Strathmore
UNIVERSITY

**STRATHMORE INSTITUTE OF MATHEMATICAL SCIENCES (SIMS)**
**MASTER OF SCIENCE IN STATISTICAL SCIENCES**
**END OF SEMESTER EXAMINATION**
**STA 8102: STATISTICAL INFERENCE**

**DATE**: Monday 23rd Aug, 2021 **TIME**: 3 Hours

---

## INSTRUCTIONS

1. This examination consists of **FOUR** questions.
2. Answer Question **ONE (COMPULSORY)** and any other **TWO** questions.
3. You may use a **SIMPLE CALCULATOR**. No **MOBILE PHONES** in the exams room.

---

## Question One (20 Marks)

(i) Define Test statistics as used in hypothesis testing. (2 marks)

(ii) Let $X_1, X_2, \cdots, X_n$ be a random sample from a Geometric ($\theta$) distribution, where $\theta$ is unknown. Find the maximum likelihood estimator of $\theta$ based on this random sample. (3 marks)

(iii) Explain the properties of point estimators. (3 marks)

(iv) Let $X_1, \ldots, X_n$ be a s.r.s. of a r.v. with mean $\mu$ and variance $\sigma^2$. Consider the following estimators of $\mu$:

$$\hat{\mu}_1 = \frac{X_1 + X_2}{2}, \quad \hat{\mu}_2 = \frac{X_1}{4} + \frac{X_2 + \ldots + X_{n-1}}{2(n-2)} + \frac{X_n}{4}, \quad \hat{\mu}_3 = \bar{X}.$$

(a) Which one is unbiased for $\mu$? (3 marks)

(b) Which one is consistent in probability for $\mu$? (5 marks)

(v) A manufacture of cell phone batteries wants to estimate the useful life of it's battery (in thousand of hours.) The estimate is the be within 0.10 (100) hours. Assume 1 95 percent level of confidence and that the standard deviation of the useful life of the battery is 0.90 (900 hours). Determine the required sample size. (4 marks)

**Question Two (20 Marks)**

(i) Let $X$ be a random variable defined by

$$X = \theta + W,$$

where $W \sim N\left(0, \frac{1}{9}\right)$.

(a) Consider the test $H_0 : \theta = \theta_0 = 0$ vs. $H_1 : \theta = \theta_1 = 1$. Find the likelihood ratio test statistic. (5 marks)

(b) Let $X = x$. Design a level 0.05 test ($\alpha = 0.05$) to decide between $H_0$ and $H_1$. (5 marks)

(ii) Consider a population where $X \sim$ Bin$(6; \pi)$ is observed, i.e. $X$ follows a binomial distribution. We wish to test:

$$H_0 : \pi = 0.50 \text{ vs. } H_1 : \pi = 0.75.$$

The chosen statistical test rejects $H_0$ if the observed value of $X$ is $x \geq 5$.

(a) Calculate the size of the test. (3 marks)

(b) Calculate the power of the test. (3 marks)

(iii) An airline wants to evaluate the depth perception of its pilots over the age of 50. A random sample of $n = 14$ airline pilots over the age of 50 are asked to judge the distance between two markers placed 20 feet apart at the opposite end of the laboratory. The sample data listed here are the pilots' error (recorded in feet) in judging the distance.

| | | | | | | |
|---|---|---|---|---|---|---|
| 2.7 | 2.4 | 1.9 | 2.6 | 2.4 | 1.9 | 2.3 |
| 2.2 | 2.5 | 2.3 | 1.8 | 2.5 | 2.0 | 2.2 |

Use the sample data to test the hypothesis that the average error in depth perception for the company's pilots over the age of 50 is 2.00 at $\alpha = 0.05$ confidence level on $\mu$. (4 marks)

**Question Three (20 Marks)**

(i) The compressive strengths of 40 test cubes of concrete samples with the sample mean and sample standard deviation of 60.14 and $5.02 N/mm^2$, respectively. We also assume that the compressive strengths are normally distributed. To facilitate the application, let us assume that the estimated standard deviation of $5.02 N/mm^2$ is the true on known value.

(a) Construct a 95% confidence interval for the population mean $\mu$. (3 marks)

(b) Construct an upper one-sided 99% confidence limit for the population variance. (3 marks)

(c) Construct a 95% two-sided confidence limit for the population variance. (3 marks)

(ii) The summary statistics given below from two catalysts types in which 8 samples in the pilot plant are take from each are being analyzed to determine how they affect the mean yield of a chemical process. Specifically, the 1st catalyst is currently in use, but the 2nd catalyst is acceptable.

| Observation number | Catalyst 1 | Catalyst 2 |
|---|---|---|
| 1 | 91.50 | 89.19 |
| 2 | 94.18 | 90.95 |
| 3 | 92.18 | 90.46 |
| 4 | 95.39 | 93.21 |
| 5 | 91.79 | 97.19 |
| 6 | 89.07 | 97.04 |
| 7 | 94.72 | 91.07 |
| 8 | 89.21 | 92.75 |

(a) Construct a confidence interval for difference between the mean yields. Use $\alpha = 0.05$, and assume equal variances. (3 marks)

(b) Construct a confidence interval for the ratio variance of yields. Use $\alpha = 0.05$. (3 marks)

(iii) In a sample of 200 residents of Georgetown country, 120 reported they believe that country real estate taxes were too high. Developed a 95% confidence interval that proportioned of residents who believe the tax rate is too high. Would it be reasonable to conclude that the majority of the taxpayers feel that the taxes are too high. (5 marks)

## Question Four (20 Marks)

(i) (a) Suppose we have a random sample $X_1, X_2, \cdots, X_n$ for which the probability density (or mass) function of each $X_i$ is $f(x_i; \theta)$, describe (briefly) how the maximum likelihood point estimator (MLE) is obtained. (3 marks)

(b) Discuss the optimality properties of the MLE. (3 marks)

(ii) Suppose that $X$ is a random variable with mean $\mu$ and variance $\sigma^2$. Let $X = \{X_1, X_2, \cdots, X_n\}$ be a random variable of size $n$ from the population represented by $X$.

Show that $S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$ is biased estimator of $\sigma^2$. (7 marks)

(iii) Suppose that $X$ is a discrete random variable with the following probability mass function: where $0 \leq \theta \leq 1$ is a parameter. The following 10 independent observations

| X | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| P(X) | $\frac{2\theta}{3}$ | $\frac{\theta}{3}$ | $\frac{2(1-\theta)}{3}$ | $\frac{(1-\theta)}{3}$ |

were taken from such a distribution: (3,0,2,1,3,2,1,0,2,1). What is the maximum likelihood estimate of $\theta$. (7 marks)

***END***