



KENYATTA UNIVERSITY
UNIVERSITY EXAMINATIONS 2016/2017
CITY CAMPUS

FIRST SEMESTER EXAMINATION FOR THE DEGREE OF MASTER OF SCIENCE
(FOOD, NUTRITION AND DIETETICS/FASHION, DESIGN AND MARKETING)

HCU 801: STATISTICS

DATE: THURSDAY 1ST DECEMBER 2016

TIME: 5.30 P.M. - 8.30 P.M.

INSTRUCTIONS: ANSWER QUESTION ONE AND ANY OTHER TWO QUESTIONS

QUESTION ONE (30 MARKS) COMPULSORY

- a) Explain the characteristics of parametric and non-parametric methods (5 marks)
- b) The costs of milk corresponding to the volume are tabulated below:

Volume of milk (ml)	0	400	800	1200	1600	2000	2400
Cost in KSh.	0	40	78	118	155	195	235

- i. Determine using suitable statistical method the strength of relationship between the variables, volume of milk and cost of milk. (3 marks)
- ii. Using suitable statistical method, obtain the regression equation for the best line of fit. (6 marks)
- iii. Plot the points on a graph paper provide using the above data, and draw the best line of fit. (3 marks)
- iv. What would be the cost of 1.85 liters of milk? (1 mark)
- v. State the assumptions made in getting the regression equation for the best line of fit and comment on their validity. (3 marks)
- vi. Calculate the 95% confidence interval for the slope of the regression line. (3 marks)

c) Explain the following:

- i. Standard error of the mean (1 mark)
- ii. Degree of freedom (1 mark)
- iii. Normal distribution curve and its characteristics (2 marks)
- iv. Coefficient of determination (1 mark)

0.7 x 1.85

QUESTION TWO

a) Distinguish between the following statistical terms:

i. Gross and systematic errors (2 marks)

ii. Accuracy and precision (2 marks)

b) Consider the following results obtained by two methods, ~~X~~ and ~~Y~~ for the level of butterfat in milk samples.

Milk samples	1	2	3	4	5	6	7	8	9	10
% butterfat using method A	5.0	5.6	4.0	5.2	6.2	5.4	6.2	4.4	5.2	5.3
% butterfat using method B	4.5	4.5	4.3	5.0	5.4	5.6	5.2	4.0	4.8	5.0

Using suitable statistical non-parametric test determine whether the results obtained by method A is significantly higher than those obtained by method B at $P \leq 0.01$ level of significance. (5 marks)

c) Consider the following distribution of data:

13.5, 11.8, 12.5, 7.2, 12.3, 14.5, 13.1

Convert the above data to natural logarithm values and test whether the smallest value after conversion to natural logarithm values is an outlier at the 95% confidence level. (3 marks)

d) Using the data in part (c) above, obtain the 95% confidence level within which the population mean is certain to lie. (3 marks)

e) Consider the following levels of fructose in the honey samples from different countries determined by two methods (X and Y).

	Level of fructose (%)						
Countries	Zambia	Kenya	Tanzania	Britain	Canada	Israel	Uganda
Method X	42.5	41.5	45.8	43.5	40.5	47.8	42.9
Method Y	45.5	46.5	44.8	47.5	46.5	45.5	45.8

Using suitable parametric method, determine whether there is any significant difference in the level of fructose in the honey samples from the countries indicated in the above table obtained using methods X and Y at the 95% confidence level. (5 marks)

QUESTION THREE

a) In a class of 200 MSc. Students, 60% are women. The students were asked to indicate whether they would like to do a PhD programme after completing their MSc course. 55% of women and 60% of men indicated that they would like to do their PhD course after completion of their MSc course. The rest of the students indicated that they would not like to undertake PhD programme after completion of their MSc course. Using suitable statistical test, determine

whether the decision to undertake PhD programme by the MSc students is independent of gender at the $p=0.01$ level of significance. (6 marks)

- b) The following tabulated results show the level of vitamin C in orange juice sample at different temperatures (25°C , 50°C and 80°C)

Level of vitamin C (mg/L)		
25°C	50°C	80°C
15.5	14.2	9.5
17.5	15.2	13.0
15.8	13.6	10.5
16.5	13.0	7.5
16.0	14.2	8.9

Determine whether there is significant difference in the results for vitamin C in the orange samples at temperatures 25°C , 50°C and 80°C at $P \leq 0.05$ level of significance. (10 marks)

- c) The performance of 300 MSc students in an examination conformed to the normal distribution, with a mean of 52 and standard deviation of 3.8. Calculate:
- The % probability of the students obtaining a pass mark of 40% (2 marks)
 - The number of students who obtained at least 42% in the examination. (2 marks)

QUESTION FOUR

- a) The level of silver in a photographic waste before and after treatment are tabulated below:

	Concentration of silver (mg/L)						
Before treatment	12.5	13.5	12.5	14.5	13.8	14.2	12.7
After treatment	10.2	9.5	12.8	11.5	7.2	9.2	7.2

Using suitable non-parametric method, establish whether treatment significantly reduced the level of silver in photographic waste at the 95% confidence level. (8 marks)

- b) Consider the following distribution of data:

Class limit	52-68	43-51	29-42	22-28	11-21	4-20
Frequency	3	4	8	6	5	4

- Calculate
 - Mean (2 marks)
 - 70^{th} percentile (2 marks)
 - Quartile deviation (3 marks)
 - Standard deviation (3 marks)
- Draw a histogram and frequency polygon using above data on a graph paper provided using common axes. (2 marks)