PROBABILITY AND STATISTICAL METHODS
(MATH 2111)

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) A bag contains 8 items of which 2 are defective. A man selects 3 items at random. Then the expected number of defective items is
(a) $\frac{3}{4}$ (b) $\frac{1}{4}$ (c) $\frac{5}{14}$ (d) $\frac{3}{28}$.

(ii) A fair die is rolled 60 times. Let the random variable X be the number of sixes that appear. The best upper bound on the probability of rolling 30 or more sixes is
(a) $\frac{1}{6}$ (b) $\frac{2}{3}$ (c) $\frac{1}{3}$ (d) $\frac{1}{4}$.

(iii) Let X be a binomially distributed random variable with parameters 8 and $\frac{1}{2}$. Then the value of $Var(-3X-5)$ is
(a) -6 (b) -18 (c) 18 (c) 6.

(iv) The characteristic function $\varphi(t)$ of exponential distribution with parameter λ is
(a) $\frac{\lambda}{\lambda-it}$ (b) $\frac{\lambda}{\lambda-t}$ (c) $\frac{\lambda t}{\lambda-t}$ (d) $\frac{i\lambda}{\lambda-it}$.

(v) For a 99% confidence interval, the confidence coefficient is
(a) 0.001 (b) 0.099 (c) 0.01 (d) 0.99.

(vi) If the two bivariates (x, y) and (u, v) are such that $x = -2u + 4$ and $y = 3v - 6$ then
(a) $r_{xy} = r_{uv}$ (b) $r_{xy} = -6r_{uv}$
(c) $r_{xy} = 6r_{uv}$ (d) $r_{xy} = -r_{uv}$.

(vii) Following is a transition probability matrix for a Markov chain with states $\{0, 1, 2\}$,

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

then
(a) all states are aperiodic
(b) state 0 is aperiodic
(c) all states have period 3
(d) state 0 has period 2 but other states are aperiodic.

(viii) If $n = 10, \sum d^2 = 280$, then the rank correlation coefficient is
(a) 2.8 (b) 0.28 (c) -0.7 (d) 0.7.

(ix) If $H_1: \mu < 60$ be an alternative hypothesis, then the null hypothesis is
(a) $H_0: \mu > 60$ (b) $H_0: \mu \geq 60$
(c) $H_0: \mu \leq 60$ (d) $H_0: \mu = 60$.

(x) If u + 3x = 5, 2y - v = 7 and correlation coefficient of x and y is 0.12, then what is the correlation coefficient of u and v?
(a) 0.12 (b) -0.12 (c) 1 (d) 0.

Group – B

2. (a) Suppose a book of 585 pages contains 43 typographical errors. If these errors are randomly distributed throughout the book, what is the probability that 10 pages selected at random, will be free from errors?
[(MATH2102.1, MATH2102.2)(Apply/IOCQ)]

(b) A fair die is tossed 720 times. Use Chebyshev's inequality to find a lower bound for the probability of getting 100 to 140 sixes.
[(MATH2111.1, MATH2111.2)(Understand/LOCQ)]

(c) A bag contains 5 red balls, 4 white and 3 blue balls. A ball in drawn at random from the bag, its colour is noted, and then the ball is replaced. Find the probability that out of 6 balls selected in this manner 3 are red, 2 are white and 1 is blue.
[(MATH2111.1, MATH2111.2)(Remember/LOCQ)]

5 + 4 + 3 = 12

3. (a) The lifetime of a certain brand of an electric bulb may be considered as a random variable with mean 1200 hours and standard deviation 240 hours. Using central limit theorem, determine the probability that the average lifetime of 60 bulbs exceeds 250 hours. [(MATH2102.1, MATH2102.2)(Apply/IOCQ)]

(b) The life length X of an electronic component follows an exponential distribution. There are two processes by which the component may be manufactured. The expected life length of the component is 100 hours if process-I is used to manufacture, while it is 150 hours if process-II is used. The cost of manufacturing a single component by process-I is Rs 10 while it is Rs 20 for process-II. Moreover, if the component lasts less than the guaranteed life of 200 hours, a loss of Rs 50 is to be borne by the manufacturer. Which process is advantageous to the manufacturer?
[(MATH2111.1, MATH2111.2)(Evaluate/HOCQ)]

6 + 6 = 12

Group – C

4. (a) Three fair coins are tossed. Let X denotes the number of heads on the first two coins, and let Y denotes the number of heads on the last two coins.
(i) Find the joint probability mass function of X and Y.
(ii) Are the random variables X and Y independent? Justify.
(iii) Find the conditional distribution of Y given X $=$ 1.
[(MATH2111.1, MATH2111.2, MATH2111.3)(Understand/LOCQ)]

(b) Every two years, a man trades his car for a new car. If he has a Maruti, he trades it for a Tata Tiago. If he has a Tata Tiago, he trades it for a Santro. However if he has a Santro, he as just likely to trade it for a new Santro as to trade it for a Tata Tiago or a Maruti. In 2020 he bought his first car, which was a Santro.
(i) Find the probability that in 2024 he will have a Tata Tiago.
(ii) In the long run, how often will he have a Santro.
[(MATH2111.1, MATH2111.2, MATH2111.3)(Analyze/IOCQ)]

6 + 6 = 12

5. (a) If X and Y have a joint p.d.f.

$$f(x,y) = \begin{cases} k, & \text{for } x^2 + y^2 \leq 4 \\ 0, & \text{otherwise}' \end{cases}$$

where k is a constant. Find the value of k and $P(X^2 + Y^2 > 1)$. Also find the marginal p.d.f. of Y. [(MATH2111.1, MATH2111.2, MATH2111.3)(Analyze/IOCQ)]

(b) The three state Markov chain is given by the transition probability matrix $P = \begin{pmatrix} 0 & 2/3 & 1/3 \\ 1/2 & 0 & 1/2 \\ 1/2 & 1/2 & 0 \end{pmatrix}$. Prove that the chain is irreducible and all the states are aperiodic and non-null recurrent. Find also the steady state distribution of the chain.
[(MATH2111.1, MATH2111.2, MATH2111.3)(Analyse/IOCQ)]

6 + 6 = 12

Group – D

6. (a) The profits y (lakh Rs.) of a certain company in the x-th year of its life are given by

x:	1	2	3	4	5
y:	2.18	2.44	2.78	3.25	3.83

Fit a second-degree parabola $y = a + bx + cx^2$ by the method of least squares.
[(MATH2111.3, MATH2111.4, MATH2111.5, MATH2111.6)(Apply/IOCQ)]

(b) Calculate the first four central moments and hence find the measures of skewness and kurtosis for the following distribution:

Class interval:	0 – 10	10 – 20	20 – 30	30 – 40
Frequency:	1	3	4	2

Also analyze the nature of the distribution.
[(MATH2111.3, MATH2111.4, MATH2111.5, MATH2111.6)(Evaluate/HOCQ)]

6 + 6 = 12

7. (a) Find the moment generating function of Poisson distribution with parameter λ. Hence use it to find $\mathrm{Var}(X^2 + 1)$.
[(MATH2111.3, MATH2111.4, MATH2111.5, MATH2111.6)(Evaluate/HOCQ)]

(b) Out of two regression lines given by $x + 4y + 3 = 0$ and $4x + 9y + 5 = 0$, which one is the regression line of "y on x"? Find the mean of x and y. Find the correlation coefficient between x and y. Estimate also the value of x when $y = 1.2$.
[MATH2111.3. MATH2111.4, MATH2111.5, MATH2111.6)(Analyse/IOCQ]

6 + 6 = 12

Group – E

8. (a) A Statistics professor has taken five tests. A student scored 70, 75, 65, 80 and 95 respectively in five tests. The professor decides to determine his grade by randomly selecting a sample of three test scores. Construct the sampling distribution for this process. Determine the mean and variance of the sampling distribution. Clearly state the theorem you used.

[(MATH2111.4, MATH2111.5, MATH2111.6)(Understand/LOCQ)]

(b) To test the hypothesis that eating fish makes one smarter, a random sample of 12 persons take a fish oil supplement for one year and then are given an IQ test. Here are the results: 116, 111, 101, 120, 99, 94, 106, 115, 107, 101, 110, 92.

Test using the following hypotheses, report the test statistic and summarize your conclusion.

$$H_0: \mu = 100,\ H_1: \mu > 100$$

[(MATH2111.4, MATH2111.5, MATH2111.6)(Analyze/IOCQ)]

6 + 6 = 12

9. (a) A machine produces a component of a product with a standard deviation of 1.6 cm in length. A random sample of 64 components was selected from the output and this sample has a mean length of 90 cm. The customer will reject the part if it is either less than 88 cm or more than 92 cm. Does the 95% confidence interval for the true mean length of all the components produced ensure acceptance by the customer?

[(MATH2111.4, MATH2111.5, MATH2111.6)(Evaluate/HOCQ)]

(b) (i) If T is an unbiased estimator of the population parameter θ, then prove that $\sqrt{T}$ is biased estimator of $\sqrt{\theta}$.

(ii) Define consistent estimator.

[(MATH2111.4, MATH2111.5, MATH2111.6)(Understand/LOCQ)]

6 + 6 = 12

Cognition Level	LOCQ	IOCQ	HOCQ
Percentage distribution	26.04	48.96	25

Course Outcome (CO):

After the completion of the course students will be able to
MATH 2111.1 Articulate the axioms (laws) of probability.
MATH 2111.2 Compare and contrast different interpretations of probability theory and take a stance on which might be preferred.
MATH 2111.3 Formulate predictive models to tackle situations where deterministic algorithms are intractable.
MATH 2111.4 Summarize data visually and numerically.
MATH 2111.5 Assess data-based models.
MATH 2111.6 Apply tools of formal inference.

*LOCQ: Lower Order Cognitive Question; IOCQ: Intermediate Order Cognitive Question; HOCQ: Higher Order Cognitive Question.