AFRICA UNIVERSITY
(A United Methodist-Related Institution)

"Investing in Africa's Future"

FACULTY OF HEALTH SCIENCES

2015 FIRST SEMESTER EXAMINATIONS

COURSE CODE: SPH521

COURSE TITLE: 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

SECTION A

QUESTION 1

Distinguish between the following

a. Population and sample
b. Standard deviation and Standard error
c. type 1 and type 2 errors in hypothesis testing
d. Null and alternative hypotheses
e. Nominal and Ordinal scales [10]

QUESTION 2

a Discuss the advantages and disadvantages of personal interviews. [4]

b. A physician claims that more than 10% of pregnant women smoke while pregnant. A survey of 400 randomly selected pregnant women revealed that 60 of them smoked while pregnant. Test the physician's claim at the 0.05 level of significance. [8]

QUESTION 3

A concerned parents group determined the number of commercials shown in each of five children's progress over a period of time.

Number of Commercials. X	5	6	7	8	9
Probability, P(X)	0.2	0.25	a	0.10	0.07

Find the i). the value of a
iii). E(X) =mean
iv). variance and
v). standard deviation for the distribution shown. [8]

QUESTION 4

(a) A sample of 10 individuals is selected for participation in a study of cardiovascular risk factors. The following data represent the ages of the enrolled individuals measured in yeas (continuous variable). The data are as follows:

85 83 82 79 77 76 73 71 68 65

(i) Find the sample mean, standard deviation and standard error of the mean. Interpret the standard deviation of the mean [8]

(ii) Explain when you would use the median instead of the mean as a measure of central tendency? [2]

SECTION B

QUESTION 5

You want to determine the mean blood pressure among government employees. In order to do this, you measure the blood pressure of 200 employees. Use the descriptive statistics below to determine a 95% Confidence Interval around the mean.

N = 300; $\bar{x}$ = 137mmHg; s = 15mmHg. [5]

b. Interpret the 95% CI you calculated in (a) above. [3]

c. You record gestational age at birth for live births in the past month at three primary health facilities in the region. The following statistics were obtained:

n = 250 $\bar{x}$ = 39.5 weeks s = 10.2

(i) Calculate the standard error of the mean. [3]

(ii) Find the 95% confidence interval [6]

(iii) State and Interpret the interval [3]

QUESTION 6

An AU MPH student on an internship visits a prominent weight loss center and receives permission to examine the records (randomly selected) of 8 black women. She is interested in determining whether the healthy eating plan recommended by the dietician was effective in encouraging patients to lose weight. The following table gives the weights (lbs) for 10 subjects measured both at baseline and after 18 months. Do these data provide sufficient evidence, at the 0.05 level of significance, to indicate that the diet regimen is effective for weight loss? [20]

Subject	Baseline	After 6 months
1	340	243
2	265	251
3	297	279
4	325	279
5	290	253
6	303	276
7	287	242
8	279	255
9	285	260
10	320	220

QUESTION 7

a. What is the probability of rolling a pair of dice and obtaining

i). a total score of 10 or more? [2]

ii). A total score more than 6 but less than 8 [2]

b. A box contains three black balls, two white balls and four green balls. A ball is selected randomly and then placed back in the box. A second ball is selected randomly. What is the probability that:

i) both balls are black? [3]

ii) the first ball is black and the second ball is green? [3]

iii) the two balls are of the same colour [3]

c. List the properties of the normal distribution. [4]

QUESTION 8

a. In a cross-sectional survey administered to a random sample of 100 attendees of a local health fair, the following 2x2 table was constructed after reviewing the data:

Current Smoker	**Diabetes**		**Totals**
	Yes	**No**	
Yes	50	25	75
No	20	5	25
Totals	70	30	100

We want to determine, using a level of significance (α) of 0.05, if the risk of having diabetes in the surveyed population is related to smoking. [10]

The price of a stand and the size of a stand in Harare are of interest to a real estate agent which collected the information below.

Price (in $thousnds)	42	50	55	45	52	49	47
Size (m^2)	250	800	600	300	650	500	400

(i) State the independent and dependent variables

(ii) Calculate the Pearson product moment's correlation coefficient.

(iii) Comment on the relationship between price and size. [10]

QUESTION 9

The relationship between forced expiratory volume (FEV), which is measured in litres, and age, which is measured in years, is evaluated in a random sample of 200 men between the ages of 20 and 60. A simple linear regression analysis is performed to predict FEV from age. The following results are published in a paper.

Results of Simple Linear Regression Analysis

	Regression Coefficient	**Standard Error**
Intercept	2.6	.30
Age	0.72	.005

a) Interpret the coefficient of age in words. [3]

b) Give a 95% confidence interval for the coefficient of age. Write a sentence explaining the interpretation of this confidence interval. [8]

c) Given the above results, can you ascertain whether the linear relationship between FEV and age is strong? Why or why not? [3]

d) Suppose above results are used to compare 60 year old men to 50 year old men – what would be the estimated average difference in FEV between the two groups of men? [3]

e) Would it make sense to use the above results to estimate the average FEV levels for men 80 years of age? Why or why not? [3]