UNIVERSITY OF RUHUNA

Faculty of Engineering

End-Semester 4 Examination in Engineering: December 2018

Module Number: IS4301 **Module Name: Probability and Statistics**

[Three Hours]

[Answer **all questions**, each question carries fourteen marks]

Q1. a) Briefly explain the followings.

i Descriptive Statistics

ii Inferential Statistics

[2.0 Marks]

b) Classify the following variables as either categorical or numerical.

i Length (in hours) of baseball games

ii Colors of paint in a paint company's inventory

iii Ranks of personnel in the military

iv Age of students entering a college

[2.0 Marks]

c) A tire manufacturer wants to determine the inner diameter of a certain grade of tire. Ideally, the diameter would be 570 mm. The data are as follows.

$$572, 572, 573, 568, 569, 575, 565, 570$$

i Find the sample mean and the median.

ii Find the sample variance, the standard deviation, and the range.

iii Using the calculated statistics in parts i) and ii), comment on the quality of the tires.

[6.0 Marks]

d) An experiment is proposed to test the three types of antimissile systems. From the design point of view, each of these systems has an equally likely chance of detecting and destroying an incoming missile within a range of 250 miles with a speed ranging up to nine times the speed of sound. However, in actual practice it has been observed that the precisions of these antimissile systems are not the same; that is, the first system will usually detect and destroy the target 10 of 12 times, the second will detect and destroy it 9 of 12 times, and the third will detect and destroy it 8 of 12 times. It is observed that a target has been detected and destroyed. What is the probability that the antimissile defense system was of the third type?

[4.0 Marks]

Q2. a) Briefly explain the followings.
  i Random variable
  ii Probability distribution

[2.0 Marks]

b) The length of time that an individual talks on a long-distance telephone call has been found to be of a random nature. Let X be the length of the talk; assume it to be a continuous random variable with probability density function given by

$$f(x) = \begin{cases} \alpha e^{-(1/5)x} & ; \quad x > 0 \\ 0 & ; \quad elsewhere \end{cases}$$

  i Find the value of α that makes $f(x)$ a probability density function.
  ii Find the probability that this individual will talk
    a) between 8 to 12 minutes.
    b) less than 8 minutes.
    c) more than 12 minutes.

[5.0 Marks]

c) The moment-generating function of the random variable X is given by,

$$M_X(t) = E(e^{tX})$$

Then

$$\frac{d^r M_X(t)}{dt^r}\Big|_{t=0} = \mu_r' \quad ; \quad \text{where} \quad \mu_r' = E(X^r), \; r = 1, 2, 3, \ldots$$

Let X be a standard normal random variable has a probability density function

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \; , \quad -\infty < x < \infty$$

  i Find the moment-generating function of X.
  ii Find the mean and the variance of the random variable X.
  iii Given that $M_Y(t) = e^{bt} M_X(at)$ where $Y = aX + b$. Hence, by using the answer in part i), find the moment-generating function of a normal random variable Y.

[7.0 Marks]

Q3. a) Let $\hat{\theta}_1$ and $\hat{\theta}_2$ be two unbiased estimators of θ. Let $\hat{\theta}_3$ is a convex combination of $\hat{\theta}_1$ and $\hat{\theta}_2$ such that

$$\hat{\theta}_3 = a\hat{\theta}_1 + (1 - a)\hat{\theta}_2 \; ; \; 0 \le a \le 1$$

  i Show that $\hat{\theta}_3$ is an unbiased estimator of θ.
  ii If $\hat{\theta}_1$ and $\hat{\theta}_2$ are independent, and variance of $\hat{\theta}_1$ and $\hat{\theta}_2$ are σ_1^2 and σ_2^2 respectively, then find the variance of $\hat{\theta}_3$.
  iii How should the constant a be chosen in order to minimize the variance of $\hat{\theta}_3$?

[5.0 Marks]

b) A chemist has two different methods for measuring the concentration level C of a chemical solution. Methods A and B produce measurements X_A and X_B respectively and those are distributed as follows.

$$X_A \sim N(C, 2.97) \quad , \quad X_B \sim N(C, 1.62)$$

i Find 99.7% confidence intervals for the measurements X_A and X_B separately. Hence explain that how does the Chemist select more accurate method for measuring the concentration level C?

ii With the knowledge of estimation theory, explain how the Chemist arrives at an optimum point estimate of the concentration level C?

[4.0 Marks]

c) Consider a set of independent data observations $x_1, x_2, \dots, x_n$ that have a gamma distribution with $k = 5$ and unknown parameter $\lambda > 0$ has a probability density function

$$f(x) = \begin{cases} \dfrac{\lambda^k x^{k-1} e^{-\lambda x}}{\Gamma(k)} & ; \ x \geq 0 \\ 0 & ; \ x < 0 \end{cases}$$

where $\Gamma(k) = (k-1)\Gamma(k-1)$, $\Gamma(1) = 1$ and $E(X) = \frac{k}{\lambda}$.

i Use Maximum Likelihood Estimation Method to find estimators for λ.

ii Show that the estimator of λ is an unbiased estimator.

[5.0 Marks]

Q4. a) An experimenter is interested in the hypothesis testing problem

$$H_0 : \mu = 3.0\ mm \qquad \text{versus} \qquad H_1 : \mu \neq 3.0\ mm$$

where μ is the average thickness of a set of glass sheets. Suppose that a sample of $n = 21$ glass sheets is obtained and their thicknesses are measured. Suppose that the sample mean is $3.04\ mm$ and the sample standard deviation is $0.124\ mm$.
(Assume that the thickness of a glass sheet has a normal distribution)

i What is the critical region that the experimenter accept the null hypothesis with a size $\alpha = 0.10$?

ii What is the critical region that the experimenter reject the null hypothesis with a size $\alpha = 0.01$?

iii Determine whether the null hypothesis accepted with $\alpha = 0.10$ and $\alpha = 0.01$.

iv Write down an expression for the p value to make the decision at $\alpha = 0.10$ level of significance.

[7.0 Marks]

b) In a test of the ability of a certain polymer to remove toxic wastes from water, experiments were conducted at three different temperatures. The data in the following table give the percentages of the impurities that were removed by the polymer in 21 independent attempts. Test the hypothesis that the polymer performs equally well at all three temperatures at the

i 5 percent level of significance

ii 1 percent level of significance.

Low Temperature	Medium Temperature	High Temperature
42	36	33
41	35	44
37	32	40
29	38	36
35	39	44
40	42	37
32	34	45

[7.0 Marks]

Q5. a) Infrared spectroscopy is often used to determine the natural rubber content of mixtures of natural and synthetic rubber. For mixtures of known percentages, the infrared spectroscopy gave the following readings.

Percentage	0	20	40	60	80	100
Reading	0.734	0.885	1.050	1.191	1.314	1.432

i Plot a scatter diagram to see if a linear relationship is indicated.

ii Find the least squares estimates of the regression coefficients.

iii If a new mixture gives an infrared spectroscopy reading of 1.15, estimate its percentage of natural rubber.

[8.0 Marks]

b) The percent survival of a certain type of animal semen, after storage, was measured at various combinations of concentrations of three materials used to increase chance of survival. The summarized data in the usual matrix notation are given by the least squares estimating equations, $(X'X)b = X'y$ and the inverse matrix, $(X'X)^{-1}$ as follows.

$$\begin{bmatrix} 13 & 59.43 & 81.82 & 115.40 \\ 59.43 & 394.7255 & 360.6621 & 522.0780 \\ 81.82 & 360.6621 & 576.7264 & 728.3100 \\ 115.40 & 522.0780 & 728.3100 & 1035.9600 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} 377.5 \\ 1877.567 \\ 2246.661 \\ 3337.780 \end{bmatrix}$$

$$(X'X)^{-1} = \begin{bmatrix} 8.0648 & -0.0826 & -0.0942 & -0.7905 \\ -0.0826 & 0.0085 & 0.0017 & 0.0037 \\ -0.0942 & 0.0017 & 0.0166 & -0.0021 \\ -0.7905 & 0.0037 & -0.0021 & 0.0886 \end{bmatrix}$$

By using the above relations, estimate the multiple linear regression equation.

[6.0 Marks]