INTERNATIONAL UNIVERSITY OF EAST AFRICA
'Learning to Succeed'

FACULTY OF SCIENCE AND TECHNOLOGY

END OF SEMESTER EXAMINATIONS - AUGUST 2025

PROGRAMME: BIT
YEAR/SEM: YEAR 1/SEMESTER 2
COURSE CODE: BIT1208
NAME: BASIC STATISTICS

DATE:
TIME:

INSTRUCTIONS TO CANDIDATES:

SECTION A

Attempt all (40 Marks, Write the correct alternative in the booklet provided)

Question 1:

1. The population mean is called _____.

   A. Discrete variable

   B. Parameter

   C. Sampling unit

   D. Continuous variable

2. A researcher is curious about the IQ of students at IUEA. The entire group
   students is an example of a:
   A. Parameter
   B. Population
   C. Statistic
   D. Sample

3. A researcher studies the factors that determine the number of children future couples
   decide to have. The variable 'number of children' is a :
   A. Continuous variable
   B. Discrete variable
   C. Categorical variable
   D. Ordinal variable

4. The positive square root of the variance of a distribution is known as _____.
   A. Standard deviation
   B. Mean deviation
   C. Absolute deviation
   D. Relative deviation

5. The branch of statistics concerned with the procedure and methodology for obtaining valid conclusion is called
   A. Inferential statistics
   B. Descriptive statistics
   C. Sample Statistics

D. Deductive Statistics

6. The probability of obtaining at least a tail when two coins are tossed is given by

   A. ¼

   B. 1

   C. 1/2

   D. 3/4

7. A set of scores is presented in a histogram. The histogram shows a series of bars that tend
   to decrease in height from left to right. What is the shape of the distribution?
   A. Leptokurtic
   B. Normal
   C. Negativity skewed
   D. Positively skewed

8. What is the mean for the following scores: 2, 5, 4, 1, 8?

   A. 6

   B. 8

   C. 7

   D. 4

9. What is the median for the following scores: 2, 5, 4, 1, 8?
   A. 3.5
   B. 4.5
   C. 4
   D. 7

10. What is the variance for the following set of scores: 100, 100, 100, 100, 100.
    A. 10
    B. 100
    C. 0
    D. 5

11. Normally distributed data are normally referred to as:
A. Skewed
B. Asymmetrical
C. Bell-shaped
D. Peaked

12. Take the formula Z=(X-mean)/standard deviation. What does this formula calculate?
A. Confidence interval.
B. Variance.
C. Standard error of the mean.
D. Standard score.

13. If the scores on a test have a mean of 26 and a standard deviation of 4, what is the z-score for a score of 18?

A. 1

B. -1

C. 2

D. -2

14. The median is always:
A. The most frequently occurring score in a data set
B. The difference between the maximum and minimum scores.

C. The same as the mean
D. The middle score when results are ranked in order of magnitude

15. The seminar rooms in the library are identified by the letters A to H. A researcher records
the number of classes held in each room during the first semester. What kind of graph
would be appropriate to present the frequency distributions of these data?
A. Histogram
B. Scatterplot
C. Bar chart
D. Box plot

16. A population has a mean=50 and standard deviation=10. If these

scores are transformed into z-scores, the
population of z-scores will have a mean and standard deviation of:
A. mean=0 and standard deviation=1
B. mean=50 and standard deviation=10
C. mean=1 and standard deviation=0
D.

mean=0 and standard deviation=10

17. A scatterplot shows:
A. Scores on one variable plotted against scores on a second variable.
B. The average value of groups of data.
C. The frequency with which values appear in the data.
D. The proportion of data falling into different categories

18. A Pearson correlation of r=-0.6 indicates
A. An increase in X is accompanied by an increase in Y; the relationship is strong.
B. An increase in X is accompanied by a decrease in Y; the relationship is moderate.
C. An increase in X is accompanied by a decrease in Y; the relationship is strong.
D. An increase in X is accompanied by an increase in Y; the relationship is moderate.

19. In a linear regression equation, Y=a + bX, what is the b denote?
A. The intercept with the Y-axis.
B. The regression coefficient, the slope of the line.
C. The correlation coefficient, the strength of the line
D. The score on the variable X.

20. The paasche index number is based on:

A. Base Year Quantity
B. Current Year Quantity
C. Average of Base and Current Year Quantity
D. none of the above

Question 2:

Find the mode of the call received on 7 consecutive day 11,13,13,17,19,23,25

A. 11

B. 13

C. 17

D. 23

2. Calculate the range of the given sets of data 7,47,8,42,47,95,42,96,2

A. 6

B. 94

C. 71

D. 84

3. The Mean of a random variable K is given by equation

A. E(K)

B. $(EK)^2$

C. $E^2 - K^2$

D. None of these

4. Find the Mean of a constant k

A. K

B. k/2

C. k^2

D. 1

5. Find the Variance of the constant 'K'

A. 1

B. 0

C. K^2

D. K/2

6. What is the standard deviation for the following set of scores: 1, 1, 1,

1, 1.
A. 0
B. 1
C. 4
D. 25

7. The median is always:
A. The most frequently occurring score in a data set
B. The middle score when results are ranked in order of magnitude
C. The same as the mean
D. The difference between the maximum and minimum scores.

8. What is the mean for the following scores: 2, 5, 4, 1, 8?

A. 6

B. 4

C. 7

D. 8

9. What is the median for the following scores: 2, 5, 4, 1, 8?
A. 3.5
B. 4
C. 4.5
D. 7

10. A researcher studies the factors that determine the number of children future couples
decide to have. The variable ‘number of children’ is a :
A. Discrete variable
B. Continuous variable
C. Categorical variable
D. Ordinal variable

11. The seminar rooms in the library are identified by the letters A to H. A researcher records
the number of classes held in each room during the first semester. What kind of graph
would be appropriate to present the frequency distributions of these data?
A. Histogram
B. Scatterplot

C. Bar chart
D. Box plot

12. A researcher is curious about the IQ of students at IUEA. The entire group
students is an example of a:
A. Parameter
B. Statistic
C. Population
D. Sample

13. A set of scores is presented in a histogram. The histogram shows a series of bars that tend
to decrease in height from left to right. What is the shape of the distribution?
A. Leptokurtic
B. Positively skewed
C. Negativity skewed
D. Normal

14. Normally distributed data are normally referred to as:
A. Bell-shaped
B. Asymmetrical
C. Skewed
D. Peaked

15. If the scores on a test have a mean of 26 and a standard deviation of 4, what is the z-score for a score of 18?

A. 1

B. -1

C. -2

D. 2

16. A scatterplot shows:
A. The frequency with which values appear in the data.
B. The average value of groups of data.
C. Scores on one variable plotted against scores on a second variable.
D. The proportion of data falling into different categories

17. In a linear regression equation, Y=a + bX, what is the b denote?

A. The regression coefficient, the slope of the line.

B. The intercept with the Y-axis.

C. The correlation coefficient, the strength of the line

D. The score on the variable X.

18. The population mean is called _____.

A. Discrete variable

B. Parameter

C. Sampling unit

D. Continuous variable

19. The branch of statistics concerned with the procedure and methodology for obtaining valid conclusion is called

A. Descriptive statistics

B. Inferential statistics

C. Sample Statistics

D. Deductive Statistics

20. The probability of obtaining at least a head when two coins are tossed is given by

A. ¼

B. ¾

C. 2/4

D. ½

SECTION B

Attempt 4 (60 marks, each question carries 15 marks)

Question 1:

The following sample consists of the lengths of leaves in centimeters collected from the compound of IUEA: 10, 6, 17, 15, 23, 7, 6, 9, 10, 6, 13, 12.

Find the following:

(i) mean length of leaves

(ii) Median length of leaves

(iii) Mode

(iv) Variance

(v) Standard deviation

Question 2:

From the following data compute Laspeyre's and Paasche's index numbers for the current year:

[Figure from the original paper could not be retrieved]

data-filename="image.png" style="--tw-scale-x: 1; --tw-scale-y: 1; --tw-pan-x: ; --tw-pan-y: ; --tw-pinch-zoom: ; --tw-scroll-snap-strictness: proximity; --tw-gradient-from-position: ; --tw-gradient-via-position: ; --tw-gradient-to-position: ; --tw-ordinal: ; --tw-slashed-zero: ; --tw-numeric-figure: ; --tw-numeric-spacing: ; --tw-numeric-fraction: ; --tw-ring-inset: ; --tw-ring-offset-width: 0px; --tw-ring-offset-color: #fff; --tw-ring-color: rgb(59 130 246 / 0.5); --tw-ring-offset-shadow: 0 0 #0000; --tw-ring-shadow: 0 0 #0000; --tw-shadow: 0 0 #0000; --tw-shadow-colored: 0 0 #0000; --tw-blur: ; --tw-brightness: ; --tw-contrast: ; --tw-grayscale: ; --tw-hue-rotate: ; --tw-invert: ; --tw-saturate: ; --tw-sepia: ; --tw-drop-shadow: ; --tw-backdrop-blur: ; --tw-backdrop-brightness: ; --tw-backdrop-contrast: ; --tw-backdrop-grayscale: ; --tw-backdrop-hue-rotate: ; --tw-backdrop-invert: ; --tw-backdrop-opacity: ; --tw-backdrop-saturate: ; --tw-backdrop-sepia: ; --tw-contain-size: ; --tw-contain-layout: ; --tw-contain-paint: ; --tw-contain-style: ; width: 442px;">

Question 3:

A random variable X has a pdf given by

[Figure from the original paper could not be retrieved]

style="width: 442px;" data-filename="image.png">

Find

a. Value of d

b. E(X)

c. Var(X)

d. Standard deviation (X)

Question 4:

Calculate the Pearson’s correlation coefficient of the following data and comment on your result:

[Figure from the original paper could not be retrieved]

style="width: 841px;" data-filename="image.png">

Question 5:

(a). The average number of acres burned by forest and range fires in a large New Mexico County is 4,300 acres per year, with a standard deviation of 750 acres. The distribution of the number of acres burned is normal. What is the probability that between 2,500 and 4,200 acres will be burned in any given year?

(b). Most graduate schools of business require applicants for admission to take the Graduate Management Admission Council’s GMAT examination. Scores on the GMAT are roughly normally distributed with a mean of 527 and a standard deviation of 112. What is the probability of an individual scoring above 500 on the GMAT?

Question 6:

A company has the following data on its sales during the last year in each of its regions and the corresponding number of salespersons employed during this time:

[Figure from the original paper could not be retrieved]

style="width: 442px;" data-filename="image.png">

(i) Develop a regression model for the data.

(ii) forecast the number of sales that would be expected next year in regions that employed (a) 14 salespersons; and (b) 25 salespersons.

Question 7:

The probability of winning a game is 4/5. If ten games are played, what is the

(i) Mean number of successes

(ii) Variance and standard deviation

(iii) Probability of at least 8 successes.

(iv) Probability of at most 3 successes.

Question 8:

The table below shows the prices of some food items and their corresponding weights in years 2000, 2005 and 2010.

Commodity	Unit	Price(in UgShs)			Weight
		2000	2005	2010	
Matooke	Bunch	17,000	15,000	20,000	5
Meat	Kilogram	7,500	8,000	8,500	4
Posho	Kilogram	3,000	2,800	2,600	3
Beans	Kilogram	3,200	3,000	2,800	2
Vegetables	Kilogram	1,000	1,500	2,000	1

Using **2000** as the base year, calculate the:

(a) Simple aggregate price index for **2005**. Comment on your result.

(b) Price relative for each food item for **2010**.

(c). Weighted average price index for **2010**.

Question 9:

A box contains 12 transistors of type A and 18 of type B, all identical in appearance. If one transistor is taken at random, tested and returned to the box, and a second transistor then treated in the same manner, determine the probability that:

(a) the first is type A and the second type B

(b) both are type A

(c) neither is type A**.**

Question 10:

(a) The average annual rainfall in a certain region is 1200 mm with a standard deviation of 150 mm. Assuming the distribution of rainfall is normal, what is the probability that the annual rainfall is between 1000 mm and 1400 mm?

(b) The average score on a standardized test is 500 with a standard deviation of 100. If the scores are normally distributed, what is the probability of an individual scoring above 600?

Question 11:

A mass-produced article can exhibit two types of fault, A or B. Past records indicate that 1 per cent of production has fault A, 2 per cent has fault B. Further, the presence of fault A has no

effect on whether fault B is present or not. Find the following probabilities:

(a) P (an article has both faults);

(b) P (an article has no fault);

(c) P (an article has only fault A);

(d) P (an article has precisely one fault);

(e) P (an article has fault A, given that it does not have fault B).

Question 12:

As part of its recruitment procedures, a company awards applicants ratings from A (excellent) to E (unsatisfactory) for their interview performance, and marks out of 100 for a written test. The results for five interviewees are as follows.

[Figure from the original paper could not be retrieved]

style="width: 378px;" data-filename="image.png">
Calculate the Spearman’s rank correlation coefficient for this data, and comment on its value