MID TERM PREPARATION MEGA FILE BY MASTERS

STA301

(STATISTICS & PROBABILITY)

Harmonic means is the reciprocal of?

Arithmetic mean

2. How many percent corresponds to Q1 if you will interpret it? **Page 65**

25th

3. What is the mode in the word MATHEMATICS?

T, A, M

4. The middle value of an ordered array of numbers is the?

Median

If 10% added to each value of variable, the geometric mean of new variable is
added by?

10%

7. The total number of observations, which are below a certain value are known
as?

Cumulative frequency

8. The sum of deviations of the observation from their mean is?

Always Zero

9. In a commutative frequency polygon, the cumulative frequency of each class is plotted against?

Upper class boundary

Masters

10. If an observation in the date set is zero, then its geometric mean will be?

Zero

11. Data arrangement in order of magnitude is called?

Arrayed date

the most frequent value in the data is called?

Mode

14. The mode value from raw data can be obtained by the help of?

Dot plot

15. The value of the middle term in a ranked (ordered) data set is called the?

Median

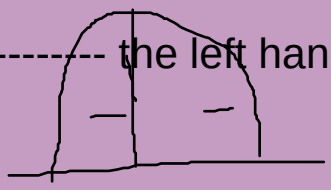
16. A list of 7 pulse rates is: 70,64,80,74,92,96,98 what is the median for this list?

74

17. Relationship among the averages? **Page 75**

AM ≥ GM ≥ HM

18. If we place a vertical mirror in the center of ---------- the left hand side will be the mirror image of the right hand side?



Symmetrical distribution

19. Which one is the formula of mid-range? **Page 75**

(X0+xm)/2

20. What is the mode for the following set of data (1,m,d,n,2,d,2,d,s,5,5,7)

2,d,5

21. What is the mode in the word TOMATO?

T and O

22. Cumulative frequency distribution is called a _____type of a cumulative frequency distribution. **Page 37**

Masters

Less than

A histogram is consists of a set of adjacent rectangles whose bases are marked off by? **Page 44**

Class Boundaries

24. The median is?

The middle point

25. A frequency polygon is constructed by plotting frequency of the class interval and the?

Class marks

26. If the sum of reciprocal of the values of a data is 256.7 and the n=23, then harmonic mean will be?

0.0895

27. In case of an open-ended class? **Page 63**

A median can-not be computed

29. The mean of a distribution is 23 the median is 24 and the mode is 25.5. , It is most likely that this distribution is?

Negatively skewed

30. Which of the following comes first to make frequency distribution?

Range

31. Histogram and histogram are?

Not same

The _____ is often the preferred measure of central tendency if the data are severely skewed?

Median

33. The median of the values are 1, 2,4,5,7?

4

34. As the number of observations and classes increase, the shape of a frequency polygon?

Masters

Tend to smooth

36. Which one is the formula of mid quartile range**? Page 80**

Q3+Q1 / 2

37. If any of the value in the data set is negative then it is impossible to compute?

Geometric Mean

38. Find the range of these values 24, 23, 20,15,17,19?

9

It is recommended that the number of classes in a frequency distribution be between?

5 and 20

41. The number of times each value appears is called the values?

Frequency

42. For the given data 20, 13, 27, 0,-8, G.M will be?

0

43. Quartiles are?

Points on a distribution which split it into 4 equal- sized portions

44. The number of classes in a frequency distribution generally should be?

Between five and twenty

45. In a cumulative frequency polygon, the cumulative frequency of each class is plotted against.

Upper class boundary

46. The number of classes in a frequency distribution is obtained by dividing the range of variable by the?

Class interval

47. If you correct the mid points of rectangles in a histogram by a series of lines that also touches the X-axis from both ends. You will get?

Frequency polygon

48. Frequency polygon is constructed by plotting frequencies against their?

Class mark

49. Which of the following measures of average is not based on the values given in the data set? Arithmetic Mean

50. A series of data with exclusive classes along with the corresponding frequencies is called?

Continuous frequency distribution

51. A tabular management for classifying data into different groups is called? Frequency distribution.

52. Which of the following is a true statement about the median?

It is the middle value of the data values.

53. Which of the following average gives information about most common value in the data?

Mean

54. If any of the value in the data set is zero then it is not possible (i.e. impossible) to compute?

Harmonic Mean

55. The sum of squared deviations from mean is? Zero

56. A frequency curve touches X-axis? Yes

57. Harmonic mean is better than other means if the data are for _____. Speed and rates

58. If in a set of discrete values, 50 percent values are greater than 25, then
q2 is? 75

60. For a certain distribution, A.M=136.75, median=148.37, and mode=152.80, then the distribution will be?

Masters

Negatively skewed

61. The total number of observations, which are below a certain value are known as?

Quartile deviation

62. Grouping error is less serious in case of?

Positively skewed

63. Half of the observations are always larger than?

Median

64. The ---- is the value you calculate when you want the arithmetic average?

Mean

66. Histogram can be drawn only for:

Continuous frequency distribution

67. If a distribution has two modes then this distribution is called_____?

Bi model

68. A histogram consists of a set of adjacent rectangles whose bases are marked off by class boundaries along?

X-axis

69. Which of the display of the data is divided into two parts?

Stem and leaf display

70. The mean is obtained by?

Dividing the sum of all observations by their number

71. The extremely positive skewed curve is also known as?

Reverse J-shaped curve

Masters

STA301 MIDTERM CURRENT PAPER SPRING 2022

Q1) Define statistics

The word "Statistics" which comes from the Latin words status, meaning a political state, originally meant information useful to the state, for example, information about the sizes of population sand armed forces. But this word has now acquired different meanings.

Q2) Find Class mark & relative frequency

Class Limit	Frequency	Relative Frequency	%age Frequency
30.0 – 32.9	2	2/30 = 0.067	6.7
33.0 – 35.9	4	4/30 = 0.133	13.3
36.0 – 38.9	14	14/30 = 0.467	4.67
39.0 – 41.9	8	8/30 = 0.267	26.7
42.0 – 44.9	2	2/30 = 0.067	6.7
	30		

For any query feel free to ask princealvi9222@gmail.com

STA301 MIDTERM CURRENT PAPER SPRING 2022

Class (Mileage Rating)	Frequency (No. of Cars)
30.0 – 32.9	2
33.0 – 35.9	4
36.0 – 38.9	14
39.0 – 41.9	8
42.0 – 44.9	2
Total	30

The first step is to compute the mid-point of every class.
(You will recall that the concept of the mid-point has already been discussed in an earlier lecture.)

CLASS-MARK (MID-POINT):

The mid-point of each class is obtained by adding the sum of the two limits of the class and dividing by 2. Hence, in this example, our mid-points are computed in this manner:

30.0 plus 32.9 divided by 2 is equal to 31.45,
33.0 plus 35.9 divided by 2 is equal to 34.45,

Class (Mileage Rating)	Class-mark (Midpoint) X
30.0 – 32.9	31.45
33.0 – 35.9	34.45
36.0 – 38.9	37.45
39.0 – 41.9	40.45
42.0 – 44.9	43.45

Q3) Write Formula for mean Deviation

$$M.D. = \frac{\sum |d|}{n}$$

Q4) FIVE-NUMBER SUMMARY

A five-number summary consists of X0,Q1, Median, Q3, and Xm ; It provides us quite a good idea about the shape of the distribution. If the data were perfectly symmetrical

The Five-Number Summary:

X_0	Q_1	$\tilde{X}$	Q_3	X_m

For any query feel free to ask princealvi9222@gmail.com

STA301 MIDTERM CURRENT PAPER SPRING 2022

Q5) 5 values were given find Median

EXAMPLE-1

The average number of floors in the buildings at the centre of a city:
5, 4, 3, 4, 5, 4, 3, 4, 5, 20, 5, 6, 32, 8, 27
Arranging these values in ascending order, we obtain
3, 3, 4, 4, 4, 4, 5, 5, 5, 5, 6, 8, 20, 27, 32
Picking up the middle value, we obtain the median equal to 5.

EXAMPLE -4:

The number of passengers traveling on a bus at six different times during a day is as follows:
5, 14, 47, 34, 18, 23
Find the median.

Solution:
Arranging the values in ascending order, we obtain
5, 14, 18, 23, 34, 47
As before, a slight complication has arisen because of the fact that there are even numbers of observations in the series and, as such, there are two middle values. As before, we take the arithmetic mean of the two middle values.
Hence we obtain:

Median:

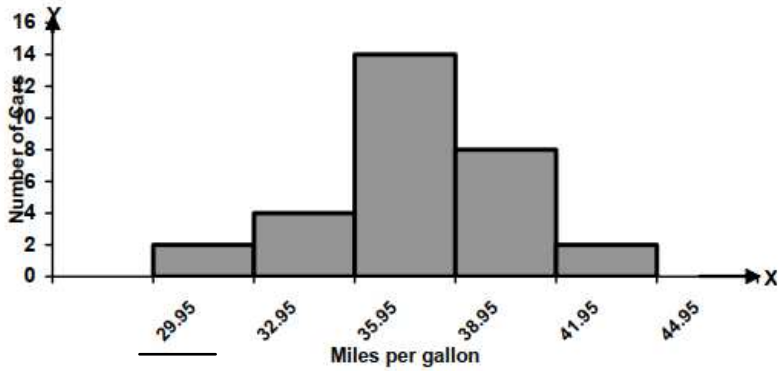
$$\tilde{X} = \frac{18 + 23}{2} = 20.$$

Q6) Histogram

HISTOGRAM:

A histogram consists of a set of adjacent rectangles whose bases are marked off by class boundaries along the X-axis, and whose heights are proportional to the frequencies associated with the respective classes.
It will be recalled that, in the last lecture, we were considering the mileage ratings of the cars that had been inspected by the Environment Protection Agency. Our frequency table came out as shown below:

Class Limit	Class Boundaries	Frequency
30.0 – 32.9	29.95 – 32.95	2
33.0 – 35.9	32.95 – 35.95	4
36.0 – 38.9	35.95 – 38.95	14
39.0 – 41.9	38.95 – 41.95	8
42.0 – 44.9	41.95 – 44.95	2
	Total	30



This diagram is known as the histogram, and it gives an indication of the overall pattern of our frequency distribution.

For any query feel free to ask princealvi9222@gmail.com

STA301 MIDTERM CURRENT PAPER SPRING 2022

Q7) Coefficient standard deviation

EXAMPLE

Life (in Hundreds of Hours)	No. of Bulbs f	Mid-point x	fx	fx^2
0 – 5	4	2.5	10.0	25.0
5 – 10	9	7.5	67.5	506.25
10 – 20	38	15.0	570.0	8550.0
20 – 40	33	30.0	990.0	29700.0
40 and over	16	50.0	800.0	40000.0
	100		2437.5	78781.25

Therefore, standard deviation:

$$S = \sqrt{\left\{\frac{78781.25}{100} - \left(\frac{2437.5}{100}\right)^2\right\}}$$

mean = sum of fx / sum of f

= 2437.5 / 100

= 24.37

=13.9hundredhours
= 1390 hours

As far as the graphical representation of the standard deviation is concerned, a horizontal line segment is drawn below the X-axis on the graph of the frequency distribution --- just as in the case of the mean deviation.

COEFFICIENT OF S.D

$$= \frac{\text{Standard Deviation}}{\text{Mean}}$$

=1390 / 24.37

= 57.03

Q8) Sample space page 140 lecture 17

SAMPLE SPACE

A set consisting of all possible outcomes that can result from a random experiment (real or conceptual), can be defined as the sample space for the experiment and is denoted by the letter S. Each possible outcome is a member of the sample space, and is called a sample point in that space. Let us consider a few examples:

EXAMPLE-1

The experiment of tossing a coin results in either of the two possible outcomes: a head (H) or a tail (T). (We assume that it is not possible for the coin to land on its edge or to roll away). The sample space for this experiment may be expressed in set notation as S = {H, T}. 'H' and 'T' are the two sample points.

EXAMPLE-2

The sample space for tossing two coins once (or tossing a coin twice) will contain four possible outcomes denoted by
S = {HH, HT, TH, TT}.
In this example, clearly, S is the Cartesian product A × A, where A = {H, T}.

EXAMPLE-3

The sample space S for the random experiment of throwing two six-sided dice can be described by the Cartesian product A × A, where A = {1, 2, 3, 4, 5,6}. In other words, S = A × A = {(x, y) | x ∈ A and y ∈ A}, Where x denotes the number of dots on the upper face of the first die, and y denotes the number of dots on the upper face of the second die. Hence, S contains 36 outcomes or sample points, as shown below:

S = {(1, 1), (1, 2), (1, 3), (1, 5), (1, 6),
(2, 1), (2, 2), (2, 3), (2, 5), (2, 6),
(3, 1), (3, 2), (3, 3), (3, 5), (3, 6),
(4, 1), (4, 2), (4, 3), (4, 5), (4, 6),
(5, 1), (5, 2), (5, 3), (5, 5), (5, 6),
(6, 1), (6, 2), (6, 3), (6, 5), (6, 6)}

For any query feel free to ask princealvi9222@gmail.com

STA301 MIDTERM CURRENT PAPER SPRING 2022

Q9) Why Venn diagram is important? page 142

VENN DIAGRAM

A diagram that is understood to represent sets by circular regions, parts of circular regions or their complements with respect to a rectangle representing the space S is called a Venn diagram, named after the English logician John Venn (1834-1923). The Venn diagrams are used to represent sets and subsets in a pictorial way and to verify the relationship among sets and subsets

Q10) Compute sampling error. Values were given lecture 2 page 8

Next, we compute the sample mean age.
Adding the 10 values and dividing by 10, we obtain:
Ages of students selected in the sample (in years):
14, 15, 16, 15, 16, 16, 17, 15, 16, 16
Hence the sample mean age is: 15.6, comparing the sample mean age of 15.6 years with the population mean age of 15.785 years, we note that the difference is really quite slight, and hence the sampling error is equal to

Sampling Error

$$\bar{X} - \mu = 15.6 - 15.785$$
$$= -0.185 \text{ years}$$

Q11) Define probability

The definition was formulated by the French mathematician P.S. Laplace (1949-1827) and can be very conveniently used in experiments where the total number of possible outcomes and the number of outcomes favorable to an event can be DETERMINED.

Q12) Define Geometric Mean

GEOMETRIC MEAN

The geometric mean, G, of a set of n positive values X1, X2,…,Xn is defined as the positive nth root of their product.

Q13) Find AU(BUC)=(BUA)UC

EXAMPLE

Let A = {1, 2, 3, 4} and B = {3, 4, 5, 6}
Then A ∪ B = {1, 2, 3, 4, 5, 6}

Q14) Find Cumulative frequency

AGE X	No. of Students f	Cumulative Frequency cf
13	6	6
14	61	67
15	270	337
16	491	828
17	153	981
18	15	996
19	4	1000
	1000	

For any query feel free to ask princealvi9222@gmail.com

STA301 MIDTERM CURRENT PAPER SPRING 2022

Q15) Difference between Ordinal scale or nominal scale

NOMINAL SCALE

The classification or grouping of the observations into mutually exclusive qualitative categories or classes is said to constitute a nominal scale.

ORDINAL OR RANKING SCALE

It includes the characteristic of a nominal scale and in addition has the property of ordering or ranking of measurements.

Q16) Scales ki examples di thin, unke age likhna tha konsi type hy

NOMINAL SCALE

The classification or grouping of the observations into mutually exclusive qualitative categories or classes is said to constitute a nominal scale. For example, students are classified as male and female. Number 1 and 2 may also be used to identify these two categories. Similarly, rainfall may be classified as heavy moderate and light. We may use number 1, 2 and 3 to denote the three classes of rainfall. The numbers when they are used only to identify the categories of the given scale carry no numerical significance and there is no particular order for the grouping.

ORDINAL OR RANKING SCALE

It includes the characteristic of a nominal scale and in addition has the property of ordering or ranking of measurements. For example, the performance of students (or players) is rated as excellent, good fair or poor, etc. Number 1, 2, 3, 4 etc. are also used to indicate ranks. The only relation that holds between any pair of categories is that of "greater than" (or more preferred).

INTERVAL SCALE

A measurement scale possessing a constant interval size (distance) but not a true zero point, is called an interval scale. Temperature measured on either the Celsius or the Fahrenheit scale is an outstanding example of interval scale because the same difference exists between 20o C (68o F) and 30o C (86o F) as between 5o C (41o F) and 15o C (59o F). It cannot be said that a temperature of 40 degrees is twice as hot as a temperature of 20 degree, i.e. the ratio 40/20 has no meaning. The arithmetic operation of addition, subtraction, etc. is meaningful.

RATIO SCALE

It is a special kind of an interval scale where the sale of measurement has a true zero point as its origin. The ratio scale is used to measure weight, volume, distance, money, etc. The, key to differentiating interval and ratio scale is that the zero point is meaningful for ratio scale.

Q17) Find Range

(MILES PER GALLON)		
36.3	42.1	44.9
30.1	37.5	32.9
40.5	40.0	40.2
36.2	35.6	35.9
38.5	38.8	38.6
36.3	38.4	40.5
41.0	39.0	37.0
37.0	36.7	37.1
37.1	34.8	33.9
39.9	38.1	39.8

For any query feel free to ask princealvi9222@gmail.com

STA301 MIDTERM CURRENT PAPER SPRING 2022

CONSTRUCTION OF A FREQUENCY DISTRIBUTION

Step-1

Identify the smallest and the largest measurements in the data set.
In our example:

Smallest value (X0) = 30.1,
Largest Value (Xm) = 44.9,

Step-2

Find the range which is defined as the difference between the largest value and the smallest value
In our example:

Range = Xm – X0
= 44.9 – 30.1
= 14.8

Q18) Find Quartile deviation

QUARTILE DEVIATION

The quartile deviation is defined as half of the difference between the third and first quartiles i.e.

$$Q.D. = \frac{Q_3 - Q_1}{2}$$

EXAMPLE

The shareholding structure of two companies is given below:

	Company X	Company Y
1st quartile	60 shares	165 shares
Median	185 shares	185 shares
3rd quartile	270 shares	210 shares

The quartile deviation for company X is

$$\frac{270-60}{2} = 105 \text{ Shares}$$

For company Y, it is

$$\frac{210-165}{2} = 22 \text{ Shares}$$

Q19) 5 Statements were given we have to find which ~~one~~ discrete and continuous

DISCRETE AND CONTINUOUS VARIABLES

A quantitative variable may be classified as discrete or continuous. A discrete variable is one that can take only a discrete set of integers or whole numbers, which is the values, are taken by jumps or breaks. A discrete variable represents count data such as the number of persons in a family, the number of rooms in a house, the number of deaths in an accident, the income of an individual, etc.

A variable is called a continuous variable if it can take on any value-fractional or integral—within a given interval, i.e. its domain is an interval with all possible values without gaps. A continuous variable represents measurement data such as the age of a person, the height of a

For any query feel free to ask princealvi9222@gmail.com

plant, the weight of a commodity, the temperature at a place, etc.
A variable whether countable or measurable, is generally denoted by some symbol such as X or Y and Xi or Xj represents the ith or jth value of the variable. The subscript i or j is replaced by a number such as 1,2,3, … when referred to a particular value.

Q20) Stem leaf find krna that frectional points diay Howy thy

EXAMPLE

Construct a stem-and-leaf display for the data of mean annual death rates per thousand at ages 20-65 given below:
7.5, 8.2, 7.2, 8.9, 7.8, 5.4, 9.4, 9.9, 10.9, 10.8, 7.4, 9.7, 11.6, 12.6, 5.0, 10.2, 9.2, 12.0, 9.9, 7.3, 7.3, 8.4, 10.3, 10.1, 10.0, 11.1, 6.5, 12.5, 7.8, 6.5, 8.7, 9.3, 12.4, 10.6, 9.1, 9.7, 9.3, 6.2, 10.3, 6.6, 7.4, 8.6, 7.7, 9.4, 7.7, 12.8, 8.7, 5.5, 8.6, 9.6, 11.9, 10.4, 7.8, 7.6, 12.1, 4.6, 14.0, 8.1, 11.4, 10.6, 11.6, 10.4, 8.1, 4.6, 6.6, 12.8, 6.8, 7.1, 6.6, 8.8, 8.8, 10.7, 10.8, 6.0, 7.9, 7.3, 9.3, 9.3, 8.9, 10.1, 3.9, 6.0, 6.9, 9.0, 8.8, 9.4, 11.4, 10.9

Using the decimal part in each number as the leaf and the rest of the digits as the stem, we get the ordered stem-and-leaf display shown below:

STEM AND LEAF DISPLAY

Stem	Leaf
-3	9
4	6 6
5	0 4 5
6	0 0 2 2 5 5 6 6 6 8 9
7	1 3 3 3 4 4 5 6 7 7 8 8 8 9
8	1 1 2 4 6 6 7 7 8 8 9 9
9	0 1 2 3 3 3 3 4 4 4 6 7 7 9 9
10	0 1 1 2 3 3 4 4 6 6 7 8 8 9 9
11	1 4 4 6 6 9
12	0 1 4 5 6 8 8
14	0

Q21) "Mathematics" word tha us ka btnaa tha k kitny words formed hoty

lecture 17
page 136

EXAMPLE

How many different (meaningless) words can be formed from the word 'committee'?
In this example:
n = 9 (because the total number of letters in this word is 9)
n1 = 1 (because there is one c)
n2 = 1 (because there is one o)
n3 = 2 (because there are two m's)
n4 = 1 (because there is one i)
n5 = 2 (because there are two t's)
and
n6 = 2 (because there are two e's)
Hence, the total number of (meaningless) words (permutations) is:

$$P = \frac{n!}{n_1!\, n_2!\, ...n_k!}.$$

$$= \frac{9!}{1!\ 1!\ 2!\ 1!\ 2!\ 2!}$$

$$= \frac{9\times8\times7\times6\times5\times4\times3\times2\times1}{1\times1\times2\times1\times1\times2\times1\times2\times1}$$

$$= 45360$$

STA301 MIDTERM CURRENT PAPER SPRING 2022

Q22) What is the relationship between 75th percentile and third quartile?

The 50th percentile is the same as the median, the 25th percentile is the same as the 1st quartile, the 75th percentile is the same as the 3rd quartile, the 40th percentile is the same as the 4th decile, and so on.

We know that the 75th percentile is the third quartile (Q3). The third quartile (Q3) is similarly the "median" of the values to the right of the true median. We can see that 80 is the middle number of the numbers to the right of the true median, so 80 is the 75th percentile and the third quartile (Q3).

Q23) Set A was given and have to find no.of subsets.

SUBSETS

A set that consists of some elements of another set, is called a subset of that set.
For example, if B is a subset of A, then every member of set B is also a member of set A.
If B is a subset of A, we write:
$B \subset A$ or equivalently:
$A \supset B$
'B is a subset of A' is also read as 'B is contained in A',
or 'A contains B'.

EXAMPLE

If $A = \{1, 2, 3, 4, 5, 10\}$
and $B \{1, 3, 5\}$
then ~~$B \subset A$~~,
i.e. B is contained in A.

EXAMPLE;

Consider the set $A = \{1, 2, 3\}$.
All possible subsets of this set are:
ϕ, {1}, {2}, {3}, {1, 2}, {1, 3}, {2, 3} and {1, 2, 3}
Hence, there are $23 = 8$ subsets of the set A.

Q24) Mean, median and standard deviation were given have to find coefficient of skewnes.

Pearson's Coefficient of Skewness

$$= \frac{3\,(\text{mean} - \text{median})}{\text{standard deviation}}$$

For any query feel free to ask princealvi9222@gmail.com

STA301 MIDTERM CURRENT PAPER SPRING 2022

	Children of Manual Workers	Children of Non-Manual Workers
Mean	8.50 years	8.50 years
Standard deviation	3.61 years	3.61 years
Median	8.50 years	9.16 years
Q_1	6.00 years	5.50 years
Q_3	11.00 years	10.83 years
Quartile deviation	2.50 years	2.66 years

The Pearson's Coefficient of Skewness is calculated for each of the two categories of children, as shown below:
Pearson's Coefficient of Skewness (Modified):

Ages of Children of Manual Workers

$$\frac{3(8.50-8.50)}{3.61}$$

$=0$

Ages of Children of Non-Manual Workers

$$\frac{3(8.50-9.16)}{3.61}$$

$=-0.55$

Q25) Define Negatively skewed curve

Negatively skewed frequency curve is the one for which the left tail is longer than the right tail.



Q26) Find coefficient of dispersion page 78

COEFFICIENT OF DISPERSION

$$=\frac{\frac{1}{2}(\text{Range})}{\text{Mid}-\text{Range}}$$

$$=\frac{\frac{X_m-X_0}{2}}{\frac{X_m+X_0}{2}}=\frac{X_m-X_0}{X_m+X_0}$$

Q27) A fair coin is tossed find the possible outcomes having two Heads kch is taran ka tha

EXAMPLE-2

A fair coin is tossed three times. What is the probability that at least one head appears?

SOLUTION

The sample space for this experiment is

S = {HHH, HHT, HTH, THH, HTT, THT, TTH, TTT}

and thus the total number of sample points is 8 i.e. n(S) = 8. Let A denote the event that at least one head appears. Then

$\overline{A}$ = {HHH, HHT, HTH, THH, HTT, THT, TTH}

Therefore n(A) = 7.

Hence

$$P(A)=\frac{n(A)}{n(S)}=\frac{7}{8}.$$

For any query feel free to ask princealvi9222@gmail.com

STA301 MIDTERM CURRENT PAPER SPRING 2022

Q28) P(A),p(B) , p(A intersection B) given thy P(A/B) find krna tha

EXAMPLE-2

At a certain elementary school in a Western country, the school-record of the past ten years shows that 75% of the students come from a two-parent home and that 20% of the students are low-achievers and belong to two-parent homes. What is the probability that such a randomly selected student will be a low achiever GIVEN THAT he or she comes from a two-parent home?

SOLUTION

Let A denote a low achiever and B a student from a two-parent home. Applying the *relative frequency definition* of probability, we have

$P(B) = 0.75$ and $P(A \cap B) = 0.20$.

Thus, we obtain

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{0.20}{0.75} = 0.27$$

Q29) Find C. V

COEFFICIENT OF VARIATION

$$C.V. = \frac{S}{\overline{X}} \times 100$$

EXAMPLE-1

Suppose that, in a particular year, the mean weekly earnings of skilled factory workers in one particular country were $ 19.50 with a standard deviation of $ 4, while for its neighboring country the figures were Rs. 75 and Rs. 28 respectively.

From these figures, it is not immediately apparent which country has the GREATER VARIABILITY in earnings. The coefficient of variation quickly provides the answer:

For country No. 1:

$$\frac{4}{19.5} \times 100 = 20.5 \text{ per cent,}$$

And for country No. 2:

$$\frac{28}{75} \times 100 = 37.3 \text{ per cent.}$$

Q30) Moment ratio find M1,m2,M3,m4 given thy

MOMENT RATIOS:

$$b_1 = \frac{(m_3)^2}{(m_2)^3} \text{ and } b_2 = \frac{m_4}{(m_2)^2}$$

Q31) Find Geometric means

For any query feel free to ask princealvi9222@gmail.com

STA301 MIDTERM CURRENT PAPER SPRING 2022

EXAMPLE

Find the geometric mean of numbers:

45, 32, 37, 46, 39, 36, 41, 48, 36

Solution:
We need to compute the numerical value of

$$= \sqrt[9]{45\times32\times37\times46\times39\times36\times41\times48\times36}$$

But, obviously, it is a bit cumbersome to find the ninth root of a quantity. So we make use of logarithms, as shown below:

X	log X
45	1.6532
32	1.5052
37	1.5682
46	1.6628
39	1.5911
36	1.5563
41	1.6128
48	1.6812
36	1.5563
	14.3870

$$\log G = \frac{\sum \log X}{n}$$

$$= \frac{14.3870}{9} = 1.5986$$

$$\text{Hence } G = \text{anti-log } 1.5986$$

$$= 39.68$$

Q32) Find first two moments of following data

EXAMPLE

Calculate the first four moments about the mean for the following set of examination marks: 45, 32, 37, 46, 39, 36, 41, 48 & 36.
For convenience, the observed values are written in an increasing sequence. The necessary calculations appear in the table below:

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$(x_i - \bar{x})^3$	$(x_i - \bar{x})^4$
32	– 8	64	– 512	4096
36	– 4	16	– 64	256
36	– 4	16	– 64	256
37	– 3	9	– 27	81
39	– 1	1	– 1	1
41	1	1	1	1
45	5	25	125	625
46	6	36	216	1296
48	8	64	512	4096
360	0	232	186	10708

Now $\bar{x} = \frac{\sum x_i}{n} = \frac{360}{9} = 40$ marks.

For any query feel free to ask princealvi9222@gmail.com

Therefore

$$m_1 = \frac{\sum(x_i - \bar{x})}{n} = 0$$

$$m_2 = \frac{\sum(x_i - \bar{x})^2}{n} = \frac{232}{9} = 25.78 \text{ (marks)}^2$$

$$m_3 = \frac{\sum(x_i - \bar{x})^3}{n} = \frac{186}{9} = 20.67 \text{ (marks)}^3$$

$$m_4 = \frac{\sum(x_i - \bar{x})^4}{n} = \frac{10708}{9} = 1189.78 \text{ (marks)}^4$$

Q33) Identify the example of secondary and primary data page 6

PRIMARY AND SECONDARY DATA

Data that have been originally collected (raw data) and have not undergone any sort of statistical treatment, are called PRIMARY data. Data that have undergone any sort of treatment by statistical methods at least ONCE, i.e. the data that have been collected, classified, tabulated or presented in some form for a certain purpose, are called SECONDARY data

Example of primary data is the national census data collected by the government while an example of secondary data is the data collected from online sources. The secondary data collected from an online source could be the primary data collected by another researcher

Q34) Disadvantage of range

Range has two serious disadvantages.

- It ignores all the INFORMATION available from the intermediate observations.
- It might give a MISLEADING picture of the spread in the data.

Q35) Write data limits (x+2s) and (x+3s) by empirical rule

According to this empirical rule:

- Approximately 68% of the measurements will fall within 1 standard deviation of the mean, i.e. within the interval ($\overline{X}$ – S, $\overline{X}$ + S)
- Approximately 95% of the measurements will fall within 2 standard deviations of the mean, i.e. within the interval ($\overline{X}$ – 2S, $\overline{X}$ + 2S).
- Approximately 100% (practically all) of the measurements will fall within 3 standard deviations of the mean, i.e. within the interval ($\overline{X}$ – 3S, $\overline{X}$ + 3S)

Q36) Find Permutation

EXAMPLE

A club consists of four members. How many ways are there of selecting three officers: president, secretary and treasurer? It is evident that the order, in which 3 officers are to be chosen, is of significance. Thus there are 4 choices for the first office, 3 choices for the second office, and 2 choices for the third office. Hence the total number of ways in which the three offices can be filled is $4 \times 3 \times 2 = 24$.

The same result is obtained by applying the rule of permutations:

$$^4P_3 = \frac{4!}{(4-3)!} = 4 \times 3 \times 2 = 24$$

STA301 MIDTERM CURRENT PAPER SPRING 2022

Q37) Find Mean

EXAMPLE:

Information regarding the receipts of a news agent for seven days of a particular week are given below

Day	Receipt of News Agent
Monday	£ 9.90
Tuesday	£ 7.75
Wednesday	£ 19.50
Thursday	£ 32.75
Friday	£ 63.75
Saturday	£ 75.50
Sunday	£ 50.70
Week Total	£ 259.85

Mean sales per day in this week:

= £ 259.85/7 = £ 37.12 (To the nearest penny).

Q38) QUANTITATIVE AND QUALITATIVE VARIABLES

Variables may be classified into quantitative and qualitative according to the form of the characteristic of interest. A variable is called a quantitative variable when a characteristic can be expressed numerically such as age, weight, income or number of children. On the other hand, if the characteristic is non-numerical such as education, sex, eyecolour, quality, intelligence, poverty, satisfaction, etc. the variable is referred to as a qualitative variable. A qualitative characteristic is also called an attribute. An individual or an object with such a characteristic can be counted or enumerated after having been assigned to one of the several mutually exclusive classes or categories.

Q39) Define multiple bar chart

A multi-bar chart is a bar chart in which multiple data sets are represented by drawing the bars side by side in a cluster. The multi-bar layer is created by using XYChart.

Q40) Define Negative skewed

Negative skew refers to a longer or fatter tail on the left side of the distribution, while positive skew refers to a longer or fatter tail on the right. These two skews refer to the direction or weight of the distribution.

Q41) Find Harmonic Mean

$$H.M. = \frac{n}{\sum\left(\frac{1}{X}\right)}$$

Q42) Find Mode

EXAMPLE:

Suppose that the marks of eight students in a particular test are as follows:

2, 7, 9, 5, 8, 9, 10, 9

Obviously, the most common mark is 9. In other words, **Mode** = 9.

For any query feel free to ask princealvi9222@gmail.com

Virtual University Overseas Examination System V2.1 (x86)

Statistics and Probability (STA301)

Questio 26 (Marks: 5) **Attempted 26 Total 2**

The following table gives the probability distribution of the random variable X where

X represents the number of Spots showing when a six-sided symmetric die is rolled.

Number of Spots Xi	P(Xi)

Answe

The following table gives the probability distribution of the random variable X where

X represents the number of Spots showing when a six-sided symmetric die is rolled.

Number of Spots Xi	P(Xi)
1	1/6
2	1/6
3	1/6
4	1/6
5	1/6
6	1/6

i): $E(X)$

ii): $Var(X)$

Calculate

1) E(X)

2(Var(X)

Virtual University Overseas Examination System V2.1 (x86)

Statistics and Probability (STA301)

Questio 2 (Marks: 1) **Attempted 21 Total**

In a symmetric distribution, what percent of values are less than the median?

Choice

30%

50%

correct ✓

70%

90%

Virtual University Overseas Examination System V2.1 (x86)

Statistics and Probability (STA301) bc

Questio 1 (Marks: 1) Attempted 21 Total 26

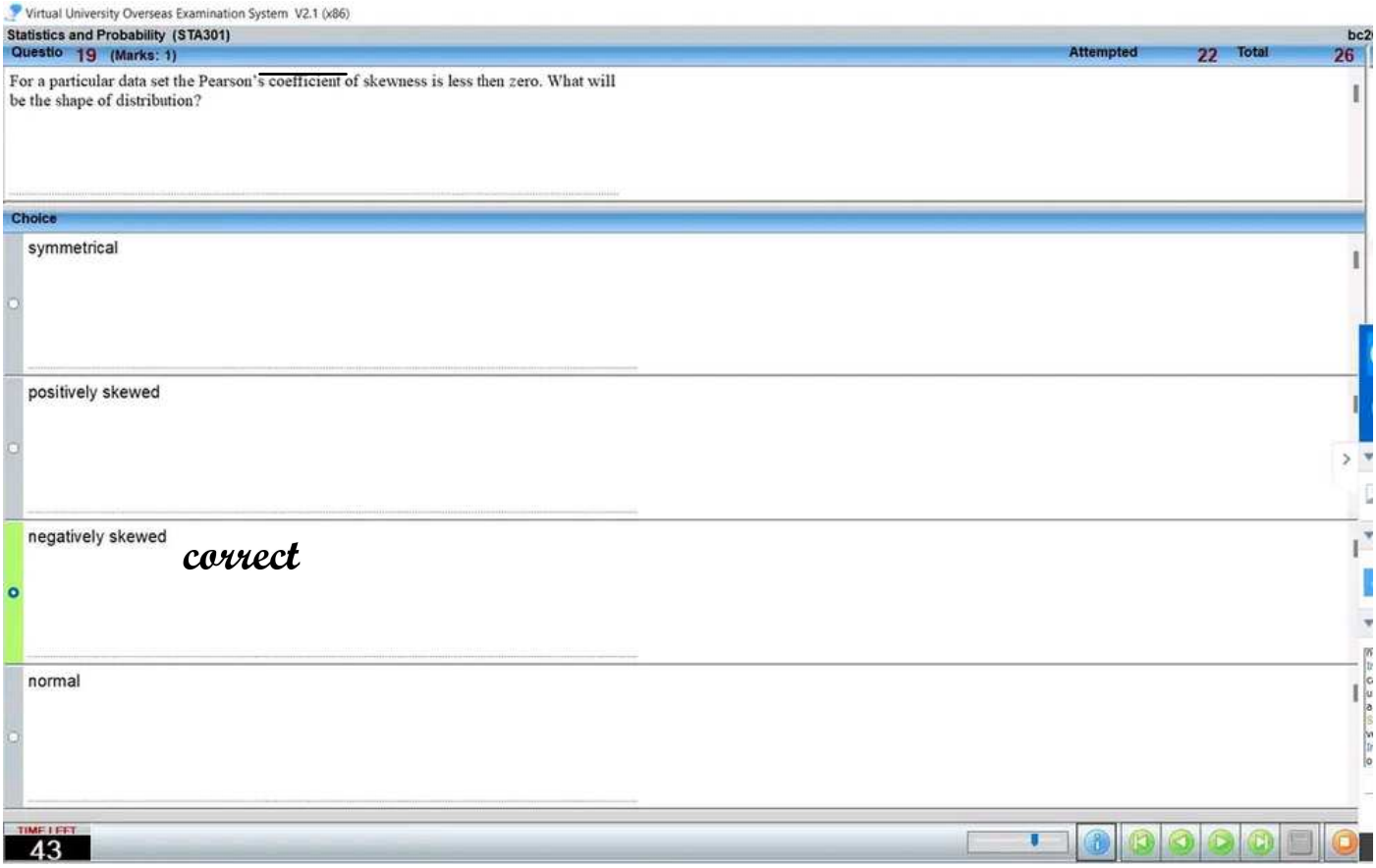
First moment about origin is always equals to:

Choice

- Mean
- Variance
- Standard Deviation
- Zero

correct

TIME LEFT


Virtual University Overseas Examination System V2.1 (x86)
Statistics and Probability (STA301)
Questio 19 (Marks: 1)
Attempted 22 Total 26
For a particular data set the Pearson's coefficient of skewness is less then zero. What will be the shape of distribution?
Choice
symmetrical
positively skewed
negatively skewed
correct
normal
TIME LEFT
43

Virtual University Overseas Examination System V2.1 (x86)

Statistics and Probability (STA301)

Questio 6 (Marks: 1) Attempted 22 Total 26

The range of the value of correlation coefficient is from:

is a measure of the strength or the degree of relationship between two RANDOM variables.

Choice

0 to 1

0 to ∞

- 1 to +1

correct

+1 to ∞

Statistics and Probability (STA301)

Questio 18 (Marks: 1) **Attempted 22** **Total 26**

The class width from the following frequency distribution is:

Class Intervals	Frequency
17-20	17
21-24	7

Choice

4

5

3

2

Virtual University Overseas Examination System V2.1 (x86)

Statistics and Probability (STA301)

Questio 5 (Marks: 1) Attempted 22 Total 2

If a player well shuffles the pack of 52 playing card, then the probability of a black card from 52 playing cards is:

black card= 26 queen =4 suits =13card in each suits jack =4

red card = 26 king =4 Ace =A =4 face cards =12

Choice

- 1/52
- 13/52
- 4/52
- 26/52 *correct*

Virtual University Overseas Examination System V2.1 (x86)

Statistics and Probability (STA301)

Questio 17 (Marks: 1) **Attempted 22 Total 2**

The extremely positive skewed curve is also known as: page 35

Choice

- Reverse J-shaped curve *correct*
- Frequency curve
- U-shaped curve
- J-shaped curve

Virtual University Overseas Examination System V2.1 (x86)

Statistics and Probability (STA301)

Questio 4 (Marks: 1) Attempted 22 Total 26

A set which is the sub-set of every set is

Choice

Universal Set

Super Set

Empty Set

correct

Power Set

TIME LEFT

Virtual University Overseas Examination System V2.1 (x86)

Statistics and Probability (STA301)

Questio 16 (Marks: 1) **Attempted 22 Total**

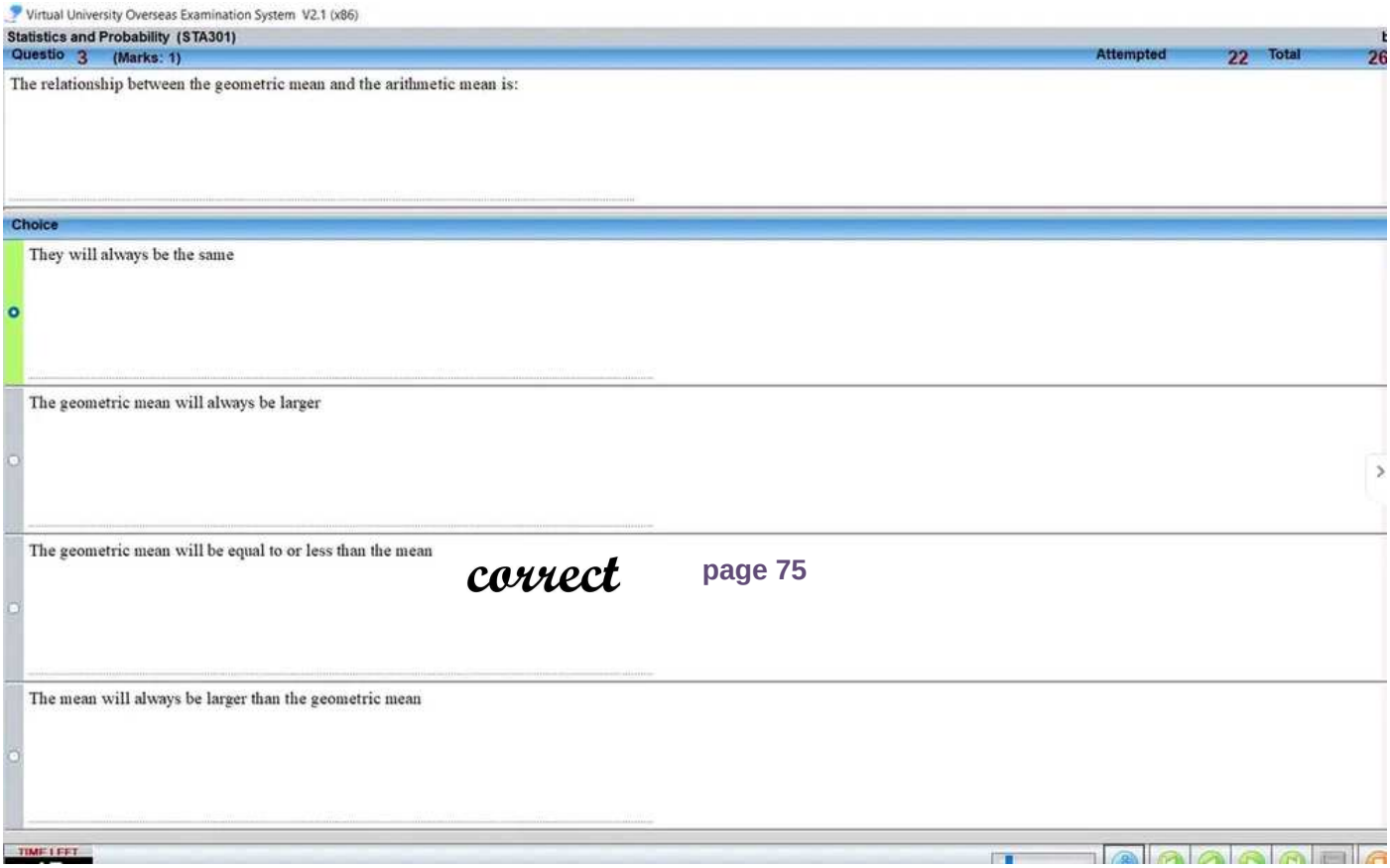
Chebychev's Inequality is more appropriate if the distribution is:

Choice

- Symmetric

  correct

- U-shaped
- J-shaped
- Moderately skewed



correct page 75

Virtual University Overseas Examination System V2.1 (x86)

Statistics and Probability (STA301) bc2

Questio 23 (Marks: 3) Attempted 26 Total 26

Average mark of the students in a class is 35. A sample of 9 students showed the following marks.

45 32 37 46 39 36 41 48 36

Calculate sampling error?

Answe

n = 9

mean = 35 = x

$\frac{\sum \int x}{\Sigma x}$ IS THE FORMULA

The ordered data is

32, 36 , 36 , 37, 39 , 41 , 45 , 46 ,48

Therefore ∑x = 360

∫x = 360*9

= 3240

where mean is 3240/35

92.57

TIME LEFT

Virtual University Overseas Examination System V2.1 (x86)

Statistics and Probability (STA301)

Questio 11 (Marks: 1) Attempted 22 Total 26

What is the mode in the word STATISTICS:

Choice

S

T

I

S and T *correct*

TIME LEFT 44

Virtual University Overseas Examination System V2.1 (x86)

Statistics and Probability (STA301)

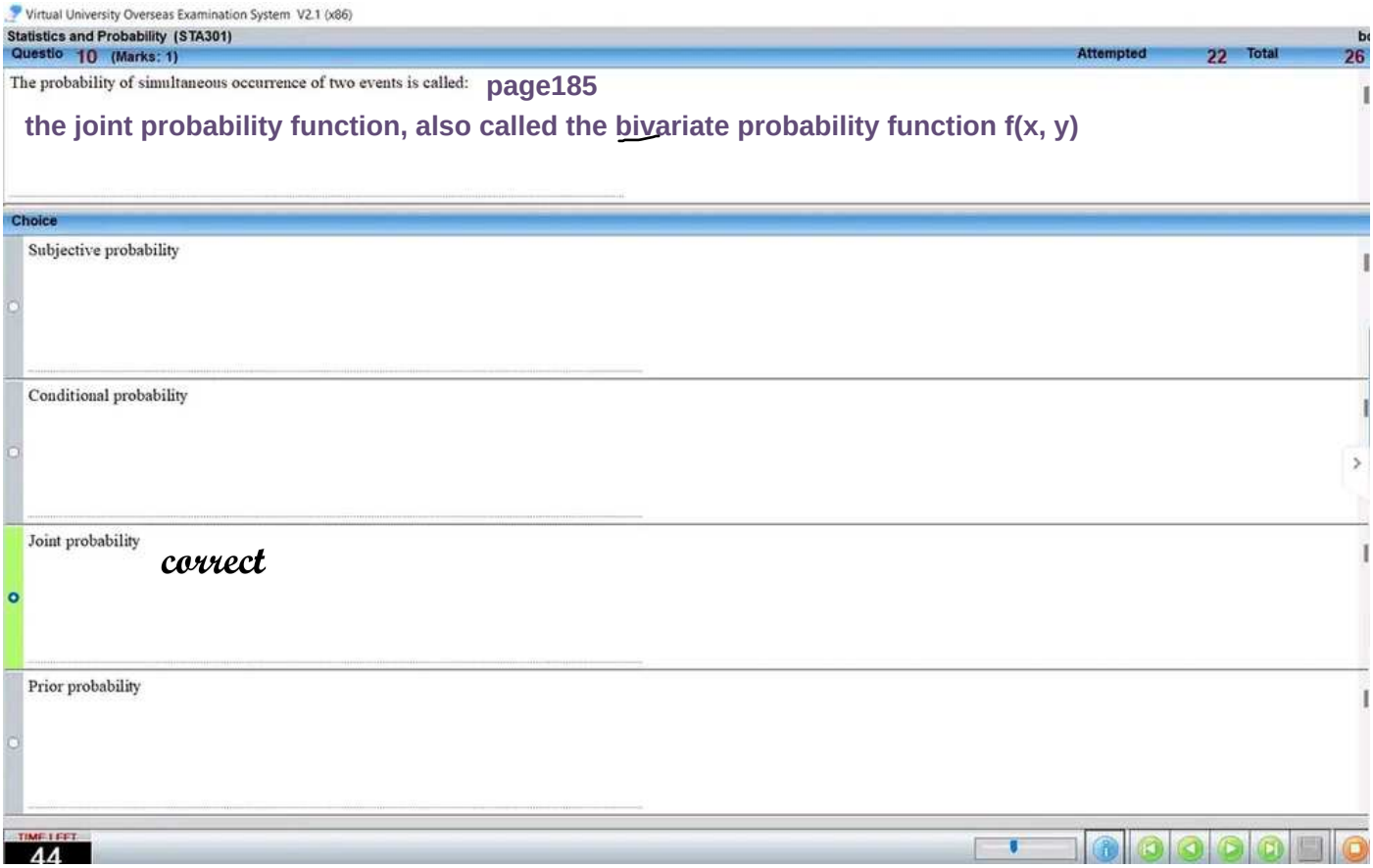
Questio 22 (Marks: 2) **Attempted 22 Total 2**

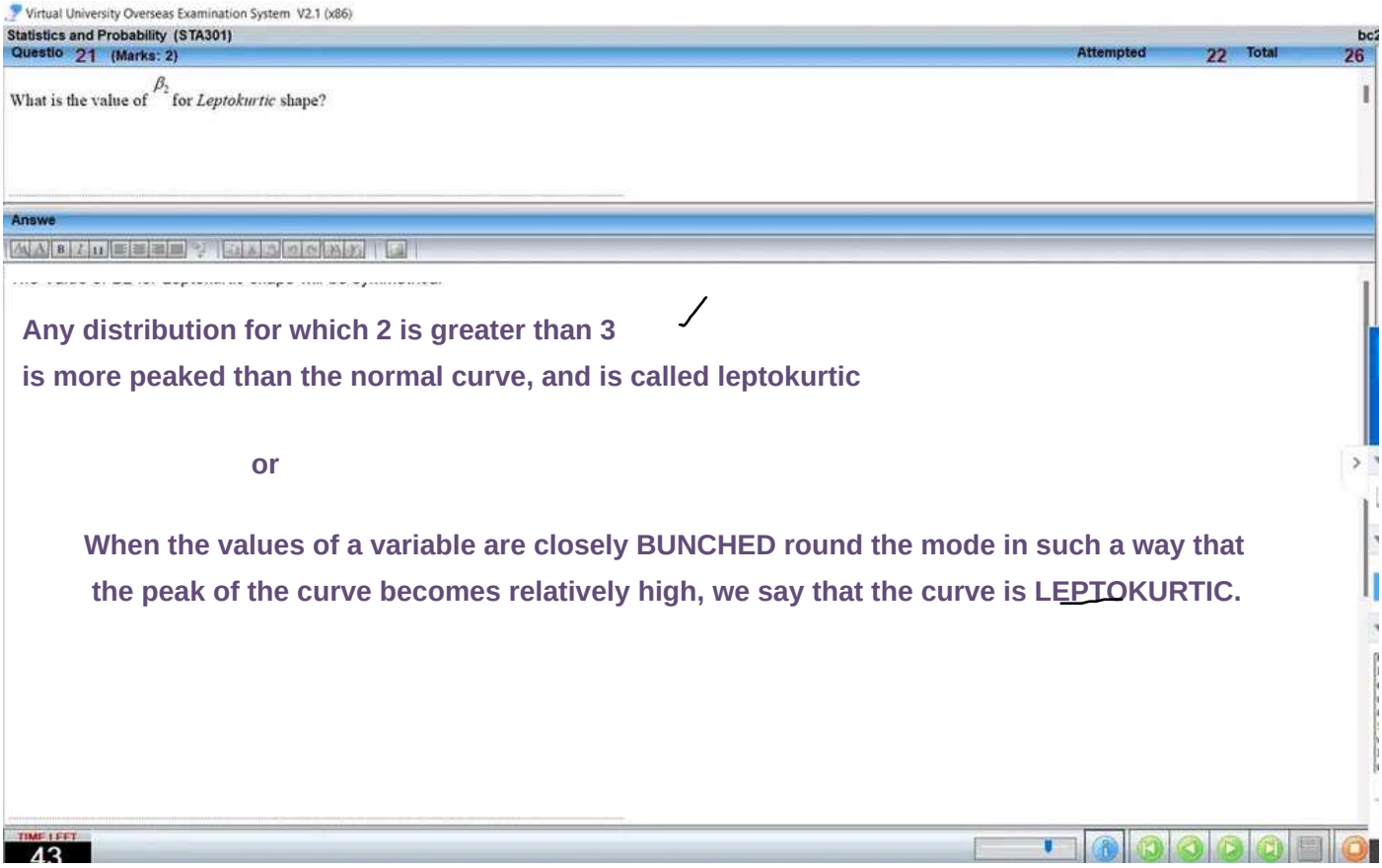
Write down the name of FOUR commonly used measures of dispersion.

Answe

The Four Commonly used measure of dispersion are
1) Standard Deviation
2) Quartile Deviation
3) Median Deviation
4) Mode Deviation

TIME LEFT
41


Virtual University Overseas Examination System V2.1 (x86)
Statistics and Probability (STA301)
Questio 10 (Marks: 1)
Attempted 22 Total 26
The probability of simultaneous occurrence of two events is called:
page185
the joint probability function, also called the bivariate probability function f(x, y)
Choice
Subjective probability
Conditional probability
Joint probability
correct
Prior probability
TIME LEFT
44


Virtual University Overseas Examination System V2.1 (x86)
Statistics and Probability (STA301)
Questio 21 (Marks: 2)
Attempted 22 Total 26
What is the value of β_2 for Leptokurtic shape?
Answe
Any distribution for which 2 is greater than 3
is more peaked than the normal curve, and is called leptokurtic
or
When the values of a variable are closely BUNCHED round the mode in such a way that the peak of the curve becomes relatively high, we say that the curve is LEPTOKURTIC.
TIME LEFT
43

Virtual University Overseas Examination System V2.1 (x86)

Statistics and Probability (STA301) bc

Questio 8 (Marks: 1) Attempted 22 Total 26

The probability of drawing a spade card is:

Choice

1/52

4/52

13/52

correct

26/52

TIME LEFT

Virtual University Overseas Examination System V2.1 (x86)

Statistics and Probability (STA301)

Questio 20 (Marks: 1) Attempted 22 Total 26

From the following given data calculate the sampling error?
population mean= 3 and sample mean= 4

Choice

1 *correct*

2

3

4

Virtual University Overseas Examination System V2.1 (x86)

Statistics and Probability (STA301)

Questio 7 (Marks: 1) Attempted 22 Total 26

What is the result of the following factorial expression?
(9-4)!

Choice

362880

120 *correct*

24

6

Virtual University Overseas Examination System V2.1 (x86)

Statistics and Probability (STA301) bc2

Questio 15 (Marks: 1) Attempted 22 Total 26

In a Pie diagram, the sector of a circle is obtained by:

Choice

$\frac{component\ part}{total} \times 100$

$\frac{component\ part}{total} \times 180$

$\frac{component\ part}{total} \times 300$

$\frac{component\ part}{total} \times 360$ correct

TIME LEFT 43

42°C

Virtual University Overseas Examination System V2.1 (x86)

Statistics and Probability (STA301)

Questio 25 (Marks: 5) Attempted 26 Total 26

There are 200 members in a saving group of the company. The number of savings certificate held by them are shown in the following table.

No. of certificates held	No. of members	Class boundaries	Relative frequency
1 – 50	10		10
51 – 100	15		25
101 – 150	30		55
151 – 200	40		95
201 – 300	105		200

Virtual University Overseas Examination System V2.1 (x86)

Statistics and Probability (STA301)

Questio 13 (Marks: 1) Attempted 22 Total

Under which condition standard deviation could assume a negative value:

Choice

When more than half of the data values are negative

If all the data values are same

The standard deviation cannot be negative *correct*

When all the data values are negative

TIME LEFT

Virtual University Overseas Examination System V2.1 (x86)

Statistics and Probability (STA301)

Questio 24 (Marks: 3) Attempted 26 Total 26

Arrange the given values in the sequence of Five Number summary.

$X_0 = 5, X_m = 25, Q_2 = 15, Q_1 = 10, Q_3 = 20$

Answe

The sequence of Five Number summary for the above values is
5, 10, 15, 20 , 25 as
Xmin , Q1 , Q2 , Q3, Xmax is the 5 number series

Virtual University Overseas Examination System V2.1 (x86)

Statistics and Probability (STA301) bc

Questio 12 (Marks: 1) Attempted 22 Total 26

If $\sum_{i=1}^{5}(X_i - 20) = 0, then \bar{X} = ...$

Choice

0

20

correct

5

25

TIME LEFT 44

Virtual University Overseas Examination System V2.0 (x86)

Statistics and Probability (STA301)

Question: 1 (Marks: 1) **Attempted Questions: 4** **Total Questions: 23**

A five-number summary consists of :

.....

Choices:

X_0, Q_1, Median, Q_3, and X_m *correct*

.....

X_m, Q_1, Mean, Q_3, and X_0

.....

X_m, Q_1, Mode, Q_3, and X_0

.....

X_0, Q_1, Median, Q_2, and X_m

.....

Virtual University Overseas Examination System V2.0 (x86)

Statistics and Probability (STA301)

Question: 3 (Marks: 1) Attempted Questions: 4 Total Questions: 23

In a multiplication theorem P (A∩B) equals:

.....

Choices:

P (A) P (B) *correct*

.....

P (A) + P (B)

.....

P (A) * P (B|A)

.....

P(B\A)*P(B)

.....

The probability of drawing a 'white' ball from a bag containing 4 red, 8 black and 3 white balls is:

Choices:

0

$\frac{3}{15}$ *correct*

$\frac{1}{12}$

$\frac{1}{2}$

Statistics and Probability (STA301)

Question: 5 **(Marks: 1)** **Attempted Questions: 4** **Total Questions: 23**

If $\hat{y}$ is the predicted value for a given x-value and b is the y-intercept then the equation of a regression line for an independent variable x and a dependent variable y is:

Choices:

$\hat{y} = mx + b$, where m = slope

correct

$x = \hat{y} + mb$, where m = slope

$\hat{y} = x/m + b$, where m = slope

$\hat{y} = x + mb$, where m = slope

Statistics and Probability (STA301)

Question: 6 (Marks: 1) **Attempted Questions: 5** **Total Questions: 23**

If A and B are mutually exclusive events then $P(A \text{ or } B)$ equals:

they cannot both occur at the same time

.....

Choices:

$P(A) + P(B) - P(A \text{ and } B)$

.....

$P(A) \times P(B)$

.....

$P(A) + P(B)$ *correct*

.....

$P(A|B) + P(B|A)$

.....

Virtual University Overseas Examination System V2.0 (x86)

Statistics and Probability (STA301)

Question: 7 **(Marks: 1)** **Attempted Questions: 6** **Total Questions: 23**

The simultaneous occurrence of two events is called:

.....

Choices:

- Prior probability

  .....

- Subjective probability

  .....

- Conditional probability

  .....

- Joint probability

  correct

  .....

Virtual University Overseas Examination System V2.0 (x86)

Statistics and Probability (STA301)

Question: 8 (Marks: 1) Attempted Questions: 7 Total Questions: 23

First moment about origin is always equals to:

Choices:

Mean

Variance

Standard Deviation

Zero

correct

Virtual University Overseas Examination System V2.0 (x86)

Statistics and Probability (STA301)

Question: 9 **(Marks: 1)** **Attempted Questions: 8** **Total Questions: 23**

Which one of the following measurement does *not* divide a set of observations into equal parts?

Choices:

- quartiles
- deciles
- percentiles
- standard deviations

correct

Virtual University Overseas Examination System V2.0 (x86)

Statistics and Probability (STA301)

Question: 10 (Marks: 1) **Attempted Questions: 9** **Total Questions: 23**

If the first and third quartiles are 22.16 and 56.36 respectively, then the quartile deviation is:

Q3-Q1/2

Choices:

17.1 *correct*

34.2

51.3

50.5

Virtual University Overseas Examination System V2.0 (x86)

Statistics and Probability (STA301)

Question: 11 (Marks: 1) **Attempted Questions: 10** **Total Questions: 23**

The height of a student is 60 inches. This is an example of .....?

.....

Choices:

- Continuous data

  .....

- Qualitative data

  .....

- Categorical data

  .....

- Discrete data

  correct

  .....

Virtual University Overseas Examination System V2.0 (x86)

Statistics and Probability (STA301)

Question: 12 **(Marks: 1)** **Attempted Questions: 11** **Total Questions: 23**

Which one of the following statements is true regarding a sample?

Choices:

- It is a part of population

correct

- It must contain at least five observations
- It refers to descriptive statistics
- It produces True value

Statistics and Probability (STA301)

Question: 13 (Marks: 1) **Attempted Questions: 12** **Total Questions: 23**

Which one of the following graphs is used for a time series data?

Choices:

- Histogram
- Historigram
- Frequency curve
- Frequency polygon

Virtual University Overseas Examination System V2.0 (x86)

Statistics and Probability (STA301)

Question: 14 **(Marks: 1)** **Attempted Questions: 13** **Total Questions: 23**

Which of the following comes first to make frequency distribution.

.....

Choices:

- No. of Groups

  .....

- Class interval

  .....

- Range

  correct

  .....

- Tally marks

  .....

Virtual University Overseas Examination System V2.0 (x86)

Statistics and Probability (STA301)

Question: 15 (Marks: 1) **Attempted Questions: 14** **Total Questions: 23**

The average which is defined as the reciprocal of the arithmetic mean of the reciprocals of the values is called:

.....

Choices:

Geometric Mean

.....

Harmonic Mean

correct

.....

Mode

.....

Median

.....

Virtual University Overseas Examination System V2.0 (x86)

Statistics and Probability (STA301)

Question: 16 (Marks: 1) **Attempted Questions: 15** **Total Questions: 23**

If a car is running at the rate of 15km/hr during first 30km, at 20km/hr during the second 30km, which type of average will be used to find the average speed/hr: page 73

Choices:

- Weighted Mean
- Harmonic Mean *correct*
- Arithematic Mean
- Geometric Mean

Differentiate between simple and compound events. page 140

SIMPLE & COMPOUND EVENTS

An event that contains exactly one sample point is defined as a simple event. A compound event contains more than one sample point, and is produced by the union of simple events.

Two events are casted: E_1 is the event that a 6 appears on at least one die, E_2 is the event that a 5 appear on exactly one die. Are E_1 and E_2 independent?

the outcome of one event affects the probability of the other event, E1 and E2 are dependent events

How many distinct four-digits number can be performed from the following integers 1,2,3,4,5,6 if each integer is used only once?

Total number of distinct four-digit numbers = 6 * 5 * 4 * 3 = 360

there are 360 distinct four-digit numbers that can be formed using the integers 1, 2, 3, 4, 5, and 6, where each integer is used only once.