STRATHMORE INSTITUTE OF MATHEMATICAL SCIENCES

BACHELOR OF BUSINESS SCIENCE (ACTURIAL SCIENCE, FINANCE, FINANCIAL ECONOMICS)

BSM 2111: STATISTICAL INFERENCE EXAM

Date: 26th July 2019 Time: 2 Hours

Instruction: Question 1 COMPULSORY and ANY other TWO.

Question 1 (30 Marks)

(a) Let $X_1, X_2, ...X_n$ is a random sample from a distribution with probability density function $f(x;\theta)$, where θ is an unknown parameter. $T(X_1, X_2,..., X_n)$ is a function of $X_1, X_2, ...X_n$ which is to be used to estimate θ. Explain the meaning of the following:

(i) $T(X_1, X_2,..., X_n)$ is an efficient unbiased estimator of θ (2 Marks)

(ii) $T(X_1, X_2,..., X_n)$ is an unbiased estimator of θ (1 Mark)

(iii) $T(X_1, X_2,..., X_n)$ is a consistent estimator of θ (2 Marks)

(b) A financial model is given by the following probability distribution function

$$f(X) = \frac{1}{\theta}\exp(-x/\theta), \qquad x > 0 \qquad .$$

(i) Show that $\overline{X}$ is an unbiased estimator of θ (4 Marks)

(ii) Using the factorization criterion, show that $\overline{X}$ is a sufficient statistic estimator for θ (3 Marks)

(c) Claims department of an insurance firm audited 150 customers' payment claims. Out of these, 12% were found to be defective

(i) estimate the standard error of the proportion of defective parts (2 Marks)

(ii) Construct the 99% confidence interval for the proportion of defective claims (3 Mark)

(iii) Interpret meaning 99% confidence interval in part b (ii) above (1 Mark)

(d) Several life insurance firms have policies geared to working class citizens. To get more information about this group, a major insurance firm interviewed working class individuals to find out the type of life insurance they preferred, if any. The table below shows the preferences by gender.

Gender	Preferred a term policy	Preferred a whole-life policy	No preference
Adult Male	160	30	10
Adult Female	140	120	40

Is there evidence that the life insurance preference of working class citizens depends on their sex? Use $\alpha = 0.05$ (5 Marks)

(f) Let $X_1, X_2, X_3, ..., X_n$ be a random sample from a uniform distribution from $(0, \quad \theta)$

(i) Find the moment estimator for θ (5 Marks)

(ii) Given the following observations, 10.6, 12.3, 4.8, 11.0, 12.37, determine the moment estimator for θ (2 Marks)

Question 2 (20 Marks)

(a) The number of students who were absent during Monday morning lesson has a Poisson distribution with mean λ (> 0) and is independent for different classes. A researcher wants to estimate λ and examines the number of absent students $x_i > 0$ for i = 1, 2, …, n.)

(i) Find the Likelihood estimators of λ based on $x_1, x_2, ..., x_n$ (7 Marks)

(ii) Show that the estimator for λ is consistent (3 Mark

(b) A financial analyst of a supermarket has found out that daily total sales (X) is approximately normally distributed with an unknown value of mean, μ and a standard deviation of 1 million dollars. If the corresponding distribution is defined by

$$f(x) = \begin{cases} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-\mu)^2}, & -\infty \le x \le \infty \\ 0, & elsewhere \end{cases} \quad , \text{ where } \pi = \tfrac{22}{7}.$$

(i) Determine the likelihood function for the above financial model (3 Marks)

(ii) Show that the maximum likelihood estimator for the population mean μ is the sample mean $\overline{X}$ (4 Marks)

(iii) If data gathered on daily total sales were obtained as;

1.2, 2.0, 2.5, 3.1, 1.0, 2.1, 3.5, 5.1, 4.0, 2.8.

Find the estimate of the mean from the sampled data. (2Marks)

Question 3 (20 Marks)

(a) consider a linear regression model defined by $Y = \alpha + \beta x + e$, where $E(e) = 0$. Suppose n independent observations $y_1, y_2, y_3, ..., y_n$ are made show that

$$\hat{\beta} = \frac{n\sum x_i y_i \sum y_i^{\,2} - \sum x_i y_i}{n\sum x_i^{\,2} - \left(\sum x_i\right)^2}$$ is unbiased for β (5 Marks)

(b) Use the matrix notation to fit a straight line to n=4 data points given below by the method of least square

X	-2	0	-1	2
Y	0	1	2	2

(5 Marks)

(c) An insurance company has two products to launch. The finance department has four estimates to compute each being independent of one another with zero mean error and constant variance σ^2. The department computes cost for launching the two products separately, C_1 and C_2, and also the sum and the difference of these two costs.
Required: Determine the least square estimators of the unknown costs C_1 and C_2. (10 Marks)

Question 4 (20 Marks)

(a) Suppose $x_1, x_2, x_3, ..., x_m$ is a random sample from a normal population with mean μ_1 and known variance σ^2. Also let $y_1, y_2, y_3, ..., y_n$ be another random sample from a normal population with mean μ_2 and known variance σ^2. Assume that these samples are independent of each other. Construct a confidence interval for the difference of the means μ_1 and μ_2.

(10 Marks)

(b) Periodically, bank customers are asked to evaluate financial consultants and services. Higher ratings on the client satisfaction survey indicate better service, with 7 the maximum service rating. Independent samples of service ratings for two financial consultants are summarized below

Statistic	Consultant A	Consultant B
Sample size	16	10
Sample mean	6.82	6.25
Sample variance	0.64	0.64

(i) Determine the 95% confidence interval for the difference between the two population means (5 Marks)

(ii) Use 0.05 level of significance to determine whether consultant A has a higher population mean service rating. (5 Marks)

Question 5 (20 marks)

(a) Let $x_1, x_2, ..., x_n$ denote a random sample from a normal distribution with mean μ and variance δ^2. Two unbiased estimators of δ^2 are given as:

$\hat{\sigma}_1{}^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2$ and $\hat{\sigma}_2{}^2 = \frac{1}{2}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2$. Determine the efficiency of $\hat{\sigma}_1{}^2$ relative to $\hat{\sigma}_2{}^2$.

(5 Marks)

(b) Differentiate between type I and type II errors in hypotheses tests (2 Marks)

(c) The amount of time, X in minutes a bank teller spends with a customer is known to have an exponential distribution with an average amount of time of 4 minutes. The density function is defined below:

$$f(x) = \begin{cases} \frac{1}{4}e^{-\frac{1}{4}x}, x > 0 \\ 0, elsewhere \end{cases}$$

Find the probability that a clerk spends four to five minutes with a randomly selected customer

(5 Marks)

(d) A production line operates with a mean filling weight of 3kg per container. Over-filling or under-filling presents a serious problem and when detected requires the operator to shut down the production line to re-adjust the filling mechanism. The summary statistics from weighing 36 such containers are given below:

Descriptive Statistics: Volume of filling

Variable	Count	Mean	Total SE Mean	StDev	Minimum	Maximum
Volume	36	2.9200	0.0283	0.1700	2.5400	3.2100

(i) Formulate the null and alternative hypothesis (2 Mark)

(ii) Using the critical value approach, should the filling process be stopped? Use $\alpha = 0.05$ (6 Marks)