

Test Code: STA (Short Answer Type) 2015

Junior Research Fellowship for Research Course in Statistics

The candidates for research course in Statistics will have to take two short-answer type tests – STA and STB. Each test is of two-hour duration. Test STA will have about 10 questions of equal value, set from selected topics in Mathematics and Statistics at the undergraduate level. Test STB will have roughly 8 questions of equal value, on topics in Statistics at Master's level.

Syllabus for STA

Mathematics

Functions and relations. Matrices – determinants, eigenvalues and eigenvectors, solution of linear equations, and quadratic forms.

Calculus and Analysis – sequences, series and their convergence and divergence; limits, continuity of functions of one or more variables, differentiation, applications, maxima and minima. Integration, definite integrals, areas using integrals, ordinary linear differential equations.

Statistics

(a) *Probability*: Basic concepts, elementary set theory and sample space, conditional probability and Bayes theorem. Standard univariate and multivariate distributions. Transformations of variables. Moment generating functions, characteristic functions, weak and strong laws of large numbers, convergence in distribution and central limit theorem. Markov chains.

(b) *Inference*: Sufficiency, minimum variance unbiased estimation, Bayes estimates, maximum likelihood and other common methods of estimation. Optimum tests for simple and composite hypotheses. Elements of sequential and non-parametric tests. Analysis of discrete data - contingency chi-square.

(c) *Multivariate Analysis*: Standard sampling distributions. Order statistics with applications. Regression, partial and multiple correlations. Basic properties of multivariate normal distribution, Wishart distribution, Hotelling's T^2 and related tests.

(d) *Design of Experiments*: Inference in linear models. Standard orthogonal and non-orthogonal designs. Analysis of general block designs. Factorial experiments. One and two-way ANOVA.

(e) *Sample Surveys*: Simple random sampling, Systematic sampling, PPS sampling, Stratified sampling. Ratio and regression methods of estimation. Non-sampling errors, Non-response bias.

Sample Questions : STA

1. Suppose $f : [0, \infty) \rightarrow \mathbb{R}$ is a continuous function such that

$$\lim_{x \rightarrow \infty} f(x) = 2014.$$

Show that f is uniformly continuous.

2. Find the set of all possible real numbers t such that

$$x^2 + y^2 + z^2 - 2txy - tyz \geq 0$$

for all $x, y, z \in \mathbb{R}$.

3. Suppose V and W are independent Uniform(0,1) random variables. Let Z denote the number of positive roots of

$$x^2 - 2Vx + W = 0.$$

Find $E(Z)$.

4. Suppose (X, Y) is a bivariate random vector such that Y has an exponential distribution with mean 1 and for every $y > 0$, the conditional distribution of X given $Y = y$ is uniform over $(0, y)$. Let $Z = Y - X$. Find the mean and variance of Z .
5. From a population of size N with auxiliary size measure x ($x > 0$), a sample of size n ($2 < n < N$) is drawn using the following rule :
- i) the first sample unit is drawn using probability proportional to size (x), and,
 - ii) the other $(n - 1)$ sample units are drawn using Simple Random Sampling Without Replacement from the remaining $(N - 1)$ units.

Denote by $\pi(i)$ and $\pi(i, j)$ the inclusion probability of unit i and joint inclusion probability of units i and j , $i \neq j$, respectively.

Show that

(a) $\pi(i) > \pi(j) \Leftrightarrow x_i > x_j$ for all $i \neq j$,

(b) $\pi(i, k) > \pi(j, k) \Leftrightarrow x_i > x_j$ for all distinct i, j, k .

6. Let X_1, \dots, X_n be iid random variables with $X_1 \sim N(\mu, \sigma^2)$. Let $\bar{X} = \sum_{i=1}^n X_i/n$, $R = \max_{1 \leq i \leq n} X_i - \min_{1 \leq i \leq n} X_i$. Show that \bar{X} and R are independently distributed.
7. Suppose X_1, \dots, X_n are independent Poisson (λ) variables, where $\lambda > 0$ is unknown. Find the Uniformly Minimum Variance Unbiased Estimator of $e^{-\lambda}$.

8. Assume that the length of a phone-call (in minutes) of an individual follows an exponential distribution with an unknown parameter $\lambda > 0$ with density function $f(x) = \lambda e^{-\lambda x}$, $x > 0$. However, when the phone company calculates the length of a phone-call, it always considers the nearest integer greater than or equal to the actual length. For example, a 22.09 minutes long phone-call will have a call-length of 23 minutes in the phone company records. Suppose you have the data on the lengths of n independent phone-calls T_1, T_2, \dots, T_n of that individual as reported by the phone company. Based on this data, compute the maximum likelihood estimator of λ .
9. Suppose X_1, \dots, X_n are independent having uniform $(\theta, 2\theta)$ distribution, where $\theta > 0$ is unknown. Let $V_n = \frac{1}{2}X_{(n)}$ and $U_n = X_{(1)}$, where $X_{(1)}$ and $X_{(n)}$ are the smallest and largest among X_1, \dots, X_n , respectively.
- (a) Show that the estimators U_n and V_n are consistent for θ .
- (b) Find the ratio of the Mean Squared Errors of the two estimates above.
10. Suppose you have only one observation X with probability mass function p_θ , $\theta \in \{0, 1\}$ and p_0, p_1 are given by

x	1	2	3	4	5
$p_0(x)$	0.01	0.04	0.02	0.84	0.09
$p_1(x)$	0.05	0.20	0.60	0.14	0.01

- (a) Find a level- α most powerful test for $H_0 : \theta = 0$ vs. $H_1 : \theta = 1$ when (i) $\alpha = 1\%$ and (ii) $\alpha = 5\%$.
- (b) Are the most powerful tests unique in (i) and (ii) of (a)? Justify your answer.
- (c) Find the power of the most powerful tests in (a).

For more sample questions, visit
<http://www.isical.ac.in/~deanweb/JRFSTATSQ.html>