I.E.S./I.S.S. Examination => 2011

E-JTT-L-TUA

STATISTICS – I

Time Allowed : Three Hours | *Maximum Marks : 200*

INSTRUCTIONS

Candidates should attempt FIVE questions in ALL including Questions No. 1 and 5 which are compulsory. The remaining THREE questions should be answered by choosing at least ONE question each from Section A and Section B.

The number of marks carried by each question is indicated against each.

Answers must be written only in ENGLISH.

(Symbols and abbreviations are as usual)

Any essential data assumed by candidates for answering questions must be clearly stated.

A normal distribution table and a 't' table are attached with this question paper.

SECTION A

1. Answer any *five* of the following : *8×5=40*

(a) Verify the following identities :

(i) $P(A \cup B) = P(A) + P(B \cap A^c)$

(ii) $P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(B \cap C) - P(A \cap C) + P(A \cap B \cap C)$

(b) A fair die is thrown until a 6 appears. Specify the sample space. What is the probability that it must be thrown at least 3 times ?

(c) $P(A) = 1/3$ and $P(B^c) = 1/4$. Can A and B be disjoint ? Explain.

(d) Show that $E(X - a)^2$ is minimized for $a = E(X)$, assuming that the first 2 moments of X exist.

(e) Let

$$f(x) = \frac{1}{\Gamma(n)\,\beta^n}\, x^{n-1}.\, e^{-x/\beta}\ ,\ x > 0,\ n > 0,\ \beta > 0.$$

Show that f(x) is a probability density function. Obtain V(X).

(f) Show that $f(x) = \frac{a}{x^3}$, $a < x$, $(a > 0)$ is a probability density function for an appropriate value of a. Upto what order do the moments of this p.d.f. exist ?

2. (a) Let $E|X|^n < \infty$. Prove that $E|X|^k < \infty$ for $k \leq n$.

(b) Show that the sum of two independent Poisson random variables with parameters λ and μ respectively is a Poisson random variable with parameter $\lambda + \mu$.

(c) Let the joint p.d.f. of (X, Y) be

$$f(x, y) = e^{-y},\quad 0 < x < y < \infty.$$

Obtain the probability $P(X + Y \leq 1)$.

(d) Let $X_1, X_2, ..., X_n$ be a sequence of i.i.d.r.v.s with $E(X_i) = 0$ and $V(X_i) = 1$. Show that the sequence

$$S_n^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2$$ tends to 1 in probability.

10×4=40

3. (a) Let X_1 and X_2 be two independent exponentially distributed random variables with the same mean θ. Define $V = \max(X_1, X_2)$ and $W = \min(X_1, X_2)$. Show that $V - W$ and $2W$ are independent and identically distributed random variables.

(b) Let X be a positive valued random variable. Prove that

$$P(X \geq r) \leq \frac{E(X)}{r}, \; r > 0.$$

Hence deduce the Chebychev's inequality.

(c) Let X have the continuous c.d.f. F(x). Define $U = F(x)$. Show that both $-\log U$ and $-\log(1-U)$ are exponential random variables.

(d) Obtain the median and the quartiles of the Cauchy distribution with p.d.f.

$$f(x) = \frac{1}{\pi(1+x^2)}, \quad -\infty < x < \infty.$$

10×4=40

4. (a) Let Z be a random variable with p.d.f. f(z). Let z_α be its upper α^{th} quantile. Show that if X is a random variable with p.d.f. $\frac{1}{\sigma} f\left(\frac{x-\mu}{\sigma}\right)$ then $\sigma z_\alpha + \mu$ is the upper α^{th} quantile of X.

(b) Let (X, Y) have a bivariate distribution with finite moments upto order 2. Show that

(i) $E(E(X|Y)) = E(X)$, and

(ii) $V(X) \geq V(X|Y)$.

(c) Show that a c.d.f. can have at most a countable number of jumps.

(d) Consider the following bivariate p.m.f. of (X, Y) :

$$p(0, 10) = p(0, 20) = \frac{2}{18};$$

$$p(1, 10) = p(1, 30) = \frac{3}{18};$$

$$p(1, 20) = \frac{4}{18}; \quad p(2, 30) = \frac{4}{18}.$$

Obtain the conditional mass functions $p(y|x = 2)$, and $p(y|x = 1)$. *10×4=40*

SECTION B

5. Answer any ***five*** of the following : *8×5=40*

(a) Tabulate the exact null distribution of Wilcoxon rank sum statistic for $n_1 = n_2 = 3$.

(b) Let (X, Y) have the uniform distribution over the range $0 < y < x < 1$. Obtain the conditional mean and variance of X given Y = y.

(c) Let the temperature before and after administration of aspirin be

Patient	Before	After
1	100·0	98·1
2	102·1	97·2
3	100·6	98·6
4	100·1	99·1
5	101·5	97·6
6	102	98·6
7	99·9	98·2
8	102·7	98·1
9	100·4	98·2
10	100·8	97·1

Test by the sign test, whether aspirin is effective in reducing temperature. What is the p-value of the calculated statistic ?

(d) Use Simpson's rule with five ordinates to compute an approximation to π with the help of the integration of the function $(1 + x^2)^{-1}$ from 0 to 1.

(e) Use mathematical induction to prove

$$f_{x+nh} = \sum_{i=0}^{\infty} \binom{n}{i} \Delta^i f_x .$$

(f) Let $D_n = \sup_{-\infty < x < \infty} |F_n(x) - F(x)|$ be the Kolmogorov statistic. Show that

$$D_n^+ = \max_{0 \le i \le n} \ \sup_{x_{(i)} \le x < x_{(i+1)}} \left\{ \frac{i}{n} - F(x) \right\}$$

where $x_{(1)}, x_{(2)}, \ldots, x_{(n)}$ are the order statistics of the data.

6. (a) Let X and Y be two random variables with correlation coefficient 0·9. Can a third random variable Z have correlation coefficient – 0·9 with both X and Y ? Give reasons for your answer.

(b) Show that the square of the one sample t-statistic has the F-distribution. What are its degrees of freedom ?

(c) Define the correlation ratio η_1 of X on Y. If ρ is the usual correlation coefficient between X and Y, then show that $0 \le \rho^2 \le \eta_1^2 \le 1$.

(d) Prove that the sum of two independent chi-squared random variables is also chi-squared.

10×4=40

7. (a) Apply the Wilcoxon two-sample test to the following data on the first breakdown times of two brands of computers :

Brand A 98, 102, 47, 85, 99, 140, 130

Brand B 95, 125, 160, 155, 148.

Use 1·96 as the critical point for the appropriate test.

(b) Describe a test of independence of two normal random variables based on r, the sample correlation coefficient using the t-distribution. If n = 10 and r = 0·9, then carry out the test.

(c) Show that the best predictor of Y, in terms of minimum MSE, is linear in X, if (X, Y) have bivariate normal distribution.

(d) Explain the Wald – Wolfowitz run test for randomness in a sequence of two types of symbols. Find $E_{H_0}(R)$ where R denotes the number of runs of elements of one kind. *10×4=40*

8. (a) The following are the frequencies in the given intervals :

(0 – 2]	(2 – 5]	(5 – 10]	(10 – 15]	(15 – 18]	(18 – 20]
43	85	151	112	72	34

Draw the histogram of this data. Calculate the mean of the data from the frequency table.

(b) Let $X_1, X_2, \ldots, X_n$ be independent $N(0, \sigma^2)$ random variables. Obtain the mean and variance of $\sum_{i=1}^{n} X_i^2$. What is its probability distribution ?

(c) From a bivariate data set of 4995 observations, the following quantities have been calculated :

$\Sigma x = -2353$, $\Sigma y = -1400$, $\Sigma x^2 = 9508$

$\Sigma y^2 = 70802$, $\Sigma xy = 8805$.

Obtain the estimated linear regression of X on Y.

(d) Explain how to carry out the chi-squared test for $H_0 : \sigma^2 = \sigma_0{}^2$ on the basis of a random sample $X_1, X_2, \ldots, X_n$ from $N(\mu, \sigma^2)$ population. *10×4=40*