



KENYATTA UNIVERSITY
UNIVERSITY EXAMINATIONS 2011/2012
SECOND SEMESTER EXAMINATION FOR THE DEGREE OF BACHELOR
OF SCIENCE IN BIOCHEMISTRY AND BIOTECHNOLOGY

SBC 308: BIostatistics and Research Methodology

DATE: Monday, 26th March, 2012

TIME: 2.00 p.m. – 4.00 p.m.

INSTRUCTIONS:

Answer All Questions in Section A and B and Any TWO questions in section C

Section A: Compulsory (20 marks) (Each question is one mark)

1. A numerical value used as a summary measure for a sample, such as sample mean, is known as a?
 - a. population parameter
 - b. sample parameter
 - c. sample statistic
 - d. population mean
 - e. None of the above answers is correct.

2. Since the population size is always larger than the sample size, then the sample statistic
 - a. can never be larger than the population parameter
 - b. can never be equal to the population parameter
 - c. can never be zero
 - d. can never be smaller than the population parameter
 - e. None of the above answers is correct.

3. μ is an example of a
 - a. population parameter
 - b. sample statistic
 - c. population variance
 - d. mode
 - e. None of the above answers is correct.

4. The sum of the percent frequencies for all classes will always equal
- one
 - the number of classes
 - the number of items in the study
 - 100
5. In a five number summary, which of the following is not used for data summarization?
- the smallest value
 - the largest value
 - the median
 - the 25th percentile
 - the mean
6. The following data show the number of hours worked by 200 statistics students.

<u>Number of Hours</u>	<u>Frequency</u>
------------------------	------------------

0 - 9	40
10 - 19	50
20 - 29	70
30 - 39	40

- The class width for the distribution is?
 - is 9
 - is 10
 - is 11
 - varies from class to class
 - None of the above answers is correct.
 - The number of students working 19 hours or less
 - is 40
 - is 50
 - is 90
 - can not be determined without the original data
7. The difference between the largest and the smallest data values is the
- variance

- b. interquartile range
 - c. range
 - d. coefficient of variation
 - e. None of the above answers is correct.
8. A Biostatistics professor asked students in a class their ages. On the basis of this information, the professor states that the average age of all the students in the university is 21 years. This is an example of
- a. a census
 - b. descriptive statistics
 - c. an experiment
 - d. statistical inference
 - e. None of the above answers is correct.
9. The standard deviation of a sample of 100 observations equals 64. The variance of the sample equals
- a. 8
 - b. 10
 - c. 6,400
 - d. 4,096
 - e. None of the above answers is correct.
10. The variance of a sample of 81 observations equals 64. The standard deviation of the sample equals
- a. 0
 - b. 4096
 - c. 8
 - d. 6,561
11. In computing descriptive statistics from grouped data,
- a. data values are treated as if they occur at the midpoint of a class
 - b. the grouped data result is more accurate than the ungrouped result
 - c. the grouped data computations are used only when a population is being analyzed
 - d. All of the above answers are correct.

12. When should measures of location and dispersion be computed from grouped data rather than from individual data values?

- a. as much as possible since computations are easier
- b. only when individual data values are unavailable
- c. whenever computer packages for descriptive statistics are unavailable
- d. only when the data are from a population

13. _____ is a procedure for deriving a mathematical relationship, in the form of an equation between a single metric dependent variable and a single metric independent variable.

14. _____ are hypothesis testing procedures that assume that the variables of interest are measured on at least an interval scale.

- a. parameter tests
- b. parametric tests
- c. nonparametric tests
- d. none of the above

15. The _____ is a symmetric bell-shaped distribution that is useful for small ($n < 30$) testing.

16. The degrees of freedom for the t statistic to test hypothesis about one mean are _____

- a. n
- b. $n - 1$
- c. $n_1 + n_2$
- d. $n_1 + n_2 - 2$

17. The _____ is a statistical test of the equality of the variance of two populations.

18. Is a statistical procedure for analyzing associative relationships between a metric dependent variable and one or more independent variables.

- a. regression
- b. partial correlation coefficient
- c. ANOVA
- d. Product movement correlation

19. When each member of a population has an equally likely chance of being selected, this is called:

- a. A nonrandom sampling method
- b. A quota sample

- c. A snowball sample
- d. An Equal probability selection method

20. Sampling in qualitative research is similar to which type of sampling in quantitative research?

- a. Simple random sampling
- b. Systematic sampling
- c. Quota sampling
- d. Purposive sampling

Section B: Answer all questions (30 marks)

- 1) a. Distinguish between an experiment and an observational study [2 Marks]
 b. State and describe 3 principles of experimentation [3 marks]

- 2) Show that if there are n events which are mutually exclusive, then the probability of happening of any one of them is equal to the sum of the probabilities of the happening of separate events, i.e.

$$P(E_1 + E_2 + E_3 + \dots + E_n) = P(E_1) + P(E_2) + P(E_3) + \dots + P(E_n)$$
 where E_i 's are events and $P(E_i)$'s are the respective probabilities [5 marks]

- 3) Show that given a dataset from standard normal distribution $\mu=0$ and variance $\sigma^2=1$ [5 marks]

- 4) The following data relates to the plasma volumes of eight healthy adult males 2.75, 2.86, 3.37, 2.76, 2.62, 3.49, 3.05, 3.12 litres. Obtain the following
 - a) Mean [1 marks]
 - b) Median [1 marks]
 - c) Variance [1 marks]
 - d) Standard deviation [1/2 mark]
 - e) Coefficient of variation explaining its importance [1.5 marks]

- 5) Show that the mean and variance of a Poisson distribution are the same ($\mu = \sigma^2 = \lambda$) [5 marks]

- 6) Describe the salient characteristics of an objective [5 marks]

Section C: (20 marks) Answer any two questions

- 1) A firm wishes to compare four programs for training workers to perform a certain manual task. Twenty new employees are randomly assigned to the training programs, with 5 in each program. At the end of the training period, a test is conducted to see how quickly trainees can perform the task. The number of times the task is performed per minute is recorded for each trainee, with the following results:

Observation	Program1	Program2	Program3	Program4
1	9	10	12	9
2	12	6	14	8
3	14	9	11	11
4	11	9	13	7
5	13	10	11	8

Estimate the treatment effects for the four programs.

[10 Marks]

- 2) With the aid of a diagram describe the cyclic nature of a research process

[10 marks]

- 3) a) State any four properties of the student t-distribution

[4 marks]

- b) A random sample for the testing of fasting blood sugar of 10 boys was given in mg/dl.

70, 120, 110, 101, 88, 83, 95, 107, 100, 98 Do the data support the assumption of population mean of 100 mg/dl?. Obtain the 95% confidence interval for the mean of fasting blood sugar among the 10 boys.

[6 marks]