STRATHMORE INSTITUTE OF MATHEMATICAL SCIENCES (SIMS)
BBS FENG & BBS ACT & BBS FE
END OF SEMESTER EXAMINATION
BSM 2111: STATISTICAL INFERENCE

DATE: Tuesday, 25th July 2023 **TIME**: 2 Hours

INSTRUCTIONS

1. This examination consists of **FIVE** 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 (30 Marks)

(i) Explain and give examples of three types of random sampling. (6 marks)

(ii) One source of water pollution is gasoline leakage from underground storage tanks. In Mombasa, a random sample of $n = 74$ gasoline stations is selected and the tanks are inspected; 10 stations are found to have at least one leaking tank. Calculate a 95 percent confidence interval for p, the population proportion of gasoline stations with at least one leaking tank. (4 marks)

(iii) Consider the density function

$$f(x) = \begin{cases} (p+1)x^p, & 0 \leq x \leq 1 \\ 0, & \text{otherwise} \end{cases}$$

where p is greater than -1. Compute the expected value of X. (3 marks)

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

(v) Using the identity (5 marks)

$$(\hat{\theta} - \theta) = \left(\hat{\theta} - E[\hat{\theta}]\right) + \left(E[\hat{\theta}] - \theta\right) = \left(\hat{\theta} - E[\hat{\theta}]\right)$$

show that

$$\text{MSE}[\hat{\theta}] = E\left[(\hat{\theta} - \theta)^2\right] = \text{Var}[\hat{\theta}] + \left(\text{Bias}[\hat{\theta}]\right)^2$$

(vi) After taking a refresher course, a salesman found that his sales (in dollars) on 9 random days were 1280, 1250, 990, 1100, 880, 1300, 1100, 950 and 1050. Does the sample indicate that the refresher course had the desired effect, in that his mean sale is now more than 1000 dollars? Assume $\sigma = 100$, and the probability of erroneously saying that the refresher course is beneficial should not exceed 0.01. Also assume that the sales are normally distributed. (5 marks)

(vii) 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 θ. (4 marks)

Question Two (20 Marks)

(i) Assuming that $X_i \sim N(\mu, \sigma^2)$, which of the statistics below are unbiased estimators of μ?

$$(a)\quad \hat{\mu}_1 = \frac{X_1 + X_2 + X_3 + X_4}{4} \quad (b)\quad \hat{\mu}_2 = \frac{2(X_1 + X_2)}{6} + \frac{X_3 + X_4}{6} \quad (c)\quad \hat{\mu}_3 = \frac{X_1 - X_2 + X_3 - X_4}{4}.$$

Among all the unbiased estimators, which one is the most efficient? Which one is the most consistent among all the three estimators? (7 marks)

(ii) Suppose $X_1, X_2, \cdots, X_n$ are i.i.d random variables with density function $f(x|\sigma) = \frac{1}{2\sigma} \exp\left(-\frac{|x|}{\sigma}\right)$.

(a) Find the maximum likelihood estimator for σ. (3 marks)

(b) Show that MSE of $\hat{\sigma}$ is equal to its variance. (3 marks)

(iii) Let $X_1, X_2, \ldots, X_n$ be gamma random variables with parameters α and θ so that the probability density function is:

$$f(x_i) = \frac{1}{\Gamma(\alpha)\theta^{\alpha}} x^{\alpha-1} e^{-x/\theta}$$

what are the method of moments estimators of α and θ? (7 marks)

Question Three (20 Marks)

(i) The head of the Statistics department of a certain university is interested in the difference in writing scores between freshman Statistics students who are taught by different teachers. The incoming freshmen are randomly assigned to one of two Statistics teachers and are given a standardized writing test after the first semester. We take a sample of eight students from one class and nine from the other class. Is there a difference in achievement on the writing test between these two classes? (5 marks)

Class1	35	51	66	42	37	46	60	55	53
Class2	52	87	76	62	81	71	55	67	

(ii) A university has found over the years that out of all the students who are offered admission, the proportion who accept is 0.70. After a new director of admissions is hired, the university wants to check if the proportion of students accepting has changed significantly. Suppose they offer admission to 1200 students and 888 accept. Is this evidence at the $\alpha = .05$ level that there has been a real change from the status quo? How about at the 0.02 level? (3 marks)

(iii) In an experiment in breeding mice, a geneticist has obtained 120 brown mice with pink eyes, 48 brown mice with brown eyes, 36 white mice with pink eyes and 13 white mice with brown eyes. Theory predicts that these types of mice should be obtained in the ratios 9 : 3 : 3 : 1. Test the compatibility of the data with theory, using a 5% critical value. (6 marks)

The following table records the observed frequencies in its first row and the frequencies expected under the null hypothesis H0 in its second row:

Observed	120	48	36	13
Expected	122	41	41	14

(iv) With the current level of communication resources for an online bookstore and their projected growth over the next 6 months, a company will be able to provide satisfactory service if the average connection time per customer is no more than 13.5 minutes. Based on a random sample of 45 connections yielding s a sample mean of 15.3 minutes with a sample standard deviation of 6.7 minutes, would you recommend that the company upgrades their communication resources? (Perform a one-sided test at a 5% significance level.) (6 marks)

Question Four (20 Marks)

(i) A certain student did a survey of 40 small town coffee shops and 49 big city coffee shops, and established that the mean price of a large cup of coffee is \$3.75 and in the big cities it is \$4.50. The population standard deviation in small towns is known to be 1.20, and in big cities the population standard deviation is known to be 0.98. Construct a confidence interval for the difference of their two means, and draw conclusions from it. (6 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

Construct a confidence interval for the ratio variance of yields. Use $\alpha = 0.05$. (8 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 μ. (6 marks)

Question Five (20 Marks)

(i) A continous random variable X is uniformly distibuted over the interval $[-2, 7]$.

(a) Write down fully the probability density function $f(x)$ of X. (3 marks)

(b) Find $E(X^2)$ (4 marks)

(c) Find $P(-0.2 < X < 0.6)$. (3 marks)

(ii) A discrete random variable X has the probability function

$$P(X = x) = \begin{cases} kx & x = 2, 4, 6 \\ k(x-2) & x = 8 \\ 0 & \text{otherwise,} \end{cases}$$

where k is a constant.

(a) Find the value of k. (2 marks)

(b) Find the exact value of $E(X)$ and $E(X^2)$. (4 marks)

(c) Calculate $\text{Var}(3 - 4X)$ giving your answer to 3 significant figures. (4 marks)

END