

**KABARAK**



**UNIVERSITY**

**UNIVERSITY EXAMINATIONS**

**2010/2011 ACADEMIC YEAR**

**FOR THE DEGREE OF BACHELOR OF COMPUTER**

**SCIENCE**

**COURSE CODE: COMP 328**

**COURSE TITLE: RESEARCH METHODS IN COMPUTER**

**STREAM: Y3S2**

**DAY: TUESDAY**

**TIME: 2.00 – 4.00 P.M.**

**DATE: 15/03/2011**

---

**INSTRUCTIONS:**

1. Answer question **ONE** and any other two questions
2. Show all your working and be neat

**PLEASE TURN OVER**

**QUESTION ONE (30 marks)**

- a) Outline research procedure cycle one has to follow for a credible study as taught to you in this course **[5 marks]**
- b) Assume you have conducted a research for your academic dissertation/project, briefly outline what should appear in each chapter of the university research booklet **[5 marks]**
- c) Mention four types of sample survey designs normally used for collecting data. State the advantages and disadvantages of the design methods you have mentioned. **[6 marks]**
- d) Distinguish between Karl Pearson’s correlation coefficient and Spearman’s rank correlation coefficient, showing clearly when one selected over the other **[ 4 marks]**
- e) What are the uses of correlation and regression as taught to you in this course **[6 marks]**
- f) Which hypotheses does chi-squares test **[ 4 marks]**

**QUESTION TWO (20 MARKS)**

An investigator wants to investigate overheating of our old desktops in Lab 3. He measures the heat (Y) against time (seconds) on 21 randomly selected computers as follows:

- i) Fill in the appropriate spaces and confirm the additions in the relevant columns **[2 marks]**

X	$(x - \bar{x})=x$	$x^2$	Y	$(y - \bar{y})=y$	$y^2$	xy
49			46			
53			57			
62			119			
58			90			
50			44			
63			131			
59			104			
57			100			
52			53			
51			65			
53			89			
51			70			
56			85			
60			96			
65			131			
61			113			
52			69			
52			66			
56			96			
58			111			
54			69			
$\sum=117$ 2		$\sum x^2$ =429.238095	$\sum=180$ 4		$\sum y^2$ =14207.8095	$\sum xy=2327.$ 619

- ii) Test whether there is an association between X and Y using Pearson correlation at  $\alpha = 0.05$  [ 3 marks]
- iii) State a simple linear regression model of y on x clearly describing all terms shown and the relevant assumptions [ 3marks]
- iv) Estimate parameters of the linear regression line using the data and test hypothesis for intercept and slope at  $\alpha = 0.05$  [5 marks]
- v) Construct a  $100(1-\alpha)\%$  confidence intervals for  $\beta_0$  and  $\beta_1$  [3marks]
- vi) Complete the analysis of variance table below: [2marks]

Source of variation	SS	DF	MS	F
Regression on x				
Residual				
Total				

- vii) Predict  $Y_0$  given  $X_0 = 67$  [marks 2]

### QUESTION THREE (20 MARKS)

- a) Consider a sample of size 150 with mean=12.3 and standard deviation=4.2, test the hypothesis of the following using level of significance of 5% [Use critical  $z=-1.65$ ]  
 $H_0: \mu=13.1$  [6 marks]  
 $H_1: \mu < 13.1$
- b) A computer software firm is testing two alternative introductory programming packages to see if one is easier to learn than the other. Below are test scores achieved by students using the two different packages. Do a paired t test to see if one package is easier at  $\alpha = 0.05$ . [6 marks]

Package A	68	72	92	58	81	89	95	76	89	92	75	64
Package B	72	69	92	64	85	88	94	86	94	98	84	70

- c) A survey was conducted using the methods indicated in the table below. Fill in on advantages of using the methods with “Yes” or “No”

	Personal Interview	Group Interview	Postal Survey	Telephone Survey
Are sure that correct respondent is interviewed				
Can be sure respondent is part of sampling frame				
Subjects selected independently				
Interviewers trained to deal with problems, obtain information for particular questionnaire				
Can make explanations and answer questions				
Can probe for details				
Opportunity to motivate respondents				
Opportunity to edit response				
Opportunity to evaluate responses				
Some control over respondent's task performance				
Opportunity for direct quality control of interviewing/data entry process				
Opportunity to reflect on responses				
Data can be entered directly into computer				
Can use visual aids				
Possibility of recording errors by interviewer				
Invasion of privacy				

[8 marks]

**QUESTION FOUR (20 MARKS)**

a) A researcher compared demographic characteristics of users and nonusers of extension services. Calculate the appropriate measures of association at 5% significant level for the table below. [Use Critical  $\chi_{0.025}^2(4) = 11.14$  ]

[10 marks]

Age	Nonusers	Users	Total
18-29	2160	804	2964
30-39	1227	619	1846
40-49	1157	585	1742
50-59	1045	609	1654
60+	1424	768	2192
Total	7013	3385	10398

b) *In an experiment on two farms, the number of Calliandra seedlings that survived is given below*

Survival of calliandra seedlings on two farms

Farm type	surviving	Not Surviving
Farm 1	10	8
Farm 2	15	4

i) Calculate the odds of each farm type.

ii) Calculate the odds ratio and comment on the answer

**[10 marks]**