

**STATISTICS AND INFORMATION THEORY
(MTH3122)**

Time Allotted : 2½ hrs

Full Marks : 60

Figures out of the right margin indicate full marks.

Candidates are required to answer Group A and any 4 (four) from Group B to E, taking one from each group.

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

Group – A

1. Answer any twelve:

12 × 1 = 12

Choose the correct alternative for the following

- (i) Let X and Y be two random variables each taking three values $-1, 0, 1$ and having joint probability distribution

$Y \backslash X$	X	-1	0	1
-1	0	0	0	$1/3$
0	0	0	0	0
1	0	$1/3$	$1/3$	$1/3$

Then $E(X|Y = 1)$ is

- (a) 0 (b) 1 (c) $\frac{1}{2}$ (d) $\frac{1}{3}$.
- (ii) For a random variable X with mean 0, $P(-2\sigma < X < 2\sigma) \geq$
 (a) $\frac{1}{4}$ (b) $\frac{3}{4}$ (c) $\frac{1}{2}$ (d) $\frac{9}{16}$.
- (iii) If the joint *p. d. f.* is given by $f(x, y) = \begin{cases} cx(1-x), & 0 \leq x \leq y \leq 1 \\ 0, & \text{otherwise} \end{cases}$, then the value of c is
 (a) 1 (b) 6 (c) 12 (d) 21.
- (iv) The measurements of spread or scatter of the individual values around the central point is called:
 (a) Measure of dispersion (b) Measure of central tendency
 (c) Measure of skewness (d) Measure of kurtosis.
- (v) First order central moment for any distribution is
 (a) 1 (b) -1 (c) 0 (d) 2.
- (vi) If the two bivariates (x, y) and (p, q) are such that $x = -4p + 6$ and $y = -3q - 2$, then
 (a) $r_{xy} = r_{uv}$ (b) $r_{xy} = -12r_{uv}$ (c) $r_{xy} = -r_{uv}$ (d) $r_{xy} = 12r_{uv}$.
- (vii) If the variance of \bar{X} equals 25 and the sample size $n = 6$, the population variance is
 (a) 4.167 (b) 2.041 (c) 150 (d) 12.247.

- (viii) For a 99% confidence interval, the confidence coefficient is
 (a) 0.001 (b) 0.099 (c) 0.01 (d) 0.99.
- (ix) Which one of the following inequalities is correct?
 (a) $H(X|Y) \geq H(X)$ (b) $H(Y|X) \leq H(X)$
 (c) $H(X|Y) \leq H(X)$ (d) $H(X|Y) \geq H(Y|X)$.
- (x) An event has two possible outcomes with probability $P_1 = 1/2$ and $P_2 = 1/64$. Calculate the rate of information with 16 outcomes per second is
 (a) 24/3 (b) 64/6 (c) 38/4 (d) 1.

Fill in the blanks with the correct word

- (xi) If the population standard deviation σ is unknown, and the sample size is small, the confidence interval for the population mean μ is based on _____ distribution.
- (xii) If regression coefficients, $b_{xy} = -0.4$ and $b_{yx} = -0.9$, then correlation coefficient r_{xy} is _____.
- (xiii) If the moment generating function of a random variable X be $M_X(t)$, then the moment generating function of a random variable $Y = 2X - 5$ is _____.
- (xiv) A fair coin is tossed 30 times. Let the random variable X be the number of heads that appear. The best upper bound on the probability of getting 20 or more heads is _____.
- (xv) If $n = 8$ and $\sum d^2 = 28$, then the rank correlation coefficient is _____.

Group - B

2. (a) Use Chebyshev's inequality to find how many times a fair coin must be tossed in order that the probability that the ration of number of heads to the number of tosses will lie between 0.45 and 0.55 will be at least 0.95. [[MATH3122.1 & MATH3122.2] (Apply/IOCQ)]
- (b) Suppose that the moment generating function of the random variable X is $M_X(t) = \left(\frac{1+3e^t}{4}\right)^{10}$. What is the mean and variance of X ? [[MATH3122.1 & MATH3122.2] (Evaluate/HOCQ)]

6 + 6 = 12

3. (a) Following is the joint probability distribution of X and Y :

Y X	1	2	3
0	$\frac{1}{10}$	$\frac{3}{10}$	$\frac{1}{10}$
2	$\frac{1}{5}$	$\frac{1}{10}$	$\frac{1}{5}$

- (i) Is it a valid distribution? Give reason.
- (ii) Find the marginal probability mass function of X and Y .
- (iii) Are X and Y independent? Justify. [[MATH3122.1 & MATH3122.2] (Apply/IOCQ)]
- (b) A fair die is thrown 720 times. Use Chebyshev's inequality to obtain a lower bound for the probability of getting 91 to 149 sixes. [[MATH3122.1 & MATH3122.2] (Apply/IOCQ)]

6 + 6 = 12

Group - C

4. (a) By using the method of least squares fit a straight line to the following data:

x	5	10	15	20	25
y	16	19	23	26	30

[[MATH3122.3 & MATH3122.6](Remember/LOCQ)]

- (b) Classify all the states of the Markov chain $\{X_n: n \geq 0\}$ with the following one-step transition probability matrix

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 1/2 \\ 0 & 1/2 & 1/2 \end{bmatrix}$$

Is this Markov chain irreducible? Justify your answer.

[[MATH3122.3 & MATH3122.6](Evaluate/HOCQ)]

8 + 4 = 12

5. (a) Calculate the first four moments of the following distribution about the mean and hence find the skewness and kurtosis and interpret about the distribution:

x	0	1	2	3	4	5	6	7	8
f	1	8	28	56	70	56	28	8	1

[[MATH3122.3 & MATH3122.6](Analyse/IOCQ)]

- (b) A student's study habits are as follows: If he studies one night, he is 70% sure not to study the next night. On the other hand, if he does not study one night, he is 60% sure not to study the next night as well as well. In the long run how often does he study?

[[MATH3122.3 & MATH3122.6](Evaluate/HOCQ)]

6 + 6 = 12

Group - D

6. (a) A random sample of size 36 is taken from a population of size 100 having mean $m = 22.40$ and standard deviation 0.048. Find the mean and standard deviation of the sampling distribution of sample means if the sampling is done (i) without replacement (ii) with replacement. [[MATH3122.2, MATH3122.3 & MATH3122.6](Analyse/IOCQ)]

- (b) To test the unbiasedness of a die it is thrown six times and is accepted that the unbiasedness if not more than one six is obtained. Find the probability of type-I error. [[MATH3122.2, MATH3122.3 & MATH3122.6](Apply/IOCQ)]

6 + 6 = 12

7. (a) A person eats at the same fast-food restaurant every day. Suppose the time X between the moment he enters the restaurant and the moment he is served his food is normally distributed with mean 4.2 minutes and standard deviation 1.3 minutes.

(i) Find the probability that when he enters (any day) the restaurant it will take at least 5 minutes until he is served.

(ii) Find the probability that average time until he is served in eight randomly selected visits to the restaurant will be at least 5 minutes.

[[MATH3122.2, MATH3122.3 & MATH3122.6](Analyse/IOCQ)]

- (b) If T_1 and T_2 be two statistics with expectation $E(T_1) = 2\theta_1 + 3\theta_2$ and $E(T_2) = \theta_1 + \theta_2$, find unbiased estimators of the parameters θ_1 and θ_2 .

[[MATH3122.2, MATH3122.3 & MATH3122.6]](Understand/LOCQ)]

6 + 6 = 12

Group - E

8. (a) A discrete source transmits message x_1, x_2 and x_3 with the probabilities 0.3, 0.4 and 0.3. The source is connected to the channel y_1, y_2 and y_3 by the following conditional probability matrix:

$$P(Y|X) = \begin{matrix} & y_1 & y_2 & y_3 \\ x_1 & \begin{pmatrix} 0.8 & 0.2 & 0 \end{pmatrix} \\ x_2 & \begin{pmatrix} 0 & 1 & 0 \end{pmatrix} \\ x_3 & \begin{pmatrix} 0 & 0.3 & 0.7 \end{pmatrix} \end{matrix}$$

Calculate the mutual information between X and Y . [[MATH3122.4 & MATH3122.5]](Apply/IOCQ)]

- (b) Use Jensen's inequality to find the appropriate inequalities between

(i) $E(X \log X)$ and $E(X) \log(E(X))$, $x \geq 0$

(ii) $E(\sin X)$ and $\sin(E(X))$, $0 \leq x \leq \pi$.

[[MATH3122.4 & MATH3122.5]](Analyse/IOCQ)]

6 + 6 = 12

9. (a) A communication system has a source alphabet $= \{x_1, x_2, x_3\}$ and a receiver alphabet $Y = \{y_1, y_2, y_3\}$. The joint probability distribution $P(X, Y)$ is given by the following matrix:

$$P(X, Y) = \begin{pmatrix} \frac{1}{8} & \frac{1}{16} & \frac{1}{16} \\ \frac{1}{16} & \frac{1}{8} & \frac{1}{16} \\ \frac{1}{4} & \frac{1}{8} & \frac{1}{8} \end{pmatrix}$$

Where the rows correspond to the input symbols X and the columns correspond to the output symbols Y .

Calculate $H(X), H(Y), H(X|Y), H(X, Y)$ and $I(X; Y)$. [[MATH3122.4, MATH3122.5]](Apply/IOCQ)]

- (b) Let the random variable X have four possible outcomes $\{a, b, c, d\}$. Consider three distributions on this random variable

Symbols	$p(x)$	$q(x)$	$r(x)$
a	0.4	1/4	1/2
b	0.3	1/4	1/5
c	0.2	1/4	1/5
d	0.1	1/4	1/10

Calculate $D(p||q), D(p||r)$ and $D(q||r)$.

[[MATH3122.4, MATH3122.5]](Remember/LOCQ)]

6 + 6 = 12

Cognition Level	LOCQ	IOCQ	HOCQ
Percentage distribution	20.83	62.50	16.67