Frequently Asked Questions

1. How once can study the relationship between tri variate or multivariate distribution?

Answer:

Suppose in a trivariate or multi-variate distribution we are interested in the relationship between two variables only. There are two alternatives viz.,

1. We consider only those two members of the observed data in which the other members have specified values or
2. We may eliminate mathematically the effect of other variates on two variates.

2. What do you mean by partial correlation?

Answer:

The correlation between only two variates eliminating the linear effect of other variates in them is called the partial correlation.

3. What do you mean by Partial Regression

Answer:

The regression between only two variates eliminating the linear effect of other variates in them is called the partial regression.

4. Write the expression for plane of regression of X_1 on X_2 and X_3.

Answer:

$$\frac{X_1}{\sigma_1}\omega_{11} + \frac{X_2}{\sigma_2}\omega_{12} + \frac{X_3}{\sigma_3}\omega_{13} = 0$$

5. What do you mean by coefficient of partial correlation?

Answer:

Sometimes the correlation between two variables X_1 and X_2 may be partly due to the correlation of a third variable, X_3 with both X_1 and X_2. In such a situation, one may want to know what the correlation between X_1 and X_2 would be if the effect of X_3 on each of X_1 and X_2 were eliminated. This correlation is called the partial correlation and the correlation coefficient between X_1 and X_2 the linear effect of X_3 on each of them has been eliminated is called the partial correlation coefficient.

6. Derive an expression for plane of regression.

Answer:

The equation of the plane of regression of X_1 on X_2 and X_3 is

$X1 = b_{12.3}\, X_2 + b_{13.2}\, X_3$

The constants b in the above equation are determined by the principle of least square, ie., by minimizing the sum of the squares of residuals, namely,

$S = \Sigma X_{1.23}{}^2 = \Sigma(X_1 - b_{12.3}\, X_2 + b_{13.2}\, X_3)^2$, the summation being extended to the given values of the variables.

Here we make N observations on each of the variables X_1, X_2 and X_3.

$$\frac{dS}{db_{12.3}} = 0 = -2\Sigma X_2 (X_1 - b_{12.3} X_2 - b_{13.2} X_3)$$

$$\frac{dS}{db_{13.2}} = 0 = -2\Sigma X_3 (X_1 - b_{12.3} X_2 - b_{13.2} X_3)$$

ie., $\Sigma X_2 X_{1.23} = 0$ and $\Sigma X_3 X_{1.23} = 0$

$$\Rightarrow \Sigma X_1 X_2 - b_{12.3} \Sigma X_2^{\ 2} - b_{13.2} \Sigma X_2 X_3 = 0$$

$$\Sigma X_1 X_3 - b_{12.3} \Sigma X_2 X_3 - b_{13.2} \Sigma X_3^{\ 2} = 0$$

Since Xi's are measured from their respective means, we have,

$$\sigma_i^{\ 2} = \frac{1}{N} \Sigma X_i^{\ 2}, Cov(X_i, X_j) = \frac{1}{N} \Sigma X_i X_j$$

And

$$r_{ij} = \frac{Cov(X_i, X_j)}{\sigma_i \sigma_j} = \frac{\Sigma X_i X_j}{N \sigma_i \sigma_j}$$

Hence we get,

$r_{12}\, \sigma_1\, \sigma_2 - b_{12.3}\, \sigma_2^{\ 2} - b_{13.2}\, r_{23}\, \sigma_2 \sigma_3 = 0$

$r_{13}\, \sigma_1\, \sigma_3 - b_{12.3}\, \sigma_2\, \sigma_3 - b_{13.2}\, r_{23}\, \sigma_3^{\ 2} = 0$

Solving the above equations for $b_{12.3}$ and $b_{13.2}$, we get,

$$b_{12.3} = \frac{\begin{vmatrix} r_{12}\sigma_1 & r_{23}\sigma_3 \\ r_{13}\sigma_1 & \sigma_3 \end{vmatrix}}{\begin{vmatrix} \sigma_2 & r_{23}\sigma_3 \\ r_{23}\sigma_2 & \sigma_3 \end{vmatrix}} = \frac{\sigma_1}{\sigma_2} \frac{\begin{vmatrix} r_{12} & r_{23} \\ r_{13} & 1 \end{vmatrix}}{\begin{vmatrix} 1 & r_{23} \\ r_{23} & 1 \end{vmatrix}}$$

Similarly,

$$b_{13.2} = \frac{\sigma_1}{\sigma_3} \frac{\begin{vmatrix} 1 & r_{12} \\ r_{23} & r_{13} \end{vmatrix}}{\begin{vmatrix} 1 & r_{23} \\ r_{23} & 1 \end{vmatrix}}$$

If we write $\omega = \begin{vmatrix} 1 & r_{12} & r_{13} \\ r_{21} & 1 & r_{23} \\ r_{31} & r_{32} & 1 \end{vmatrix}$

and ω_{ij} is the cofactor of the element in the ith row and jth column of ω we have,

$$b_{12.3} = \frac{\sigma_1 \omega_{12}}{\sigma_2 \omega_{11}}, b_{13.2} = \frac{\sigma_1 \omega_{13}}{\sigma_3 \omega_{11}}$$

Substituting the values in the plane of regression, mentioned in the beginning, we get the required equation of the plane of regression of X_1 on X_2 and X_3 as,

$$X_1 = -\frac{\sigma_1 \omega_{12}}{\sigma_2 \omega_{11}} X_2 - \frac{\sigma_1 \omega_{13}}{\sigma_3 \omega_{11}} X_3 \Rightarrow \frac{X_1}{\sigma_1} \omega_{11} + \frac{X_2}{\sigma_2} \omega_{12} + \frac{X_3}{\sigma_3} \omega_{13} = 0$$

7. Write an expression for the error of the estimate or residual.
Answer:

$X_{1.23} = X_1 - b_{12.3} X_2 - b_{13.2} X_3 = X_1 - e_{1.23}$ is called the error of estimate or residual.

8. Write the properties of residual.
   Answer:
   1. The sum of the product of any residual of order zero with any other residual of higher order is zero, provided the subscript of the former occurs among the secondary subscripts of the later.
   2. The sum of the product of any two residuals in which all the secondary subscripts of the first occur among the secondary subscripts of the second is unaltered if we omit any or all of the secondary subscripts of the first. Conversely, the product sum of any residual of order 'p' with a residual of order p plus q, the 'p' subscripts being the same in each case is unaltered by adding to the secondary subscripts of the former any or all the q additional subscripts of the later.
   3. The sum of the product of two residuals is zero if all the subscripts (primary as well as secondary) of the one occur among the secondary subscripts of the other.

9. Obtain the expression for the partial correlation coefficient.
   Answer:

The partial correlation coefficient between X_1 and X_2, is denoted by

$$r_{12.3} = \frac{Cov(X_{1.3}, X_{2.3})}{\sqrt{Var(X_{1.3})V(X_{2.3})}}$$

We have, $Cov(X_{1.3}, X_{2.3}) = \frac{1}{N}\Sigma X_{1.3}X_{2.3} = \frac{1}{N}\Sigma X_1 X_{2.3} = \frac{1}{N}\Sigma X_1(X_2 - b_{23}X_3)$

$$= \frac{1}{N}\Sigma X_1 X_2 - b_{23}\frac{1}{N}\Sigma X_1 X_3 = r_{12}\sigma_1\sigma_2 - r_{23}\frac{\sigma_2}{\sigma_3}(r_{13}\sigma_1\sigma_3) = \sigma_1\sigma_2(r_{12} - r_{13}r_{23})$$

$$V(X_{1.3}) = \frac{1}{N}\Sigma X_{1.3}^2 = \frac{1}{N}\Sigma X_{1.3}X_{1.3}$$

$$= \frac{1}{N}\Sigma X_1 X_{1.3} = \frac{1}{N}\Sigma X_1(X_1 - b_{13}X_3)$$

$$= \frac{1}{N}\Sigma X_1^2 - b_{13}\frac{1}{N}\Sigma X_1 X_3 = \sigma_1^2 - r_{13}\frac{\sigma_1}{\sigma_3}r_{13}\sigma_1\sigma_3$$

$$= \sigma_1^2(1 - r_{13}^2)$$

Similarly we get,

$$V(X_{2.3}) = \sigma_2^2(1 - r_{23}^2)$$

$$r_{12.3} = \frac{\sigma_1\sigma_2(r_{12} - r_{13}r_{23})}{\sqrt{\sigma_1^2(1 - r_{13}^2)\sigma_2^2(1 - r_{23}^2)}}$$

$$= \frac{(r_{12} - r_{13}r_{23})}{\sqrt{(1 - r_{13}^2)(1 - r_{23}^2)}}$$

10. Obtain an expression for partial correlation coefficient using the property of regression coefficient.

Answer

We know that correlation coefficient is nothing but the geometric mean of the regression coefficients. Hence first we find the regression coefficients.

We have,

$0=\Sigma X_{2.3}\ X_{1.23} = \Sigma X_{2.3}\ (X_1 - b_{12.3}\ X_2 - b_{13.2}\ X_3)$

$= \Sigma\ X_1\ X_{2.3} - b_{12.3}\ \Sigma\ X_{2.3}\ X_2 - b_{13.2}\ \Sigma\ X_{2.3}\ X_3$

$= \Sigma\ X_{1.3}\ X_{2.3} - b_{12.3}\ \Sigma\ X_{2.3}\ X_{2.3}$

Implies,

$$b_{12.3} = \frac{\Sigma X_{1.3} X_{2.3}}{\Sigma X_{2.3}^{\ 2}}$$

Form this it follows that $b_{12.3}$ is coefficient of regression of $X_{1.3}$ on $X_{2.3}$.

Similarly, $b_{21.3}$ is the coefficient of regression of $X_{2.3}$ on $X_{1.3}$.

Since correlation coefficient is the geometric mean between regression coefficients, we have,

$r_{12.3}^{\ 2} = b_{12.3}\ b_{21.3}$

But by definition,

$$b_{12.3} = -\frac{\sigma_1}{\sigma_2}\frac{\omega_{12}}{\omega_{11}}, b_{21.3} = -\frac{\sigma_2}{\sigma_1}\frac{\omega_{21}}{\omega_{22}}$$

$$r_{12.3}^{\ 2} = \left(-\frac{\sigma_1}{\sigma_2}\frac{\omega_{12}}{\omega_{11}}\right)\left(-\frac{\sigma_2}{\sigma_1}\frac{\omega_{21}}{\omega_{22}}\right) = \frac{\omega_{12}^{\ 2}}{\omega_{11}\omega_{22}}$$

$$\Rightarrow r_{12.3} = -\frac{\omega_{12}}{\sqrt{\omega_{11}\omega_{22}}}$$

The negative sign being taken since the sign of regression coefficients is the same as that of $-\omega_{12}$.

Substituting the values of ω_{12}, ω_{11} and ω_{22} in the above expression, we get,

$$r_{12.3} = \frac{(r_{12} - r_{13}r_{23})}{\sqrt{(1-r_{13}^{\ 2})(1-r_{23}^{\ 2})}}$$

11. Write the expression for partial correlation coefficient between X_1 and X_3 when we eliminate the linear effect of X_2.

Answer:

$$r_{13.2} = \frac{(r_{13} - r_{12}r_{32})}{\sqrt{(1-r_{12}^{\ 2})(1-r_{32}^{\ 2})}}$$

12. Give the expression of coefficient of partial correlation between X_2 and X_3 when we eliminate the linear effect of X_1.

Answer:

$$r_{23.1} = \frac{(r_{23} - r_{21}r_{31})}{\sqrt{(1-r_{21}^{\ 2})(1-r_{31}^{\ 2})}}$$

13. Give a generalized expression for the coefficient of partial correlation.

Answer:

In the case of n variables, X_1, X_2, …X_n the partial correlation coefficient $r_{12.34\ldots n}$ between X_1 and X_2 (after the linear effect of X_3, X_4, …X_n on them has been eliminated), is given by,

$r_{12.34\ldots n}{}^2 = b_{12.3}\, b_{21.3}$

But we have,

$$b_{12.34..n} = -\frac{\sigma_1}{\sigma_2}\frac{\omega_{12}}{\omega_{11}}, b_{21.34...n} = -\frac{\sigma_2}{\sigma_1}\frac{\omega_{21}}{\omega_{22}}$$

$$r_{12.34...n}{}^2 = \left(-\frac{\sigma_1}{\sigma_2}\frac{\omega_{12}}{\omega_{11}}\right)\left(-\frac{\sigma_2}{\sigma_1}\frac{\omega_{21}}{\omega_{22}}\right) = \frac{\omega_{12}{}^2}{\omega_{11}\omega_{22}}$$

$$\Rightarrow r_{12.34...n} = -\frac{\omega_{12}}{\sqrt{\omega_{11}\omega_{22}}}$$

The negative sign being taken since the sign of the regression coefficient is the same as that of $-\omega_{12}$

14. What is the use of partial correlation coefficient?

Answer:

Partial correlation coefficient helps in deciding whether to include or not an additional independent variable in regression analysis

15. If $r_{12,3}$ =0, whether the variables are uncorrelation?

Answer:

If $r_{12,3}$ =0, we have then $r_{12} = r_{13}\, r_{23}$. It means that r_{12} will not be zero if X_3 is correlated with both X_1 and X_2. Thus, although X_1 and X_2 may be uncorrelated when effect of X_3 is eliminated, yet X_1 and X_2 may appear to be correlated because they carry the effect of X_3 on them.