THE OPEN UNIVERSITY OF SRI LANKA

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

APPLIED MATHEMATICS-LEVEL 05

PCU3141/PCE5141/PCU1142/PCE3142/PSU1182/PSE3182 – Bio Statistics

FINAL EXAMINATION 2015/2016

Duration: Two Hours.

Date: 11.01.2017	**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.**
- **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 present three values that you read as an example.**
- **Non-programmable calculators are permitted.**

1. A researcher is interested in finding out the spinning speed of the machine and the type of raw material that produces yarn with maximum strength. Resources are available to measure the strength of 420 samples of yarn produced at pre-determined spinning speeds using specified raw material type. Two spinning speeds labeled $S1$ and $S2$ and three raw material types $M1$, $M2$ and $M3$ are to be examined. The strength of yarn produced may depend on the operator of the machine. The data are to be collected under all three machine operators in the company.

   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

2. The manufacturer of a newly improved water purifying equipment claims that the expected lifetime of a machine exceeds two years. The following summary statistics were computed from 25 lifetimes measured to the nearest year.

Sample mean = 2.1, sample standard deviation = 0.4

i) Write down the null and the alternative hypotheses you would examine to test the validity of the manufacturer's claim.

ii) Using an appropriate statistical test, examine the validity of the manufacturer's claim and clearly state the findings.

iii) Cleary state the assumptions involved in the test you used in part (ii).

3. 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 sample mean and explain what it measures in relation to this study.

ii) Calculate the sample median.

iii) Based on the measures calculated in parts (i) and (ii), what can you say about the distribution of the data?

iv) Construct a suitable graph that can be used to find the percentiles of the data.

v) Using the graph constructed in part (v), find the 10^{th} percentile and explain what it measures in relation to this study.

4. In a study to investigate the effect of heat treatment on anti-termite activity, percentage mass loss due to termite attacks were calculated on each of 40 wood blocks after 10 days of exposure to one of the two heat treatments as described below. The following table summarises the data collected in the study. The researcher is interested in finding out whether the population means of percentage mass losses under the two heat treatments are equal or not.

Description	Heat Treatment Applied	
	Treatment1	Treatment2
Temperature	60 °C	135 °C
Sample size	20	20
Sample average of mass loss percentage	2.1	12.3
Standard deviation	7.9	8.4

Assume that the population variances of percentage mass loss under the two heat treatments are equal and the sample mean of each of the two populations follow a normal distribution with unknown means and standard deviations.

i) State whether the design used in this study is a paired sample design or an independent sample design. Give reasons for your answer.

ii) Write down the null and the alternative hypotheses you would test to address the researcher's objective.

iii) Write down the test statistic you would use to test the hypothesis stated in part (iii).

iv) Test the hypothesis and clearly state the findings.

5. In a study to estimate the number of cattle in a farm infected with a disease, a researcher wants to select 30 cattle from the 550 cattle in the farm for observation. Suppose the researcher seeks your advice on how to do the sampling.

i) Clearly describe how you advise the researcher and the assumptions you make.

ii) In relation to this study explain the following terms:

a) Sampling error

b) Bias

c) Sampling variation

iii) If a sample of size 30 contains three infected cattle, estimate the total number of infected cattle in the farm.

6. The data collected on visits to two sites are presented below, using a code in which the letter indicates the bird species and the number indicates the number of birds observed of that species. For instance, the code *A*10 indicates that 10 birds are observed of species *A*. If no birds of a species are observed on a visit, the relevant letter is not included for that visit. The habitat types of the two sites are different.

Site visited	Number of days	Data collected on each visit	Type of weather
Site1	2	Day1:A10, B12,C31,D47,E15,F1,G62,H19	cold
		Day2: A13,B18,C32,D43,E16,G67,H14	cold
Site2	3	Day1:A4,B16,C21,D44,E11,G21,H9	sunny
		Day2:A3,B15,C20,D42,E10,G22,H9	sunny
		Day3:A5,B15,C22,D53,E11,G21,H10	sunny

State whether each of the following statements is true or false based on this study. In each case, give reasons.

a) Site2 has less number of birds compared to Site1.

b) Site1 has a rare species.

c) Species G is most observable in the cold weather.

d) The researcher has collected only nominal data.

e) Mode cannot be used as a measure of location to describe the data collected in Site1 under cold weather.

f) Histogram is an appropriate graph for summarizing the data collected in Site1 under cold weather.