



PAN AFRICA CHRISTIAN UNIVERSITY

SCHOOL OF HUMANITIES AND SOCIAL SCIENCES

**END OF SEMESTER EXAMINATION FOR THE DEGREE OF MASTER
OF ARTS IN MARRIAGE AND FAMILY THERAPY**

SEPTEMBER-DECEMBER 2020 TRIMESTER

CAMPUS: ROYSAMBU

DEPARTMENT: PSYCHOLOGY

COURSE CODE: MFT 601A

COURSE TITLE: STATISTICS FOR RESEARCH IN MFT

EXAM DATE: THURSDAY 17TH DECEMBER 2020

TIME: 8.00-11.00AM

DURATION: 3 HRS

INSTRUCTIONS

- Read all questions carefully before attempting.
- Answer **QUESTION ONE AND ANY OTHER THREE**
- Write your **student number** on the answer booklet provided.

QUESTION ONE [COMPULSORY] – 10 MARKS

Provide the most accurate choice in each of the following questions

- a) A numerical value derived from sample data is known as a?
 - A. Parameter
 - B. Sample measure
 - C. Statistic
 - D. Unit
- b) A coefficient of correlation is computed to be -0.95 means that
 - A. The relationship between the two variables is weak
 - B. The relationship between the two variables is strong and positive
 - C. The relationship between the two variables is strong and negative
 - D. The correlation coefficient cannot have this value
- c) The percent of the total variation of the dependent variable Y explained by the set of independent variables X is measured by
 - A. Coefficient of Correlation
 - B. Coefficient of Skewness
 - C. Coefficient of Determination
 - D. Standard Error of Estimate
- d) In the model $Y = mX + a$, Y is also known as the:
 - A. Predictor variable
 - B. Independent variable
 - C. Dependent variable
 - D. Explanatory variable
- e) A magazine conducts a survey and asks its readers to cut the questionnaire from the magazine, fill it and send it via mail. This type of sampling technique is called?
 - A. Purposive Sampling
 - B. Snowball Sampling
 - C. Sequential Sampling
 - D. Convenience Sampling
- f) You asked five of your classmates about their height. On the basis of this information, you stated that the average height of all students in your university is 67 inches. This is an example of:
 - A. Descriptive statistics
 - B. Inferential Statistics
 - C. Parameter
 - D. Population
- g) Which one is the not measure of dispersion.
 - A. The Range
 - B. Percentile
 - C. Inter-Quartile Range
 - D. Variance

- h) The height of a student is 60 inches. This is an example of
- A. Qualitative data
 - B. Categorical data
 - C. Continuous data
 - D. Discrete data
- i) In uni-modal distribution, if the mode is less than mean, the distribution is:
- A. Symmetrical
 - B. Normal
 - C. Positively Skewed
 - D. Negatively Skewed
- j) If the mean is 25 and the standard deviation is 5 then C.V (Coefficient of variation) is
- A. 100%
 - B. 25%
 - C. 20%
 - D. None of these

QUESTION TWO

Consider the data set 2, 2, 3, 6, 10.

- a) Compute the mode, median, and mean. **[3 marks]**
- b) Multiply each data value by 5 and compute the mode, median, and mean. **[3 marks]**
- c) Comparing the results obtained in parts (a) and (b), comment on how you think the mode, median, and mean are affected when each data value in a set is multiplied by the same constant C? **[1 mark]**
- d) Suppose you are given that the mode, median and mean heights of a random sample of students is 70 inches, 68 inches and 71 inches respectively, and further given that 1 inch = 2.54 cm, calculate the values of the mode, median, and mean in centimeters. **[3 marks]**

QUESTION THREE

There are 4 types of measurement scales used in collecting statistical data: Nominal, ordinal, Interval and Ratio. Identify the statistical measurement scale you would use to collect the data on the following statistical issues. **[10**

marks]

- i. Places from where students are travelling to take a statistics course
- ii. Students' scores on a Biology test
- iii. Number of cars
- iv. Calendar year
- v. Amount of calories in a small Yogurt
- vi. Shades of lipstick available in a beauty shop
- vii. Your shoe size
- viii. Your age
- ix. Country of your birth
- x. Postal code

QUESTION FOUR

In a certain school, the mean performance was 500 marks and the standard deviation was 100 marks. You are further informed that the performance was normally distributed.

- a) Suppose that a student scored 465 marks in the exam, what is the Z score that corresponds to this raw score? **[2 marks]**
- b) Explain the meaning of this Z score value. **[1 mark]**
- c) Use the standard normal tables to determine the percentage of the students scored above 465 marks. **[2 marks]**
- d) Determine the percentage of students who took the exam scored between a 465 and 500? **[2 marks]**
- e) Assume for the moment that the student who scored 465 marks retook the exam and posted a Z score of +0.35, representing a marked improvement over his previous

score, Calculate the marks scored in the retaken examination.
marks]

[3

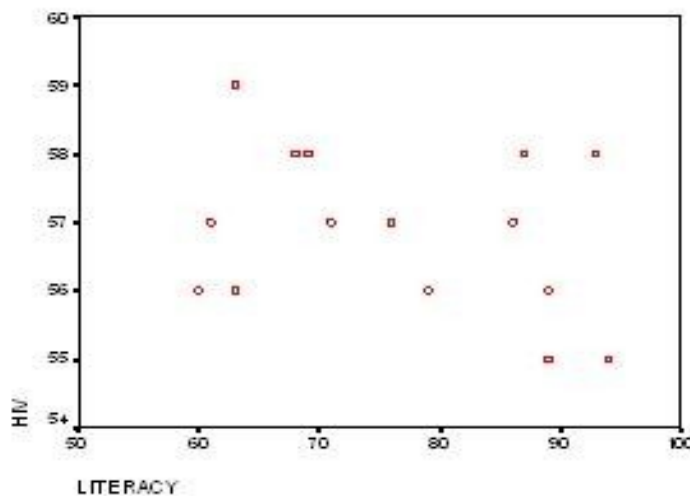
QUESTION FIVE

The following data was collected by Population Reference Bureau (PRB) on the relationship between women's literacy and the prevalence of HIV in Sub-Saharan Africa. Data were assembled for 15 countries in Sub-Saharan Africa through the year 2015 as shown below.

Country	% female literacy	% Female with HIV
Angola	59	63
Botswana	58	93
DRC	57	61
Djibouti	56	79
Eritrea	56	60
Madagascar	58	68
Malawi	57	71
Namibia	55	94
Nigeria	58	87
Rwanda	57	76
Sudan	58	69
Swaziland	55	89
Tanzania	56	89
Togo	56	63

Zambia	57	86
--------	----	----

- a) A researcher advances the hypothesis that women in countries with lower female literacy rates tend to see a greater percentage of the female population infected with HIV. Identify the independent and the dependent variables here. **[2 marks]**
- b) Using the above data, a scatter diagram was constructed as shown below:



Very generally, explain what this scatter diagram signifies about the relationship between female literacy rates and the percentage of females with HIV. **[2 marks]**

- c) Several statistics were calculated from these data for both the independent and the dependent variables as below:

$$\bar{X} = 56.87; \bar{Y} = 76.53$$

$$\sigma_X = 1.19; \sigma_Y = 12.35$$

$$\text{Cov}_{X,Y} = -4.06$$

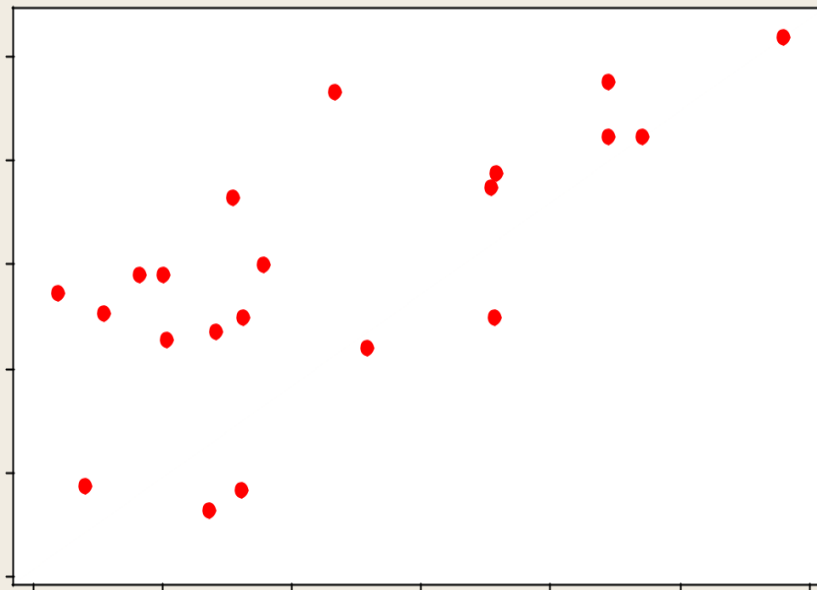
Identify each of the statistic relating to the independent and dependent variables above. **[2 marks]**

- d) Without doing any calculations, identify the direction of the relationship between percentage female literacy and the percentage of females with HIV from the statistics provided in part c. **[2 marks]**
- e) Using the statistics provided in part c, calculate the value of Pearson's correlation coefficient, r , and provide a one-sentence interpretation of this value. **[2 marks]**

QUESTION SIX

The following table presents data collected from 21 countries on Annual Per Capita Cigarette Consumption ("Cigarette"), and Deaths from lung cancer per 100,000 persons of age 35-64 ("Deaths").

Country	Cigarette	Deaths
United States	3900	259.9
Canada	3350	211.6
Australia	3220	238.1
New Zealand	3220	211.8
United Kingdom	2790	194.1
Switzerland	2780	124.5
Ireland	2770	187.3
Iceland	2290	110.5
Finland	2160	233.1
West Germany	1890	150.3
Netherlands	1810	124.7
Greece	1800	41.2
Austria	1770	182.1
Belgium	1700	118.1
Mexico	1680	31.9
Italy	1510	114.3
Denmark	1500	144.9
France	1410	144.9
Sweden	1270	126.9
Spain	1200	43.9
Norway	1090	136.3



Scatterplot of Deaths vs Cigarette

Death

Cigarettes

Further, the regression analysis provided the following output:

Regression Analysis: Deaths versus Cigarette

The regression equation is

Deaths = 29.5 + 0.0557 Cigarette

Predictor	Coef.	SE Coef	t	p-value
Constant	29.45	29.48	1.00	0.330
Cigarette	0.05568	0.01288	4.32	0.000

S = 46.5558 R-Sq = 49.6% R-Sq(adj) = 46.9%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	40484	40484	18.68	0.000
Residual Error	19	41181	2167		
Total	20	81666			

- a) Based on the scatterplot of Deaths versus Cigarette, what would you say is the nature of the relationship between the 2 variables? **[2 Marks]**

- b) Turning to the fitted regression equation, provide a statistical interpretation of the cigarette coefficient of 0.0557 and the constant 29.5 **[3 marks]**
- c) State whether or not the explanatory variable in this study was statistically significant and justify **[2 marks]**
- d) Use the obtained regression equation to estimate the number of deaths per 100,000 people for a country with a per capita cigarette consumption of 5520 **[3 marks]**