Frequently Asked Questions

1. Which type of relationship is measured from correlation coefficient?

Answer:

Correlation is the degree to which two or more quantities are linearly associated.

2. Define coefficient of correlation.

Answer:

Correlation is obtained by dividing the covariance by the product of the two standard deviations. i.e. $\rho_{XY} = \dfrac{Cov(X,Y)}{SD(X).SD(Y)} = \dfrac{\sigma_{XY}}{\sigma_X \sigma_Y}$

3. Give any 4 examples of correlation analysis.

Answer:

Correlation analysis may answer the questions like

- How long will someone live?
- Fluctuation in the stock market.
- Chances of someone becoming a criminal.
- Chances of a surgery prolonging a cancer patient's life.

4. What do you mean by causation?

Answer:

Causation means 'cause and effect' relationship.

5. What is the relation between correlation and causation?

Answer:

No discussion of correlation would be complete without a discussion of causation.
Suppose there exist correlation between two variables without causation, then that correlation is called non-sense correlation or spurious correlation.

6. How to interpret the value of coefficient of correlation?

Answer:

The following points are the accepted guidelines for interpreting the correlation coefficient:

- 0 indicates no linear relationship.
- +1 indicates a perfect positive linear relationship: as one variable increases in its values, the other variable also increases in its values via an exact linear rule.
- -1 indicates a perfect negative linear relationship: as one variable increases in its values, the other variable decreases in its values via an exact linear rule.
- Values between 0 and 0.3 (0 and -0.3) indicate a weak positive (negative) linear relationship via a shaky linear rule.
- Values between 0.3 and 0.7 (-0.3 and -0.7) indicate a moderate positive (negative) linear relationship via a fuzzy-firm linear rule.
- Values between 0.7 and 1.0 (-0.7 and -1.0) indicate a strong positive (negative) linear relationship via a firm linear rule.

7. What is the necessity to know the strength of relationship between two variables?

Answer:

It is important to know the strength of relationship because it allows you to make predictions.
The main point of knowing the strength is that a strong relationship allows you to make much more accurate predictions than a weak relationship. This ability to make accurate predictions is critical in a great many professional settings.

8. Name some of the fields, where correlation analysis is used to make predictions.

Answer:

Psychologists, medical professionals, business executives, stock brokers, military leaders, law enforcement agents are all interested in being able to make predictions. The concept of correlation provides a tool that helps people make predictions and to do so with some amount of accuracy.

9. Write the properties of correlation coefficient.

Answer:

- Following are the properties of correlation coefficient:
    - Correlation coefficient is independent of unit of measurement of the variables.
    - If correlation is present, then coefficient of correlation would lie between ±1.
    - Correlation coefficient is independent of change of origin and scale.
    - If X and Y are random variables and a, b, c, d are any numbers provided that $a \neq 0, c \neq 0$ then $r(aX+b, cY+d) = \frac{ac}{|ac|} r(X,Y)$

10. The variables X and Y are connected by the equation aX+bY+c=0. Show that the correlation between them is -1 if the signs of a and b are alike and +1 if they are different.

Answer:

$aX+bY+c=0 \Rightarrow aE(X)+bE(Y)+c=0 \Rightarrow a[X-E(X)]+b[Y-E(Y)]=0$

$$\Rightarrow X - E(X) = -\frac{b}{a}[Y - E(Y)]$$

$$\therefore Cov(X,Y) = E[X - E(X)][Y - E(Y)] = -\frac{b}{a}[Y - E(Y)]^2 = -\frac{b}{a}\sigma_Y^2$$

$$\sigma_X^2 = E[X - E(X)]^2 = \frac{b^2}{a^2}\sigma_Y^2$$

$$r = \frac{Cov(X,Y)}{\sigma_X \sigma_Y} = \frac{-\frac{b}{a}\sigma_Y^2}{\sqrt{\sigma_Y^2}\sqrt{\frac{b^2}{a^2}\sigma_Y^2}} = \frac{-\frac{b}{a}\sigma_Y^2}{\left|\frac{b}{a}\right|\sigma_Y^2}$$

=+1, if b and a are of opposite sign

-1, if b and a are of same sign.

11. What is the importance of correlation analysis?

Answer:

1. Most of the variables in economics and business are show relationship. For example, price and supply, income and expenditure etc. The correlation analysis helps the investigator to know the degree and direction of such relationship between variables. These days, correlation analysis finds application in various fields including the field of life science.
2. Once the correlation is established between the two variables, regression analysis helps us to estimate value of dependent variable for the given value of independent variable.
3. Correlation analysis together with regression analysis helps us to understand the behaviour of various social and economic variables.
4. The effect of correlation is to reduce the range of uncertainty in our predictions.

12. If X and Y are uncorrelated random variables, and $r(aX + bY, bX + aY) = \frac{1+2ab}{a^2+b^2}$
Find r(X,Y), the coefficient of correlation between X and Y.

Answer:

Since X and Y are standardised random variables, we have

E(X)=E(Y)=0

And V(X)=V(Y)=1⇒E(X^2)=E(Y^2)=1

And cov(X,Y)=E(XY)⇒E(XY)=r(X,Y)$\sigma_X\sigma_Y$=r(X,Y)

Also we can write

$$r(aX+bY, bX+aY) = \frac{E(aX+bY)(bX+aY) - E(aX+bY)E(bX+aY)}{[V(aX+bY)V(bX+aY)]^{1/2}}$$

$$= \frac{E(abX^2 + a^2XY + b^2YX + abY^2) - 0}{\left\{\begin{matrix}[a^2V(X) + b^2V(Y) + 2abCov(X,Y)] \\ \times [b^2V(X) + a^2V(Y) + 2baCov(X,Y)]\end{matrix}\right\}^{1/2}}$$

$$= \frac{ab.1 + a^2r(X,Y) + b^2r(X,Y) + ab.1}{\left\{[a^2+b^2+2abr(X,Y)][b^2+a^2+2bar(X,Y)]\right\}^{1/2}}$$

$$= \frac{2ab + (a^2+b^2)r(X,Y)}{[a^2+b^2+2abr(X,Y)]}$$

Given that

$$r(aX + bY, bX + aY) = \frac{1+2ab}{a^2+b^2}$$

Hence, equating the right hand sides, we get,

$$\frac{1+2ab}{a^2+b^2} = \frac{2ab + (a^2+b^2)r(X,Y)}{[a^2+b^2+2abr(X,Y)]}$$

$$\Rightarrow (1+2ab)[a^2+b^2+2abr(X,Y)] = (a^2+b^2)[2ab + (a^2+b^2)r(X,Y)]$$

$$\Rightarrow (1+2ab)(a^2+b^2) + 2abr(X,Y)(1+2ab) = (a^2+b^2)^2r(X,Y) + 2ab(a^2+b^2)$$

$$\Rightarrow (a^4+b^4+2a^2b^2-2ab-4a^2b^2)r(X,Y) = a^2+b^2$$

$$\Rightarrow r(X,Y) = \frac{a^2+b^2}{(a^2-b^2)^2-2ab}$$

13. Write the merits of correlation coefficient.

Answer:

1. It gives a precise quantitative value indicating the degree of relationship existing between the two variables.
2. It measures the direction as well as relationship between the two variables.
3. Further in regression analysis it is used for estimating the value of dependent variable from the known value of the independent variable.

14. Write the demerits of correlation coefficient.

Answer:

1. Extreme items affect the value of the coefficient of correlation.
2. Its computational method is difficult as compared to other methods.
3. It assumes the linear relationship between the two variables, whether such relationship exist or not.

15. Let the random variable have marginal density f(x)=1, -1/2 <x<1/2 and the conditional density of Y bef(y|x)=1, x<y<x+1,-1/2<x<0
=1, -x<y<1-x, 0<x<1/2
Show that variables X and Y are uncorrelated.

Answer:

We have

$$E(X) = \int_{-1/2}^{1/2} x.f(x)dx = \int_{-1/2}^{1/2} x.1.dx = 0$$

If f(x,y) is the joint pdf of X and Y then
f(x,y)=f(x).f(y|x)==f(y|x), since f(x)=1

$$E(XY) = \int_{-1/2}^{0}\int_{x}^{x+1} xydxdy + \int_{0}^{1/2}\int_{x}^{x-1} xydxdy$$

$$= \int_{-1/2}^{0} x\left[\int_{x}^{x+1} ydy\right]dx + \int_{0}^{1/2} x\left[\int_{x}^{x-1} ydy\right]dx$$

$$= \frac{1}{2}\int_{-1/2}^{0} x(2x+1)dx + \int_{0}^{1/2} x(1-2x)dx = 0$$

$\therefore$ Cov(X, Y)=E(XY)-E(X)E(Y)=0$\Rightarrow$r(X,Y)=0. Hence, the variables are uncorrelated.