



# MERU UNIVERSITY OF SCIENCE AND TECHNOLOGY

P.O. Box 972-60200 – Meru-Kenya.

Tel: +254 (0)799529958, +254 (0)799529959, +254 (0)712524293

Website: [www.must.ac.ke](http://www.must.ac.ke) Email: [info@must.ac.ke](mailto:info@must.ac.ke)

---

## University Examinations 2022/2023

FIRST YEAR, SECOND SEMESTER SPECIAL/SUPPLEMENTARY EXAMINATION FOR THE  
DEGREE OF BACHELOR OF SCIENCE IN DATA SCIENCE

SECOND YEAR, SECOND SEMESTER SPECIAL/SUPPLEMENTARY EXAMINATION FOR  
THE DEGREE OF BACHELOR OF SCIENCE IN COMPUTER SCIENCE

SECOND YEAR, SECOND SEMESTER SPECIAL/SUPPLEMENTARY EXAMINATION FOR  
THE DEGREE OF BACHELOR OF SCIENCE IN COMPUTER TECHNOLOGY

### SSA 3120: PROBABILITY AND STATISTICS

**DATE: AUGUST 2023**

**TIME: 2 HOURS**

---

**INSTRUCTIONS:** *Answer question one and any other two questions*

---

#### QUESTION ONE

- a) Define the following terms
- |                      |          |
|----------------------|----------|
| (i) Statistics       | (1 mark) |
| (ii) Sampling design | (1 mark) |
| (iii) Data           | (1 mark) |
| (iv) Random variable | (1 mark) |
- b) State and explain the two major types of variables (4 marks)
- c) Fill in the blanks in the table and round your answer to two decimal places for the Relative Frequency and Cumulative Relative Frequency cells (4 marks)



Years	Frequency	Relative Frequency	Cumulative Relative Frequency
<5	25		
5 -10	30		
>	45		

d) Describe how you might draw a random sample of 30 students from a lecture class of 200 students (3 marks)

e) The midterm grades on a chemistry exam graded on a scale of 0-100 were:

62	64	65	65	68	70	72	72	74	75	75	75	76	78	78	81
83	83	84	85	87	88	92	95	98	98	100	100	740			

(i) Do you see any outliers in this data? If so how would you address the situation? (3 marks)

(ii) Calculate the average midterm grade on the chemistry exam (3 marks)

f) (i) Events A and B are independent if  $P(A) = 0.3$  and  $P(B)=0.5$  find  $P(A \text{ and } B)$  (3 marks)

(ii) C and D are mutually exclusive events. If  $P(C) = 0.18$  and  $P(D) = 0.03$ . find  $P(C \text{ or } D)$  (3 marks)

g) The pmf of a discrete random variable random variable w is given by the table below

w	-3	-2	-1	0	1
$P(w=w)$	0.1	0.25	0.3	0.15	d

Find the  $P(-1 < w < 1)$  (3 marks)

## QUESTION TWO

a) We are interested in the number of times a teenager must be reminded to do his or her chores each week. A survey of 40 mothers was conducted the results of the survey are

x	$P(x)$
0	$\frac{2}{40}$
1	$\frac{5}{40}$
2	-
3	$\frac{14}{40}$
4	$\frac{7}{40}$
5	$\frac{4}{40}$



- (i) Find the probability that a teenager is reminded two time (2 marks)
- (ii) Find the expected number of times a teenager is reminded to do his or her chores (3 marks)
- b) The following are the ages of six employees of an insurance company 46, 36, 48, 51, 49, 52  
For these data (10 marks)
- (i) Calculate the sample mean, mode and median
- (ii) Calculate the sample variance and standard deviation
- (iii) Calculate the value of the upper and lower qualities and the interquartile range
- (iv) Sketch the box plot
- c) In a marathon the median finishing time was 3:35:04 (three hours, 35 minutes and four seconds). You finished in 3:34:10. Interpret the meaning of the median time and discuss your time in relation to it (5 marks)

### QUESTION THREE

- a) You conduct a poll of 30 students to see how many classes they are taking this term. Your results are
- 1; 1; 1; 1
- 2; 2; 2; 2; 2
- 3; 3; 3; 3; 3; 3; 3
- 4; 4; 4; 4; 4; 4; 4; 4; 4; 4
- 5; 5; 5; 5
- Construct a histogram for this data (6 marks)
- b) The marks obtained by 8 applicants for a clerical post in mathematics and Accounting test are given in the table below:

Applicant	A	B	C	D	E	F	G	H
Math	15	20	28	12	40	60	20	80
Accounting	40	30	50	30	20	40	25	60

Calculate Spearman's rank correlation coefficient with adjustment for ties and interpret it (7 marks)



- c) You have a jar full of marbles, 50 are red, 25 are blue and 15 are yellow. Assume you draw one marble at random for each and replace it before the next trial. Let
- $P(R)$  = the probability of drawing and red marble
- $P(B)$  = the probability of drawing a blue marble
- $P(Y)$  = the probability of drawing a yellow marble
- (i) Find  $P(B)$  (2 marks)
- (ii) Which is more likely, drawing and a red marble or a yellow marble. Justify (3 marks)
- d) Your daughter brings home scores showing that she scored in the 80th percentile in math and the 76<sup>th</sup> percentile in reading for her grade. Interpret these scores (2 marks)

#### QUESTION FOUR

- a) The data below represents the marks scored by 100 students in a class

Marks	0 – 19	20 – 29	30 – 49	50 – 59	60 – 79	80-89	90 - 100
Frequency	10	15	5	25	20	15	10

- Construct a cumulative frequency curve and use it to estimate the median (5 marks)
- b) For a particular set of data  $x_1, y_2, \dots, x_{100}$  it was found that  $\sum xi = 550, \sum xi^2 = 3250$
- (i) Investigate the mean and the variance of the distribution (4 marks)
- (ii) Find the coefficient of variation (2 marks)
- c) Five randomly selected students took a math aptitude test before they began their statistics course the statistics department has three questions
- (i) What linear regression equation best predicts statistics performance based on math aptitude scores? (5 marks)
- (ii) If a student made an 80 on the aptitude test, what grade would we expect her to make in statistics? (1 mark)
- (iii) How well does the regression equation fit the data? (3 marks)

In the table below  $x_i$  column shows scores on the aptitude test similarly the  $y_i$  column shows statistics grades



<b>Student</b>	<b><math>x_i</math></b>	<b><math>y_i</math></b>
1	95	85
2	85	95
3	80	70
4	70	65
5	60	70
Sum	390	385
Mean	78	77

### QUESTION FIVE

- a) Discuss various probability and non-probability sampling methods (10 marks)
- b) Discuss the basic rules of probability (10 marks)

