

THE OPEN UNIVERSITY OF SRI LANKA

B.Sc. /B.Ed. Degree Programme, Continuing Education Programme

APPLIED MATHEMATICS-LEVEL 05

ADU5318/PCU1142/ /PCU3141/PSU1182 – Bio Statistics

FINAL EXAMINATION 2017/2018

Duration: Two Hours.



Date: 03.04.2019

Time: 1.30p.m- 3.30pm

Instructions:

- **This question paper consists of 06 questions. Answer only four questions.**
- **Statistical Tables are provided. When reading values, you may use the closest degrees of freedom given in the table.**
- **In all tests, use the significance level as 0.05.**
- **If the random number table is to be used, you are required to clearly indicate how to read the values and as an illustration present three values that you read.**
- **Non-programmable calculators are permitted.**

1. A researcher is interested in estimating the number of persons residing in an area, diagnosed with thyroid malfunction problems. Out of the 3200 persons in the area, 1800 are females. Of the 1800 females, 500 are less than 15 years of age, 600 are in the age group of 15 to 45 and the rest are older. Among the males, 400 are less than 15 years of age, 550 are in the age group of 15 to 45 years and the rest are older. The researcher suspects that thyroid malfunction problems are more common among females and in both gender groups to differ across the three age groups. The researcher has resources to sample 800 persons for data collection.

Suppose the researcher seeks your advice to design this study.

- i) Clearly describe how you advise. If you use the random number table, clearly state how you read the values.
- ii) Explain the following terms in relation to this study:
 - a) sampling unit
 - b) population
 - c) confounding

2. A researcher is interested in estimating the total number of trees in a banana cultivation, infected with a disease. The following table summarises the number of trees examined and the number of infected trees found on each of 10 days in which data were collected.

Day number	Number of trees examined	Number of infected trees found
1	50	5
2	45	6
3	50	7
4	65	10
5	60	10
6	70	15
7	75	20
8	45	7
9	50	12
10	50	8

- i) Estimate the percentage of trees in the cultivation that are infected with the disease.
- ii) Suppose in the first four days, the researcher had sampled from the north side of the cultivation and on the other six days, the researcher had sampled from the south side. Construct a suitable graphical summary that can be used to highlight the differences in the two sides with respect to the extent in which trees are infected.
- iii) Clearly state the findings from the graphical summary constructed in part (ii).
- iv) Giving reasons for your answer, state whether you agree with the following statement made by a student on the comparison of the two sides. *The fact that the researcher had sampled more from the south side, will add a bias for any comparison of the two sides with respect to the spread of the disease.*

3. A manufacturer claims that the jam content in bottles labeled 100ml produced by his company exceeds the labeled volume. To test the validity of this claim, jam content of 25 bottles were measured. The sample mean and standard deviation were 101 ml and 4 ml.

- i) Explain the terms hypothesis and *an assumption*, using this study as an example.
- ii) Clearly describing your notation, write down the null and the alternative hypotheses you would examine to test the validity of the manufacturer's claim.
- iii) Using an appropriate statistical test, examine the validity of the manufacturer's claim using a 5% significance level and clearly state the findings in relation to this study.
- iv) Clearly state the assumptions needed for the test carried out in part (iii).

4. A researcher is interested in finding out the amount and the type of fertilizer that can give the highest expected yield of bean plants. He expects to study this with two varieties of bean plants (V_1, V_2), using two fertilizers (F_1, F_2) and study the effects under two fertilizer amounts (0mg, 2mg) per plot per month. Yield of each plot is to be measured at the end of three months of applying the fertilizer consecutively for three months. Similar 40 experimental plots and similar 100 bean plants of each variety are available for this study. The researcher plans to plant 5 plants in each plot. Plants in each plot has to be of the same variety and has to be kept under the same experimental condition.

- i) Propose a suitable design for this study.
- ii) State whether the study is an exploratory study or a confirmatory study. Give reasons for your answer.
- iii) Explain the following terms in relation to this study:
 - a) Interaction
 - b) Experimental unit

5. A large lecture hall with multimedia facility has 300 chairs arranged in 30 rows with 10 chairs in each row. Suppose 300 students attended a lecture conducted in this lecture hall using multimedia. In order to assess the satisfaction of students about the lecture, a researcher wants to randomly sample ten rows out of the 30 rows and to get the views of all the students in the selected rows.

- i) If the researcher seeks your advice on how to select students using the random number table given to you, clearly explain how you advise.
- ii) Suppose the researcher wants to record the data on the following variables.

V_1 : gender of the student coded as 1: male; 2: female

V_2 : age of the student as of 01st of January 2019 recorded to the nearest year coded as

1: less than 25; 2: 25 or more but less than 30; 3: 30 or more

V_3 : number of years since first registration for the degree

V_4 : level of overall satisfaction on the lecture coded as

1: extremely satisfied; 2: somewhat satisfied; 3: not much satisfied and 4: not satisfied at all

V_5 : Clarity of the sound coded as 1: clear; 2: not clear

- a) Classify the variables as qualitative or quantitative.
 - b) Classify the quantitative variables as discrete or continuous.
 - c) Classify the variables as nominal, ordinal, interval or ratio.
- iii) State whether each of the following statements is true or false in relation to this study. In each case, give reasons for your answer.

- a) The population in this study is a finite population.
- b) The researcher has done stratified sampling.
- c) If the researcher had increased the sample size, the sampling error will be reduced.
- d) If the researcher had sampled all the students, there will not be non-sampling error.

6. The following table summarises the yield of paddy per acre, measured to the nearest bushel, of a group of farmers selected for a research study.

Yield (bushels per acre)	Number of farmers
15 – 19	5
20 – 24	16
25 – 29	48
30 – 34	36
35 – 39	5

- i) Calculate the relative cumulative frequency corresponding to the third class interval and explain what it measures in relation to this study.
- ii) Construct a suitable graph that can be used to find the percentiles of the data.
- iii) Using the graph constructed in part (ii),
 - a) find the 10th percentile and explain what it measures in relation to this study.
 - b) Calculate the first quartile of the data and describe what it measures in relation to this study.
- iv) Calculate the inter-quartile range of the data.
- v) For this data, describe a weakness in using range as a measure of dispersion as opposed to the inter-quartile range.

***** Copyrights reserved *****