



KENYATTA UNIVERSITY

UNIVERSITY EXAMINATIONS 2011/2012

SECOND SEMESTER EXAMINATION FOR THE DEGREE OF BACHELOR OF
SCIENCE (SOFTWARE ENGINEERING)

SSE 313: STATISTICS & EMPIRICAL METHODS

DATE: WEDNESDAY, 18TH APRIL 2012

TIME: 11.00 A.M. – 1.00 P.M.

INSTRUCTIONS: Answer Question 1 and any other two questions.

Q1. (a) Let X be a random variable with the following probability distribution

x	-3	6	9
$p(x)$	$\frac{1}{6}$	$\frac{1}{2}$	$\frac{1}{3}$

Obtain the mean and variance of X

(3marks)

(b) The continuous random variable X has pdf

$$f(x) = \begin{cases} \frac{1}{5}x + k, & 0 \leq x \leq 3 \\ 0, & \text{otherwise} \end{cases}$$

where k is a constant.

Find (i) the value of k and determine

(ii) $P(1 < X < 5)$

(4marks)

(c) Scores on the SAT verbal test in recent years follow approximately the normal distribution with mean 505 and standard deviation of 110. How high must a student score in order to be placed in the top 20% of all the students taking the SAT

(4marks)

(d) It is known that 5% of switches from a large consignment of switches are bad.

If an engineer chooses a sample of 15 switches from a shipment of 10,000

switches what is the probability that no at least two switches are bad? (4marks)

(e) People entering a museum are counted by the warden on duty at the door. The number of people entering in any time interval of length t minutes has a Poisson distribution with mean $0.5t$. The warden leaves the front door unattended for three minutes.

Calculate to 3 decimal places that

(i) no one

(ii) two or more people will enter the museum during this time

(5marks)

(f) Complete the following analysis of variance table for the randomized block design

Source of variation	Degrees of freedom	Sum of squares	Mean sum of squares	Variance ratio
blocks	--	42.09	--	--
treatments	3	--	--	--
error	--	--	1.59	
total	23	81.86		

Test the hypothesis that the treatment effects are equal to zero. (5marks)

(g) Define reliability in terms of probability and for engineering purposes. What is the main key element of this definition? (2marks)

(h) What is cost estimation? Give any two common types of cost estimation. (3marks)

Q2 (a). The following are the scores obtained in a personality test by samples of eight married and unmarried women.

Unmarried(X)	4.4	3.4	3.8	4.6	4.0	3.9	3.5	3.6
Married(Y)	3.6	3.8	3.2	3.5	3.1	3.5	3.5	3.6

Calculate the correlation coefficient between the two test scores and test whether they are independent at 5% level of significance given that

$$\sum x = 31.2, \quad \sum x^2 = 122.94 \quad \sum y = 27.8 \quad \sum y^2 = 96.96 \quad \sum xy = 108.28$$

(7marks)

(b) When a standard lubricant is used on the bearings of a wheel, the length of time for which the wheel will spin in an experiment has a mean of 210s and standard deviation of 9s. A new 'improved' lubricant is tried out and five spins take time in seconds:

219 214 196 229 212

Test, assuming that the standard deviation is unchanged, whether new lubricant has decreased the spinning at 1% level of significance (7marks)

- (c). The owner of a big farm agrees to purchase the product of a factory, if the produced items have a standard deviation of less than 0.5 square mm in their length. To make sure of the specifications, the buyer selects a sample of 18 items from this lot. The length of each item was measured and the following were obtained; $\sum x = 331.69$ $\sum x^2 = 6112.64$. Should the buyer purchase the lot? Test at 5% level of significance. Assume that lengths follows a normal distribution (6marks)

Q3(a). Test scores in mathematics, in two samples of students of the same class from two schools, A and B, were as given below:

A	65	45	63	47	84	37	48	52	95	26	68
B	57	39	44	76	49	83	91	68			

Can it be concluded that on average, the performance of students in the two schools is at par. What assumptions do you make? Use $\alpha = 0.05$ (10marks)

- (b). Six different machines are being considered for use in manufacturing a certain product. The machines are being compared with respect to the tensile strength of the product. A random sample of four items from each machine is used to determine whether or not the mean tensile strength varies from machine to machine. The following are measurements in kg per square cm

machines						
	1	2	3	4	5	6
	35.1	32.8	40.4	29.2	33.5	18.3
	33.8	38.4	32.5	33.4	34.8	16.3
	31.6	37.2	35.4	46.1	35.3	17.6
	35.6	30.7	38.7	37.8	41.1	20.1
total	136.1	139.1	147	146.5	144.7	72.3

Analyze the data and draw relevant conclusions (take $\alpha = 0.05$) (10 marks)

Q4 (a) The following table gives the gain yields of 4 varieties of wheat W_1, W_2, W_3 and W_4 arranged in five randomized blocks.

Varieties	1	2	3	4	5	Total
W_1	15.3	25.1	20.2	21.2	18.5	100.3
W_2	10.6	18.7	16.0	18.6	16.3	80.2
W_3	11.5	18.4	17.4	12.6	19.7	79.6
W_4	10.0	18.0	11.8	18.8	11.7	70.3
Total	47.4	80.2	65.4	71.2	66.2	

At 5% level of significance test whether

- (i) the four varieties of wheat differ in their yields
- (ii) blocks are significantly different (12marks)

(b) The following data show the place of wedding ceremony against length of courtship, for a random sample of 140 samples.

Place of ceremony	Length of courtship		
	< 6 months	6 -12	> 12 months
Registry office	39	20	14
Church	19	13	35

Can it be said that place of marriage is independent of length of courtship? Use 5% level of Significance (8marks)

Q5. (a) **Q4.** The hardness, Y , of egg shells laid by a certain breed of chicken is thought to be linearly related to the amount, X , of a certain food supplement put in their diet. The model assumed linear is $y = \alpha + \beta x + e$ where $e \sim N(0, \sigma^2)$. Using the data below, test the hypothesis test whether there is a linear relationship between X and Y , i.e. $H_0 : \beta = 0$ against $H_1 : \beta \neq 0$ at 5% level of significance.

X	0.12	0.21	0.34	0.61	0.13	0.17	0.21	.034	0.71	0.64
Y	1.95	1.75	1.24	0.95	0.82	0.76	1.75	0.66	0.98	0.7

(12marks)

(b) The lengths of mass-produced components are normally distributed with mean 15cm and standard deviation 0.03cm. Each component cost Ksh8 to produce. A component shorter than 14.95cm has to be scrapped, whereas a component larger than 15.05cm can be trimmed at a further cost of Ksh3. components having lengths between these limits are satisfactory and cost the manufacturer nothing extra. Acceptable components are then sold at Ksh15.

Of 1000 components made, how many of them would you expect to

- (i) be scrapped
- (ii) be acceptable after trimming
- (iii) be immediately acceptable.

What is the expected total profit for the 1000 components sold? (8marks)

