



PAN AFRICA CHRISTIAN UNIVERSITY

MASTER OF ARTS IN MARRIAGE AND FAMILY THERAPY

END OF TERM EXAMINATION

DEPARTMENT:

COURSE CODE: MFT 601A

COURSE TITLE: STATISTICS FOR RESEARCH IN MFT

EXAM DATE: APRIL 2021

TIME: 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) Which one of the following is NOT appropriate for studying the relationship between two quantitative variables?
- A. Scatterplot
 - B. Bar chart
 - C. Correlation
 - D. Regression
- b) A numerical value derived from sample data is known as a?
- A. Parameter
 - B. Sample measure
 - C. Statistic
 - D. Unit
- c) 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
- d) 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
- e) A regression between foot length (dependent variable in cm) and height (explanatory variable in inches) for 33 students resulted in the following regression equation:
 $\hat{y} = 10.9 + 0.23x$
What is the estimated average foot length for students who are 70 inches tall?
- A. 27 cm
 - B. 28 cm
 - C. 29 cm
 - D. 30 cm
- f) 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
- g) 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
- h) Correlation and regression are concerned with
- A. the relationship between two categorical variables.
 - B. the relationship between two quantitative variables.
 - C. the relationship between a quantitative explanatory variable and a categorical response variable.
 - D. the relationship between a categorical explanatory variable and a quantitative response variable.
- i) 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
- j) In uni-modal distribution, if the mode is less than mean, the distribution is:
- A. Symmetrical
 - B. Normal
 - C. Positively Skewed
 - D. Negatively Skewed
- k) 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

A study was conducted to determine the level of annual income per family in Kenya among the top-earning class. 100 families in that earning class were randomly selected and analyzed. The following result was obtained:

Variable	Output	Measure
Mean	76.5411	Sh.
Median	21.97	Sh.
Mode	13.01	Sh.
Standard Deviation	4.5	Sh.
Variance	20.25	Square Sh.
Kurtosis	22.2632	
Skewness	4.596	
Range	1096.01	Sh.
Maximum	1105	Sh.
Minimum	8.99	Sh.
Sum	7654.11	Sh.
Count	100	No.

Write a brief report interpreting the meaning of each of the above highlighted descriptive statistics **[10 marks]**

QUESTION THREE

- a) 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 **[6 marks]**

- i. Attitudes by different people
- ii. Size of a piece of land in Acres
- iii. Shoe size
- iv. Blood group/type
- v. Political party affiliation
- vi. Order of finish in a race or a contest

- b) Describe the meaning and statistical application of the Snowballing sampling techniques **[4 marks]**

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. **[3 marks]**

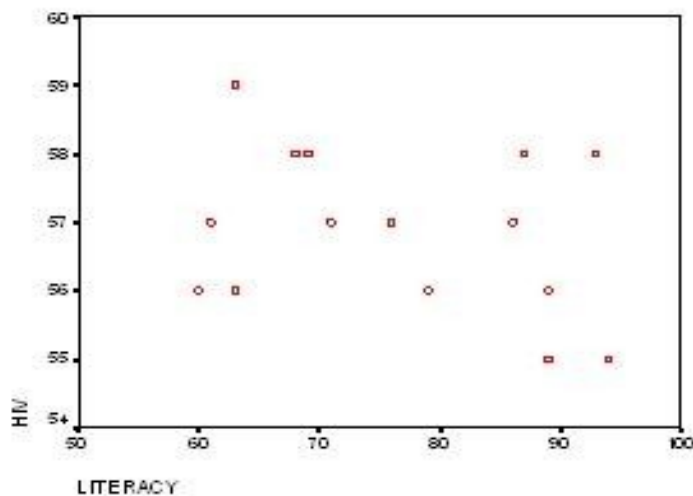
QUESTION FIVE

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]**

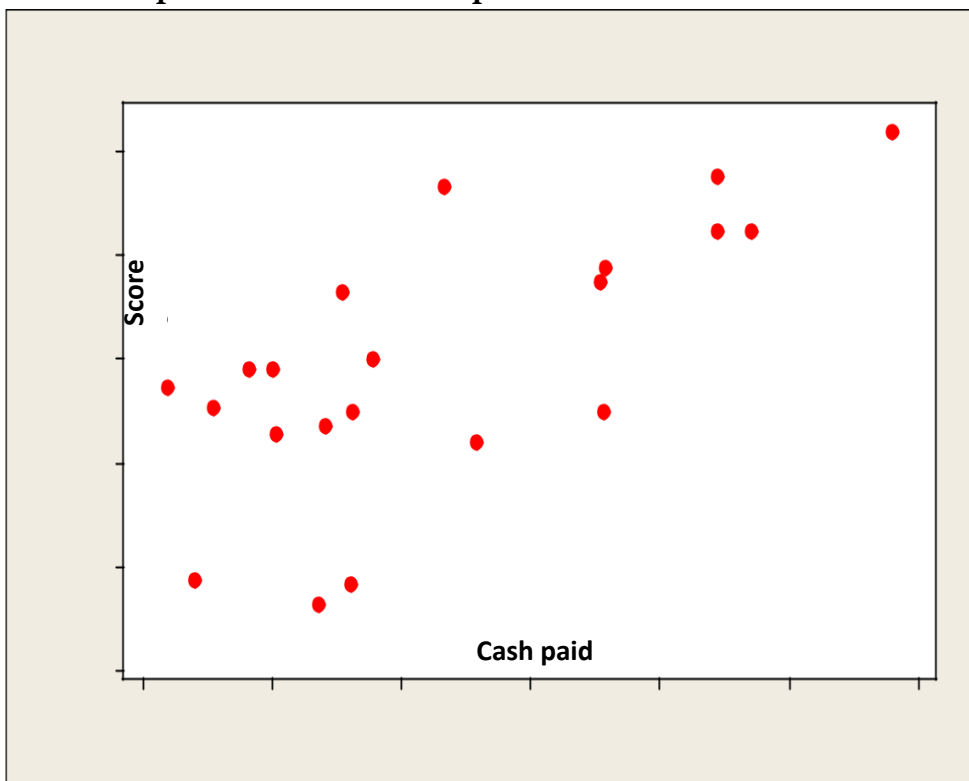
- d) Without doing any calculations, identify the direction of the relationship between percentage female literacy and the percentage of females with HIV from the statistics provided in part c. **[2 marks]**
- e) Using the statistics provided in part c, calculate the value of Pearson's correlation coefficient,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

QUESTION SIX

The National body has recently begun a controversial program in which high school students are paid cash incentives for passing an end-of-year standardized test. Suppose we conduct a study, in which end-of-year test scores (y) are measured on a scale of 0–100 and the amount of the cash incentive offered to the student (x) is measured in dollars from \$0 to \$500.

A scatterplot of the 96 observations in the sample is shown below, along with the output from statistical software.

Scatterplot of Scores vs Cash paid



Regression analysis output:

Regression Analysis: Score versus Cash paid

The regression equation is

$$\text{Scores} = 67.51 + 0.0148 \text{ Cash}$$

Predictor	Coef.	SE Coef	t	p-value
Constant	67.513	2.509	1.00	0.330
Cash	0.014762	0.00886	4.32	0.0952

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			

- i. Based on the scatterplot, what would you say is the nature of the relationship between the 2 variables? [2 marks]
- ii. Turning to the fitted regression equation, provide a statistical interpretation of the cash coefficient of 0.0148 and the constant 67.513 [4 marks]
- iii. State whether or not the explanatory variable in this study was statistically significant and justify [2 marks]
- iv. Use the obtained regression equation to estimate the scores when the student is offered \$200 [2 marks]