SARDAR PATEL UNIVERSITY
M.Sc. (IV Semester) Examination
2019
Monday, 18th March
2:00 pm to 5:00 pm
STATISTICS COURSE No. PS04CSTA21/PS04CAST21/ PS04CSTA01
(Computer Oriented Statistical Methods)

Note: Figures to the right indicate full marks of the questions. (Total Marks: 70)

1   Attempt all, write correct answers   08

(i)   The value of X_2 according to the LCG $(13X_i +1)$mod 16 with $X_0=5$ is _____
   a) 9   b) 10
   c) 11   d) 12

(ii)   Inverse transform function for p.d.f. $f(x) = x/2 \; 0<x<2$ in usual notation is _____
   a) a constant   b) $2\sqrt{u}$

   c) $\sqrt{u}/2$ d) $\pm 2\sqrt{u}$

(iii) Given a U(0, 1) uniform number u, a geom(u) deviate generated as per discrete inverse transform method is _____.
   a) 1   b) 2

   c) 3   d) 4

(iv)   The mathematical result explaining Gibbs sampler method is known as _____

       a) SLLN Theorem   b) Hull Dobell Theorem
       c) Hammersely-Clifford Theorem d) Birkhoff Ergodic Theorem

(v)   In Principal Component Analysis, the principal component loading is measured by
       a) $g_{ij}\sqrt{l_j}/\sqrt{S_{ii}}$ b) $g_{ij}l_j/\sqrt{S_{ii}}$
       c) $g_{ij}^2\sqrt{l_j}$ d) g_{ij}/S_{ii}

(vi)   In Factor analysis, if the KMO is greater than 0.6 and less than or equal to 0.7 we say that the Factor Analysis is _____.
       a) unacceptable   b) meritorious

       c) mediocre   d) None of these

(vii) In linear discriminant analysis if number of groups is 5 and number of explanatory variables are 3 then _____ linear discriminant score function can be obtained.

       (a) 3   (b) 4
       (c) 5   (d) 6

(viii) In _____ method if the initial assignment of an object to a cluster is inaccurate, it may be reallocated to another cluster.

       a) Agglomerative   b) Divisive
       c) Adjoining   d) K-means Clustering

2   Attempt ANY 7, each carries 2 marks   14
(a)   Define LCG and combined LCG. Give an example of a standard LCG.

(b) State and prove the theorem of acceptance-rejection method.

(c) Why do we need bootstrap method? How it works?

(d) Define Cox proportional hazard model with essential features and assumptions.

(e) In usual notation, obtain mean and variance of Principal Component.

(f) Define MSA and give its application.

(g) In context of Canonical Correlation, define redundancy coefficients of the variable and give an application.

(h) Explain briefly, how to test linear discriminant function.

(i) Write only one difference between single linkage method, complete linkage method and average linkage method.

3(a) Derive the Box-Muller method and an acceptance – rejection method of generating   **06**
standard normal deviates. Explain the procedure of converting the standard normal deviates into the generation of nonstandard normal deviates.

3(b) Show that If U and V are iid U[0, 1] deviates and $U^{1/a} + V^{1/b} \leq 1$ then   **06**
$U^{1/a} / (U^{1/a} + V^{1/b})$ follows Beta (a, b) distribution, a, b >0.

OR

Give algorithms for generating random deviate that <u>make use of</u> standard normal deviates. (at least two).

4(a) What are MCMC methods? Stating two names of this method, explain one of them   **06**
with mathematical result behind it.

4(b) What is importance sampling? Obtain Monte Carlo estimate of the integral   **06**
$\int_0^\infty \int_{\frac{1}{2}}^1 \int_0^{\frac{2}{3}} (x + y)e^z dxdydz$ so that it has minimum variance.

OR

Describe giving example, the variance reduction technique based on use of control variate.

5(a) Discuss step by step procedure of performing Principal Component Regression   **06**
Analysis.

5(b) Define canonical correlation coefficient. Stating the objective, discuss Canonical   **06**
Correlation Analysis.

OR

Explain Cluster Analysis. How is it different from other multivariate techniques you know for example principal component analysis and others?

6(a) Distinguish between linear discriminant analysis and binary logistic regression.   **06**
Explain procedure for performing the former.

6(b) (i)Explain the impact of ignoring some components from Principal Component   **06**
Analysis. (ii) Discuss various distance measures used in cluster analysis for quantitative variables.

OR

Explain factor analysis including the difference between the confirmatory factor analysis and exploratory factor analysis.

— X —

2