



KENYATTA UNIVERSITY

UNIVERSITY EXAMINATIONS 2010/2011

INSTITUTIONAL BASED PROGRAMME (IBP) DECEMBER SESSION

EXAMINATION FOR THE DEGREE OF BACHELOR OF SCIENCE

HFN 321: INTRODUCTION TO STATISTICS

DATE: Thursday 28th April, 2011

TIME: 4.30 p.m. – 6.30 p.m.

INSTRUCTIONS:

Answer question **ONE** and any other **TWO** questions.

Use the following information where appropriate:

$$t_{tab}(at\ n - 1 = 4) = 2.7764 \text{ at } 95\% \text{ confidence level.}$$

$$t_{tab}(at\ df = 8) = 2.3060 \text{ at } 95\% \text{ confidence level}$$

$$F_{tab}(at\ df\ 4,4) = 9.364 \text{ at the } 95\% \text{ confidence level}$$

$$t_{tab}(at\ n = 5) \text{ SIGNTEST} = 0 \text{ at } 95\% \text{ confidence level}$$

$$Q_{tab}(at\ n = 4) = 0.835 \text{ at the } 95\% \text{ confidence level}$$

QUESTION ONE (30 Marks)

- (a) Distinguish between the following terms
- (i) Sample and population (2 Marks)
 - (ii) Gross error and systematic error (2 Marks)
 - (iii) Bar chart and histogram (2 Marks)
 - (iv) Accuracy and precision (2 Marks)
 - (v) Continuous and random variables. (2 Marks)
- (b) A half litre packet of milk measured by three people (Ted, Juliet and Edith) using measuring cylinder. The results obtained are tabulated below.

	Volume of milk (ml)		
Ted	455	460	455
Juliet	450	460	470
Edith	450	449	451

(i) Provide descending order of

(a) Precision (2 Marks)

(b) Accuracy (3 Marks)

Of results obtained by the three people, giving reasons for your answer.

(ii) Establish whether the results obtained by Edith have evidence of systematic error at the 95% confidence level.

$t_{cel} (df=2) = 4.3027$ at the 95% confidence level. (4 Marks)

(c) Consider the following data

10.5, 11.2, 9.3, 10.5, 9.8, 9.3, 10.8

(i) Calculate the

(a) Mean (1 Mark)

(b) Median (1 Mark)

(c) Mode (1 Mark)

(ii) If the frequency of each value in the data above is tripled, calculate the new mean, mode and median. (3 Marks)

(iii) If the highest value in the above data is replaced by 151.2 calculate the new mean, mode and median. Comment on your answer as compared to the answer you obtained in part (i). (5 Marks)

QUESTION TWO (20 Marks)

(a) Consider the following distribution of data

Class limit	31-36	25-30	19-24	13-18	7-12
Frequency	8	14	10	6	2

Calculate:

(i) 78th percentile (3 Marks)

(ii) Mode (2 Marks)

- (iii) Mean (2 Marks)
- (iv) Variance (4 Marks)
- (b) Draw a frequency polygon using the tabulated data above. (2 Marks)
- (c) The weight of children aged 2 to 5 are normally distributed with mean value of 8.5kg and standard deviation of 2.0kg.
- (i) Find the percentage of such weights between 6.5 and 12.5kg. (2 Marks)
- (ii) Find the fraction of such children with the weights above 10.5. (2 Marks)
- (iii) Among 400 children aged 2 to 5 years, how many are expected to weigh between 4.5kg and 14.5kg. (3 Marks)
- Given: For $Z = 1$ area under the curve = 0.3413
 For $Z = 2$ area under the curve = 0.4772
 For $Z = 3$ area under the curve = 0.4987
 for the normal distribution curve

QUESTION THREE (20 Marks)

- (a) Consider the following results obtained for the sale of sugar in one of the supermarkets in Nairobi.

Amount of sugar (kg)	0	1	2	3	4	5
Cash (KShs)	0	200	390	590	790	990

- (i) Using suitable statistical method, determine whether the relationship between the dependent variable and independent variable given in the above table is linear. (3 Marks)
- (ii) Work out the regression equation for the best line of fit, using suitable statistical method. (8 Marks)
- (iii) List three assumptions used in getting the best line (of fit). Comment on the validity of these assumptions. (4 Marks)

- (iv) Calculate the amount of sugar that would have been sold if KSh650 had been paid for it. (2 Marks)

(b) Indicate the scale of measurement of the following

- (i) Height of children
- (ii) Freezing point of water
- (iii) Rank in the Police Force
- (iv) Sex
- (v) Grade in Exam
- (vi) Velocity of light.

(3 Marks)

QUESTION FOUR (20 Marks)

- (a) Five samples of honey had the following tabulated % glucose using Method A and Method B.

Sample	% Glucose	
	Method A	Method B
A	35.1	34.7
B	34.0	33.5
C	33.5	33.5
D	34.5	34.7
E	33.2	32.9

Using suitable

- (i) Parametric method (8 Marks)
- (ii) Non-parametric method (4 Marks)

Establish whether there was significant difference between the results obtained using the two Methods (A and B) at the 95% confidence level.

- (b) Consider the following level of butterfat in the milk sample, determined four times:
4.5, 4.0, 3.8, 5.6.

Is the largest value on outlier at the 95% confidence level. (3 Marks)

(c) Distinguish between the following

- (i) Parametric and non-parametric methods. (3 Marks)
- (ii) Mean and mode. (2 Marks)