(Knowledge for Development)

KIBABII UNIVERSITY

UNIVERSITY EXAMINATIONS

2019/2020 ACADEMIC YEAR

FOURTH YEAR SECOND SEMESTER

SPECIAL/ SUPPLEMENTARY EXAMINATION

FOR THE DEGREE OF BACHELOR OF SCIENCE

(MATHEMATICS)

COURSE CODE: STA 452

COURSE TITLE: CATEGORICAL DATA ANALYSIS (CDA)

DATE: 05/02/2021 **TIME:** 11 AM – 1 PM

INSTRUCTIONS TO CANDIDATES

Answer Question One and Any other TWO Questions

TIME: 2 Hours

This Paper Consists of 5 Printed Pages. Please Turn Over.

QUESTION ONE: (30 marks)

Consider the data below pertaining to gun registration and death penalty.

	Death Penalty	
Gun Registration	Favour	Oppose
Favour	884	376
Oppose	411	96

(a) Is Gun Registration independent from Death Penalty? Formulate a suitable test to enable you make an appropriate conclusion (6 marks)

(b) Obtain the estimated, i) probability of those who oppose death penalty in the group that favours gun registration (2 marks)

ii) probability of those who favour death penalty in the group that favours gun registration (2 marks)

iii) odds for those opposed to death penalty for those favouring Gun registration (3 marks)

iv) odds for those opposed to death penalty for those opposed to Gun registration (3 marks)

Find the odds ratio, $\hat{\theta}$ and give an interpretation for it (4 marks)

(c) (i) Define the following

I) Concordant pairs (2 marks)

II) Discordant pairs (2 marks)

(ii) Explain how you would use the Concordant and Discordant pairs to measure the Strength of Association between two Categorical variables (6 marks)

(d) Use the fact that $\frac{\ln(\hat{\theta})}{SE(\ln(\hat{\theta}))}$ has a distribution that is approximately standard normal for large values of the sample size n to acertain whether there is an evidence of a strong association between Gun registration and Death penalty. (5 marks)

QUESTION TWO: (20 marks)

In a certain cohort study, it was noted that per year, the proportion who died from lung Cancer was 0.0013 for cigarette smokers and 0.0002 for non-smokers. The proportion who died from coronary heart disease was 0.00779 for smokers and 0.00313 for non-smokers.

(a) Describe the association of smoking with each of lung cancer and heart disease using the relative risk. Interpret (5 marks)

(b) Describe the associations using the difference of proportions.Interpret. (5 marks)

(c) Describe the associations using the odds ratio.Interpret (7 marks)

(d) Which response is more strongly related to cigarette smoking in terms of the reduction in the number of deaths that would occur with elimination of cigarettes (3 marks)

QUESTION THREE: (20 marks)

The number of licensor companies classified by the proportion of foreign profits derived from

Licence agreements were as follows:

Proportion of profit	**Licensor Type** Dominant Product Diversified		**Total**
Less than 5%	1	6	7
5% or more	7	6	13
Total	8	12	20

(a) Formulate a hypothesis to investigate whether the proportion of profit is independent of the Licensor types (4 marks)

(b) Test the hypotheses in (a) above at 5% level of significance, taking into consideration, Yate's correction factor (12 marks)

(c) Comment on the relationship between proportion of profit and licence types (4 marks)

QUESTION FOUR: (20 marks)

Consider the following data on the job status of students working within the university. The students are categorized as freshmen or sophomores and required to confirm whether they had a job or not.

	Job Status	
	No	Yes
Freshmen	35	22
Sophomores	21	24

(a) Calculate the odds ratio, $\hat{\theta}$ (5 marks)

(b) The data in the Table below is of records of accidents in 2018 compiled by the ministry of Transport in the Republic of Kenya.

Safety equipment in use	**Injury**	
	Fatal	Non Fatal
None	1600	262,527
Seat belt	710	432,369

Work out the number of

(i) Concordant pairs, (4 marks)
(ii) Discordant pairs and hence (4 marks)
(iii) Obtain the value of gamma, $\hat{\gamma}$ (5 marks)

Comment on the association between the safety equipment in use and the nature of injury (2 marks)

QUESTION FIVE: (20 marks)

The following Table gives the effect of Measles immunizing vaccination on 12 months Care and cohort infants.

	Suffered from Measles, (A)	Did not suffer from Measles, (a)	Total
Vaccinated, (B)	17 (AB)	58 (aB)	60 (B)
Not- Vaccinated, (b)	73 (Ab)	16 (ab)	104 (b)
Total	90 (A)	74 (a)	164 (N)

(i) Is the above data consistent?
Hint: Use ultimate class frequencies to investigate this. (4 marks)

(ii) Evaluate the coefficient of association Q between vaccination and the suffering of patients from measles (12 marks)

(iii) Comment on the effect of vaccination in terms of its ability to control measles (4 marks)