Strathmore
UNIVERSITY

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

DATE: 15th December, 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) It is estimated that a particular flight is profitable if the average occupation rate during a year is at least a 60%. An airline is interested in determining whether it is profitable to keep a particular flight operative. For that, they record the occupation rates of 120 random flights scattered around the year, resulting a mean occupation rate of 58% and a standard deviation of 11%. Considering that the occupation rates (in proportion) have an approximate normal distribution, is there enough evidence to cancel the flight because it is not profitable? Employ a significance level of $\alpha = 0.10$. (4 marks)

(ii) Let $X_1, X_2, \ldots, X_n$ be a s.r.s from a Weibull distribution, whose p.d.f is given by

$$f(x;\theta) = \frac{2x}{\theta} e^{-\frac{x^2}{\theta}}, \qquad x > 0.$$

Find a minimal sufficient statistics for θ. (3 marks)

(iii) Show how Bayesian approach to statistical inference can be used to determine an estimate of μ. Assume normal distribution. (3 marks)

(iv) 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. (3 marks)

(v) Assume that X represents a single observation of the r.v. with p.d.f.

$$f(x;\theta) = \begin{cases} \theta x^{\theta-1} & 0 < x < 1, \\ 0 & \text{otherwise.} \end{cases}$$

Find the UMP test at a significance level $\alpha = 0.05$ for testing (4 marks)

$$H_0 : \theta = 1 \quad \text{vs.} \quad H_1 : \theta = 2.$$

(vi) Let $X_1, X_2, \ldots, X_n$ be a s.r.s of a $U(\theta_1, \theta_2)$ r.v with $\theta_1 < \theta_2$. Find a sufficient statistics for (θ_1, θ_2). (3 marks)

Question Two (20 Marks)

(i) Civil engineers have found that the ability to see and read a sign at night depends in part on its "surround luminance;" i.e., the light intensity near the sign. It is believed that the mean surround luminance is 10 candela per m^2 in a large metropolitan area. The data below are $n = 30$ measurements of the random variable X, the surround luminance (in candela per m^2). The 30 measurements constitute a random sample from all the signs in the large metropolitan area in question:

10.9,	1.7,	9.5,	2.9,	9.1,	3.2,	9.1,	7.4,	13.3,	13.1
6.6,	13.7,	1.5,	6.3,	7.4,	9.9,	13.6,	17.3,	3.6,	4.9
13.1,	7.8,	10.3,	10.3,	9.6,	5.7,	2.6,	15.1,	2.9,	16.2

Based on past experience, the engineers assume a normal population distribution (for the population of all signs) with known population variance $\sigma^2 = 20$. From this data what conclusions should we draw about the hypothesized mean surround luminance, at the $\alpha = 0.05$ significance level? (5 marks)

(ii) A labor union registers the number of complains that are filled per week by the workers of two different shifts in the production line of a factory. One hundred independent observations about the number of complains for both shifts gave the means $\bar{X} = 20$ and $\bar{Y} = 22$. Assume that the number of complains per week of the i-th shift has a Pois(θ_i) distribution, $i = 1, 2$. The labour union wants to test if the average number of complaints per week of both shifts is significantly different or not at a significance level $\alpha = 0.01$. Use the likelihood ratio test. (6 marks)

(iii) A standard insecticide is used to control a particular insect pest in soy beans. The probability with which this insecticide eliminates the insect from an individual plant infested with the insect is 0.7. A new insecticide is developed which is known not to decrease the probability of insect elimination; it is desired to determine if the new insecticide performs better than the standard. An experiment is to be conducted in which 500 randomly selected plants (known to be infested with the insect) are sprayed with the new insecticide. After a fixed period of time, the plant for which the new insecticide eliminate the insect will be counted. The experiment

will reject the null hypothesis that the insecticide has no effect if the insecticide eliminates the insect in 375 or more plants.

(a) State symbolically the null and alternative hypotheses. (1 mark)

(b) What is the α-level of the proposed test? (2 marks)

(c) If in fact the success rate of the insecticide is 0.8, what is the power of the test? (3 marks)

(d) When the experiment was actually carried out, it was found that new insecticide eliminates insect in 395 plants. Find the P-value of the test. (3 marks)

Question Three (20 Marks)

(i) 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)

(ii) 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? (5 marks)

(iii) Let $X_1, X_2, \ldots, X_n$ be a s.r.s from a distribution with mean μ and variance σ^2. Consider the following estimators of μ:

$$\hat{\mu}_1 = \frac{X_1 + 2X_2 + 3X_3}{6}, \quad \hat{\mu}_2 = \frac{X_1 + 4X_2 + X_3}{6}, \quad \hat{\mu}_3 = \frac{\frac{3}{2}X_1 + \frac{1}{2}X_2 + X_3 + \cdots + X_n}{n}.$$

(a) Which ones are unbiased? (3 marks)

(b) Among the unbiased ones, which is the most effective? (3 marks)

(c) Which of them is consistent in squared mean? (3 marks)

Question Four (20 Marks)

(i) The feeding habits of two species of net-casting spiders are studied. The species, the deinopis and menneus, coexist in eastern Australia. The following data were obtained on the size, in millimeters, of the prey of random samples of the two species:

dinopis	12.9	10.2	7.4	7.0	10.5	11.9	7.1	9.9	14.4	11.3
menneus	10.2	6.9	10.9	11.0	10.1	5.3	7.5	10.3	9.2	8.8

(a) What is the difference, if any, in the mean size of the prey (of the entire populations) of the two species? (4 marks)

(b) Estimate, with 95% confidence, the ratio of the two population variances. (4 marks)

(ii) A gunpowder manufacturer developed a new formula that was tested in eight bullets. The resultant initial velocities, measured in feet per second, were

$$\begin{array}{cccc} 3005, & 2925, & 2935, & 2968, \\ 2995, & 3005, & 2937, & 2905. \end{array}$$

Assuming that the initial velocities have normal distribution with $\sigma = 39$ feet per second, find a confidence interval at level $\alpha = 0.05$ for the initial mean velocity of the bullets that employ the new gunpowder. (3 marks)

(iii) A random sample of 20 nominally measured 2mm diameter steel ball bearings is taken and the diameters are measured precisely. The measurements, in mm, are as follows:

$$\begin{array}{cccccccccc} 2.02, & 1.94, & 2.09, & 1.95, & 1.98, & 2.00, & 2.03, & 2.04, & 2.08, & 2.07 \\ 1.99, & 1.96, & 1.99, & 1.95, & 1.99, & 1.99, & 2.03, & 2.05, & 2.01, & 2.03. \end{array}$$

Assuming that the diameters are normally distributed with unknown mean, μ, and unknown variance, σ^2. find a two-sided 95% confidence interval for the variance, σ^2. (4 marks)

(iv) In a biomedical experiment, we would like to estimate the population mean remaining life μ of healthy rats that are given a certain dose of a toxic substance. Suppose that we would like to write a 95 percent confidence interval for μ with a margin of error equal to $E = 2$ days. From past studies, remaining rat lifetimes have been approximated by a normal distribution with standard deviation $\sigma = 8$ days. How many rats should we use for the experiment? (2 marks)

(v) Acute exposure to cadmium produces respiratory distress and kidney and liver damage (and possibly death). For this reason, the level of airborne cadmium dust and cadmiumoxide fume in the air, denoted by X (measured in milligrams of cadmium per m^3 of air), is closely monitored. A random sample of $n = 35$ measurements from a large factory are given below:

$$\begin{array}{ccccccc} 0.044, & 0.030, & 0.052, & 0.044, & 0.046, & 0.020, & 0.066, \\ 0.052, & 0.049, & 0.030, & 0.040, & 0.045, & 0.039, & 0.039, \\ 0.039, & 0.057, & 0.050, & 0.056, & 0.061, & 0.042, & 0.055, \\ 0.037, & 0.062, & 0.062, & 0.070, & 0.061, & 0.061, & 0.058, \\ 0.053, & 0.060, & 0.047, & 0.051, & 0.054, & 0.042, & 0.051. \end{array}$$

Find a 99 percent confidence interval for μ, the mean level of airborne cadmium. (3 marks)

END