**University of Kelaniya – Sri Lanka**
**Centre for Distance & Continuing Education**
**Bachelor of Science (General) External**
**Second year second semester examination - 2019 (2024 February)**
**(New Syllabus)**
**Faculty of Science**

**Statistics**
**STAT 27533 – Inferential Statistics**

No. of Questions: Five (05) No. of Pages: Three(03) Time: **Two & half (2 1/2) Hours.**

Answer **Four (04)** questions only.

---

1. *(a)* Briefly describe the following methods used in point estimation:
   *(i)* Method of Moments,
   *(ii)* Method of Maximum Likelihood.

   *(b)* Let $X_1, X_2, \dots, X_n$ represent a random sample taken from the distribution with probability density function

$$f(x;\theta) = \frac{1}{\theta} x^{(1-\theta)/\theta}; \quad 0 < x < 1, \qquad 0 < \theta < \infty.$$

   *(i).* Show that the mean of the above distribution is $\frac{1}{\theta+1}$.
   *(ii).* Find the method of moment estimator for the parameter $\theta$.
   *(iii).* Show that the maximum likelihood estimator of $\theta$ is $\hat{\theta} = \frac{-1}{n}\sum_{i=1}^{n} \ln X_i$.

2. *(a)* Define the following terms used in estimation theory:
   (i) An Unbiased Estimator,
   (ii) A Sufficient statistic.

   *(b)* (i) State the Factorization theorem.
   (ii) Let $X_1, \dots, X_n$ be a random sample of size $n$ from a distribution with the probability density function

$$f(x;\theta) = \frac{1}{2}\theta^3 x^2 e^{\theta x}; \quad 0 < x < 1.$$

   Find a sufficient statistic for $\theta$ using the factorization theorem.

(c) Let $X_1, \ldots, X_n$ be a random sample from the distribution with probability mass function $f(x;p) = p(1-p)^{x-1}\,; x = 0,1,2,\cdots \; where \; 0 < p \leq 1.$

(i) Show that $f(x;p)$ belongs to the exponential family.
(ii) Hence find the sufficient statistic for $p$.
(iii) Find an unbiased estimator for $1/p$ using the sufficient statistic obtained in part (ii).

3. (a) Let $X_1, \ldots, X_n$ be iid Bernoulli $(p)$. We are interested in estimating, $p$ (probability of success). Let $T(\mathrm{X}) = X_1$.

*(i)* Show that $T(\mathrm{X})$ is an unbiased estimator for estimating $p$.
*(ii)* Show that Bernoulli distribution belongs to the exponential family. Hence find a sufficient statistic for $p$.
*(iii)* Use Rao Blackwell theorem to find a better unbiased estimator than $T(\mathbf{X})$ for estimating $p$.
*(iv)* Is the answer found in *(iii)* a uniformly minimum variance unbiased estimator (UMVUE)? Explain.

(b) Let $X_1, X_2, \ldots, X_n$ be a random sample taken from the distribution,

$$f(x;\theta) = \frac{1}{\theta}e^{-\frac{x}{\theta}};\; 0 < x < \infty, 0 < \theta < \infty.$$

*(i)* Show that $\bar{X}$ is an unbiased estimator for $\theta$.
*(ii)* Find the Cramer-Rao Lower bound for $\theta$.
*(iii)* Is $\bar{X}$ the uniformly minimum variance unbiased estimator for$\theta$? Explain your answer.

4. (a) If the population standard deviation, $\sigma$ is known, Construct a $100\,(\alpha)$ % confidence interval for the population mean, $\mu$ of all samples of size $n$ using the pivotal quantity method. State the assumptions.

(b) A survey was conducted by the World Wild Fund (WWF) organization in order to measure the average daily food intake by Brown bears. A random sample of 36 observations gave an average of 82.6 pounds. Construct and interpret a 90% confidence interval for the average daily food consumption, $\mu$, of all Brown bears. Assume that the population standard deviation is 12.6 pounds.

(c) Having a consistent tensile strength is a key fact that affects the quality of synthetic fiber. A sample of 10 pieces of fiber is taken and the tensile strength (in kg) of each fiber is tested. Assume that the sample is taken randomly from a normally distributed population with mean $\mu$ and variance$\sigma^2$. Following sample statistics were obtained, $\bar{x} = 46.32kg\,,\; s^2 = 9.48kg$. Construct a 95% confidence interval for the population variance.

5. *(a)* Briefly describe, Type I error and Type II error as used in hypothesis testing.

*(b) (i)* Let $X_1, \dots, X_n$ be a random sample from the population having the distribution $\aleph(\mu_X, \sigma^2)$ and $Y_1, \dots, Y_m$ be a random sample from the population having the distribution $\aleph(\mu_Y, \sigma^2)$. Write the test statistics to test the null hypothesis $H_0: \mu_X - \mu_Y = 0$ against the alternative hypothesis $H_0: \mu_X - \mu_Y < 0$.

*(ii)* A botanist is interested in comparing the growth response of dwarf pea stems against two different levels of the hormone indoleacetic acid (IAA). Using 16-day-old pea plants, the botanist obtains 5-mm sections and floats these sections on solutions with different hormone concentrations to observe the effect of the hormone on the growth of the pea stem. Let $X$ and $Y$ denote, respectively, the independent growths that can be attributed to the hormone during the first 26 hours after sectioning for $(0.5)\,(10)^{-4}$ and $10^{-4}$ levels of concentration of IAA. She measured the growths of pea stem segments, in millimeters, for $n = 11$ observations of $X$:

0.8 1.8 1.0 0.1 0.9 1.7 1.0 1.4 0.9 1.2 0.5.

She did the same with $m = 13$ observations of $Y$:

1.0 0.8 1.6 2.6 1.3 1.1 2.4 1.8 2.5 1.4 1.9
2.0 1.2.

Determine the null and alternative hypotheses.
Test the hypotheses at 5% significance level using the critical value approach and interpret the results.

*(c)* The number of typing errors in a newspaper can be modeled by a Poisson distribution with mean $\mu$. Each of the 31 students in a journalism course are allocated at random a past edition of the newspaper and asked to find the number of typing errors it contains. It has been hypothesized by the instructor of the course that there is an average of 5 typing errors in each edition of the newspaper. The editor of the newspaper has objected, claiming that the mean number of errors is less than 5 per edition.

Using the Neyman-Pearson lemma, find the most powerful test with significance level $\alpha$ to test the hypotheses, $H_0 : \mu = 5$ versus $H_1 : \mu = 3$.