



**2015 FIRST SEMESTER EXAMINATIONS**

**COURSE CODE:** SHS 203

**COURSE TITLE:** INTRODUCTION TO EPIDEMIOLOGY & BIostatISTICS

**DATE:** NOVEMBER-DECEMBER 2015

**TIME:** 3 HOURS

---

***INSTRUCTIONS***

---

**Answer ALL Questions in Section A and ANY 3 questions from Section B**

---

The mark allocation for each question is indicated at the end of the question

---

Credit will be given for logical, systematic and neat presentations.

---

# **SECTION A**

## **QUESTION 1**

Distinguish between

- i. Descriptive and Inferential statistics
- ii. Parameter and Statistic
- iii. Population and Sample
- iv. Data and Variable
- v. Nominal and Ordinal scales
- vi. Interval and Ratio scales [12]

## **QUESTION 2**

- a. Define Epidemiology and state its FUNCTIONS [8]
- b. Define following terms and give the formulae for calculating each
  - i. Attack rate [2]
  - ii. Secondary attack rate [2]

## **QUESTION 3**

- a. List the properties of the normal distribution. [4]
- b. What is the difference between Simple Random Sampling and Systematic Random Sampling? [4]
- c. Discuss the advantages and disadvantages of postal interviews. [4]
- d. Describe in detail the steps you take when sampling [4]

## SECTION B

### QUESTION 4

Use the table below (dataset collected from a *Baby-Weighing clinic*) to find measures of central tendency for the variables age and length (ie the mean, mode and median) then measures of dispersion for the variable *Age* and length ie find the maximum, minimum, range, and inter-quartile range. Then calculate the variance, the standard deviation, and the standard error.

[20]

Age (months)	Length (cm)
21	77
34	87
23	84
30	92
27	85
24	82
31	87
26	85
22	85
32	86

### QUESTION 5

- a). Define the terms
- i. Point Source [3]
  - ii. Intermittent Source [3]
  - iii. Continuous Source [3]
  - iv. Propagated Source [3]
  - v. annual death rate [2]
  - vi. crude death rate [3]
  - vii. infant mortality rate (ratio) [3]

### QUESTION 6

- a). A discrete random variable X has the following probability distribution.

X	1	2	3	4
P(X=x)	0.1	0.3	a	0.2

- i. Find the value of a [2]
  - ii. Find a).  $P(X > 2)$  [2]
  - b).  $P(2 \leq X \leq 3)$  [2]
  - iii. Find  $E(X)$ . [2]
  - iv. Find  $E(X^2)$ . [2]
  - v. Find  $\text{Var}(X)$  [2]
  - vi. Hence find Standard Deviation and Standard Error for the distribution. [2]
- Distinguish between a Cohort and Case-Control study design [6]

- b). For the following situations, state the study design and the reason.
- i. A study in which children are randomly assigned to receive either a newly formulated vaccine or the currently available vaccine, and are followed to monitor for side effects and effectiveness of each vaccine. [2]
  - ii. The Zimbabwe Women's Health Study, in which researchers enrolled 41,837 women in 1986 and collected exposure and lifestyle information to assess the relationship between these factors and subsequent occurrence of cancer [2]
  - iii. MoHCC investigators conducted a study to compare measles-mumps-rubella (MMR) vaccine history among 1,294 children with pervasive development disorder (e.g., autism and Asperger's syndrome) and 4,469 children without such disorders. (They found no association.) [2]
- c). Define and describe the three essential characteristics that are examined to study the cause(s) for disease in analytic epidemiology. [4]
- d). Define and describe the three essential characteristics of disease that we look for in descriptive epidemiology [4]

### QUESTION 7

- (i) A mammogram detects 250 positives for breast cancer, of which 180 are incorrect and 70 are correct.
  - a. Construct a 2 x 2 table to measure the diagnostic performance of a mammogram for breast cancer [2]
  - b. What is the sensitivity, specificity of the mammogram [4]
  - c. What is the false positive and false negative rate? [4]
  - d. What is the positive and negative predictive value? [4]
  - e. What is the accuracy of the mammogram? [2]
- (ii) What is the difference between Cluster and Stratified sampling? [4]

### QUESTION 8

- a. Among females in South Africa between the ages of 18 and 74 years, diastolic blood pressure is normally distributed with a mean  $\mu = 77$  mm Hg and a standard deviation  $\sigma = 11.6$  mm Hg.
  - i. What is the probability that a randomly selected woman has a diastolic blood pressure of  $< 60$  mm Hg? [2]
  - ii. What is the probability that has a diastolic blood pressure  $> 90$  mmHg? [3]
  - iii. What is the probability that she has a diastolic blood pressure between 60 and 90 mm Hg? [3]
- b. Pulmonary function values were measured from a sample of 12 individuals with asthma. The mean FVC values for the sample was  $\bar{x} = 4.49$  liters and the standard deviation was  $s = 0.83$  liters.
  - i. Compute a 95% confidence interval for the true population mean  $\mu$  for FVC. Interpret this interval in words. [5]

- ii. Compute a 90% confidence interval for the true population mean  $\mu$  for FVC. How does the width of the interval change? [5]
- iii. In order to compute confidence intervals, what assumption is made about the underlying distribution of FVC? [2]