THE HONG KONG POLYTECHNIC UNIVERSITY

Department of Applied Mathematics

Subject Code:	AMA 563	Subject Title:	Principles of Data Science
Programme:	PgS in Construction and Environment (04001) MSc/PgD in Applied Mathematics for Science and Technology (63022) Master of Science in Data Science and Analytics (63027)		
Session:	Semester 1, 2021/2022		
Date:	7th December 2021	Time:	19:00 - 22:00
Time Allowed:	3 hours		

This question paper has **6** pages (including this page).

Instruction to Candidates: This question paper has **5** questions.

Attempt all questions.

Subject examiners: Dr. Stan Yip

DO NOT TURN OVER THE PAGE UNTIL YOU ARE TOLD TO DO SO

1a.) Let X_1 and X_2 have the joint pmf $p(x_1, x_2) = x_1 x_2/18$, $x_1 = 1, 2, 3$ and $x_2 = 1, 2$, zero elsewhere.

i.) Find the joint pmf of $Y_1 = X_1 X_2$ and $Y_2 = X_2$. (3 marks)

ii.) Find the marginal pmf of Y_1. (3 marks)

iii.) Find the mean, variance and moment generating function of Y_1. (4 marks)

1b.) Let $X_1, \dots, X_n$ be a random sample from a Beta distribution with parameters $\alpha = \theta$ and $\beta = 1$.

i.) Find the maximum likelihood estimator (MLE) of θ and the MLE of $\tau(\theta) = \frac{\theta}{1+\theta}$. (7 marks)

ii.) Is your MLE of θ a sufficient statistic? (3 marks)

2.) Let $X_1, \ldots, X_n$ be independent identically distributed random variables from a half normal $\mathrm{HN}(\mu, \sigma^2)$ distribution with pdf,

$$f(x) = \frac{2}{\sqrt{2\pi}\sigma} \exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right),$$

where $\sigma > 0$ and $x > 0$ and μ is real. Assume that μ is known.

a.) Find the maximum likelihood estimator of σ^2. (5 marks)

b.) What is the maximum likelihood estimator of σ? (3 marks)

c.) What is the uniformly most powerful with significance level α test for $H_0 : \sigma_2^2 = 1$ vs $H_1 : \sigma_2^2 = 4$? (8 marks)

d.) Assuming $\mu = 0$, find the expectation of the half normal distribution by substituting $w = \frac{x^2}{2\sigma^2}$. (4 marks)

3. Let $X_1, \ldots, X_n$ be independent identically distributed random variables with pdf

$$f(x|\theta) = \begin{cases} \frac{2x}{\theta} e^{-x^2/\theta} & x > 0, \\ 0 & x \leq 0. \end{cases}$$

a.) Find the MLE of θ. (3 marks)

b.) Show that X_1^2 is an unbiased estimator of θ by substituting $W = X^2$ or otherwise. (4 marks)

c.) Show that the distribution of X is a member of exponential family of distribution. (3 marks)

d.) Find the Cramer-Rao lower bound for the variance of an unbiased estimator of θ. (5 marks)

e.) Find the uniformly minimum variance unbiased estimator (UMVUE) of θ. (5 marks)

4a. Let $X_1, \ldots, X_n$ be independent identically distributed random variables from a $N(\mu, \sigma^2)$ distribution where the variance σ^2 is known. We want to test $H_0 : \mu = \mu_0$ against $H_1 : \mu \neq \mu_0$.

i.) Find the likelihood ratio test of size α for $H_0 : \theta = \theta_0$ vs $H_A : \theta \neq \theta_0$. (6 marks)

4b. Let $X_1, \ldots, X_n$ be a random sample from Bernoulli(p), where $p \in (0, 1)$ is unknown. Let $Y = \sum_{i=1}^{n} X_i$

i.) Find the maximum likelihood estimator of p. (4 marks)

ii.) Show that $\hat{p} = Y/n$ is an unbiased estimator of p. (3 marks)

iii.) Find the Fisher information $I(p)$. (4 marks)

iv.) Find the variance of $\hat{p}$ and show that it is a consistent estimator of p. (3 marks)

5a. A continuous random variable X follows a gamma distribution with parameters $\theta > 0$ and $\alpha > 0$ ($X \sim \Gamma(\alpha, \theta)$) if its probability density function is

$$f(x) = \frac{1}{\Gamma(\alpha)\theta^{\alpha}} x^{\alpha-1} e^{-x/\theta}, \quad , \text{for} x > 0$$

i.) By substituting $y = x\left(\frac{1}{\theta} - t\right)$, find the moment generating function of X. (7 marks)

ii.) From the result of part (i), show that $\sum_{i=1}^{m} X_i \sim \Gamma(m\alpha, \theta)$. (3 marks)

5b. Let X_1, X_2 be a random sample of size $n = 2$ from the distribution having pdf $f(x; \theta) = (1/\theta)e^{-x/\theta}$, $0 < x < \infty$, 0 elsewhere. We reject $H_0 : \theta = 2$ and accept $H_1 : \theta = 1$ if the observed values of X_1, X_2, say x_1, x_2, are such that

$$\frac{f(x_1; 2)f(x_2; 2)}{f(x_1; 1)f(x_2; 1)} \leq \frac{1}{2}.$$

Here $\Omega = \{\theta : \theta = 1, 2\}$.

i.) Find the significance level of the test. (6 marks)

ii.) Find the power the test when H_0 is false. (4 marks)