

University of Ferrara

Degree Course in “Economics, Markets and Management”

STATISTICAL METHODS for ECONOMICS and BUSINESS – 6 June 2016

Q01

Which of the following matrices is the transposed of $A = \begin{pmatrix} 3 & 2 \\ 7 & -4 \end{pmatrix}$?

$$\begin{pmatrix} 0.15 & 0.08 \\ 0.27 & -0.12 \end{pmatrix}$$

(a)

$$\begin{pmatrix} -3 & -2 \\ -7 & 4 \end{pmatrix}$$

(b)

$$\begin{pmatrix} 3 & 7 \\ 2 & -4 \end{pmatrix}$$

(c)

Q02

Which of the following values is the determinant of $A = \begin{pmatrix} 3 & 2 \\ 7 & -4 \end{pmatrix}$?

- a) -0.038.
- b) -26.
- c) -1.

Q03

The inverse of a square matrix B is the square matrix C such that...

- a) All the elements of C are the inverse of the corresponding elements of B.
- b) The element in the i-th row and in the j-th column of C is equal to the element in the j-th row and in the i-th column of B, for every i and every j.
- c) None of the two previous statements is true.

Q04

Let us consider the multiple linear regression model $Y_i = \beta_0 + \beta_1 X_{i1} + \dots + \beta_k X_{ik} + \epsilon_i$. The least square method for the estimation of the coefficients $\beta_0, \beta_1, \dots, \beta_k$ consists in determining the values $\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_k$ that minimize...

- a) The sum of the squares of the differences between the observed and the predicted values of Y.
- b) The sum of the squares of the distances of the dots with coordinates $(X_{i1}, X_{i2}, \dots, X_{ik}, Y_i)$ from the hyperplane of the equation $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_{i1} + \dots + \hat{\beta}_k X_{ik}$
- c) None of the two previous statements is true.

Q05

Let us consider the following results of a multiple linear regression analysis and let $\alpha=0.10$ be the significance level:

	Coefficients	p-value
Intercept	5.76	0.502
X1	2.28	0.018
X2	8.56	0.278
X3	-5.29	0.421

Which of the following statements is true?

- a) Among the three explanatory variables, only X1 significantly affects the dependent variable of the model.
- b) Among the three explanatory variables, only X1 and X2 significantly affect the dependent variable of the model.
- c) None of the three explanatory variables significantly affect the dependent variable of the model.

Q06

In case of one large Variance Inflation Factor (VIF), e.g. in case of one VIF greater than or equal to 5, which of the following actions should be done?

- a) Remove one or more explanatory variables.
- b) Use the scatter plot of residuals versus dependent variable to check the model assumptions.
- c) Use the normal probability plot of residuals to check the assumption of normality.

Q07

Given the following observed values $x_1=5, x_2=3, x_3=10, x_4=9, x_5=5, x_6=8$, what's the minimum and the maximum for the transformed values $z_u=x_u/\max(x_1,\dots,x_6), u=1,\dots,6$.

- a) 0 and 1.
- b) 0.3 and 1.
- c) 1 and 6.

Q08

According to the NPC methodology, which of the following functions is a suitable combination for the aggregation of the k (transformed) informative variables: $y_u = \Psi(z_{u1}, z_{u2}, \dots, z_{uk})$?

- a) $y_u = \sum_{v=1}^k w_v \cdot (z_{uv} - 0.5)^2$.
- b) $y_u = -\sum_{v=1}^k w_v \cdot \left(\frac{1}{z_{uv}}\right)$.
- c) none of the previous combinations satisfies the required properties.

Q09

Let us consider the following data related to the expenditure for scientific research (millions of euros), for the regions in the north east of Italy, by type of economic activity. The weights reflect the degrees of importance of the activities at national level according to the policy makers.

	Manufacturing	Research centers	Other
Weight	0.5	0.3	0.2
Veneto	300	15	10
Trentino A.A.	19	2	1
Friuli V.G.	92	1	5

By applying rescaling method and additive aggregation to compute the global score, what are the values of the composite index of expenditure for research for the three regions under study?

- a) Veneto: 1.00, Trentino A.A.: 0.25, Friuli V.G.: 0.13.
- b) Veneto: 1.00, Trentino A.A.: 0.33, Friuli V.G.: 0.66.
- c) Veneto: 1.00, Trentino A.A.: 0.02, Friuli V.G.: 0.22.

Q10

In Factor Analysis, the rotation of the factors (e.g. varimax method), is useful for...

- a) ...a simplification for the computation of the communality.
- b) ...the interpretation of the results (factor loadings).
- c) None of the two previous statements is true.

Q11

In case of k numeric response variables, how many Principal Components can be extracted?

- a) A number less than k .
- b) k .
- c) An unlimited number.

Q12

Let us consider the following output of a Factor Analysis related to the first three factors:

Factor	Eigenvalue	% Variance	Cumulative %
1	3.728	38	38
2	1.504	22	60
3	0.889	8	68

According to method based on the eigenvalues, how many factors should be considered to reduce the number of informative variables by replacing the original variables with factors?

- a) 2
- b) 3
- c) it cannot be said

Q13

Which of the following R codes must be applied in order to apply a non hierarchical method in Cluster Analysis?

- a) `hclust(d,method)`.
- b) `kmeans(x,centers,iter.max)`.
- c) none of the two previous codes.

Q14

In order to compute the distances between clusters, which of the following methods is preferable to avoid the chain effect?

- a) Farthest neighbour.
- b) Single linkage.
- c) Nearest neighbour.

Q15

In Cluster Analysis, in the presence of k variables, which of the following distances (between two units) is more affected by large differences between one or more of the k pairs of values?

- a) Euclidean distance.
- b) Manhattan distance.
- c) The effect on the two indices is always the same .