Frequently Asked Questions

1. What do you mean by Covariance?

Answer:

Covariance is a measure of how much two random variables change together. i.e. Covariance provides a measure of the strength of the correlation between two or more sets of random variates .

2. Define covariance between two random variables X and Y.

Answer:

Covariance of two random variables X and Y each with sample size n is defined by the expectation value. i.e. If X and Y are two random variables, then covariance between them is defined as,

Cov(X,Y)=E[{(X-E(X)}{Y-E(Y)}]
=E[XY-YE(X)-XE(Y)+E(X)E(Y)]
=E(XY)-E(Y)E(X)-E(X)E(Y)+E(X)E(Y)
=E(XY)-E(X)E(Y).

It can be also denoted by $\sigma_{xy.}$

3. If X and Y are two random variables, a and b are constants then, show that Cov(aX, bY)= ab.Cov(X, Y).

Answer:

Cov(aX, bY)=E[{(aX-E(aX)}{bY-E(bY)}]
=E[{(aX-aE(X)}{bY-bE(Y)}]
=abE[{(X-E(X)}{Y-E(Y)}]
= ab.Cov(X, Y).

4. If X and Y are two random variables, a and b are constants then, show that Cov(aX, bY)=Cov(X, Y).

Answer:

Cov(X+a, Y+b)= E[{(X+a-E(X+a)}{Y+b-E(Y+b)}]
= E[{(X+a-a-E(X)}{Y+b-b-E(Y)}]
= E[{(X-E(X)}{Y-E(Y)}]=Cov(X, Y)

5. If X and Y are two random variables, a, b, c and d are constants then, show that Cov(aX+b, cY+d)s=ac.Cov(X, Y).

Answer:

Cov(aX+b, cY+d)=E[{(aX+b-E(aX+b)}{cY+d-E(cY+d)}]
=E[{(aX+b-a-bE(X)}{Y+c-c-dE(Y)}]
=ac.E[{(X-E(X)}{Y-E(Y)}]=ac.Cov(X, Y)

6. Show that variance can be shown as a particular case of covariance.

Answer:

If we consider Y=X then

Cov(X, X)=E[{(X-E(X)}{X-E(X)}]=E[X-E(X)]^2=V(X)

7. State and prove property of symmetry of Covariance.

Answer:

Property of Symmetry states that Cov(X,Y)=Cov(Y,X)

Cov(X,Y)=E[{(X-E(X)}{Y-E(Y)}]=E[{(Y-E(Y)}{X-E(X)}]=Cov(Y, X)

8. Let X, Y and Z are real valued random variables and let a, b and c are constants, then show that Cov(aX+bY, cZ)= acCov(X, Z)+bcCov(Y, Z).

Answer:

Cov(aX+bY, cZ)=E[{aX+bY-E(aX+bY)}{cZ-E(cZ)}]
=E[{aX-aE(X)+bY-bE(Y)}{cZ-cE(Z)}]
= E[{aX-aE(X)}{cZ-cE(Z)}]+E[{bY-bE(Y)}{cZ-cE(Z)}]
=acE[{X-E(X)}{Z-E(Z)}]+bcE[{Y-E(Y)}{Z-E(Z)}]
=acCov(X, Z)+bcCov(Y, Z)

9. If X1, X2, …,Xn, Y1, Y2, …Ym are random variables, then show that

$$Cov(\sum_{i=1}^{n} X_i, \sum_{j=1}^{m} Y_j) = \sum_{i=1}^{n}\sum_{j=1}^{m} Cov(X_i, Y_j)$$

Answer:

We know that $Cov(\Sigma a_i X_i, cY)=c\Sigma a_i Cov(X_i, Y)$

Suppose we take all ai's and c to be 1,

$Cov(\Sigma X_i, Y)=\Sigma Cov(X_i, Y)$

Now consider,

$$Cov(\sum_{i=1}^{n} X_i, \sum_{j=1}^{m} Y_j) = \sum_{i=1}^{n} Cov(X_i, \sum_{j=1}^{m} Y_j) = \sum_{i=1}^{n} Cov(\sum_{j=1}^{m} Y_j, X_i)$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{m} Cov(Y_j, X_i) = \sum_{i=1}^{n}\sum_{j=1}^{m} Cov(X_i, Y_j)$$

10. If X, Y, U and V are real valued random variables and a, b, c and d are constants, then show that Cov(aX+bY, cU+dV)=ac Cov(X, U)+ ad Cov(X, V)+bc Cov(Y, U)+ bd Cov(Y, V)

Answer:

Consider

Cov(aX+bY, cU+dV)=E[{aX+bY-E(aX+bY)}{cU+dV-E(cU+dV)}]
=E[{a(X-E(X))+b(Y-E(Y)}{c(U-E(U))+d(V-E(V))}]
=E[ac{X-E(X)}{U-E(U)}]+E[ad{X-E(X)}{V-E(V)}]
+E[bc{Y-E(Y)}{U-E(U)}]+E[bd{Y-E(Y)}{V-E(V)}]
= ac Cov(X, U)+ ad Cov(X, V)+bc Cov(Y, U)+ bd Cov(Y, V)

11. When you can say that two variables are independent?

Answer:

If *x* and *y* are independent, then their covariance is zero. This follows because under independence, E(XY)=E(X)E(Y)

12. If covariance is zero, is it going to imply that the variables are independent?

Answer:

Let *x* be uniformly distributed in [-1, 1] and let $y = x^2$. Clearly, *X* and Y are dependent, but
$Cov(X, Y)=Cov(X, X^2)$
$=E(X.X^2)-E(X)E(X^2)$
$=E(X^3)- E(X)E(X^2)$
$=0-0.E(X^2)=0$

In this case, the relationship between *y* and *x* is non-linear, while correlation and covariance are measures of linear dependence between two variables. Still, as in the example, if two variables are uncorrelated, that does not imply that they are independent.

13. Write the limitation of Covariance.

Answer:

Since the number representing covariance depends on the units of the data, it is difficult to compare covariances among data sets having different scales. A value that might represent a strong linear relationship for one data set might represent a very weak one in another.

14. Let X and Y be 2 random variables such that V(X)=2 and Cov(X, Y)=1. Obtain the Cov(5X, 2X+3Y).

Answer:

Let us use the property,

Cov(aX+bY, cZ) and Cov(X, X) =V(X).

Consider

Cov(5X, 2X+3Y)=Cov(5X, 2X)+Cov(5X, 3Y)=10V(X)+15Cov(X, Y)=10.2+15.1=35

15. Two random variables X and Y have the following joint probability density function. f(x,y)=(4-x-y)/8;0≤x≤2, 0≤y≤2. Find Cov(X, Y).

Answer: Cov(X, Y)=E(XY)-E(X)E(Y)

Before we find Expectation and variance, we find marginal distribution of X and Y. Marginal distribution of X is given by,

$$f(x) = \frac{1}{8}\int_0^2 (4-x-y)dy = \frac{1}{8}(6-2x) = \frac{1}{4}(3-x);0 \le x \le 2$$

Marginal distribution of Y is given by,

$$f(y) = \frac{1}{8}\int_0^2 (4-x-y)dx = \frac{1}{8}(6-2y) = \frac{1}{4}(3-y);0 \le y \le 2$$

Now, let us obtain Expectation.

$$E(x) = \int_0^2 xf(x)dx = \frac{1}{4}\int_0^2 x(3-x)dx = \frac{5}{6}$$

$$E(y) = \int_0^2 yf(y)dx = \frac{1}{4}\int_0^2 y(3-y)dy = \frac{5}{6}$$

$$E(XY) = \int_0^2\int_0^2 xyf(,y)dxdy = \frac{1}{8}\int_0^2 y\left[\int_0^2 x(4-x-y)dx\right]dy$$

$$= \frac{1}{8}\int_0^2 y\left(\frac{16}{3}-2y\right)dy = \frac{2}{3}$$

Hence, Cov(X,Y)= $^2/_3 - (^5/_6)(^5/_6) = -^1/_{36}$