

THE OPEN UNIVERSITY OF SRI LANKA
B.Sc./ B.Ed. DEGREE, CONTINUING EDUCATION PROGRAMME
FINAL EXAMINATION 2006/2007
PSU 1182/ PSE 3182/PSZ 4182 – BIO STATISTICS
DURATION: TWO HOURS



DATE: 01-12 - 2006

TIME: 9.30am – 11.30am

ANSWER FOUR QUESTIONS ONLY.

Non-programmable calculators are permitted. Statistical tables are provided.

1. A researcher is interested in finding out the effects of exercise and diet on reducing the cholesterol levels of persons with inflated cholesterol levels. For the study, the researcher has selected three conditions, exercise daily, exercise every other day and exercise twice a week. Also, the researcher is interested in two diet menus say *A* and *B*. The researcher plans to collect information from 200 volunteers who are willing to take the diets assigned to them and to engage in exercise as proposed by the researcher. Out of the 200 volunteers, 90 are with cholesterol levels just above normal, 80 are with moderately high levels and 30 are with extremely high levels. In the first group of 90, 20 are below 30 years and the rest are above 30 years. Of the second group of 80, 40 are below 30 years and the rest are above 30 years. Of the third group of 30, 18 are below 30 years and 12 are above 30 years.
 - i) Suggest a suitable design for this study. If you use the random number table, clearly describe how you read the values.
 - ii) In relation to this study, explain the following terms.
 - a) interaction
 - b) random variation
 - c) replication
2. State whether the following statements are true or false. If the statement is false, describe why it is false.
 - a) All ordinal data are qualitative.
 - b) Bar charts are useful in summarising information about the existence of multiple modes in interval data.
 - c) Third quartile is useful in conveying information about the presence of extremely large values in data sets.
 - d) Frequency polygons are useful in summarising information in nominal data.
 - e) Mean absolute deviance is a measure of random variation in the data.

3. i) Briefly explain the following terms using an example for each case.
- Non sampling error
 - Bias
- ii) The Department of Agriculture is interested in investigating the pesticide use of farmers growing paddy in three areas of Sri Lanka. A previous survey has indicated that there are 70, 120 and 60 farmers in the three areas. Suppose the three areas are very similar with respect to the pesticide use of farmers. However, pesticide use of one farmer may be different from that of another farmer. The researcher has resources to collect information from 40 farmers if all farmers are selected from the same area. However, due to travelling costs if farmers are selected from different areas he is only able to collect information from 20 farmers. Since areas are very similar with respect to the pesticide use, the researcher has decided to randomly select 40 farmers from the area that is closest to his residence.
- Since farmers are randomly selected, a student claims that the researcher has collected a simple random sample of farmers. Do you agree with the claim of the student? Give reasons for your answer.
 - If you were to advice on sample selection, would you recommend collecting a simple random sample, a stratified random sample or a cluster sample? Give reasons for your answer.

4. Water with dissolved Oxygen level below 5mg per litre can be harmful to living organisms. A researcher, in an experiment to find out whether the water quality in a particular area is harmful, measured the dissolved Oxygen levels in 30 water samples and the data are given below.

5.1	4.9	5.2	5.3	5.2	4.7	4.7	4.8	4.8	4.5
4.9	4.9	4.9	5.1	5.3	5.3	4.7	5.1	5.0	5.0
5.1	4.4	4.4	4.2	4.3	5.1	4.7	4.7	4.7	4.8

- Clearly describe the null and the alternative hypotheses you would test to meet the researcher's objective.
- Suggest a suitable test statistic that you would use to test the null hypothesis stated in part (i).
- Test the hypothesis stated in part (i) using a 10% significance level. Clearly state your findings.
- In relation to this study, explain the following terms.
 - critical region
 - significance level

5. The following table summarises the ages of 100 randomly selected persons employed in the Colombo district when they were first diagnosed with diabetes.

Age limit	Number of persons
20 - 29	5
30 - 39	44
40 - 49	33
50 - 59	18

- i) Compute the first quartile and describe what it measures.
 - ii) Compute the sample mean and describe what it measures in relation to this study.
 - iii) Estimate the cumulative frequency corresponding to the age of 45 years and describe what it measures.
 - iv) Estimate the percentage of persons who were at least 45 years of age when they were first diagnosed with diabetes.
6. A researcher is interested in finding out whether the type of noise in the environment has any significant effect on the time to react in a danger. Ten persons were randomly selected for the study. The time to react in a quite environment and in a noisy environment was measured with each of these ten persons. The results are presented in the accompanying table.

Person number	identification	Time to react (in seconds) in the	
		quite environment	noisy environment
	1	10	12
	2	8	11
	3	14	16
	4	11	15
	5	10	11
	6	12	14
	7	15	17
	8	11	13
	9	11	14
	10	12	14

- i) Classify the design as a paired sampling design or an independent sampling design. Give reasons for your answer.
- ii) Using a 5% significance level, test whether the time to react in quite and noisy environments are the same or not. Clearly state your findings.

**** Copyrights reserved *****