CONFIDENTIAL

UNIVERSITI TUN HUSSEIN ONN MALAYSIA

FINAL EXAMINATION
SEMESTER I
SESI 2015/2016

COURSE NAME	:	STATISTICAL INFERENCE
COURSE CODE	:	BWM 20503
PROGRAMME	:	2 BWQ
EXAMINATION DATE	:	DECEMBER 2015 / JANUARY 2016
DURATION	:	3 HOURS
INSTRUCTION	:	ANSWER **FIVE (5)** QUESTIONS ONLY

THIS EXAMINATION PAPER CONSISTS OF **FIVE** (5) PAGES

CONFIDENTIAL

Q1 An independent and identically distributed sample $X_1, X_2, \ldots, X_n$ is taken from a discrete distribution F with probability function $f_\theta(x) = (\theta - 1)^{x-1}\theta^{-x}$ for $x = 1, 2, 3, \ldots$ where $\theta \in \Omega = (1, \infty)$ is unknown parameter. For this distribution,

$$E(X) = \theta \text{ and } Var(X) = \theta(\theta - 1).$$

(a) Define the Score statistics (or Score function), $S(X; \theta)$.

(2 marks)

(b) For sample $s = (x_1, x_2, \ldots, x_n)$, show that $\bar{x}$ is a minimal sufficient statistics.

(5 marks)

(c) Find the maximum likelihood estimate $\hat{\theta}$ in terms of $\bar{x}$.

(5 marks)

(d) Find expression for the bias, variance and mean square error of $\hat{\theta}$.

(3 marks)

(e) Evaluate $\hat{\theta}$ for the sample $s = (4, 8, 10, 9, 1)$ and find numerical estimates of the bias, variance and mean square error.

(3 marks)

(f) An alternative parameterization is $\varphi = \frac{1}{\theta}$. Find the maximum likelihood estimate of φ.

(2 marks)

Q2 An independent and identically distributed sample $X_1, X_2, \ldots, X_n$ is taken from an absolutely continuous distribution with density $f_\theta(x) = \frac{x}{\theta^2}e^{\frac{-x}{\theta}}$ for $x > 0$ where $\theta \in \Omega = (0, \infty)$ is an unknown parameter. For this distribution, $E(X_i) = 2\theta$ and $Var(X_i) = 2\theta^2$. Recall that for sufficiently large samples that the distribution of the maximum likelihood estimate (MLE) $\hat{\theta}$ is approximately $N(\theta, (nI(\theta))^{-1})$. A sample with $n = 100$ has the following summary statistics: $\sum_{i=1}^{n} x_i = 191.8$, $\sum_{i=1}^{n} \log x_i = 35.4$, $\sum_{i=1}^{n} x_i^2 = 556.0$, $\min_i\{x_i\} = 0.13$ and $\max_i\{x_i\} = 6.52$.

(a) Define the Fisher Information for a single random variable, $I(\theta)$.

(2 marks)

(b) Find an expression for the maximum likelihood estimate $\hat{\theta}$ of θ and evaluate it numerically.

(5 marks)

(c) Find an expression for the Fisher information, $I(\theta)$ for the distribution and find a numerical estimate of it.

(5 marks)

(d) Construct a 95% confidence interval for θ using the large sample properties of the MLE.

(4 marks)

(e) In a test of the null hypothesis, H_0: $\theta = 1$ versus the two-sided alternative hypothesis, H_1: $\theta \neq 1$, find the value of the test statistics and determine the p-value is greater than or lower than 0.05.

(4 marks)

Q3 Let $X_1, X_2, \ldots, X_n$ be a random sample from a population that is uniformly distributed on the interval $(0, \theta)$, where the parameter θ is positive.

(a) Let $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$. Show that $\hat{\theta} = 2\bar{X}$ is the method of moment estimator of θ and that it is unbiased.

(5 marks)

(b) Consider the random sample 0.2, 0.3, 1.0 and 0.1 from the above population. Evaluate $\hat{\theta}$ and comment on the usefulness of the estimate.

(4 marks)

(c) Let $Y = \max \{X_i\}$. Show that the probability density function of Y is

$$g(y) = \frac{ny^{n-1}}{\theta^n}, \quad 0 < y < \theta.$$

(5 marks)

(d) An estimator $\hat{\theta} = cY$ is to be used to estimate θ, where the multiplier c is to be chosen. Show that the mean square error of $\hat{\theta}$ is minimised when $c = \frac{n+2}{n+1}$.

(6 marks)

Q4 (a) Let $X_1, X_2, \ldots, X_n$ be a random sample from a continuous population that is uniformly distributed in the interval of $(0, \theta)$, where the parameter θ is positive. You are given that this distribution has mean $\frac{\theta}{2}$ and variance $\frac{\theta^2}{12}$. Outline why the Cramer-Rao lower bound for the variance of unbiased estimators of θ does not apply in this case.

(4 marks)

(b) Let $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ and let $Y = \max\{X_i\}$. Consider the three estimator of θ as follows:

$$\hat{\theta}_1 = 2\bar{X} \ ; \ \hat{\theta}_2 = Y \ ; \ \hat{\theta}_3 = \left(\frac{n+1}{n}\right)Y$$

(i) Find the bias and variance of each of these estimators.
[Hint. The pdf of Y is $g(y) = \frac{ny^{n-1}}{\theta^n}, \ 0 < y < \theta$]

(7 marks)

(ii) Hence discuss the consistency of each of the estimators.

(4 marks)

(iii) Calculate the efficiency of $\hat{\theta}_1$ relative to $\hat{\theta}_3$. What happens to the efficiency as n approach ∞?

(5 marks)

Q5 (a) Define α and β for a statistical test of hypothesis.

(2 marks)

(b) Let $Y_1, Y_2, \ldots, Y_n$ denote a random sample from a population having a Poisson distribution with mean λ.

(i) Find the form of the rejection region for a most powerful test of

H_0: $\lambda = \lambda_0$ against H_1: $\lambda = \lambda_a$ where $\lambda_a > \lambda_0$.

(8 marks)

(ii) Let $\sum_{i=1}^{n} Y_i$ has a Poisson distribution with mean $n\lambda$. Indicate how this information can be usedto find any constants associated with the rejection region derived in **Q5(b)(i).**

(3 marks)

(iii) Is the test derived in **Q5(b)(i)**, uniformly most powerful for testing H_0: $\lambda = \lambda_0$ against H_1: $\lambda = \lambda_a$ where $\lambda_a > \lambda_0$? Explain your reason.

(2 marks)

(iv) Find the form of the rejection region for a most powerful test of H_0: $\lambda = \lambda_0$ against H_1: $\lambda = \lambda_a$ where $\lambda_a < \lambda_0$.

(5 marks)

Q6 (a) State the Neymann-Pearson Lemma.

(2 marks)

(b) Suppose X_1, X_2, …,X_n is a random sample from a normal population with mean μ and variance 16. Find the best critical region, which is the most powerful test, with a sample size of n = 16 and a significance level α = 0.05. Test the simple null hypothesis H_0: μ = 10 against the simple alternative hypothesis H_1: μ = 15.

(7 marks)

(c) A food processing company packages honey in small glass jars. Each jar is supposed to contain 10 ounces fluid of the sweet and gooey good stuff. Previous experience suggests that the volume X, the volume in fluid ounces of a randomly selected jar is normally distributed with a known variance of 2. Derive the likelihood ratio test for testing, at a significance level of α = 0.05, the null hypothesis H_0: μ = 10 against the alternative hypothesis H_A: $\mu \neq 10$.

(7 marks)

(d) An industrial designer wants to determine the average amount of time it takes an adult to assemble an "easy to assemble" toy. A sample of 16 times yielded an average time of 19.92 minutes, with a standard deviation of 5.73 minutes. Assuming normality of assembly times, provide a 95% confidence interval for the mean assembly time.

(4 marks)

- END OF QUESTION -