MCS (511)
Department of Computer Science
University of Karachi
Course Supervisor

COVARIENCE

Covariance is a measure that tells us about the linear association between the two variables.

It is used for bivariate data let's say x and y to analyses either the x and y variables are related to each other or not.

Steps:

1) Find mean of x and then variance of x

2) Find mean of y and then variance of y

3) Product of x and y

4) Covariance (xy) = $\frac{\sum(xy)}{n} - (\bar{x})(\bar{y})$

5) Coefficient of correlation (r) can also be defined as r = $\frac{Covar(xy)}{\sqrt{var(x)}.\sqrt{var(y)}}$

Sum of squares of x: SS(x) = $\sum x^2 - (\frac{(\sum x)^2}{n})$

Sum of squares of y: SS(y) = $\sum y^2 - (\frac{(\sum y)^2}{n})$

Sum of squares of xy : SS(xy) = $\sum xy - \frac{(\sum x.\sum y)}{n}$

Pearson's Correlation Coefficient: r = $\frac{SS(xy)}{\sqrt{SS(x).SS(y)}}$

Equation for the line of best fit: $\hat{y} = b_o + b_1 x$

Slope of the line of the best fit : $b_1 = \frac{SS(xy)}{SS(x)}$

y-intercept for the line of best fit: $b_o = [\sum y - b_1 . \sum x]/n$