt{12,8} \times 12,25} = --0.99$

4. the coefficient of determination $R = r^2 = 0,98$

Exercise 2: We have this table:

	y_j	6	8	10	13	$\sum$
x_i	**X|Y**	$[5-7[$	$[7-9[$	$[9-11[$	$[11-15[$	$n_{i.}$
2	$[1-3[$	0	0	2	36	38
7	$[3-11[$	0	3	12	26	41
15	$[11-19[$	1	8	35	16	60
25	$[19-31[$	10	26	22	3	61
75	$[31-59[$	22	15	6	0	43
$\sum$	$n_{.j}$	33	52	77	81	243

Determination of Covariance and Linear Coefficient

To determine the covariance $\text{Cov}(X, Y)$ between two variables X and Y, we use the formula:

$$\text{Cov}(X,Y) = \frac{1}{N}\sum_{i=1}^{5}\sum_{j=1}^{4}(x_i - \bar{X})(y_j - \bar{Y})$$

$$\text{Cov}(X,Y) = \frac{1}{N}\sum_{i=1}^{5}\sum_{j=1}^{4}n_{ij}x_iy_j - \bar{X}\bar{Y}$$

where $\bar{X}$ and $\bar{Y}$ are the means of X and Y respectively.

Calculate the mean values:

$$\bar{X} = \frac{1}{N}\sum_{i=1}^{5}n_ix_i$$

$$\bar{X} = \frac{1}{243}\left((38\times 2) + (7\times 41) + (15\times 60) + (25\times 61) + (45\times 43)\right)$$

$$\bar{X} = 19.43$$

$$\bar{Y} = \frac{1}{243}\left((33\times 6) + (52\times 8) + (77\times 10) + (81\times 13)\right)$$

$$\bar{Y} = 10.02$$

The covariance calculation is shown below:

$$\begin{aligned}
\text{Cov}(X,Y) &= \frac{1}{N}\sum_{i=1}^{n}n_{ij}x_iy_j - \bar{X}\bar{Y}\\
&= \frac{1}{243}\Big((2\times 10\times 2) + (36\times 13\times 2) + (3\times 8\times 7) + (12\times 10\times 7)\\
&\quad + (26\times 13\times 7) + (1\times 6\times 15) + (8\times 8\times 15) + (35\times 10\times 15)\\
&\quad + (16\times 13\times 15) + (6\times 10\times 25) + (26\times 8\times 25) + (22\times 10\times 25)\\
&\quad + (3\times 13\times 25) + (22\times 6\times 45) + (15\times 8\times 45) + (6\times 10\times 45)\Big)\\
&\quad - (10.02\times 19.43)\\
&\text{Cov}(X,Y) = -26.03
\end{aligned}$$

Calculation of the Linear Correlation Coefficient

Calculate the linear correlation coefficient r, which is defined as:

$$r = \frac{\text{cov}(x,y)}{\sigma_x\sigma_y}$$

where σ_x and σ_y are the standard deviations of x and y respectively, calculated from their variances.

The variance for x is given by:

$$\text{Var}(x) = \frac{1}{N}\sum_{i=1}^{5}n_ix_i^2 - \bar{X}^2 = 202,14$$

$$\sigma_X = \sqrt{202,14} = 14.21.$$

And the variance for y is:

$$\mathrm{Var}(y) = \frac{1}{N}\sum_{j=1}^{4} n_j y_j^2 - \bar{Y}^2 = 6.02$$

$$\sigma_y = \sqrt{6.02} = 2.49.$$

Thus, the correlation coefficient r can be calculated using:

$$r = \frac{\mathrm{cov}(x,y)}{\sqrt{\mathrm{Var}(x)\mathrm{Var}(y)}} =$$

$$r = \frac{-26,03}{14,21 \times 2.49} = -0.74$$

and the coefficient of determination $R = r^2 = 0,55$

3. the regression line of Y on X

$$a = \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}(X)}, \quad b = \bar{Y} - a\bar{X}$$

$$a = \frac{-29,03}{202.14} = -0,13$$

$$b = 10.02 + 0.13 \times 19,43 = 12,55$$

So

$$Y = -0.13X + 12.55$$

4. Estimate the sleep duration for a person aged 66 years

So

$$Y = -0.13 \times 66 + 12.55 = 3,97 hours$$