



MACHAKOS UNIVERSITY

University Examinations 2022/2023 academic year

SCHOOL OF EDUCATION

DEPARTMENT OF EDUCATIONAL PSYCHOLOGY, SPECIAL NEEDS EDUCATION

FIRST YEAR SECOND SEMESTER EXAMINATION FOR

MASTER OF ARTS (EDUCATION)

ECC 803: COMPUTER APPLICATIONS IN SOCIAL SCIENCES

DATE: 13/4/2022

TIME: 2:00 – 4:00 PM

INSTRUCTIONS

Answer **Question 1** and any other **two** questions.

QUESTION ONE (30 MARKS)

- (a) Highlight four causes of asymmetry in data distribution (4 marks)
- (b) The following was a height in feet for sample of students from MSc in statistics class; 3.3, 4.5, 5.5, 5.7, 4.8 6.2, 5.6, 4.6, 4.5, 5.0. Present it in a stem-and-leaf plot (4 marks)
- (c) Differentiate the following terms as they apply in data analysis
- Discrete variable and continuous variable
 - Parametric test statistics and Non-parametric test statistics (4 marks)
- (d) A marital counsellor based in Nairobi claimed that, the contribution of marital infidelity to marriage breakdown is 65%. An independent researcher interviewed 10 randomly selected individuals who had walked out of their marriage to rate infidelity among other causes as the reason of their separations. Their responses were as follows: 66, 65, 69, 70, 69, 71, 70, 63, 64 and 68 percent. The t-test output was as given

The sample t-test output, for the test value=65

	t	df	Sig.(2-tailed)	Mean difference	95% CI of the difference	
					Lower	Upper
Infidelity rated	2.825	9	0.020	2.500	0.4979	4.5021

- Construct the hypothesis statements for the above scenario
 - Based on the t-value and the sig.(2-tailed) value make statistical conclusion
 - Interpret the 95% CI lower and upper difference values (6 marks)
- (e) Explain any **three** measurement scales for statistical data as outlined by Stevens (1946), giving two examples in each case. (6 marks)

- (f) Briefly differentiate the following terms as they apply in research;
- Structured and Semi-structured questionnaire
 - Type I error and type II error
 - Null hypothesis and alternative hypothesis
- (6 marks)

QUESTION TWO (15 MARKS)

- a) Define the term *variable* as used in statistics, giving an example. (2 marks)
- b) Discuss any two types of parametric test-statistics and state when appropriate to use each (4 marks)
- (c) A random sample of 400 medical practitioners in East Africa was selected for in service clinical psychology training. As part of the training each trainee was asked to choose a country he/she preferred for an attachment from a list of three countries. The results are shown in table 1:

Table 1

Age group	Country		
	Kenya	Tanzania	Rwanda
Under 30	120	30	50
31 – 50	10	75	15
51 and above	10	30	60

Test the hypothesis that the populations were homogeneous with respect to the country they preferred for an attachment despite their age difference at 15% significance level. (9 marks)

QUESTION THREE (15 MARKS)

- a) Discuss two ways of handling an outlier during analysis stage (2 marks)
- b) Distinguish the following terms as they apply in data analysis using computers
- Computer system and computer software
 - Spread sheet and database
- (4 marks)
- c) The following are the results after conducting research on a micro finance company performance as affected by , wages, cost of capital and technology.

$$\hat{Y} = -2.478 + 0.288x_1 + 0.758x_2 + 0.085x_3 + e$$

$$Se \ (19.748) \ (0.110) \ (0.112) \ (0.168)$$

$$R^2 = 0.928 \ n = 30$$

Where ;

Y is the microfinance performance as measured by return on assets.

X_1 is wages

X_2 is cost of capital

X_3 is the technology

e is the error term

- Construct the hypotheses statement for the study (2 marks)
- Interpret the coefficient of determination. (2 marks)
- Interpret the results using regression coefficients (5 marks)

QUESTION FOUR (15 MARKS)

- a) Differentiate the following terms as they apply in scientific research
- Descriptive Statistic and inferential statistics
 - Confidence level and Significance level in hypothesis testing (4 marks)
- b) The table below is part of an output of data analyzed whose response variable was whether or not (1=yes,0=no) an undergraduate student was in a marital relationship (y), the predictor variables included the students age (x1), gender (x2), family income (x3) in Ksh'000', family size (x4) and the religion coded such that (1=Catholic,2=Protestant,3=Muslim,4=Hindu) (x5)

Variables in the equation							95% C.I. for EXP(B)	
Step1 ^a	B	S.E.	Wald	df	Sig.	Exp(B)	Lower	Upper
Family-size	1.338	.494	4.419	1	.036	.354	.135	.932
Age	1.197	.120	2.710	1	.043	.821	.649	1.038
Family Inc	.201	.173	1.342	1	.247	1.222	.870	1.716
Religion			1.085	3	.781			
Religion(1)	-.073	1.477	.002	1	.961	.930	.051	16.801
Religion(2)	1.468	1.743	1.710	1	.047	4.342	.143	132.234
Religion(3)	-19.860	11.714	.000	1	.999	.000	.000	.
Gender(F)	.645	1.343	.231	1	.631	1.906	.137	26.476
Constant	8.329	4.262	3.819	1	.051	412.576		

- With justifications highlight the variables that contributed significantly to the prediction of the relationship status of a student. (2 marks)
- Interpret the betas, $\text{Exp}(\beta)$ of the variables highlighted in (i) (5 marks)
- Write down the regression equation of the analysis output. (4 marks)

QUESTION FIVE (15 MARKS)

- a) Briefly differentiate the following terms as they apply in research
- Exploratory research design and experimental research designs
 - Quantitative Research Methodology and Qualitative Research Methodology (4 marks)
- b) A contracting farming agency collected the data below according to the listed variables in order to predict the sales of product Q in future in different counties:

X1	23	34	56	78	15	21	24
X2	34	56	78	79	90	65	14
X3	23	12	12	13	14	16	23
Y	4	5	12	23	4	5	8

Y= Sales per county in Ksh. '000,000'
X1= Amount spent on advertisement in Ksh. '000' per county
X2= Amount spent on market research in Ksh. '000' per county
X3 = Amount spent on staff development and remuneration per county in Ksh '000'
The agency used ANOVA and regression for analysis, the following is part of the results:

SUMMARY OUTPUT

<i>Regression Statistics</i>	
Multiple R	0.625373
R Square	0.391092
Adjusted R Square	-0.21782
Standard Error	7.63801
Observations	7

<i>ANOVA</i>					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance</i>
Regression	3	<i>x</i>	37.47033	<i>z</i>	0.637562
Residual	3	175.0176	<i>y</i>		
Total	6	287.4286			

- i. State the statistical model in this analysis. (1 mark)
- ii. Explain the difference between R Square and R Adjusted (2 marks)
- iii. Complete the values *x*, *y* and *z* in the ANOVA table (3 marks)
- iv. Interpret the finding in this analysis at 5% significance level. (3 marks)