INDIAN FOREST SERVICE P (EXAM)-2014

C–HENT–N–RUZUA

STATISTICS

Paper—I

Time Allowed : Three Hours

Maximum Marks : 200

QUESTION PAPER SPECIFIC INSTRUCTIONS

Please read each of the following instructions carefully before attempting questions :

There are **EIGHT** questions in all, out of which, **FIVE** are to be attempted.

Question nos. **1** and **5** are compulsory. Out of the remaining **SIX** questions, **THREE** are to be attempted selecting at least **ONE** question from each of the two Sections A and B.

Attempts of questions shall be counted in chronological order. Unless struck off, attempt of a question shall be counted even if attempted partly. Any page or portion of the page left blank in the answer book must be clearly struck off.

All questions carry equal marks. The number of marks carried by a question/part is indicated against it.

Answers must be written in **ENGLISH** only.

Unless otherwise mentioned, symbols and notations have their usual standard meanings.

Assume suitable data, if necessary and indicate the same clearly.

SECTION—A

Q. 1. **Answer the following :** **8×5=40**

Q. 1(a) (i) If A, B and C are events in a sample space, show that :

$$P(A \cup B \cup C) \le P(A) + P(B) + P(C).$$

(ii) B_1, B_2 and B_3 are mutually exclusive events with $P(B_j) = 1/3$ and $P(A \mid B_j) = j/6$, $j = 1, 2, 3$. Evaluate P(A). 8

Q. 1(b) (X, Y) is a bivariate random vector with $P(X > x, Y > y) = \exp(-x - y - xy)$; $x, y \ge 0$. Find the marginal and conditional distributions. Also find $E(Y \mid X)$. 8

Q. 1(c) Obtain the distribution for which the characteristic function is :

$$\phi(t) = \exp(-\mid t \mid).$$

If X and Y are independent random variables having the characteristic function $\phi(t)$, state the distribution of (X + Y). 8

Q. 1(d) Obtain $100(1 - \alpha)\%$ confidence interval for the ratio of variances based on two samples from $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$ when both μ_1 and μ_2 are (i) known and (ii) unknown. 8

Q. 1(e) Find the maximum likelihood estimate of α in the density :

$$f(x; \alpha) = (\alpha + 1)x^{\alpha}, \quad 0 < x < 1$$
$$= 0 \quad , \text{otherwise.}$$ 8

Q. 2(a) Identify a distribution, each, for which (i) mean < variance and (ii) mean > variance.

Among the 120 applicants for a job only 80 are actually qualified. If five of the applicants are randomly selected for an in-depth interview, find the probability that at least two of the five will be from among those qualified for the job. 10

Q. 2(b) If $X_n \xrightarrow{P} X$ and $Y_n \xrightarrow{P} Y$, discuss about the convergence of $(X_n + Y_n)$ and $X_n Y_n$. 10

Q. 2(c) X is a random variable having the binomial distribution with parameters n and θ. Show that the distribution of :

$$Z = \frac{X - n\theta}{\sqrt{n\theta(1-\theta)}}$$

approaches the standard normal distribution as $n \to \infty$. 10

Q. 2(d) A fair coin is tossed until a head appears. If X denotes the number of tosses required, find (i) the p.d.f. of X, (ii) probability generating function of X and (iii) the mean and variance of X. 10

Q. 3(a) For what value of λ, there exists an UMP test for testing :

$$H_0 : \mu = \mu_0, \lambda = \lambda_0$$

when a sample is drawn from a distribution with density :

$$f(x; \mu, \lambda) = \frac{1}{\lambda} e^{-(x-\mu)/\lambda}, \mu < x < \infty.$$ 10

Q. 3(b) (i) If T is unbiased for θ, show that T^2 is biased for θ^2.

(ii) If T is consistent for θ, show that T^2 is also consistent for θ^2. 10

Q. 3(c) Find the maximum likelihood estimate of θ based on a random sample from a distribution with density :

$$f(x) = \frac{1}{2} e^{-|x-\theta|}, -\infty < x < \infty.$$ 10

Q. 3(d) Write short notes on :

(i) Run test

(ii) Kolmogorov-Smironov test

with illustrative examples. 10

Q. 4(a) If the random variable X has an exponential distribution, show that the relationship :

$$P(X > t + s \mid X > t) = P(X > s)$$

holds for all real t, s > 0. Discuss about the utility of this result in life time data analysis. 10

Q. 4(b) Examine the validity of the statement "the initial probability vector and the t.p.m. uniquely determine a Markov chain". Explain how you will classify the states in a Markov chain. Illustrate by an example. 10

Q. 4(c) If $X_1, X_2, \dots\dots\dots, X_n$ is a random sample from a distribution with a discrete probability function :

$$p(x;\theta,p) = \begin{cases}(1-p)p^{x-\theta}, & x=\theta, \theta+1, \dots\dots\dots \\ & 0<p<1 \\ = 0 & \text{, otherwise,}\end{cases}$$

find a sufficient statistic for θ. (Assume p is known) 10

Q. 4(d) Find the MVB (minimum variance bound) estimator for the parameter of Poisson distribution and obtain the value of MVB. 10

SECTION—B

Q. 5. **Answer all of the following :** **8×5=40**

Q. 5(a) Three independent observations follow the Gauss-Markoff linear model :

$$E(X_1)=\frac{1+\theta}{4}, E(X_2)=\frac{1-\theta}{4}, E(X_3)=\frac{1}{2}.$$

Find all error functions and the error degrees of freedom. 8

Q. 5(b) Give an example to show that the normality of the conditional probability density function (pdf) does not imply the bivariate density to be normal. 8

Q. 5(c) Define cluster sampling for proportions and write down the situations where it is used. Also find the unbiased estimator of population proportion. 8

Q. 5(d) Distinguish between (i) fixed effect, and (ii) random effect models. Find the expected value of mean squares, for two-way classified data with one observation per cell under the random effect model. 8

Q. 5(e) Develop Fisher's inequality for a BIBD. 8

Q. 6(a) Consider two independent experiments leading to the observation of a 4-vector $Y^{(1)}$ and a 2-vector $Y^{(2)}$ respectively. $Y^{(1)}$ is subject to the Gauss-Markoff model $(Y^{(1)}, X^{(1)}\beta, \sigma^2 I_4)$

with $X^{(1)} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix}$ and $Y^{(2)}$ is subject to the model $(Y^{(2)}, X^{(2)}\beta, \sigma^2 I_2)$ with

$X^{(2)} = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 \end{pmatrix}$, $\beta = (\beta_1, \beta_2, \beta_3, \beta_4)'$ and $\sigma^2 > 0$ being same in both cases.

(i) Find the BLUEs $\hat{\beta}_1^{(1)}, \hat{\beta}_1^{(2)}$ and $\hat{\beta}_1$ of β_1 on the basis of $Y^{(1)}$ alone, $Y^{(2)}$ alone and $Y^{(1)}$, $Y^{(2)}$ taken together respectively.

(ii) Find the best linear combination of $\hat{\beta}_1^{(1)}$ and $\hat{\beta}_1^{(2)}$ and compare its precision with that of $\hat{\beta}_1$. 10

Q. 6(b) A car manufacturing company plans to test the average life of each of the four brands of tyres. The company uses all the four brands on randomly selected cars. The records showing the lives (hundreds of miles) of tyres are given in the following table below :

Brand 1	:	20	23	18	17	
Brand 2	:	19	15	17	20	16
Brand 3	:	21	19	20	17	16
Brand 4	:	15	17	16	18	

Test the hypothesis that the average life for each brand of tyres is the same. Assume 1% level of significance. ($F_{3, 14, .01} = 5.56$ and $F_{4, 13, .01} = 5.21$) 10

Q. 6(c) Let Y be a univariate random variable (r.v.) and **X** be a p × 1 random vector. Let $\begin{pmatrix} \mathbf{X}_\alpha \\ Y_\alpha \end{pmatrix}$,

$\alpha = 1 \ldots\ldots N$ be the data available on **X** and Y.

(i) Define the sample principal component of X and discuss their uses.

(ii) Compare the multiple correlations between Y and **X** and between Y and **U** where **U** is a p × 1 sample principal component of **X**.

(A variance-covariance matrix for **X** and Y can be taken for the above study) 10

Q. 6(d) Consider three sample observations from a trivariate normal distribution $N_3(\mu, \Sigma)$ where

$\mu = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}$. Let the observations be X_1, X_2 and X_3. Write $Y_1 = X_1 + X_2$, $Y_2 = X_2 + X_3$ and $Y_3 = X_3 + X_1$. Then obtain :

(i) the joint distribution of Y_1, Y_2 and Y_3.

(ii) the marginal distribution of Y_3.

(iii) the conditional mean and variance of Y_1 given $Y_2 = 3$ and $Y_3 = 4$.

(iv) the conditional mean and dispersion of $\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix}$ given $Y_3 = 3$. 10

Q. 7(a) When is an Incomplete Block Design said to be connected ? Show that a connected incomplete block design is balanced if and only if all the non-zero characteristic roots of c-matrix are equal. 10

Q. 7(b) Write down the need of factorial designs and explain the analysis of a 2^3 factorial design, clearly stating the assumption used. 10

Q. 7(c) Show that the variance of mean of systematic sample is : $\mathrm{Var}(\bar{y}_{sys}) = \frac{N-1}{n}\frac{S^2}{n}[1+(n-1)\rho]$, where ρ is the correlation between pairs of units that are in the same systematic sample. 10

Q. 7(d) Explain the need of probability proportion to size sampling, and obtain the variance of Horvitz-Thompson estimator. 10

Q. 8(a) Show that the prediction of X_1 by means of a linear equation in $X_2, X_3, \ldots\ldots\ldots X_{p-1}$ will be improved by including X_p as well (as an independent variable) iff $\rho_{1p.23....(p-1)}$ is non zero. (Necessary result needs to be proved). 10

Q. 8(b) The following summarized data refer to a sample of approximately 2-year old boys from Sikkim. For low land children of the same age, the height, chest and MUAC (mid uppper arm circumference) means are considered to be 90, 58 and 16 cm. respectively. Test the hypothesis that the Sikkim children also have the same means.

$$\overline{X} = (82.0, 60.2, 14.5)', S^{-1} = \frac{1}{23.14}\begin{pmatrix} 4.31 & -14.62 & 8.95 \\ -14.62 & 59.79 & -37.38 \\ 8.95 & -37.38 & 35.59 \end{pmatrix}$$

$(F_{.01, 3, 3} = 29.5)$ 10

Q. 8(c) What are the situations for which two phase sampling is used ? Find the unbiased estimator of population total for such a sampling scheme and obtain its variance. 10

Q. 8(d) Discuss the classification of Lattice designs, and the allocation of treatments within the complete block in each replication. 10