SLR-MM – 522

Seat No.	

M.Sc. (Part – II) (Semester – III) Examination, 2015
STATISTICS (Paper – XII)
Multivariate Analysis (New) (CGPA)

Day and Date : Wednesday, 18-11-2015 Max. Marks : 70

Time : 2.30 p.m. to 5.00 p.m.

Instructions : *1) Attempt* ***five*** *questions.*
2) Q.No. ***(1)*** *and Q.No.* ***(2)*** *are* ***compulsory****.*
3) Attempt ***any three*** *from Q.No.* ***(3)*** *to Q.No.* ***(7)****.*
4) Figures to the ***right*** *indicate* ***full*** *marks.*

1. A) Choose the correct alternative : **5**

1) Let X be a p×1 random vector such that $X \sim N_p(\mu, \Sigma)$, where rank $(\Sigma) = p$. Which of the following is true ?

a) $E\left[(X-\mu)'\Sigma^{-1}(X-\mu)\right] = 2p,\ V\left[(X-\mu)'\Sigma^{-1}(X-\mu)\right] = 2p$

b) $E\left[(X-\mu)'\Sigma^{-1}(X-\mu)\right] = 2p,\ V\left[(X-\mu)'\Sigma^{-1}(X-\mu)\right] = p$

c) $E\left[(X-\mu)'\Sigma^{-1}(X-\mu)\right] = p,\ V\left[(X-\mu)'\Sigma^{-1}(X-\mu)\right] = p$

d) $E\left[(X-\mu)'\Sigma^{-1}(X-\mu)\right] = p,\ V\left[(X-\mu)'\Sigma^{-1}(X-\mu)\right] = 2p$

2) Let $X_1, X_2, ..., X_n$ be a random sample of size n from p-variate normal distribution with mean vector μ and covariance matrix Σ. The distribution of mean vector $\overline{X}$ is

a) $N_p(\mu, \Sigma)$ b) $N_p(\mu, \frac{1}{n}\Sigma)$ c) $N_p(\mu, \frac{1}{n-1}\Sigma)$ d) $N_p(\frac{1}{n}\mu, \frac{1}{n}\Sigma)$

P.T.O.

3) Let A ~W_p (m, Σ) and a $\in R^p$ with $a'\Sigma a \neq 0$. Then distribution of $\frac{a'Aa}{a'\Sigma a}$ is

a) χ^2_p b) χ^2_m c) χ^2_{m-p} d) χ^2_{m-p+1}

4) Canonical correlation is a measure of association between

a) one variable and set of other variables

b) two sets of variables

c) two types of variables

d) none of these

5) Let $X_1, X_2, ..., X_n$ be a random sample of size n from p-variate normal distribution with mean vector 0 and covariance matrix Σ. The MLE of Σ is

a) $\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})(X_i - \overline{X})'$ b) $\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})(X_i - \overline{X})'$

c) $\frac{1}{n}\sum_{i=1}^{n} X_i X_i'$ d) $\sum_{i=1}^{n} X_i X_i'$

B) Fill in the blanks : **5**

1) The first pair of canonical variables have _____ correlation.

2) If X has $N_p(\mu, \Sigma)$ distribution and $Z = \Sigma^{-\frac{1}{2}}(X - \mu)$ then E (Z) = _____

3) Let A has W_p (n, Σ) distribution and C is some nonsingular matrix then distribution of CA C′ is ________

4) Hotelling's T^2 is multivariate extension of _____

5) Generalized variance is ________ of covariance matrix.

C) State whether the following statements are **True** or **false** : **4**

1) In factor analysis, original variables are expressed as linear combinations of the factors.

2) Canonical correlation can be negative.

3) If X has $N_p(\mu, \Sigma)$ distribution then all marginal distributions for any subset of X are normally distributed.

4) Wishart matrix is not symmetric.

2. a) Answer the following. **6**

i) Define Hotelling's T^2 and Mahalanobis D^2 statistics.

ii) Obtain the null distribution of Hotelling's T^2 statistic.

b) Write short notes on the following : **8**

i) Rao's U statistic.

ii) Single linkage clustering method.

3. a) Let vector X be distributed according to $N_p(\mu, \Sigma)$, show that the marginal distribution of any set of components of X is multivariate normal with means, variances and covariances obtained by taking proper components of μ and Σ respectively.

b) Let $X \sim N_p(\mu, \Sigma)$. Obtain characteristic function of X. **(8+6)**

4. a) State and prove additive property of Wishart distribution.

b) Define canonical correlations and variates. Show that canonical correlation is a generalization of multiple correlation coefficient. **(6+8)**

5. a) Define principal components. State and prove any two properties of principal components.

   b) Obtain the two principal components and percentage of variation explained by these components if $\Sigma = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}$. **(7+7)**

6. a) Develop a test for equality of mean vectors of two multivariate normal populations. State your assumptions clearly.

   b) Let $X \sim N_3(\mu, \Sigma)$. Obtain the distribution of $Y = X_1 - X_2 + X_3$. **(8+6)**

7. a) Explain discriminant function. Derive Fisher's best linear discriminant function.

   b) Describe orthogonal factor model with m common factors. **(8+6)**

SLR-MB – 629

Seat No.	

M.Sc. (Part – II) (Semester – III) Examination, 2016
STATISTICS (Paper – XII) (Old CGPA)
Multivariate Analysis

Day and Date : Thursday, 31-3-2016 Total Marks : 70
Time : 2.30 p.m. to 5.00 p.m.

Instructions : *1) Attempt **five** questions.*
*2) Q. No. (**1**) and Q. No. (**2**) are **compulsory**.*
*3) Attempt **any three** from Q. No. (**3**) to Q. No. (**7**).*
*4) Figures to the **right** indicate **full** marks.*

1. A) Choose the correct alternative :

i) Let $\overline{\underline{X}}$ be the sample mean based on a random sample of size n from $N_p(\underline{\mu}, I)$. What is the distribution of $n(\overline{\underline{X}} - \underline{\mu})' (\overline{\underline{X}} - \underline{\mu})$?

a) $N_p(\underline{\mu}, I)$ b) χ^2_p
c) χ^2_n d) $N_p(\underline{\mu}, I/n)$

ii) The sufficient statistic for $\underline{\mu}$ in $N_p(\underline{\mu}, I)$ based on a random sample X of size n is

a) X b) $XE_{n!}$
c) Both a) and b) d) Neither a) nor b)

iii) The sampling distribution of a p-variate mean vector while sampling from $N_p(\underline{\mu}, \Sigma)$ is

a) $N_p(\underline{\mu}, \Sigma)$ b) χ^2_p
c) $N_p(\underline{\mu}, \Sigma/n)$ d) $N_p(\underline{\mu}/n, \Sigma)$

P.T.O.

iv) There is no gain if PCA is performed on variance-covariance matrix Σ when

a) Σ is identity matrix
b) Σ is a diagonal matrix
c) Both a) and b)
d) Neither a) nor b)

v) Which of the following is not a dimension reduction technique ?

a) PCA
b) Factor analysis
c) Cluster analysis
d) Discriminant analysis

B) Fill in the blanks :

i) Characteristic function of $N_p(\underline{0}, I)$ is ________________

ii) ________________ is used to pictorially represent the process of clustering.

iii) ______________ is used to pictorially represent the variance explained by principal components.

iv) Canonical correlation is generalization of ________________ correlation.

v) _________________ is a non-hierarchical clustering method.

C) State whether **true** or **false** :

i) Hotelling's T^2 = Mahalanobis D^2.

ii) Large Mahalanobis distance between two populations implies small error of misclassification.

iii) If $(X_1, X_2)'$ has dispersion matrix $\begin{pmatrix} 4 & 0 \\ 0 & 5 \end{pmatrix}$ then X_1 and X_2 must be independent.

iv) Chi-square distribution is a particular case of Wishart distribution. **(5+5+4)**

2. a) i) State and prove additive property of Wishart distribution.

ii) Give two equivalent definitions of p-variate non-singular normal distribution. **(3+3)**

b) Write short notes on the following :

i) Single linkage clustering.

ii) Generalized variance. **(4+4)**

3. a) Let $\underset{\sim}{X}$ be distributed as $N_3(\underset{\sim}{0}, \Sigma)$ where $\Sigma = \begin{pmatrix} 4 & 0 & -1 \\ 0 & 5 & 0 \\ -1 & 0 & 2 \end{pmatrix}$, which of the following random variables are independent ? Explain.

   i) $(X_1, X_3)'$ and X_2

   ii) X_1 and $X_1 + 3X_2 - 2X_3$

   iii) $X_1 + X_2$ and $X_1 - X_2$

   iv) $X_1 + X_3$ and $X_1 - X_3$.

   b) Show that $\overline{\underset{\sim}{X}}$ and S are independently distributed when sampling from $N_p(\underset{\sim}{\mu}, \Sigma)$. **(7+7)**

4. a) Derive Wishart distribution in canonical case.

   b) Derive characteristic function of $W_p(f, \Sigma)$. **(7+7)**

5. a) Derive ECM rule for two class classification.

   b) Derive Hotelling's T^2 statistic through Roy's union intersection principle. **(7+7)**

6. a) Derive principal components and interpret them.

   b) Discuss the use of Hotelling's T^2 in the problem of symmetry. **(7+7)**

7. a) Derive a test for testing the need for additional variable for discrimination purpose.

   b) Explain canonical correlation and canonical variates. State and prove their properties. **(7+7)**

____

SLR-ES-378

Seat No.	

Set P

M.Sc. (Semester - III) (CBCS) Examination March/April-2019
Statistics
MULTIVARIATE ANALYSIS

Day & Date: Monday, 29-04-2019 Max. Marks: 70
Time: 03:30 PM To 06:00 PM

Instructions: 1) All questions are compulsory.
2) Figures to the right indicate full marks.

Q.1 Choose the most correct alternative. **14**

1) All of the following techniques are useful for dimension reduction except ______.
   a) clustering b) factor analysis
   c) discriminant analysis d) principal components analysis

2) The mean vector of a random vector (X_1, X_2) is (10,0), then the mean vector of (X_1+X_2,X_1-X_2) is _______.
   a) (10,0) b) (10,10)
   c) (20,20) d) (10,-10)

3) Principal components are _______.
   a) orthogonal b) uncorrelated
   c) independent d) all of these

4) A principal component analysis was run and the following eigen values were obtained: 4.731, 2.218, 0.442. How many components would you retain so that 50% of the variation present in the old variables will be explained?
   a) 1 b) 2
   c) 3 d) 0

5) Generalised variance is ______ of covariance matrix.
   a) trace b) determinant
   c) trace+ determinant d) none of these

6) Wishart distribution is a multivariate generalization of _____.
   a) Normal b) t
   c) chi-square d) F

7) Statistical techniques that focus upon bring out the structure of simultaneous relation among three or more variables are called ______ analysis.
   a) bivariate b) parametric
   c) multivariate d) non-parametric

8) A canonical correlation cannot be negative, because _____.
   a) we take only positive eigen values
   b) it is generalisation of multiple correlation
   c) we take only positive square root
   d) we rejected negative value

9) Let $A \sim W_p(\mu, \Sigma)$ and $\underline{a}$ be a (p x 1) vector which is independently distributed. Then is distributed $\frac{a'Aa}{a'\Sigma a}$ is distributed as _______.
   a) ${\chi^2}_n$ b) ${\chi^2}_p$
   c) ${\chi^2}_{n-p}$ d) ${\chi^2}_{n-p+1}$

10) Based on a random sample of size n from $N_p(\mu, \Sigma)$, the distribution of $\overline{X}$ is ____.

a) $N_p(\mu, \Sigma)$

b) $N_p\left(\mu, \frac{1}{n}\Sigma\right)$

c) $N_p\left(\mu, \frac{1}{n-1}\Sigma\right)$

d) None of these

11) Let $X \sim N_p(\mu, \Sigma)$ then Σ is ______ matrix.

a) a positive definite

b) square

c) symmetric

d) all of these

12) The first principal component is having ______ variance.

a) largest

b) least

c) average

d) none of these

13) Let $\underline{X}$ is multivariate normal, then $\underline{a}'\underline{X}$ is univariate normal, only if _____.

a) $\underline{a}$ is unit vector

b) $\underline{a}$ is zero vector

c) for all $\underline{a}$

d) none of these

14) Let $X \sim N_p(\mu, \Sigma)$ then variance of AX is ______.

a) $A\,\Sigma\,A'$

b) $A'\,\Sigma\,A$

c) $AA\,\Sigma$

d) None of these

Q.2 A) Answer the following (Any Four) **08**

1) Define Mahalanobis distance.
2) Define multivariate normal distribution.
3) Define multiple correlation coefficient.
4) State moment generating function of Wishart distribution.
5) Define Wishart distribution.

B) Writes notes on following. (Any Two) **06**

1) Principal components analysis.
2) Average linkage method of clustering.
3) Additive property of Wishart distribution.

Q.3 A) Answer the following. (Any Two) **08**

1) Let vector X = (X_1, X_2)' be distributed according to $N_p(\mu, \Sigma)$. Then find marginal distribution of X_1.
2) Show that two p-variate normal vectors X_1 and X_2 are independent iff Cov(X_1, X_2) = 0
3) If vector X is distributed according to $N_p(\mu, \Sigma)$ and A is a p x k matrix, then find distribution of AX.

B) Answer the following. (Any One) **06**

1) Define generalized variance. Derive sampling distribution of sample generalized variance when sampling from $N_p(\mu, \Sigma)$.
2) Explain what is meant by singular and non-singular normal distribution.

Q.4 A) Answer the following: (Any Two) **10**

1) Derive characteristic function of Wishart distribution in canonical case.
2) Based on following variance-covariance matrix, obtain two principal components and percentage of variance explained by these components
   $$\Sigma = \begin{bmatrix} 3 & 1 \\ 1 & 4 \end{bmatrix}$$
3) Describe canonical variable and canonical correlations. State and prove any two properties of canonical variables.

B) Answer the following. (Any One) **04**

1) Let vector X be distributed according to $N_p(\mu, \Sigma)$. Then obtain moment generating function of X.
2) Describe Roy's Union-Intersection principle. Show that Roy's Union-Intersection principle leads to Hotelling's T^2 statistic.

Q.5 Answer the following. (Any Two) **14**

a) In usual notations, for $N_p(\mu, \Sigma)$, show that X and S are maximum likely estimators of μ and Σ respectively.
b) Obtain the rule of discrimination for two normal populations Π_1 and Π_2
c) Explain method of clustering. What is meant by agglomerative clustering and divisive clustering? Also explain single linkage and complete linkage.

SLR-JS-379

Seat No.	

Set P

M.Sc. (Semester - III) (CBCS) Examination Oct/Nov-2019
Statistics
MULTIVARIATE ANALYSIS

Day & Date: Tuesday, 05-11-2019 Max. Marks: 70
Time: 03:00 PM To 05:30 PM

Instructions: 1) All questions are compulsory.
2) Figures to the right indicate full marks.

Q.1 Fill in the blanks by choosing correct alternatives given below. **14**

1) Generalised variance is _____ of covariance matrix.
   a) trace b) determinant
   c) trace+ determinant d) none of these

2) The mean vector of a random vector (X_1 X_2) is (3, 5), then the mean vector of $(X_1 + 2X_2, 2X_1 - X_2)$ is _____.
   a) (3, 5) b) (13, 5)
   c) (13,11) d) (13, 1)

3) Principal components are _____.
   a) orthogonal b) uncorrelated
   c) independent d) all of these

4) For a multivariate normal random vector, the variance-covariance matrix is always _____.
   a) square matrix b) non-negative definite
   c) symmetric d) all of these

5) If $\underline{X} \sim N_p\ (\underline{\mu}, \Sigma)$, then for a vector $\underline{a}$, the variable $\underline{a}'.\underline{X}$ follows which distribution?
   a) $N_p(\underline{\mu}, \Sigma)$ b) $N_p(\underline{\mu}, n\Sigma)$
   c) $N_p(\underline{\mu} - \frac{1}{n}\Sigma)$ d) none of these

6) The _____ distribution is a multivariate generalization of chi-square distribution.
   a) Multivariate Normal b) Hotelling's T^2
   c) Wishart distribution d) None of these

7) Statistical techniques that focus upon bringing out the structure of simultaneous relation among three or more variables are called _____ analysis.
   a) bivariate b) parametric
   c) multivariate d) non-parametric

8) A canonical correlation cannot be negative, because _____.
   a) we take only positive eigen values
   b) it is generalisation of multiple correlation
   c) we take only positive square root
   d) we rejected negative value

9) In factor analysis, if there are k variables and m factors, then _____.
   a) $k < m$ b) $m < k$
   c) $m = k$ d) none of these

10) Based on a random sample of size n from $N_p(\mu, \Sigma)$, the distribution of $\bar{X}$ is _____.
   a) $N_p(\mu, \Sigma)$
   b) $N_p(\mu, \frac{1}{n}\Sigma)$
   c) $N_p(\mu, \frac{1}{n-1}\Sigma)$
   d) none of these

11) While applying _____ clustering algorithm, the distance between two clusters is taken to be the smallest distance between observations from two clusters.
   a) single linkage
   b) average linkage
   c) complete linkage
   d) none of these

12) Among the principal components, the _____ principle component has largest variance.
   a) first
   b) last
   c) $(p/2)^{th}$
   d) none of these

13) Let $\underline{X}$ is multivariate normal, then $\underline{a}' \underline{X}$ is univariate normal, only if _____.
   a) $\underline{a}$ is unit vector
   b) $\underline{a}$ is zero vector
   c) for all $\underline{a}$
   d) none of these

14) Let $X \sim N_p(\mu, \Sigma)$ then variance of AX is _____.
   a) $A\,\Sigma\,A'$
   b) $A'\,\Sigma\,A$
   c) $AA\,\Sigma$
   d) none of these

Q.2 A) Answer the following questions. (Any Four) **08**

1) Define multivariate normal distribution.
2) State Wishart density function.
3) Define multiple correlation coefficients.
4) Define Hotelling-T^2 statistics.
5) Define variance covariance matrix.

B) Write notes. (Any Two) **06**

1) Mahalanobis distance
2) Generalised variance
3) Characteristic function of Wishart distribution

Q.3 A) Answer the following questions. (Any Two) **08**

1) Let vector $X = (X_1, X_2, \ldots, X_p)'$ be distributed according to $N_p(\mu, \Sigma)$. Then find marginal distribution of X_1.
2) Show that two p-variate normal vectors X_1 and X_2 are independent iff $Cov(X1, X2) = 0$
3) If vector X is distributed according to $N_p(\underline{\mu}, \Sigma)$ and A is a k x p matrix, then find distribution of AX.

B) Answer the following questions. (Any One) **06**

1) Write short notes on singular and non-singular normal distribution.
2) Explain the technique of principle component.

Q.4 A) Answer the following questions. (Any Two) **10**

1) State and prove additive property of Wishart distribution.
2) How one can find correlation between two multivariate normal random vectors? Explain in brief.
3) Explain in brief the idea of factor analysis.

B) Answer the following questions. (Any One) **04**

1) Derive the moment generating function of $N_p(\mu, \Sigma)$ distribution.
2) Let $A \sim W_p(\mu, \Sigma)$ and $\underline{a}$ be a (p x 1) vector which is independently distributed.
   Then obtain the distribution of $\frac{a'Aa}{a'\Sigma a}$

Q.5 Answer the following questions. (Any Two) **14**

1) Explain method of clustering. What is meant by agglomerative clustering and divisive clustering? Also explain single linkage and complete linkage.

2) Discuss the problem of discrimination for multivariate observation. Also explain costs associated with it.

3) In usual notations, for $N_p(\underline{\mu}, \Sigma)$, show that X and S are maximum likely estimators of $\underline{\mu}$ and Σ respectively.