National Testing Agency

Question Paper Name:	Applied Multivariate Analysis 09th November 2019 Shift 1
Subject Name:	Applied Multivariate Analysis
Creation Date:	2019-11-09 15:06:25
Duration:	180
Total Marks:	100
Display Marks:	Yes

Applied Multivariate Analysis

Group Number :	1
Group Id :	709597221
Group Maximum Duration :	0
Group Minimum Duration :	120
Revisit allowed for view? :	No
Revisit allowed for edit? :	No
Break time:	0
Group Marks:	100

Applied Multivariate Analysis-1

Section Id :	709597244
Section Number :	1
Section type :	Online
Mandatory or Optional:	Mandatory
Number of Questions:	10
Number of Questions to be attempted:	10
Section Marks:	20
Display Number Panel:	Yes
Group All Questions:	No

Sub-Section Number:	1
Sub-Section Id:	709597339
Question Shuffling Allowed :	No

Question Number : 1 Question Id : 70959721376 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 2 Wrong Marks : 0

In the Single Linkage method, the distance between two clusters is defined as

(A) the distance between the closest two members of the two clusters

(B) the distance between the farthest two members of the two clusters

(C) the distance between the centroids of the two clusters

(D) the average distance between two members of the two clusters

Question Number : 2 Question Id : 70959721377 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 2 Wrong Marks : 0

If in a 2 × 2 table with *a* Yes-Yes, *b* Yes-No, *c* No-Yes and *d* No-No, the *matching proportions* similarity measure is given by

(A) (a+c)/(a+b+c+d)

(B) (a+b)/(a+b+c+d)

(C) (a+d)/(a+b+c+d)

(D) (b+c)/(a+b+c+d)

Question Number : 3 Question Id : 70959721378 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 2 Wrong Marks : 0

In the *divisive hierarchical clustering* method with n units, we

(A) start with a single cluster of n units and split it into n/2 clusters of 2 units each

(B) start with n clusters of 1 unit each and merge them into 2 clusters of n/2 units

(C) start with a single cluster of n units and split it into n clusters of 1 unit each

(D) start with n clusters of 1 unit each and merge them into 1 cluster of n units

Question Number : 4 Question Id : 70959721379 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 2 Wrong Marks : 0

The discriminating function for two groups with underlying multivariate normal densities is *linear* if

(A) their means are equal

(B) their dispersion matrices are equal

(C) their means are unequal

(D) their dispersion matrices are unequal

Question Number : 5 Question Id : 70959721380 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 2 Wrong Marks : 0

Principal component analysis is primarily used

(A) as a data reduction technique

(B) to study cause-effect relationships

(C) as a grouping technique

(D) to study inter variable relationships

Question Number : 6 Question Id : 70959721381 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 2 Wrong Marks : 0

For a m-dimensional random variable, the sum of the variances of its first p (< m) principal components is

(A) greater than the total variability of the m variables

(B) equal to the total variability of any of p of the m variables (p < m)

(C) less than the variability of each of the m variables

(D) less than the total variability of the m variables

Question Number : 7 Question Id : 70959721382 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 2 Wrong Marks : 0

The part of the variance of a variable explained by the factors is known as

(A) communality

(B) uniqueness

(C) factor loading

(D) factor score

Question Number : 8 Question Id : 70959721383 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 2 Wrong Marks : 0

The dispersion matrix Σ of a factor model can be split as (where L is the matrix of loadings and Ψ is the error dispersion matrix) :

(A) $\Sigma = LL' + I$

(B) $\Sigma = L + \Psi$

(C) $\Sigma = I + L\Psi L'$, where I is the identity matrix

(D) $\Sigma = LL' + \Psi$.

Question Number : 9 Question Id : 70959721384 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 2 Wrong Marks : 0

For the pair of canonical variables (U_j, V_j), j= 1, …, p, which of the following is true :

(A) $Cov(U_j, V_j) = 0$ for all j and $Cov(U_j, V_k) = 0$ for all $j \neq k$.

(B) $Cov(U_j, V_j) \neq 0$ for all j and $Cov(U_j, V_k) = 0$ for all $j \neq k$.

(C) $Cov(U_j, V_j) = 0$ for all j and $Cov(U_j, V_k) \neq 0$ for all $j \neq k$.

(D) $Cov(U_j, V_j) \neq 0$ for all j and $Cov(U_j, V_k) \neq 0$ for all $j \neq k$.

Question Number : 10 Question Id : 70959721385 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 2 Wrong Marks : 0

If $X \sim N_p(\mu, \Sigma)$, then for a $q \times p$ matrix L, $Y = LX$ follows

(A) $N_q(\mu, \Sigma)$

(B) $N_q(L\mu, L\Sigma)$

(C) $N_q(L\mu L', L\Sigma L')$

(D) $N_q(L\mu, L\Sigma L')$

Applied Multivariate Analysis-2

Section Id :	709597245
Section Number :	2
Section type :	Offline
Mandatory or Optional:	Mandatory
Number of Questions:	10
Number of Questions to be attempted:	10
Section Marks:	30
Display Number Panel:	Yes
Group All Questions:	No

Sub-Section Number:	1
Sub-Section Id:	709597340
Question Shuffling Allowed :	No

Question Number : 11 Question Id : 70959721386 Question Type : SUBJECTIVE Display Question Number : Yes Correct Marks : 3

Distinguish between *hierarchical clustering* and *partitioning*.

Question Number : 12 Question Id : 70959721387 Question Type : SUBJECTIVE Display Question Number : Yes Correct Marks : 3

Distinguish between *Euclidean* and *City-block* Distances.

Question Number : 13 Question Id : 70959721388 Question Type : SUBJECTIVE Display Question Number : Yes Correct Marks : 3

What is the *optimum error rate* in a classification problem ? (No derivation/expression required)

Question Number : 14 Question Id : 70959721389 Question Type : SUBJECTIVE Display Question Number : Yes Correct Marks : 3

Discuss Fisher's *separator* for classifying two groups.

Question Number : 15 Question Id : 70959721390 Question Type : SUBJECTIVE Display Question Number : Yes Correct Marks : 3

Discuss how you will choose the number of principal components to be used.

Question Number : 16 Question Id : 70959721391 Question Type : SUBJECTIVE Display Question Number : Yes Correct Marks : 3

What is the contribution of the first factor to the total variability of all the variables ?

Question Number : 17 Question Id : 70959721392 Question Type : SUBJECTIVE Display Question Number : Yes Correct Marks : 3

Discuss the *weighted least squares* method of finding factor score.

Question Number : 18 Question Id : 70959721393 Question Type : SUBJECTIVE Display Question Number : Yes Correct Marks : 3

. Write down the characteristic function of a multivariate normal distribution.

Question Number : 19 Question Id : 70959721394 Question Type : SUBJECTIVE Display Question Number : Yes Correct Marks : 3

What will be the effect of standardization on the canonical correlation ?

Question Number : 20 Question Id : 70959721395 Question Type : SUBJECTIVE Display Question Number : Yes Correct Marks : 3

Find the confidence region for the population mean vector of a multivariate normal distribution.

Applied Multivariate Analysis-3

Section Id :	709597246
Section Number :	3
Section type :	Offline
Mandatory or Optional:	Mandatory
Number of Questions:	8
Number of Questions to be attempted:	5
Section Marks:	50
Display Number Panel:	Yes
Group All Questions:	No

Sub-Section Number:	1
Sub-Section Id:	709597341
Question Shuffling Allowed :	No

Question Number : 21 Question Id : 70959721396 Question Type : SUBJECTIVE Display Question Number : Yes Correct Marks : 10

Make a detailed discussion on Gower's composite measure.

Question Number : 22 Question Id : 70959721397 Question Type : SUBJECTIVE Display Question Number : Yes Correct Marks : 10

For 4 units A, B, C and D with distances $d_{AB} = 5$, $d_{AC} = 8$, $d_{AD} = 4$, $d_{BC} = 3$, $d_{BD} = 5$ and $d_{CD} = 6$, form the dendogram using agglomerative hierarchical clustering method.

Question Number : 23 Question Id : 70959721398 Question Type : SUBJECTIVE Display Question Number : Yes

Correct Marks : 10

Given the apriori probabilities, joint densities and the misclassification costs for two groups, obtain a classification rule based on the minimum expected misclassification cost.

Question Number : 24 Question Id : 70959721399 Question Type : SUBJECTIVE Display Question Number : Yes Correct Marks : 10

Discuss how you will derive the first principal component.

Question Number : 25 Question Id : 70959721400 Question Type : SUBJECTIVE Display Question Number : Yes Correct Marks : 10

Describe how, in a factor model, you will estimate the factor loadings using the principal component technique.

Question Number : 26 Question Id : 70959721401 Question Type : SUBJECTIVE Display Question Number : Yes Correct Marks : 10

Show that the first canonical correlation is larger than any multiple correlation involving a variable from the first set as response and the variables in the second as regressors.

Question Number : 27 Question Id : 70959721402 Question Type : SUBJECTIVE Display Question Number : Yes Correct Marks : 10

If $X \sim N_p(\mu, \Sigma)$, where Σ unknown, describe how you will test $H_0 : \mu = \mu_0$ (known) against $H_1 : \mu \neq \mu_0$.

Question Number : 28 Question Id : 70959721403 Question Type : SUBJECTIVE Display Question Number : Yes Correct Marks : 10

Describe the least square method of estimating the parameters of a multivariate linear model.