University of Ruhuna

Bachelor of Science General Degree

Level II (Semester II) Examination
November/December 2016

Subject: Applied / Industrial Mathematics
Course Unit: IMT223β/AMT223β/MAM2233 (Applied Probability - Information Theory)
Time: Two (02) Hours

Answer **Four (04) Questions** only. Calculators will be provided

---

1. Consider the following joint distribution $p(x, y)$ for two discrete random variables X and Y in tabular form:

Y \ X	1	2
1	0.25	0.50
2	0.25	0

(i) Find the marginal probability distributions $p(x)$ and $p(y)$ for X, Y respectively and calculate the corresponding entropies $H(X)$ and $H(Y)$.

(ii) **Using the definitions**, find joint entropy $H(X, Y)$ and mutual information $I(X; Y)$.

(iii) State and prove the chain rule for entropy (**in both forms**) and **hence** find the conditional entropies $H(Y|X)$ and $H(X|Y)$.

---

2. State the weak law of large numbers (WLLN) in probability theory.

Define "convergence in probability" of a sequence of random variables $U_1, U_2, \ldots, U_n$.

State and prove, in the usual notation, the Asymptotic Equipartition Property (AEP) theorem.

Let

$$U = \begin{cases} 1 & \text{with probability } 0.50 \\ 2 & \text{with probability } 0.25 \\ 3 & \text{with probability } 0.25 \end{cases}$$

Suppose that the random variables $U_1, U_2, \ldots, U_n$ are drawn independently and identically according to this distribution.

Find the limiting behaviour of

$$\left(\prod_{i=1}^{n} U_i\right)^{\frac{1}{n}}.$$

---

3. Obtain an expression for the channel capacity C of a binary symmetric channel (BSC) in the usual notation.

Consider a BSC with input probability distribution

$$\Pr(X = 0) = 0.4, \qquad \Pr(X = 1) = 0.6$$

and the crossover probability 0.1.

(i) Write down the channel matrix P.

(ii) Calculate the entropy $H(X)$ and conditional entropy $H(X|Y)$ and hence find the mutual information $I(X;Y)$.

(iii) Use the result in part (ii) to find the channel capacity C.

---

4. (a) Define the differential entropy of a continuous random variable X with the density function $f(x)$ and the support set S.
Find the differential entropy of

(i) a uniform distribution on $[\alpha, \beta]$, where $\alpha, \beta > 0$.

(ii) a Gaussian distribution with mean zero and variance σ^2.

(b) Define the Kullback-Leibler divergence $D(f\|g)$ for two probability density functions f and g on the same sample space Ω.
Let $\phi_1(x)$ and $\phi_2(x)$ be two Gaussian density functions with mean zero and varinces σ_1^2 and σ_2^2 respectively. Obtain an expression for $D(\phi_1\|\phi_2)$ in its simplest form.

---

5. Define an n-dimensional statistical model S of probability density functions $f(x;\theta)$ parameterized by θ, where $\theta = (\theta^1, \theta^2, \ldots, \theta^n)$.

Define the Fisher information matrix $G(\theta) = [g_{ij}(\theta)]$ of S at θ. Here, $g_{ij}(\theta)$ is the (i,j)th element of the matrix.

Explain briefly what you mean by "Fisher information".

Find the Fisher information matrix of a Gaussian distribution with mean μ and variance σ^2.

---

6. Explain what you mean by a "prefix code".

(a) State clearly the Kraft inequality for a D-ary alphabet.

(i) Consider the binary code $\{00, 100, 1100, 11100, 111100\}$.
Is it a prefix code? Justify your answer.
Does it satisfy the Kraft inequality? Justify your answer.

(ii) Discuss the possibility of finding a ternary prefix code with codewords of lengths 1, 2, 3, 3.

(iii) Discuss the existence of a binary prefix code for a 5-letter alphabet with codewords of lengths 1, 2, 3, 4, 5. If such a prefix code exists, give an example.

(b) Consider a source alphabet $A = \{d, e, f, g, h\}$ with the probability distribution $P = \{0.25, 0.25, 0.25, 0.125, 0.125\}$.
Suppose the prefix code $C = \{0, 10, 110, 1110, 11110\}$ is recommended.
Calculate the average length $\bar{\ell}$, entropy $H(P)$ and discuss the efficiency of the code.
