



KENYATTA UNIVERSITY

UNIVERSITY EXAMINATIONS 2011/2012

FIRST SEMESTER EXAMINATION FOR THE DEGREE OF MASTER OF SCIENCE IN SOCIAL STATISTICS

SMA/SS/562: SURVEY DATA ANALYSIS

DATE: Wednesday 30th November, 2011

TIME: 5.30 p.m. – 8.30 p.m.

INSTRUCTIONS

Answer question One and any other TWO questions.

Question 1 (30 Marks).

- (a) (i) Differentiate between type I and type II error as applied in hypothesis testing [2 marks]
- (ii) Define the confidence interval for an interval estimate [2 Marks]
- (b) A random sample of four items taken from a normal population with mean μ and variance 4.5 cm^2 gave the following data: 12.4 12.6 11.9 13.2 Construct the 99% confidence interval for mean μ [5 Marks]
- (c) On the basis of observations made on 39 cotton plants, the correlation of yield of cotton (X_1), number of bolls (X_2) and height (X_3) are found to be:
- $r_{12} = 0.7$, $r_{13} = 0.61$, $r_{23} = 0.4$. Find the value of
- (i) $r_{23 \cdot 1}$ [3 Marks]
- (ii) $R_{1 \cdot 23}$ [4 Marks]
- (d) A manufacturing company believes that the batteries for laptop computers produced by process A on average last longer than batteries for laptop computers produced by process B. To test this belief, random samples of laptop computer batteries produced by the two processes were tested and yielded the results below

Process	Sample size	Average lifetime (in minutes of usage)	Standard Deviation in Km
A	64	106	12
B	64	100	10

Determine whether there is evidence at 5% level of significance that the firm is correct in its belief. [6 Marks]

(e) Let X be a random variable observation $x_1, x_2, x_3, x_4, \dots, x_n$ suppose further that the observations occur with frequencies $f_1, f_2, f_3, f_4, \dots, f_n$. Show that

$$\frac{\sum f(x - \bar{x})^2}{\sum f} = \frac{\sum f(x - \bar{x})^2}{n}, \text{ where } \bar{x} = \text{mean} \quad [5 \text{ Marks}]$$

(f) Explain three assumptions in analysis of variance [3 Marks]

Question 2 (20 Marks)

(a) Two hundred randomly selected adults were asked whether TV shows as a whole are primarily entertaining, educational or a waste of time. Only one answer could be chosen. Respondents are categorized by gender. The following shows responses from adults questioned.

Gender	Responses		
	Entertaining	Educational	Waste of time
Female	52	28	30
Male	28	12	50

At 0.05 significance level, determine whether the responses and gender are related [6 Marks]

(b) The average commission charged by full service brokerage firms on a sale of common stock is \$ 84, and standard deviation is \$ 10. Fatuma has taken a random sample of 75 traders by her clients and determined that they paid an average commission of \$ 81.50. At 0.10 significance level, can Fatuma conclude that her clients' commissions are higher than the industry average? [7 Marks]

(c) A country's Association of Retired Persons reported that typical senior citizens living at high altitudes claim their blood pressure is lower than the average of 160. To test this claim, 16 senior citizens living at high altitudes were selected at random and their blood pressures checked. Their mean pressure was 151; the standard deviation of the sample was 12. Do senior citizens living at high altitudes have significantly lower blood pressure as claimed? Use 0.05 level of significance. [7 Marks]

Question 3 (20 Marks)

(a) Differentiate between correlation and regression analysis for a multivariate distribution [2 Marks]

(b) A sample survey of 5 families was taken and the figures obtained with respect to their annual savings x_1 (in Kshs 100's), annual income x_2 (in Kshs 1000's), and family size x_3 . The data is:

Family	Annual Savings (x_1)	Annual Income (x_2)	Family size (x_3)
1	10	16	2
2	4	13	6
4	5	15	5
5	8	13	3

(i) Fit the least – square regression equation of x_1 on x_2 and x_3 [8 Marks]

(ii) Estimate x_1 given that $x_2=11,000$ and $x_3=8$. [2 Marks]

(b) Construct the 95% confidence interval for the difference between two means $\mu_1 - \mu_2$ [5 Marks]

(c) A detergent manufacturing company claims that its top quality detergent is good, on average for at least 240 months before expiry. A consumer protection agency tested 50 such detergents and found that on average they last for 242 months with a standard deviation of 8 months. Determine the 97% level of significance for the population mean [3 Marks]

Question 4 (20 Marks)

(a) The following data relates to random variable X and Y

X	1.7	2.2	2.3	2.6	2.7	3.0	3.2	3.3	4.1	4.3	4.6	5.7	6.1
Y	1.3	1.8	1.6	2.0	2.1	2.2	3.0	2.6	4.1	3.7	5.0	5.8	5.3

(i) Estimate the parameter of a power function model [3 Marks]

(ii) Test $H_0 H_0 : \beta = \frac{3}{4}$ versus $H_a : \beta < \frac{3}{4}$, using a level 0.05 test [5 Marks]

(b) To test the significance of possible variation in performance in statistics test between different groups of students, a common tests was administered to a number of students in each group. The results are given below.

Group of students			
A	B	C	D
8	12	18	13
10	11	12	9
12	9	16	12
8	14	6	16
7	4	8	15

Make an analysis of variance of the data at 5% level.

[12 Marks]