# UNIVERSITI TUN HUSSEIN ONN MALAYSIA

## FINAL EXAMINATION
## SEMESTER I
## SESSION 2019/2020

| | | |
|---|---|---|
| COURSE NAME | : | STATISTICAL INFERENCE |
| COURSE CODE | : | BWB 20503 |
| PROGRAMME CODE | : | BWQ |
| EXAMINATION DATE | : | DECEMBER 2019 / JANUARY 2020 |
| DURATION | : | 3 HOURS |
| INSTRUCTION | : | ANSWER **ALL** QUESTIONS |



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

**Q1** (a) Let $X$ be a continuous random variable with density function given by

$$f_X(x) = \begin{cases} be^x & 0 < x < \ln(2) \\ 0 & \text{otherwise} \end{cases}$$

(i) Find the value of $b$.

(2 marks)

(ii) Let $Y = e^X$. Determine the density function $f_Y(y)$ of $Y$.

(5 marks)

(iii) Identify the type of continuous random variable of $Y$.

(2 marks)

(b) The density function of the continuous random variable $X$ is given by

$$f_X(x) = \begin{cases} v(x + \sqrt{x} & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

(i) Compute the constant $v$.

(2 marks)

(ii) Determine the density function of $Y = \frac{1}{X}$.

(4 marks)

(c) Assume $X$ is a uniform distribution on $(0, \theta)$ and fix $n$. For any given random sample $x = (x_1, x_2,.., x_n)$ and estimator of $\theta$ is given by

$$\hat{\theta}(x) = \text{Max}(x_1, x_2, \dots, x_n).$$

Identify the distribution function of $\hat{\theta}$ and its expected value.

(5 marks)

**Q2** Let $(X_1, X_2, \dots, X_n)$ be a random sample of identically independent distribution random variables distributed as follows;

$$f(x;\theta) = \frac{\theta 3^\theta}{x^{\theta+1}} \quad x > 3$$

(a) Show that $\sum_i \log(X_i)$ is sufficient statistics for $\theta$.

(3 marks)

(b) Determine $\hat{\theta}_{\text{MLE}}$ maximum likelihood estimator (MLE) for $\theta$ and discuss properties of this estimator.

(5 marks)

(c) Determine $\hat{\theta}_{\text{MOM}}$ method of moment estimator (MOM) for $\theta$.

(5 marks)

(d) Compute the score function and the Fisher information.

(4 marks)



(e) Specify asymptotic distribution of $\hat{\theta}_{\text{MLE}}$.

(3 marks)

**Q3.** (a) Let consider an independent random samples are selected from two population. **Table Q3**, presented selected summary statistics.

**Table Q3**. Summary statistics

| Population | Mean | Standard deviation | Sample size |
|---|---|---|---|
| A | 62.00 | 10.00 | 17 |
| B | 54.00 | 6.00 | 10 |

(i) Construct the 95% confidence interval for $\mu_A - \mu_B$.

(6 marks)

(ii) Obtain the $p$-value for alternative $\mu_A \neq \mu_B$ and derive your conclusion.

(4 marks)

(b) Mr. SS performs three simulation studies. His population is skewed to the right. For one study he has his computer generate 10,000 random sample of size $n = 10$ from the population. For each random sample, the computer calculates the 95% confidence interval for $\mu$ and checks to see whether the interval is correct. His second study is similar to the first study, but with the sample size $n = 120$. Finally, his third study also similar to the first study, but with different sample size of $n = 250$. In one of his studies, Mr. SS obtains 9,454 correct interval; in another he obtains 9,677 correct intervals; and in the remaining study he obtains 8,655 correct interval. Based on theory we learned about the relation between confidence interval and sample size, match each sample size to its number of correct intervals. Justify your answer.

(6 marks)

(c) For each of the following statements indicate it's **TRUE** or **FALSE**:

(i) An estimator $T = T(X)$ is a consistent estimator for $\theta$ if the distribution is asymptotically Gaussian.

(ii) An estimator $T = T(X)$ is a consistent estimator for $\theta$ if the more data you collect, the closer the estimate will be to the real population parameter.

(iii) A statistic $T = T(X)$ is a sufficient statistic for $\theta$ if it can be computed without knowing the value of $\theta$.

(iv) The Invariance property of MLE's implies that their variance approaches zero as the sample size increases.

(4 marks)

**Q4** (a) Let us consider three random observation $x_1 = 3$, $x_2 = 4$ and $x_3 = 8$ taken from a Geometric distribution with probability distribution function and unknown parameter $\theta$,

$$p(x;\theta) = (1-\theta)^{x-1}\theta, \quad x = 1, 2, 3, \ldots$$

The moment estimator for the geometric parameter $\theta$ is $\hat{\theta} = \frac{1}{\bar{X}}$

(i) Obtain the likelihood function for random sample size 3.

(3 marks)

(ii) Analyse the likelihood function at $\theta = 0.0, 0.2, 0.4, 0.6, 0.8$ and $1.0$.

(5 marks)

(iii) Estimate the maximum likelihood estimator for $\hat{\theta}$.

(2 marks)

(b) Let consider the continuous random variable $X$ with probability density function

$$f(x) = \frac{\theta K^{\theta}}{x^{\theta+1}}, \qquad x \geq K,$$

where $\theta > 1$ and $K > 0$, are said to follow a Pareto distribution with parameter $K$ and $\theta$, and also can be written as $X \sim \text{Pareto}(K, \theta)$. Estimate the maximum likelihood estimator of $\theta$, $\hat{\theta}$.

(5 marks)

(c) Let $X$ represent a population and $X_1, X_2, \ldots, X_n$ be a sample chosen independently with replacement, consider the

$$S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{x})^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \frac{1}{n}\sum_{i=1}^{n}X_i\right)^2$$

Determine $Var(X)$ if $S^2$ is an unbiased estimator of the population variance.

(5 marks)

**Q5** (a) A test statistic has a Binomial distribution with parameter $n = 25$ and the probability of success is $p$, $B(25, p)$. Given that

$H_0: p = 0.5$ $\qquad$ $H_1: p \neq 0.5$

(i) Find the critical region for the test statistics such that the probability in each tail is as close as possible to 0.025.

(3 marks)

(ii) Compute the probability of incorrectly rejecting $H_0$ using this critical region.

(2 marks)

(b) Mr AA claims that the weather forecast produced by local radio no better than those achieved by tossing a fair coin and predicting rain if a head is obtained or no rain if a tail is obtained. He records the weather for 30 randomly selected days. The local radio forecast is correct on 21 of these days. State your hypotheses and test Mr AA claims at 5% level of significance.

(5 marks)

(c) Let $X_i, X_2, \ldots, X_n$ is random variable taken from normal distribution with mean $\mu$ and variance $\sigma$ and the probability density function is given as

$$f(x_i|\mu) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}}$$

for $i = 1, 2, \ldots, n$. Derive the likelihood ratio test for the hypothesis

$$H_0: \mu = \mu_0 \text{ versus } H_0: \mu > \mu_0$$

and show whether it is Uniformly Most Powerful (UMP) test.

(10 marks)

**- END OF QUESTIONS -**