HPPSC Forest Range Officer

Previous Year Paper

Mains 2017 Statistics

This question paper contains 8+2 printed pages]

CODE : FRO-2017

STATISTICS

Roll No. ....

Time : 3 Hours *Maximum Marks :* 200

Note :— (*i*) **Question paper** consists of two parts viz. Part I and Part II. Each part contains four questions. The paper as a whole carries eight questions. Question Nos. 1 and 5 are compulsory. The candidates are required to attempt *three* more questions out of the remaining six questions taking at least *one* question from each part i.e. this is in addition to the compulsory question of each part. Attempt *five* questions in all. All questions carry equal marks. The parts of a question are to be attempted at one place in continuation. Answers should be brief and to the points.

(*ii*) Parts of same question must be attempted together and not to be attempted in between the answers to other questions.

P.T.O.

Part-I

1. (*i*) Prove that if A, B and C are random events in a sample space and if A, B, C are pairwise independent and A is independent of $(B \cup C)$ then A, B and C are mutually independent. 8

(*ii*) Two discrete random variables X and Y have the joint probability mass function :

$$p(x, y) = \frac{\lambda^x e^{-\lambda} p^y (1-p)^{x-y}}{y!(x-y)!} \quad \begin{array}{l} y = 0,1,2,\ldots\ldots\ldots x \\ x = 0,1,2\ldots\ldots\ldots \end{array}$$

where λ, p are constants with $\lambda > 0$ and $0 < p < 1$. Find the marginal probability mass function of X and Y and the conditional distribution of Y for a given X. 8

(*iii*) If X and Y are random variables taking real values, then prove that : 8

$$[E(XY)]^2 \leq E(X^2) \, . \, E(Y^2).$$

(*iv*) Two unbiased dice are thrown. If X is the sum of numbers showing up, prove that : 8

$$P(|X - 7| \geq 3) \leq \frac{35}{54}.$$

Compare this with actual probability.

(v) If $X_i = \begin{cases} 1 \text{ with Probability } p \\ 0 \text{ with Probability } q \end{cases}$

then prove that the distribution of the random variable $S_n = X_1 + X_2 + \dots\dots\dots + X_n$ where $X_i'.s$ are independent is asymptotically normal as $n \to \infty$ 8

2. (i) If n_1, n_2 are the sizes ; $\bar{x}_1, \bar{x}_2$ the means and σ_1, σ_2 the standard deviations of two series, then the standard deviation σ of the combined series is given by : 10

$$\sigma^2 = \frac{1}{n_1 + n_2}\left[n_1(\sigma_1^2 + d_1^2) + n_2(\sigma_2^2 + d_2^2)\right]$$

where $d_1 = \bar{x}_1 - \bar{x}$, $d_2 = \bar{x}_2 - \bar{x}$ and $\bar{x} = \frac{n_1\bar{x}_1 + n_2\bar{x}_2}{n_1 + n_2}$ is the mean of combined series.

(ii) If $b\ (r; n, p) = {}^nC_r p^r q^{n-r}$ is the binomial probability in the usual notations and if :

$B\ (k\ ; n, p) = \sum_{r=0}^{k} b(r; n, p)$ then prove that

$B\ (k\ ; n, p) = (n - k)\ {}^nC_k \int_0^q t^{n-k-1}(1-t)^k\, dt,$

where $q = 1 - p$. 15

P.T.O.

(*iii*) Explain bivariate normal distribution and obtain its moment generating function. 15

3. (*i*) Explain the terms sampling distribution and standard error. Derive the distribution of student's t-statistic and hence find its mean and variance. 15

(*ii*) How will you test the significance for the difference of standard deviations in case of large samples ?

Random samples drawn from two countries give the following data relating to the heights of the adult males :

	Country A	**Country B**
Mean heights in inches	67.42	67.25
Standard deviations	2.58	2.50
Number of samples	1000	1200

(a) Is the difference between the means significant ?

(b) Is the difference between standard deviations significant ? 15

(iii) If χ_1^2 and χ_2^2 are two independent χ^2 variates with n_1 and n_2 *d. f.* respectively, then derive the distribution of $\frac{\chi_1^2}{\chi_2^2}$. 10

4. (i) Define consistency and unbiasedness properties of an estimator. Prove that for Cauchy's distribution not sample mean but sample median is a consistent estimator of the population mean. 10

(ii) If T_1 and T_2 be two unbiased estimates of a parameter θ with variances σ_1^2 and σ_2^2 and correlation f. What is the best unbiased linear combination of T_1 and T_2 and what is the variance of such an estimator. 15

(*iii*) Find the maximum likelihood estimator (MLE) of the parameters α and λ (λ being large) of the distribution :

$$f(x;\ \alpha,\ \lambda) = \frac{1}{\overline{|\lambda}} \left(\frac{\lambda}{\alpha}\right) e^{-\lambda \frac{x}{\alpha}} x^{\lambda-1};$$

$0 \leq x < \infty$, $\lambda > 0$. You may use that for large values of λ.

$$\Psi(\lambda) = \frac{\partial}{\partial\lambda} \log \overline{|(\lambda)} = \log \lambda - \frac{1}{2\lambda} \text{ and}$$

$$\Psi'(\lambda) = \frac{1}{\lambda} + \frac{1}{2\lambda^2}.$$ **15**

Part-II

5. (*i*) Differentiate between most powerful and uniformly most powerful critical regions. State and prove Nayman-Pearson's lemma. 10

(*ii*) If $x \geq 1$ is the critical region for testing $\theta = 2$ against the alternative $\theta = 1$ on the basis of a single observation from the population $f(x, \theta) = \theta e^{-\theta x}$, $0 \leq x < \infty$. Obtain the values of type 1st and type 2nd errors. 10

(*iii*) Explain likelihood ratio test. How will you use this test for the mean of normal population. 10

(*iv*) Describe the procedure in median test when there are two independent samples. The win-loss record of a certain basketball team for their last 50 consecutive games was as follows :

WWWWWWLWWWWWWLWLWWWLLWWWW

LWWWLLWWWWWWLLWWLLLWWLWWW.

Apply run test to test the sequence of win and losses is random. 10

6. (*i*) If (x_i, y_i) are the pair of variates defined for every unit (i = 1, 2,..........., N) of the population, and $\bar{x}_n$ and $\bar{y}_n$ are the corresponding sample means of a simple random sampling of size n taken without replacement, then prove that : 15

$$\text{Cov}\ (\bar{x}_n, \bar{y}_n) = \frac{N-n}{nN} \cdot \frac{1}{N-1} \sum_{i=1}^{N} (x_i - \bar{x}_N)(y_i - \bar{y}_N)$$

P.T.O.

(ii) In stratified random sampling with given cost function of the form : $C = a + \sum_{i=1}^{K} c_i . n_i .$ where 'a' is the overhead cost and c_i is the cost per unit in the i-th stratum then prove that var $(\overline{y}_{st})$ is minimum if $n_i \propto \frac{N_i S_i}{\sqrt{C_i}}$. 10

(iii) If the population consists of a linear trend, then prove that :

$$\text{Var } (\overline{y}_{st}) \leq \text{Var } (\overline{y}_{sys}) \leq \text{Var } (\overline{y}_n)_R .$$ 15

7. (i) How will you estimate one missing value in case of Randomised Block Design (R.B.D.). Give the analysis of R.B.D. in case of missing observations. 20

(ii) Define Balanced Incomplete Block Design (BIBD). For a resolvable BIBD with parameters v, b, r, k and λ prove that : 10

$$b \geq v + r - 1.$$

(*iii*) What do you mean by factorial experiments ? Obtain the ANOVA table for 2^3 – experiment conducted in RBD with r replications. 10

8. (*i*) Explain in brief ratio estimation. In large samples, with simple random sampling, the ratio estimate $\widehat{Y}_R$ has smaller variance than the estimate $\widehat{Y} = N\overline{y}$ obtained by simple expansion. if :

$$f > \frac{1}{2} \left(\frac{S_x}{\overline{\overline{X}}}\right) \Big/ \left(\frac{S_y}{\overline{\overline{Y}}}\right)$$

where the symbols have their usual meaning. 10

(*ii*) A simple random sample of n clusters, each containing M elements, is drawn from N clusters in the population. Prove that the sample mean per element $\overline{\overline{y}}$ is an unbiased estimate of $\overline{\overline{Y}}$ with variance :

$$\text{Var}\ (\overline{\overline{y}}) = \frac{1-f}{n} \cdot \frac{NM-1}{M^2(N-1)}\ S^2\ [1 + (M - 1)\ f]$$

where the symbols have their usual meaning. 15

P.T.O.

(*iii*) What is the effect on the analysis of variance if the assumptions are not satisfied ? Give some transformations of variate to stabilise variance. 15