ADVANCED PROBABILITY AND STATISTICS
(MATH 4281)

Time Allotted : 3 hrs **Full Marks : 70**

Figures out of the right margin indicate full marks.

Candidates are required to answer Group A and any 5 (five) from Group B to E, taking at least one from each group.

Candidates are required to give answer in their own words as far as practicable.

Group – A
(Multiple Choice Type Questions)

1. Choose the correct alternative for the following: **10 × 1 = 10**

(i) The probability is p = 0.80 that a patient with a certain disease will be successfully treated with a new medical treatment. Suppose that the treatment is used on 40 patients. What is the "expected value" of the number of patients who are successfully treated?
(a) 40 (b) 20 (c) 8 (d) 32.

(ii) If we randomly choose a number from the closed interval [0,1], then the probability that the number is from the Cantor set is
(a) 1 (b) 0.5
(c) 0.2 (d) probability measure cannot be assigned.

(iii) Let X be a Poisson random variable with E(X) = ln(2). What is E[cos(πX)]?
(a) 1/4 (b) 1/2 (c) 1 (d) 2ln (2).

(iv) If a sample of size n is drawn from an infinite population with standard deviation σ, then for the sample variance S^2, which one is true?
(a) $E\left[\frac{n}{n-1}S^2\right] = \sigma^2$ (b) $E[S^2] = \sigma^2$
(c) $E\left[\frac{n}{n-1}S\right] = \sigma^2$ (d) $E\left[\frac{n-1}{n}S^2\right] = \sigma^2$

(v) If T_1, T_2 be two statistics with expectation $E[T_1] = 2\theta_1 + 3\theta_2$, $E[T_2] = \theta_1 + \theta_2$, , then the unbiased estimate of the parameter θ_1, is
(a) $3T_2 - T_1$ (b) $3T_2 + 2T_1$
(c) $2T_2 - 3T_1$ (d) $3T_2T_1$.

(vi) If we consider the closed interval $[0,1]$, then the probability that a randomly chosen number is rational, has the probability,
(a) 1 (b) 0.5
(c) 0.25 (d) probability cannot be assigned.

(vii) The set of binary strings of infinite length
(a) is uncountable
(b) is countable
(c) is finite
(d) has bijection with the set of natural numbers.

(viii) If $\lim_{n\to\infty} P\left\{\left|\frac{1}{\bar{X}+2} - \theta\right| > \varepsilon\right\} = 0$ holds for arbitrary $\varepsilon > 0$ then
(a) $\bar{X}$ is a consistent estimator of θ
(b) $\bar{X}+2$ is a consistent estimator of θ
(c) $\frac{1}{\bar{X}}$ is a consistent estimator of θ
(d) $\frac{1}{\bar{X}+2}$ is a consistent estimator of θ

(ix) If x be the test statistic and (a,b) is the region of acceptance corresponding to 3% level of significance then $P(a \le X \le b)$ is equal to (X is the corresponding random variable)
(a) 0.9 (b) 0.97 (c) 0.99 (d) 0.03.

(x) If a population random variable X has normal distribution with parameter $\mu = 5$ and $\sigma = 0.1$. Then the sample mean $\bar{X}$ (of size 25) is a normal variable with mean and s.d.,
(a) 5, 0.02 (b) 5, 0.1 (c) 5, 0.2 (d) 0, 1

Group – B

2. (a) Let A *and* B be independent events; show that A^C, B are independent, and deduce that A^C, B^C are independent.

(b) Show that the probability that *exactly* one of the events A, B occurs is $P(A) + P(B) - 2P(A \cap B)$.

6 + 6 = 12

3. (a) Let X be the number of times that a fair coin, flipped 40 times, lands heads. Find the probability that $X = 20$. Use the normal approximation

and then compare it to the exact solution.

(b) Let X be a random variable with probability density function:

$$f(x)=\begin{cases}c(1-x^2), -1\le x\le 1\\ 0 \qquad\quad ,otherwise\end{cases}$$

(i) What is the value of c?

(ii) What is the cumulative distribution function of X?

6 + 6 = 12

Group – C

4. (a) Use moment generating function to find the mean and variance of the exponential distribution given by, $f(x)=\lambda e^{-\lambda x}$, $x\ge 0$

(b) State and prove the Chebyshev's inequality.

7 + 5 = 12

5. (a) State and prove the Markov's inequality for a continuous random variable.

(b) If X, Y are independent continuous random variables, then prove that for any function g and h, $E\left[g(X)h(Y)\right]=E[g(X)]\,E[h(Y)]$.

6 + 6 = 12

Group – D

6. (a) In a partially destroyed laboratory record of an analysis of correlation data, the following results only are legible:
Variance of $x=9$.
Regression equations: $8x-10y+66=0,\ 40x-18y=214$.
What were (i) the means of x and y,
(ii) the correlation co-efficient between x and y, and
(iii) the standard deviation of y?

(b) Two regression lines of a sample are $x+6y=6$, $3x+2y=10$. Find correlation coefficient and the corresponding means of the random variables.

7 + 5 = 12

7. (a) Use the method of least squares to obtain the regression line of y on x.

(b) Find the moment generating function of the binomial distribution with parameters n, p. Hence, use it to find the mean and variance of the

binomial distribution.

5 + (4 + 3) = 12

Group – E

8. (a) A random variable X is found to have the following distribution:

x	0	1	2
$P(X=x)$	0.4	0.5	0.1

Considering all random sample of size 2 from the hypothetical population of X, construct the probability table for sample mean $\bar{X}$ and show that $\sigma_{\bar{X}}^2=\frac{1}{2}\sigma^2$, where σ^2 and $\sigma_{\bar{X}}^2$ are respectively population variance and sample variance.

(b) (i) If T is an unbiased estimator of θ, then show that $\sqrt{T}$ is a biased estimator of $\sqrt{\theta}$.

(ii) Define a consistent estimator.

6 + (4 + 2) = 12

9. (a) Find the maximum likelihood estimate for p where $(0<p<1)$ on the basis of a random sample $\{x_1, x_2, x_3, ..., x_n\}$ of size n drawn from the population X, where X has the following probability mass function $P(X=i)=p(1-p)^i$, $i=0,1,2,...$ $(0<p<1)$.

(b) A salesman is expected to effect an average sales of Rs. 3500 per day. Observing the sales of a particular salesman for 6 days, it is seen that he gives an average sale of Rs.3300 per day with standard deviation of Rs.1016.53. Using 0.05 level of significance comment on his work i.e., find whether his performance is below the standard.
[Given $t_{0.05}=2.02$ for 5 degrees of freedom]

6 + 6 = 12