

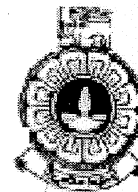
THE OPEN UNIVERSITY OF SRI LANKA

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

APPLIED MATHEMATICS-LEVEL 05

PCU3141/PCE5141/PCU1142/PCE3142/PSZ3182/PSZ4130-BIO STATISTICS

PSU1182/PSE3182
FINAL EXAMINATION 2014/2015



Duration: Two Hours.

Date: 24.10.2015

Time: 1.30p.m- 3.30p.m

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.
- Non-programmable calculators are permitted.

1. The following cumulative frequency table is constructed from the dried weights (mg) of medicinal plants measured in an experiment.

Weight range (mg)	Number of plants
10 – 15	8
16 – 21	47
22 – 27	66
28 – 33	89

The researcher is interested in finding out whether the distribution of the data has multiple peaks.

- i) What is the sample size used for this study?
- ii) Estimate the percentage of plants that have weights exceeding 25mg.
- iii) Construct a suitable graphical summary that can be used to address the researcher's objectives.
- iv) Clearly describe the findings from the graphical summary constructed in part (iii).

2. The following frequency table is constructed from the data collected from 80 persons on the age at first diagnosis of high blood sugar, measured to the nearest year.

Age at diagnosis	Number of persons
15 – 24	14
25 – 34	27
35 – 44	18
45 – 54	19
55 – 64	2

- i) Calculate the sample mean.
 - ii) Calculate the median of the data and describe what it measures in relation to this study.
 - iii) Based on the quantities computed in part (i) and part (ii), what can you conclude about the shape of the distribution of the data? Give reasons for your answer.
 - iv) Find the cumulative relative frequency corresponding to the third class interval and describe what it measures in relation to this study.
 - v) Calculate the first quartile of the data and describe what it measures in relation to this study.
3. A stock of 15000 bean seeds comprises 10000 seeds from Variety1 and 5000 seeds from Variety2. The seeds of the two varieties are packed separately in packets of 50 seeds each. The germination rates of the two varieties and how they respond to soil temperature and prior length of soaking may be different. A researcher wishes to find out the variety and the most appropriate condition to receive the highest yield. Resources are adequate to study the germination in 1000 seeds only.
- i) Clearly describe how you would advise the researcher to design this study. If you use the random number table, clearly describe how you read the values.
 - ii) State whether the study described here is observational or experimental. Give reasons for your answer.
 - iii) In relation to this study, describe the following:
 - a) Interaction
 - b) Replicate

4. From a production line, nails come out in packets of 25 each. The daily production is around 50000 nails. A quality controller plans to examine whether the production process is in control by daily inspecting 1000 nails selected from the output on each day. If the number of defectives found is not exceeding ten, the production process is taken as in control.
- i) If the quality controller seeks your advice, clearly describe how you advise to do the sampling. If you use the random number table, clearly describe how you read the values.
 - ii) A student states that if the quality controller examines all the nails in the daily production, non-sampling errors cannot happen. Do you agree with this statement? Give reasons for your answer.
 - iii) In relation to this study, clearly explain the following terms:
 - a) Sampling
 - b) Sampling unit
 - c) Population

5. The following summary statistics were computed from the heights of (cm) of 17 plants at the age of one month.

$$\sum x_i = 200; \sum x_i^2 = 2972$$

The researcher is interested in finding out whether the expected height of a randomly chosen plant is 10cm. You may assume that the heights of the plants follow a normal distribution with unknown mean and variance.

- i) Calculate the standard deviation of the data.
- ii) State the null and the alternative hypotheses you would test to address the researcher's objective.
- iii) Using the level of significance as 0.05, test the hypothesis stated in part (ii) and clearly state the findings in relation to this study.

6. In a study on the association between the carbohydrate content in the diet and the blood sugar level of females, the following variables were measured on each of 100 participants in the age group of 20 to 50 years. Three special diets were examined in the study.

V_1 : diet coded according to the carbohydrate contents as

1: low content; 2: moderate content; 3: high content

V_2 : age (to the nearest year) recorded as 1,2,3 according to

1: 20 to 29, 2: 30 to 39, 3: 40 to 50.

V_3 : engagement in exercise recorded as

1: do not exercise; 2: occasionally exercise; 3: regularly exercise

V_4 : weight measured to the nearest kilogram

V_5 : blood sugar level (mg/ dL)

- i) Classify each variable as discrete or continuous.
- ii) Classify the data on each variable as nominal, ordinal, interval or ratio.
- iii) State whether a bar chart is suitable to examine whether the data collected on age (V_2) are symmetrically distributed or not. Give reasons for your answer.
- iv) Clearly describe a suitable graph that can be used to highlight whether the association between blood sugar level (V_5) and diet (V_1) depends on the age (V_2) or not.

***** Copyrights reserved. *****