UGANDA MARTYRS UNIVERSITY

NKOZI

UNIVERSITY EXAMINATIONS

FACULTY OF SCIENCE

DEPARTMENT OF MATHEMATICS & STATISTICS

END OF SEMESTER ONE FINAL ASESSMENT

STATISTICAL INFERENCE AND DATA ANALYSIS

DATE: 5 TH DECEMBER 2014

TIME: 2:00-5:00 PM

DURATION: 3 HRS

Instructions:

1. *Carefully read through ALL the questions before attempting*
2. ***ANSWER FIVE (5) Questions*** *ONLY. (Each question carries equal marks)*
3. *No* ***names*** *should be written anywhere on the examination book.*
4. *Ensure that your* ***ID number*** *is indicated on all pages of the examination answer booklet.*
5. *Ensure your work is* ***clear*** *and* ***readable****. Untidy work shall be penalized*
6. *Any type of examination Malpractice will lead to automatic disqualification*
7. *Do not write anything on the questions paper.*

QUESTION ONE

a) Define the following terms:
   i. Inferential Statistics (2marks)
   ii. Statistic (2 marks)
   iii. Parameter (2 marks)
   iv. Interval estimate (2 marks)
   v. Confidence level (2 marks)

b) The following is a random sample of the number of tuberculosis infected people in 31 districts in Uganda.

47 95 54 33 64 4 8 57
8 90 3 49 4 44 79 80
68 7 15 21 52 6 78 109
29 80 16 50 9 48 40

Estimate the mean number of infection per district with 90% confidence (10 marks)

QUESTION TWO

a) Briefly define a Point Estimate and give three characteristics of a good estimator (5 marks)
   i. Identify the three methods of testing hypothesis (3 marks)
   ii. Describe the two types of Statistical hypothesis (2 marks)

b) A psychologist feels that playing soft music during an exam will change the results of the test. The psychologist is not sure whether the grades will be higher or lower. In the past, the mean of the scores was 60. State the hypothesis (2 marks)

c) A researcher reports that the average salary of workers in a certain company is more than $42,000. A sample of 30 workers has a mean salary of $43,260. At $\alpha = 0.05$, test the claim that workers earn more than $42,000 a year given that the standard deviation of the population is $5,230 (8 marks)

QUESTION THREE

a) Define the Maximum error of estimate (1 mark)
   i. A university dean of students wishes 2 estimate the average number of hours students spend doing assignments per week. The standard deviation is 6.2 hours. How large a sample must be selected if he wants to be 99% confident of finding whether the true mean differs from the sample mean by 1.5 hours?(3 marks)

b) State the characteristics of the t-distribution (5 marks)
   i. The data represents a sample of home fires started by candles for the past years.
   5460 3900 6090 6310 7160 8440 9930

Find the 99% confidence interval for the mean number of home fires started by the candles each year (5 marks)

c) The proportion of students in private schools is around 11%. A random sample of 450 students from a wide geographical area indicated that 55 attended private schools. Estimate the true proportion of students attending private schools with a 95% confidence. How does your estimate compare to 11%? (6 marks)

QUESTION FOUR

a) Define Confidence interval (2 marks)

i. State the central limit theorem (2 marks)

ii. The following data represents a sample of assets of 30 credit unions.

12.28	8.74	16.56	3.17	4.39	16. 85
2.89	7.92	1.24	4.78	2.17	21.58
13.19	40.22	9.16	2.42	14.64	12.24
73.25	5.01	1.91	1.47	1.06	1.42
11.59	2.27	6.69	12.77	18.13	2.76

Find the 90% confidence interval of the mean (8 marks)

b) A study of 415 university students showed that they have seen on average of 5000 hours of television. If the sample standard deviation is 900. Find the 90% confidence interval of the mean for all students. If a parent claimed that his children watched 4000 hours, would the claim be believable? (8 marks)

QUESTION FIVE

a) Define the following

i. Test value (1 mark)

ii. Level of significance (1 mark)

iii. Critical value (1 mark)

iv. Type 1 error (1 mark)

v. Type II error (1 mark)

b) State the steps involved in hypothesis testing (5 marks)

c) A researcher estimates that the average revenue of the largest business in USA is greater than $ 24 billion.

A sample of 50 companies is selected and the revenue is shown.

178 31 122 30 91 19 44 19 35 19

61 24 36 16 46 15 20 15 32 19

30 25 28 25 28 18 20 14 27 15

29 24 16 23 16 17 19 17 15 22

41 22 38 21 36 20 15 17 25 20

At $\alpha = 0.05$, is there enough evidence to support the researcher's claim? (10 marks)

QUESTION SIX

a) State the five steps involved in P-value method (5 marks)

b) A researcher claims that the average wind speed in a certain city is 8 mile per hour. A sample of 32 days has an average wind speed of 8.2 miles per hour. The standard deviation of the sample is 0.6 miles per hour. At $\alpha=0.05$, is there enough evidence to reject the claim? (5 marks)

c) A manager states that in his factory, the average number of days per year missed by employees due to illness is less than the national average of 10. The following data show the number of days missed by 40 employees last year.

0 3 7 2 3 6 9 4 5 11 1 5
12 6 7 10 8 3 0 1 5 2 4 9
5 6 8 3 4 4 3 12 2 1 3 3
7 0 15 2

Is there sufficient evidence to believe the manager's statement at $\alpha = 0.05$? Use the P-value method (10 marks)

QUESTION SEVEN

a) A researcher collects the following data and determines that there is a significant relationship between the age of the copier and its monthly cost. The regression line is $y' = 55.57 + 8.13x$. Find the Standard Error of estimate(10 marks)

MACHINE	AGE(X/yrs)	MONTHLY COST(y)
A	1	62
B	2	78
C	3	70
D	4	90
E	4	93
F	6	103

b) The director of an alumni association for a certain college wants to determine whether there is any type of relationship between the amount contributed and the years the alumnus has been out of school.

Years(x)	1	5	3	10	7	6
Contribution(y)	500	100	300	50	75	80

i. Draw the scatter plot for the variables (3 marks)

ii. Compute the value of the correlation coefficient (3 marks)

iii. Test the significance of the correlation coefficient at $\alpha = 0.05$ using P-value method (4 marks)

END