

Quiz 2 Review Questions

Fred Azizi

2023-10-10

1. Which of the following provides a measure of central location for the data?

- a. standard deviation
- b. mean**
- c. variance
- d. range

Measures of Central location
 Mean
 Median
 Mode

Measures of Relative location
 Percentiles
 Q_1 Q_3

Measures of dispersion
 sd
 var
 Range
 IQR

2. When computing the mean of a set of values x_i , the value of $\sum x_i$

- a. can never be zero
- b. can never be negative
- c. must always be positive
- d. can be any value**

$0, 0, 0, 0, 0$
 $\sum x_i = 0$

-20
 -30
 -40
 \vdots

3. In computing the mean of a sample, the value of $\sum x_i$ is divided by

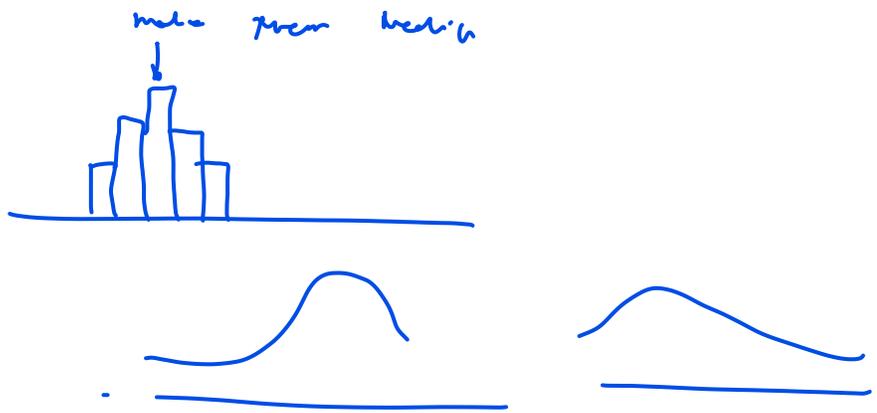
- a. n**
- b. $n - 1$
- c. $n + 1$
- d. $n - 2$

$$\bar{x} = \frac{\sum x_i}{n}$$

Sd ? var ?
 Cov ?
 $\frac{1}{n-1}$

4. The median of a sample will always equal the

- a. mode
- b. mean
- c. 50th percentile
- d. all of the above answers are correct

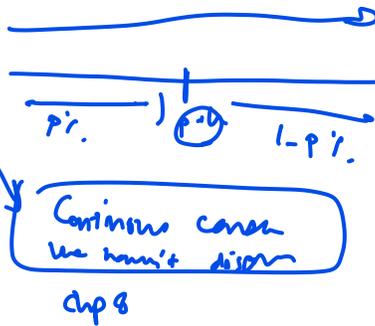


5. The median is a measure of

- a. relative dispersion
- b. absolute dispersion
- c. central location
- d. relative location

6. The p th percentile is a value such that at least p percent of the observations are

- a. less than or equal to this value
- b. less than this value
- c. more than or equal to this value
- d. more than this value



7. The difference between the largest and the smallest data values is the

- a. variance
- b. interquartile range
- c. range
- d. coefficient of variation

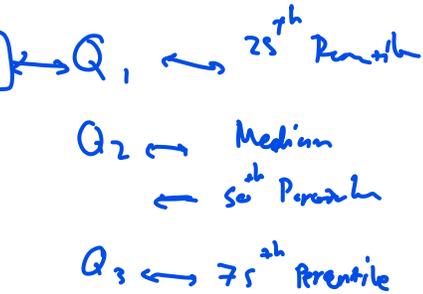
8. The first quartile

a. contains at least one third of the data elements

b. is the same as the 25th percentile

c. is the same as the 50th percentile

d. is the same as the 75th percentile



9. Which of the following is not a measure of central location?

a. mean

b. median

c. variance

d. mode

10. Which of the following is a measure of dispersion?

a. percentiles

b. quartiles

c. interquartile range

d. all of the above are measures of dispersion



11. The most frequently occurring value of a data set is called the

a. range

b. mode

c. mean

d. median

12. The interquartile range is used as a measure of variability to overcome what difficulty of the range?

- a. the sum of the range variances is zero
- b. the range is difficult to compute
- c. the range is influenced too much by extreme values
- d. the range is negative

$Range \geq 0 \xrightarrow{\text{why}} Max - Min < 0$
~~Max~~ ~~Min~~

13. The descriptive measure of dispersion that is based on the concept of a deviation about the mean is

- a. the range $\rightarrow Max - Min \rightarrow$ No Mean
- b. the interquartile range $\rightarrow Q_3 - Q_1 \rightarrow$ No Mean
- c. the absolute value of the range $\rightarrow |Max - Min| =$ No mean
- d. the standard deviation

14. The numerical value of the standard deviation can never be

- a. larger than the variance
- b. zero
- c. negative
- d. smaller than the variance

$sd > Var ?$

0.04
 $Var < 1 \rightarrow \sqrt{0.04}$
 \downarrow
 $0.2 = sd$

$\sqrt{\frac{1}{n-1} \sum (x_i - \bar{x})^2}$

$sd < Var ?$

$Var = sd$
 \uparrow \uparrow
 or or
 0 0

(0,0,0,0,0)

$Var = \sqrt{4} = 2$

e. equal to variance

15. The variance can never be

- a. zero
- b. larger than the standard deviation
- c. negative
- d. smaller than the standard deviation

16. If two groups of numbers have the same mean, then

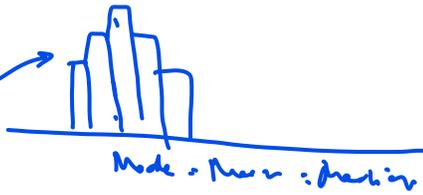
G_1 G_2
 Mean = $\frac{\sum x_1}{n_1}$ Mean = $\frac{\sum x_2}{n_2}$
 SD G_1 SD G_2
 Mod:

a. their standard deviations must also be equal \times

b. their medians must also be equal \times

c. their modes must also be equal \times

d. None of these alternatives is correct



17. The sum of deviations of the individual data elements from their mean is

a. always greater than zero

b. always less than zero

c. sometimes greater than and sometimes less than zero, depending on the data elements

d. always equal to zero

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

18. A numerical measure of linear association between two variables is the

a. variance

b. coefficient of variation

c. correlation coefficient

d. standard deviation

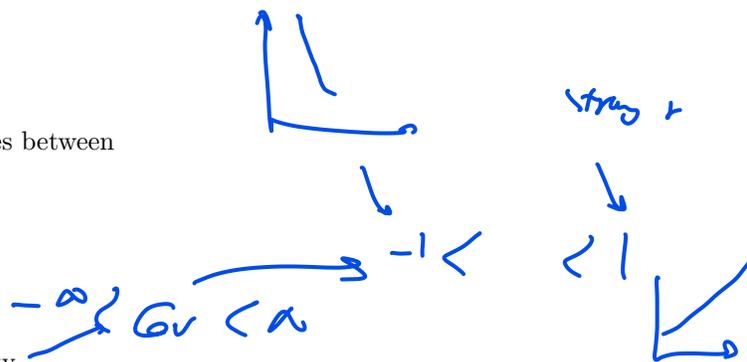
19. The coefficient of correlation ranges between

a. 0 and 1

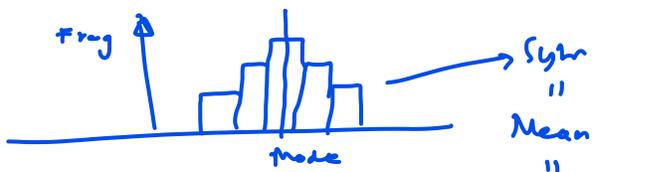
b. -1 and +1

c. minus infinity and plus infinity

d. 1 and 100



20. Since the mode is the most frequently occurring data value, it

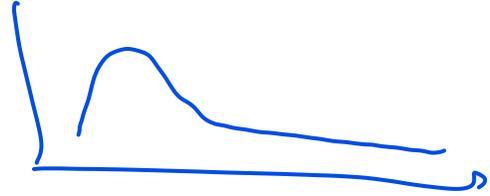


- a. can never be larger than the mean
- b. is always larger than the median
- c. is always larger than the mean

Sym
"
Mean
"
Median
"
Mode

Mode
(
Mean
)
Median

- d. None of these alternatives is correct.



21. Following observations are given for two variables.

y	x
5	2
8	12
18	3
20	6
22	11
30	19
10	18
7	9

a. Compute and interpret P_{86} .

b. Compute and interpret the correlation coefficient.

c. _____ is the relevant diagram for the data above.