



## KENYATTA UNIVERSITY

UNIVERSITY EXAMINATIONS 2011/2012

SECOND SEMESTER EXAMINATION FOR THE DEGREE OF MASTER OF  
SCIENCE (FOODS, NUTRITION AND DIETETICS)

HFN 503N: APPLIED STATISTICS

DATE: Friday 30<sup>th</sup> March 2012

TIME: 9.00a.m – 12.00p.m

INSTRUCTIONS

Answer Question ONE AND ANY OTHER TWO QUESTIONS

**QUESTION ONE - (30 MARKS)**

- a) The level of vitamin C in a fruit sample was determined by two methods (Method 1 and Method 2). The following tabulated results were obtained.

|          | Level of Vitamin C (Mg/L) |      |      |      |      |      |      |      |
|----------|---------------------------|------|------|------|------|------|------|------|
| Method 1 | 15.1                      | 14.5 | 15.0 | 15.5 | 15.7 | 15.9 | 15.5 | 16.0 |
| Method 2 | 16.5                      | 14.2 | 18.2 | 19.0 | 15.8 | 17.2 | 18.8 | 20.5 |

- i) Using suitable parametric method, establish whether the results which were obtained using Method 2 were significantly higher than those obtained using method 1 at  $p \leq 0.01$  level of significance. [12marks]
- ii) Was there any evidence of systematic error in the results obtained by Method 2 at  $p \leq 0.01$  level of assuming that Method 1 was the standard Method for determining Vitamin C. [3marks]
- iii) Converting the data obtained using Method 2 to natural log values, determine whether the largest value after this conversion is an outlier at the 95% confidence level. [3marks]
- b) Explain the following terms
- i) Degree of freedom [2marks]

- ii) Standard error of the mean [1mark]
- iii) Accuracy [1mark]
- c) Distinguish between the following terms:
- i) Histogram and frequency polygon [2marks]
- ii) Discrete and continuous variable [2marks]
- iii) Systematic and random errors [2marks]
- d) Consider the following data

| $X_i$ | Frequency |
|-------|-----------|
| 2     | 3         |
| 4     | 3         |
| 6     | 4         |
| 8     | 6         |
| 10    | 5         |

Calculate the mean. [2marks]

## QUESTION TWO (20 MARKS)

- a) Consider the following % alcohol in beer samples using two methods (Method A and Method B).

| Sample   | 1   | 2   | 3   | 4   | 5   | 6   | 7   | 8   | 9   | 10  | 11  |
|----------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Method A | 5.2 | 5.5 | 3.0 | 4.2 | 4.5 | 4.2 | 6.5 | 2.5 | 3.1 | 4.2 | 5.2 |
| Method B | 6.0 | 5.8 | 4.5 | 3.9 | 4.8 | 4.5 | 6.2 | 2.0 | 3.5 | 3.5 | 4.0 |

Determine whether there was any significant difference in the results obtained by the two methods (A and B) at the 95% confidence level using

- i) A suitable parametric method. [5marks]
- ii) A non-parametric method sum rank test. [3marks]
- b) Distinguish between parametric and non-parametric methods used for statistical analysis of data. [2marks]

c) Consider the following data:

|    |    |    |    |    |
|----|----|----|----|----|
| 28 | 52 | 70 | 48 | 18 |
| 14 | 18 | 11 | 54 | 24 |
| 15 | 40 | 28 | 63 | 22 |
| 12 | 25 | 56 | 5  | 37 |

- i) Using six classes, construct frequency distribution table using the above data, showing class limit, class boundary, class mark, frequency of each class, cumulative frequency and relative frequency. [3marks]
- ii) Calculate the following using the frequency distribution in part (i)
  - a) 72<sup>nd</sup> percentile [2marks]
  - b) Mode [2marks]
  - c) Variance [3marks]

### QUESTION THREE ( 20 MARKS)

The table below shows the amount of money paid for respective amount of beef sold in one of the supermarkets in Kahawa Sukari

|                     |   |     |     |     |      |      |      |
|---------------------|---|-----|-----|-----|------|------|------|
| Amount in Kshs      | 0 | 300 | 590 | 890 | 1190 | 1490 | 1780 |
| Amount of beef (Kg) | 0 | 2   | 4   | 6   | 8    | 10   | 12   |

- a) Indicate which variable in this question is
  - i) Dependent
  - ii) Independent variable. Give reason(s) for your answer. [2marks]
- b) Is the relationship between the dependent and independent variables in this question liner? Use suitable statistical method to establish this. [3marks]
- c) Obtain the regression equation for the best line of fit. [6marks]
- d) State the assumptions you have made in getting the regression equation in part (c) comment on the validity of these assumptions. [3marks]
- e) Plot the points on the graph paper provided and draw the best line of fit. [2marks]
- f) Calculate the error of the y-intercept. [4marks]

#### QUESTION FOUR (20 MARKS)

- a) The distribution of marks for 180 students in statistics exams showed normal distribution, with mean marks of 52 and standard deviation of 5. Calculate:
- i) The number of students who scored between 48 marks and 62 marks. [2marks]
  - ii) The number of students who obtained between 55 marks and 65 marks. [2marks]
  - iii) The number of students who failed in this examination if the pass mark was 40 marks. [1marks]
  - iv) The percentage of students who obtained higher than 65%. [1mark]
- b) In an MSc class of 150 students, 40% were women. 60% of women and 40% of men in this class indicated that they prefer taking up a teaching job after completing their course while the rest of women and men indicated that they did not. Determine whether preference of taking up a teaching by the students was dependent on gender at the 95% confidence level. [6marks]
- c) In one of the campuses of Kenyatta University with 60 academic staff, the choice of religion by the academic staff is tabulated below

| Religion    | Number of academic staff |
|-------------|--------------------------|
| Protestants | 20                       |
| Catholics   | 25                       |
| Muslims     | 10                       |
| others      | 5                        |

- Determine whether the choice of religion of preference by the academic staff is statistically significant at the 99% confidence level. [4marks]
- d) The level of silver in photographic waste before treatment and after treatment are tabulated below

|                  | Level of Silver (ppm) |      |      |      |      |      |      |      |      |
|------------------|-----------------------|------|------|------|------|------|------|------|------|
| Before Treatment | 18.5                  | 16.5 | 14.5 | 12.8 | 15.5 | 17.2 | 16.8 | 15.5 | 17.5 |
| After treatment  | 14.5                  | 15.5 | 14.8 | 10.2 | 14.0 | 17.5 | 15.6 | 14.8 | 15.5 |

Determine whether treatment significantly reduce the level of silver in the photographic waste at the 95% confidence level using suitable non-parametric method. [4marks]