Civil Services (Main) Examination, 1999

STATISTICS

(Paper I)

Time Allowed : Three Hours **Maximum Marks : 300**

QUESTION PAPER SPECIFIC INSTRUCTIONS

(Please read each of the following instructions carefully before attempting questions)

There are SIXTEEN questions divided in four Sections.
Candidate has to attempt FIVE questions in all, selecting at least ONE and not more than TWO questions from each Section.
The number of marks carried by a question/part is indicated against it.
Wherever any assumptions are made for answering a question, they must be clearly indicated.
Word limit in questions, wherever specified, should be adhered to.

Question 11 of this paper could not be traced in any available source and is not reproduced.

SECTION 'A'

1. (a) State and prove Baye's theorem. **20**

(b) An urn contains r red and w white balls, k balls are drawn from the urn and kept aside without noticing their colour. From the remaining $r+w-k$ balls one ball is drawn. What is the probability that it is red? **15**

(c) Let F be the distribution function of a random variable X. Show that F can have at most countably many discontinuities. **25**

2. (a) Prove that

$$\sum_{r=0}^{n} (e^{-\lambda}\lambda^r)/r! = 1/n! \int_{\lambda}^{\infty} e^{-x}x^n\, dx \text{ where } \lambda>0$$ **20**

(b) X_1 and X_2 are independent random variables with Poisson distribution. Find the conditional distribution of X_1 given X_1+X_2. **20**

(c) State and prove Borel-Cantelli lemma. **20**

3. (a) If X is an integer-valued random variable, prove that $E(x)=\sum_{k=0}^{\infty} P(X>k)$. Prove the corresponding result for a non-negative continuous random variable. **25**

(b) A man with n keys wants to open his door and tries the keys independently and at random. Only one key fits the door. Find the mean and variance of the number of trials required to open the door if unsuccessful keys are not eliminated from further selections. **15**

(c) Show that a distribution is symmetric about the origin iff its characteristic function is real-valued and even. **20**

4. (a) Define convergence in distribution and convergence in probability for a sequence of random variables. State possible inter-relations and illustrate. **20**

(b) State and prove Lindeberg-Levi central limit theorem. **20**

(c) $X_1, X_2, \ldots, X_n$ is a sequence of independent random variables with $P[X_k=\pm k^{\alpha}]=1/2$ where $\theta>0$. Prove that the law of large numbers holds for this sequence if $\alpha<1/2$. **20**

SECTION 'B'

5. (a) Define a complete sufficient statistic. State Lehmann-Scheffe theorem and explain how it can be used for finding uniformly minimum variance unbiased estimators. **30**

(b) Let r be the observed number of successes in n Bernoullian trials with probability of success θ. Prove that r is a complete sufficient statistic for θ. Hence or otherwise, obtain the minimum variance unbiased estimators of (i) θ and (ii) $\theta(1-\theta)$. **30**

6. (a) State and prove Cramer-Rao inequality. Give its generalization for several parameters. **30**

Optional Province · optionalprovince.com

(b) A random sample of size n is drawn from a gamma distribution with parameters α and p. Obtain estimators of α and p, using the method of moments. **15**

(c) Derive the likelihood ratio test for testing the hypothesis $\mu=\mu_0$ against $\mu\neq\mu_0$ on the basis of a random sample of size m from a normal distribution with mean μ and variance unity. **15**

7. (a) Explain the method of maximum likelihood for estimation of parameters. State important properties of maximum likelihood estimators and prove any one of them. **35**

(b) Given a random sample of size n from the p.d.f. $f(x;\mu)=e^{-(x-\mu)}$, $x>\mu$, $-\infty<\mu<\infty$ derive the UMA confidence interval for μ. **25**

8. (a) Prove that under certain conditions, to be stated, Wald's PRT for testing a simple hypothesis against a simple alternative terminates with probability one. **30**

(b) Describe Kolmogorov-Smirnov test of goodness of fit, stating necessary assumptions, for both one-sided and two-sided alternatives. **30**

SECTION 'C'

9. (a) State and prove Gauss-Markoff theorem. **30**

(b) y_1, y_2 and y_3 are uncorrelated observations with a common variance σ^2 and expectations

$$E(y_1)=\theta_1+\theta_2,\ E(y_2)=\theta_1+\theta_3 \text{ and } E(y_3)=\theta_1+\theta_2$$

(i) Show $l_1\theta_1+l_2\theta_2+l_3\theta_3$ that is estimable if and only if $l_1=l_2+l_3$. (ii) Find an unbiased estimate of σ^2. **30**

10. (a) If all correlations of order zero in a set of p variables are equal to ρ, show that (i) every partial correlation of order s is $\rho/(1+s\rho)$ and (ii) $1-R_{1.23}^2=((1-\rho)(1+2\rho))/(1+\rho)$, symbols having usual meanings. **30**

(b) $X\sim N_p(\mu,\Sigma)$ and let C be a $q\times p$ matrix of rank q. Prove that $Y=CX$ follows $N_q(C\mu,C\Sigma C')$ distribution. Explain how this can be used to find the marginal distribution of a q-component vector of X. **30**

12. (a) What are principal components? Explain the differences between principal components analysis and factor analysis. Briefly indicate how you can obtain maximum likelihood estimates of the principal components, given a random sample from $N_p(0,\Sigma)$ where Σ has all its characteristic roots distinct. **35**

(b) Define Hotelling's T^2. Show that it is invariant under non-singular linear transformations. Mention some uses of the T^2 statistic. **25**

SECTION 'D'

13. (a) Define a simple lattice design and derive its complete analysis. **30**

(b) Define a balanced incomplete block design and show that it is a balanced design. **15**

(c) Prove that a balanced incomplete block design with parameters $v=b=22$, $r=k=7$ and $\lambda=2$ cannot be constructed. **15**

14. (a) What is missing plot technique? Explain its importance. In a randomized block design one observation is missing. Estimate the missing value and indicate how you will carry out the exact analysis. **30**

(b) Explain the technique of confounding. Indicate the method of construction of a $(3^3,3^2)$ design, retaining full information on the main effects. **30**

15. (a) Discuss ratio and regression methods of estimation and compare them. **30**

(b) Describe two-stage sampling. How would you estimate the population mean and obtain its variance, using a two-stage sample if all primary units have equal number of secondary units? **30**

Optional Province · optionalprovince.com

16. **(a)** Explain proportional and optimum (Neyman) allocations in the context of stratified sampling. **15**

(b) Write a note on Randomised Response Technique, indicating situations where to use it. **15**

(c) Consider p.p.s. sampling (without replacement) of two units based on positive size measure. Show that the Yates–Grundy estimator of the variance of the Horvitz-Thompson estimator of the population total is always non-negative. **30**

Optional Province · optionalprovince.com