JARAMOGI OGINGA ODINGA UNIVERSITY OF SCIENCE AND TECHNOLOGY

SCHOOL OF MATHEMATICS AND ACTURIAL SCIENCE

UNIVERSITY EXAMINATION FOR DEGREE OF BACHELOR OF SCIENCE IN ACTURIAL SCIENCE

3rd Year 2nd SEMESTER 2017/2019 ACADEMIC YEAR

MAIN REGULAR

COURSE CODE: SAS 304

COURSE TITLE: TESTS OF HYPOTHESIS

EXAM VENUE: **STREAM: (BSc Actuarial science)**

DATE: 23/4/19 **EXAM SESSION: 12.00 – 2.00pm**

TIME: 2.00 HOURS

Instructions:

Answer question one and any other two questions.

Candidates are advised not to write on the question paper.

Candidates must hand in their answer booklets to the invigilator while in the examination room.

QUESTION ONE (20 Marks)

a) Differentiate between the following terms/phrases as used in statistical inference
   i) Type I and type II error (2 marks)
   ii) Confidence level and significance level (2 marks)

b) On a group of anaemic patients, an iron preparation was administered and haemoglobin levels of the patients before and after therapy were noted and are provided below. It is desired to find out whether there is a significant change in the haemoglobin level of the group after the therapy at $\alpha = 0.05$ level of significance. (6 marks)

Hb level in gm%

Patient number	Before therapy	After therapy
1	5.6	10.2
2	4.8	9.4
3	6.5	11.0
4	7.5	7.5
5	4.5	7.5
6	3.5	6.0
7	6.7	8.0
8	6.2	9.6
9	5.6	10.0
10	4.4	8.4
11	7.5	8.0
12	8.0	8.0

c) The following is a summary of data presented by Refrigeration Company. The company had insisted to determine whether there was a relationship between the optimum Lot size of refrigeration replacement parts (X) and labour hours (Y) required to produce the Lot.

Summary:

n=25; $\sum_{i=1}^{n} x_i$ = 1750; $\sum_{i=1}^{n} y_i$ = 7807; $\sum_{i=1}^{n} x_i\, y_i - n\bar{x}\,\bar{y}$ = 70670;

$\sum_{i=1}^{n}(x_i - \bar{x})^2$ = 19800; $\sum_{i=1}^{n}(y_i - \hat{y})^2$ = 54825

i) Obtain the regression parameter estimates (4 marks)
ii) Obtain a 95% confidence interval estimate for β_1 (3 marks)
iii) Test the hypothesis H_0: β_1 = 0 vs H_1: $\beta_1 \neq 0$ at 5% level of significance (2 marks)

iv) Obtain a 90% confidence interval estimate for the mean of the response when x_h =120 (2 marks)

v) Construct the ANOVA table (4 marks)

c) In an epidemic of gastroenteritis in an area the number of cases reported in two populations consuming water from different sources were as follows:

Source of variation	Number of people consuming water from the source	Number of cases of gastroenteritis
Tap water	800	35
Hand pump	2400	120
Total	3200	155

Test statistical difference in the proportion of cases in the two groups is significantly different at 95% confidence level. (5 marks)

QUESTION TWO (20 Marks)

Iron intake of antenatal mothers in different periods of pregnancy in 10 villages is provided below.

Antenatal period (Trimester)	Villages										Total of each trimester
	1	2	3	4	5	6	7	8	9	10	
I	11.5	19.5	18.5	12.5	18.5	16.5	26.5	18.5	16.5	24.5	182.5
II	27.0	28.0	22.0	21.0	15.0	19.5	20.0	26.0	30.0	28.5	237.0
III	28.0	30.0	26.0	30.0	24.5	28.5	26.0	30.0	27.0	25.5	275.5
Total of village	66.5	77.5	66.5	63.5	58.0	64.5	72.5	74.5	73.0	78.5	695.0

Test the hypothesis that the iron intake in the different periods is significantly not different at 95% confidence level. (20 marks)

QUESTION THREE (20 Marks)

The following data gives the age of the mother in years at the time of delivery and weight of the new born in sample deliveries.

Age of the mother (X_1)	25	36	25	28	30	22	35	30	21	20	41	35	20	20	20
Weight of the new born in Ibs (X_2)	5.0	8.0	7.0	7.5	7.5	6.0	7.0	7.0	5.2	6.1	8.0	7.0	7.0	4.5	4.0

i) Quantify the relationship between the two variables age of the mother (X_1) and weight of the new born (X_2) using parametric correlation analysis. (4 marks)

ii) Fit a regression line for the data hence provide the regression equation. (6 marks)

iii) Interpret the regression coefficients (2 marks)

iv) Test for the statistical significance of the regression coefficients at 95% confidence level. (4 marks)

v) Obtain an ANOVA table for the fitted regression model. (6 marks)

QUESTION FOUR (20 Marks)

(a) A company studied the effect of 3 different types of promotions on sales of its craters. Fifteen (n=15). Stores were selected at random for the study with 5 stores assigned to each promotional type. Data on the number of cases of product sold a previous promotional period (x) and on the current period (y) gives below.

Promotional type	1		2		3		4		5	
	Y_{i1}	X_{i1}	Y_{i2}	X_{i2}	Y_{i3}	X_{i3}	Y_{i4}	X_{i4}	Y_{i5}	X_{i5}
1	38	21	39	26	26	22	45	28	33	19
2	43	34	38	26	38	29	27	18	34	25
3	24	23	32	39	31	30	21	16	28	29

Fit an ANACOVA model to these data and assess whether there was promotional type effect on sales after adjusting for previous sales (use $\alpha = 0.05$)

QUESTION FIVE (20 Marks)

Suppose that we are interested in the factors that influence whether a political candidate wins an election. The outcome (response) variable is binary (1/0) for win /lose. The predictor variables of interest are the amount of money spent on the campaign (X_1), the amount of time spent campaigning negatively (X_2) and whether or not the candidate is incumbent (X_3). The regression coefficients are provided below;

X_1	=	0.8040
X_2	=	-0.6754
X_3	=	0.0226
Null deviance	=	499.98
Residual deviance	=	358,52

i) Write the logistic regression formula and interpret the regression coefficients. (4 marks)

ii) Obtain the odds ratio for the regression coefficients showing clearly your working formula. (6 marks)

iii) Interpret the odds ratio. Also, compare performance of incumbent and non-incumbent using the odds ratio. (6 marks)

iv) Obtain pseudo-R squared and use it to evaluate the model goodness-of-fit interpreting accordingly. (4 marks)