



TECHNICAL UNIVERSITY OF MOMBASA

FACULTY OF APPLIED AND HEALTH SCIENCES

MATHEMATICS AND PHYSICS

UNIVERSITY EXAMINATION FOR:

BACHELOR OF SCIENCE IN MATHEMATICS AND COMPUTER SCIENCE/

BACHELOR OF SCIENCE IN STATISTICS AND COMPUTER SCIENCE

AMA 4107: PROBABILITY AND STATISTICS I

END OF SEMESTER EXAMINATION

SERIES: MAY SERIES

TIME: 2 HOURS

DATE: MAY 2016

Instructions to Candidates

You should have the following for this examination

-Answer Booklet, examination pass and student ID

This paper consists of 5 questions. Attempt ONE AND ANY TWO.

Do not write on the question paper.

QUESTION ONE

- a. State any three methods of data collection (3 marks).
- b. In a series of 20 spot checks the following number of passengers were counted at a certain depot;

137	136	135	136	135
135	137	138	136	137
136	136	138	137	136
137	136	136	138	135

Using this data set determine;

- i. Mean (3marks)
- ii. Median (2marks)
- iii. Mode (1 mark)
- c. A total of 28% of males living in Old Town smoke cigarettes, 6% smoke cigars and 3% smoke both cigars and cigarettes. What percentage of males smokes neither cigars nor cigarettes? (3marks)
- d. State any three levels of measurement scales. (3 marks)

- e. In a certain school class, consisting of 60 girls and 40 boys, it is observed that 24 girls and 16 boys wear eye glasses. What is the probability that a student picked at random wears eye glasses given that the student is a boy? (3marks)
- f. State the assumptions of linear regressions (3 marks)
- g. At a certain stage of a criminal investigation, the inspector in charge is 60% convinced of the guilt of a certain suspect. Suppose now that a new piece of evidence that shows that the criminal has a certain characteristic is uncovered. If 20% of the population possesses this characteristic, how certain of the guilt of the suspect should the inspector now be if it turns out the suspect is among this group? (4marks)
- h. Using the set of data given below, construct a stem and a leaf diagram (5 marks)

40 22 50 61 30 58 51 75
 58 70 23 59 49 55 63 38
 37 57 53 41 60 57 52 77
 21 62 32 83 48 73 28 31
 48 76 64 39 56 25 42 63

QUESTION TWO

The following table shows production units (X) and costs (Y)(%) of a certain company

X(000 units)	10	4	6	9	10	8	5	7	11	12
Y (\$000)	15	11	12	19	22	20	16	13	24	20

- a. Draw a scatter diagram and interpret your diagram (3 marks)
- b. Determine the Pearson correlation coefficient between X and Y and interpret your answer (6 marks)
- c. Determine the coefficient of determination and interpret your answer (3marks)
- d. Fit a simple linear regression model between X and Y (5 marks)
- e. Predict the cost of production for the next one month when output is scheduled to be 10,000 (3 marks)

QUESTION THREE

- a. Disks of polycarbonate plastic from a supplier are analyzed for scratch and shock resistance. The results from 100 disks are summarized below

		Shock resistance	
		High	Low
Scratch resistance	High	70	9
	Low	16	5

Let A denote the event that a disk has high shock resistance, and let B denote the event that a disk has high scratch resistance. Determine;

- i. Probability of A and B (2 marks)
- ii. Probability of A or B (2 marks)
- iii. Conditional probability of A given B (2marks)

- iv. Are the events A and B independent? (3 marks)
- b. Suppose 2% of cotton fabric rolls and 3% of nylon fabric rolls contain flaws. Of the rolls used by a manufacturer, 70% are cotton and 30% are nylon. What is the probability that a randomly selected roll used by the manufacturer contains flaws? (3marks)
- c. Samples of laboratory glass are in small, light packaging or heavy, large packaging. Suppose that 2 and 1% of the sample shipped in small and large packages, respectively, break during transit. If 60% of the samples are shipped in large packages and 40% are shipped in small packages, what proportion of samples break during shipment? (3marks)
- a. Because a new medical procedure has been shown to be effective in the early detection of an illness, a medical screening of the population is proposed. The probability that the test correctly identifies someone with the illness as positive is 0.99, and the probability that the test correctly identifies someone without the illness as negative is 0.95. The incidence of the illness in the general population is 0.0001. You take the test, and the result is positive. What is the probability that you have the illness? (5marks)

QUESTION FOUR

The following are the number of babies born during a year in 60 community hospitals.

30	55	40	58	54	45	49	32	57	47
37	26	59	46	48	56	59	35	46	24
32	56	43	56	42	32	57	29	42	42
39	57	45	54	31	54	53	59	30	53
52	27	34	53	53	21	34	28	50	22
55	52	28	49	54	31	24	24	57	29

- a. Using Sturges' rule construct a frequency distribution table (6marks)
- b. Using the table in part a above, calculate;
- The lower quartile (3 marks)
 - middle quartile (3 marks)
 - upper quartile (3 marks)
- c. construct a box and a whisker plot (5 marks)

QUESTION FIVE

The following table shows the age distribution of cases of certain disease reported during a year in a particular state.

age	5-14	15-24	25-34	35-44	45-54	55-64
No of cases	5	10	20	22	13	5

Using the above data compute

- the mean (4marks)
- median (3 marks)

- c. mode (3marks)
- d. standard deviation (5marks)
- e. coefficient of variation (2marks)
- f. Pearson skewness coefficient (3marks)