ADVANCED PROBABILITY AND INFORMATION THEORY
(MATH 3222)

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 source generates information with probabilities $p_1 = \frac{1}{8}, p_2 = \frac{1}{8}, p_3 = \frac{1}{4}$ and $p_4 = 0.5$. The entropy of the system is
(a) 1 bit (b) $\frac{7}{4}$ bits (c) 2 bits (d) $\frac{3}{4}$ bits.

(ii) The conditional entropy $H\left(Y/X\right)$ is
(a) $H(X,Y) - H(X)$ (b) $H(Y) - H\left(Y/X\right)$
(c) $H(X,Y) + H(X)$ (d) $H(X,Y) - H(Y)$.

(iii) Which of the following function is a **convex** function?
(a) $\sqrt{x}$ (b) $\log x$ (c) $-x^2$ (d) $|x|$.

(iv) The value of $H\left(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\right)$ is
(a) 1 (b) 2 (c) 3 (d) 4.

(v) If X and Y are independent random variables, then
(a) $F(x,y) = F(x)F(y)$ (b) $P(X+Y) = P(X) + P(Y)$
(c) $F(x,y) = F(x) + F(y)$ (d) $F(x,y) = 0$.

(vi) Let $X \in \{0,1\}$ and $Y \in \{0,1\}$ be two independent random variables. If $P(X=0) = p$ and $P(Y=0) = q$, then $P(X+Y \geq 1)$ is
(a) $pq + (1-p)(1-q)$ (b) pq
(c) $p(1-q)$ (d) $1 - pq$.

(vii) If X and Y are two independent random variables and $M_X(t)$ and $M_Y(t)$ are their respective moment generating functions, then the moment generating functions of $X + Y$ is
(a) $M_X(t) + M_Y(t)$ (b) $M_X(t) - M_Y(t)$ (c) $M_X(t)\, M_Y(t)$ (d) $\frac{M_X(t)}{M_Y(t)}$.

(viii) For the independent random variables $X_1, X_2, \cdots, X_{50}$ if $E(X_i) = 3$ and $Var(X_i) = 2$ for all i, then $\bar{X} = \frac{X_1+X_2+\cdots+X_{50}}{50}$ approximately normal having variance
(a) 0.01 (b) 0.02 (c) 0.03 (d) 0.04.

(ix) 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.

(x) The first four moments of a distribution about the origin are $-1.5, 17, -30$ and 108. Then the third moment about the mean is
(a) 39.75 (b) 41.75 (c) 40.75 (d) 42.75.

Group- B

2. (a) Find the moment generating function of X, where X is a Binomial variate with parameters n and p. Hence, calculate mean and variance of the distribution. Furthermore Calculate $E(X^3)$. *[(MATH3222.1, MATH3222.2)(Evaluate/HOCQ)]*

(b) Suppose that the average grade on the upcoming MATH3222 examination is 70%. Give an upper bound on the proportion of students who score at least 90%. *[(MATH3222.1, MATH3222.2)(Apply/IOCQ)]*

7 + 5 = 12

3. (a) Carbon monoxide (CO) emission for a certain kind of car vary with mean 2.9 gm/ml and standard deviation 0.4 gm/ml. A company has 80 of these cars in its fleet acquired from various sources. What is the probability that the average CO emission for all 80 cars is in excess of 3? Assume that the CO emission follows Normal distribution. *[(MATH3222.1, MATH3222.2)(Apply/IOCQ)]*

(b) A fair coin is tossed three times. Let X denote the number of heads in three tossings and Y denote the absolute difference between the number of heads and number of tails. Find the joint $p.m.f$ of (X, Y) and the marginal $p.m.f$ of X and Y. Also find $P(X + Y \geq 1)$ and $P(^{X = 1}/_{Y = 2})$. *[(MATH3222.1, MATH3222.2)(Apply/IOCQ)]*

6 + 6 = 12

Group - C

4. (a) Fit a second degree parabolic curve $y = a + bx + cx^2$ by using the method of least squares to the following data:

x	1	2	3	4	5	6	7	8	9
y	2	6	7	8	10	11	11	10	9

[(MATH3222.3, MATH3222.4)(Remember/LOCQ)]

(b) Suppose a city is served by two cable TV companies, CableTV I and CableTV II. Due to their aggressive sales tactics, each year 40% of the cableTV I customers switch to cable TV II and 30% of the cable TV II customers switch to cableTV I. Express the information as a transition probability matrix of a suitable Markov chain. Furthermore, if today $\frac{1}{4}$ of the customers subscribe to CableTV I, after 2 years what percent subscribe to each company? *[(MATH3222.3 & MATH3222.4)(Understand/LOCQ)]*

7 + 5 = 12

5. (a) The intelligence quotients of 55 students are as follows:

Class Mark	4	6	8	10	12	14	16	18	20	22
Frequency	4	5	8	12	10	6	4	3	2	1

Calculate the first three raw moments and hence the central moments. Find the measure of skewness for the intelligence quotient. Is it skewed to the left or right? *[(MATH3222.3, MATH3222.4)(Analyze/IOCQ)]*

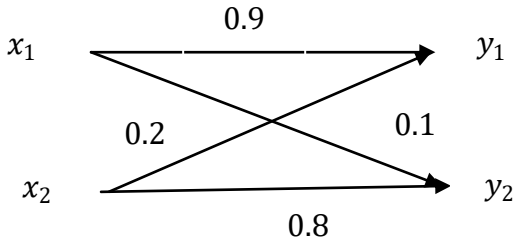
(b) Consider the Markov chain with transition matrix P and state space $\{1, 2, 3\}$, where $P = \begin{pmatrix} \frac{1}{2} & \frac{1}{3} & \frac{1}{6} \\ \frac{3}{4} & 0 & \frac{1}{4} \\ 0 & 1 & 0 \end{pmatrix}$

Find its steady state solution. *[(MATH3222.3, MATH3222.4)(Analyze/IOCQ)]*

7 + 5 = 12

Group – D

6. (a) Consider a binary channel as shown in figure given below:



Calculate the mutual information between X and Y when $P(x_1) = P(x_2) = 0.5$. *[(MATH3222.2, MATH3222.5)(Understand/LOCQ)]*

(b) Prove that $H(X) \leq \log|\chi|$, *where* $|\chi|$ denotes the number of elements in the range of X with equality if and only X has uniform distribution over χ. *[(MATH3222.2 & MATH3222.5)(Apply/IOCQ)]*

6 + 6 = 12

7. (a) A transmitter has an alphabet of four letters (x_1, x_2, x_3, x_4) and receiver has an alphabet of four letters (y_1, y_2, y_3, y_4). The joint probability matrix is

$$P(X,Y) = \begin{array}{c} \\ y_1 \\ y_2 \\ y_3 \\ y_4 \end{array} \begin{array}{c} \begin{array}{cccc} x_1 & x_2 & x_3 & x_4 \end{array} \\ \begin{bmatrix} 1/8 & 1/16 & 1/32 & 1/32 \\ 1/16 & 1/8 & 1/32 & 1/32 \\ 1/16 & 1/16 & 1/16 & 1/16 \\ 1/4 & 0 & 0 & 0 \end{bmatrix} \end{array}$$

Calculate H(X), H(Y), H(X, Y), $H(X/Y)$ and $H(Y/X)$. *[(MATH3222.2, MATH3222.5)(Understand/LOCQ)]*

(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	1/8	1/4	1/16
b	1/8	1/4	1/16
c	1/4	1/4	1/8
d	1/2	1/4	3/4

Show that triangle inequality does hold for divergence, i.e.

$D(p\,||r\,) > D(p\,||q) + D(q\,||r\,)$. *[(MATH3222.2, MATH3222.5)(Remember/LOCQ)]*

6 + 6 = 12

Group – E

8. (a) Let us consider the following joint probability mass function of (X, Y)

$$P\,(X,Y) = \begin{array}{c} \\ x_1 \\ x_2 \\ x_3 \end{array} \begin{array}{c} \begin{array}{ccc} y_1 & y_2 & y_3 \end{array} \\ \begin{bmatrix} \frac{1}{6} & \frac{1}{12} & \frac{1}{12} \\ \frac{1}{12} & \frac{1}{6} & \frac{1}{12} \\ \frac{1}{12} & \frac{1}{12} & \frac{1}{6} \end{bmatrix} \end{array}$$

(i) Find the minimum probability of error estimator $\hat{X}(Y)$ and the associated probability.

(ii) Evaluate Fano's inequality for this problem and compare.

[(MATH3222.6)(Evaluate/HOCQ)]

(b) Let $X_1, X_2, \cdots$ be $i.i.d.$ drawn according to probability mass function $p(x)$. Use Asymptotic Equipartition property to find the limit of $[P(X_1, X_2, \cdots, X_n)]^{1/n}$.

[(MATH3222.6)(Apply/IOCQ)]

6 + 6 = 12

9. (a) Let the transition probability matrix of a two-state Markov chain $X \rightarrow Y \rightarrow Z$ is given as follows: $P = \begin{pmatrix} 0.4 & 0.6 \\ 0.5 & 0.5 \end{pmatrix}$

Verify Data Processing lemma for the above problem (i. e. $I(X; Z) \leq I(X; Y)$). It is given that $p(x_1) = 0.6$ and $p(x_2) = 0.4$. *[(MATH3222.6)(Apply/IOCQ)]*

(b) Consider a sequence of i. i. d. binary random variables, $X_1, X_2, \cdots X_n$, where the probability that $X_i = 1$ is 0.7.

(i) With $n = 25$ and $\in = 0.1$, which sequences fall in the typical set $A_{\in}^{(n)}$? What is the probability of the typical set? How many elements are there in the typical set?

(ii) How many elements are there in the smallest set that has probability 0.9? (**Use Table-I**) *[(MATH3222.6)(Evaluate/HOCQ)]*

6 + 6 = 12

Cognition Level	*LOCQ*	*IOCQ*	*HOCQ*
Percentage distribution	*31.25*	*48.96*	*19.79*

Course Outcome (CO):

After the completion of the course students will be able to

MATH3222.1: Articulate the axioms (laws) of probability.

MATH3222.2: Compare and contrast different interpretations of probability theory selecting the preferred one in a specific context.

MATH3222.3: Formulate predictive models to tackle situations where deterministic algorithms are intractable.

MATH3222.4: Quantifies the amount of uncertainty involved in the value of a random variable or the outcome of a random process.

MATH3222.5: Apply the data processing inequality to data science, machine learning and social science.

MATH3222.6: Develop the concept of data compression in the process of encoding information in signal processing.

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