



# **GRETSA UNIVERSITY - THIKA**

## **UNIVERSITY EXAMINATIONS JANUARY - APRIL 2020 SEMESTER**

### **BACHELOR OF COMMUNITY DEVELOPMENT**

**COURSE CODE: CMD 303**

**COURSE TITLE: STATISTICS FOR SOCIAL SCIENCES**

**DATE: 25 SEPTEMBER 2020**

**TIME: 8.00AM -**

**11.00AM**

#### **INSTRUCTIONS TO CANDIDATES**

1. SECTION A IS **COMPULSORY**.
2. SECTION B: ANSWER ANY OTHER **THREE** QUESTIONS.
3. LOG INTO THE KENET WEB CONFERENCING WITH THE VIDEO ON USING THE LINK PROVIDED ON GOOGLE CLASSROOM STREAM SECTION EXAM'S PUBLIC CHAT FOR ALL VIRTUAL EXAMS, AND CORRECTLY POSITIONING

YOUR SMARTPHONE/LAPTOP WEBCAM, IS COMPULSORY. NONCOMPLIANCE WILL LEAD TO THE CANCELLATION OF YOUR EXAMS.

**SECTION A: COMPULSORY**

**Question One**

a) Explain the differences between the following terms:

- i. Discrete and continuous variable
- ii. Primary data and secondary data

[4 marks]

b) The following are historical data on staff salaries (dollars per pupil on 30 schools sampled in the eastern part of the United States in the early 1970s).

3.79 2.99 2.77 2.91 3.10 1.84 2.52 3.22 2.45 2.14 2.67 2.52 2.71  
 2.75 3.57 3.85 3.36 2.05 2.89 2.83 3.13 2.44 2.10 3.71 3.14  
 3.54 2.37 2.68 3.51 3.37

Compute the sample mean and sample standard deviation

[6 marks]

c) The times for 100 athletes to cover one lap of running track were recorded as given in the table below

Time	65-69	70-74	75-79	80-84	85-89	90-94	95-99
Frequency	1	7	20	25	31	10	6

Calculate

- i. Median
- ii. Mode
- iii. Mean Absolute Deviation

[8 marks]

d) The following frequency distribution summarizes the number of goals scored in Series of football matches.

Number of goals	1	A	3	4	5
Number of matches	3	4	1	2	3a

If the mean is 3.25, find

- i. The value of a
- ii. The Mean Absolute Deviation (MAD)

[3 marks]

[3 marks]

e) Let  $x$  have a normal distribution with the mean and variance unknown. You want to test the null hypothesis  $\mu = 11.1$  against a one sided alternative  $\mu > 11.1$  and you decide to use a random sample of size 25. If the random sample of size 25 yields a sample mean of  $\bar{x} = 11.2$  with standard deviation  $s = 0.39$ , what is your decision?  $\alpha = 0.005$

[6 marks]

A study was made on the amount of converted sugar in a certain process at various temperatures. The data were coded and recorded as follows;

Temperature x	Converted sugar y
1.0	8.1
1.1	7.8
1.2	8.5
1.3	9.8
1.4	9.5
1.5	8.9
1.6	8.6

- |     |      |
|-----|------|
| 1.7 | 10.2 |
| 1.8 | 9.3  |
| 1.9 | 9.2  |
| 2.0 | 10.5 |

- f) Find the least squares line to predict the amount of converted sugar from temperatures [10 marks]

**SECTION B: ANSWER ANY THREE QUESTIONS**

**Question Two**

- a) Consider the following measurement of heat producing capacity of coal produced by two mines (in millions of calories per tonne)

Mine 1 8260 8130 8350 8070 8340

Mine 2 7950 7890 7900 8140 7920 7840

Can it be concluded that the two population variances are equal? [8 marks]

- b) A research engineer for a tire manufacturer is investigating tire life for a new rubber compound and has built 16 tires and tested them to the end of life on a road test. The sample mean is 60139.7. The data is assumed to be normal with a standard deviation of 3645.96 km. The engineer would like to demonstrate that the mean life of the tires is in excess of 60000km. Formulate and test an appropriate hypotheses and draw conclusion using  $\alpha = 5\%$  [4 marks]

- g) Find the mean, mode and median of the following grouped data.

Class	30---34	35---39	40---44	45---49	50---54	55---59
frequency	1	6	10	8	2	3

[8 marks]

**Question Three**

A set of bivariate data between the measurement of air velocity (x) and evaporation of a burning fuel droplets (y), in an impulse engine was summarized as follows:

$$n = 10, \sum x_i = 250, \sum x_i^2 = 6500, \sum y_i = 300, \sum y_i^2 = 10000 \text{ and } \sum x_i y_i = 7900$$

- a) Determine the least squares regression line of (y) on (x). [9 marks]  
 b) Estimate the evaporation coefficient of a droplet when air velocity is 45 units. [2 marks]  
 c) Calculate the sample coefficient of correlation and interpret it. [5 marks]  
 d) What proportion of total variability in evaporation of the burning fuel droplets is accounted for by the regression model? [4 marks]

**Question Four**

- a) Students are randomly assigned to groups which are taught German by 3 different methods; class instruction and language laboratory, only class room instructions and only self-study in language laboratory. The following are the final exam results of samples of students from the three groups.

Method1	94	88	91	74	86	97	
Method2	85	82	79	84	61	72	80
Method 3	89	67	72	76	69		

Use the H-test at 5% level of significant to test whether the three population samples are identical. [10 marks]

- b) The following arrangement indicates whether 50 consecutive persons interviewed by a pollster are for (F) or against (A) the increase in the state fuel tax to build more roads.  
 AAA F A F AAAAA FF AA F AAAAA F AA FF AA F AAAA F A FF AAAAA F AA F  
 AAAA F

Test for randomness at 5% level of significance

[6 marks]

- c) The marks out of 40 obtained by an education class of 50 students at a University are as shown below:

marks	0-4	5-9	10-14	15-19	20-24	25-29-	30-34	35-39
Frequency	3	4	5	6	11	8	9	4

- i. Compute the median

[4 marks]

### Question Five

- a) Customer retention is an outcome that is the result of several different factors. Describe them [12 marks]
- b) Business must adhere to legal issues in order to ensure that the business activities are within the law. Unlawful business activities can lead to suits and penalties for the business and the business employees. What are the ethical and legal issues related to e-commerce? [8 marks]