Maths & Stats Pre-Sessional Tutorial

Topic 1: Sampling, measures of central tendency, and variability

---

Exercise 1.

A mortgage company randomly samples accounts of their time-share customers. State whether each of the following variables is categorical or numerical. If categorical, give the level of measurement. If numerical, is it discrete or continuous?

(a) The original purchase price of a customer's time-share unit

(b) The state (or country) of residence of a time-share owner

(c) A time-share owner's satisfaction level with the maintenance of the unit purchased (1: very dissatisfied to 5: very satisfied)

(d) The number of times a customer's payment was late

Exercise 2.

(a) Complete the following table:

Class	Absolute Frequency	Relative Frequency	Cumulative Absolute Frequency	Cumulative Relative Frequency
$0 < 10$	8			
$10 < 20$	10			
$20 < 30$	13			
$30 < 40$	12			
$40 < 50$	6			

(b) Visualize the data using a histogram.

(c) Plot the cumulative frequency polygon (OGIVE).

Exercise 3.

Students drink a lot of coffee during exam season. The following table contains the number of coffee cups sold for a random sample of n = 12 days by the Ground in Mile End campus:

60	84	65	67	75	72
80	85	63	82	70	75

(a) Describe the central tendency of the data (mean, median and mode)

(b) Compute the variance, standard deviation and range of the data

Exercise 4.

The annual percentage returns on common stocks and U.S. Treasury Bills over a 7 year period were as follows:

Stocks	4.0%	14.3%	19.0%	-14.7%	-26.5%	37.2%	23.8%
Bonds	6.5%	4.4%	3.8%	6.9%	8.0%	5.8%	5.1%

(a) Compute the mean of annual returns on stocks and T-Bills

(b) Compute the standard deviation and coefficient of variation for both assets

(c) What is the correlation between stocks and bonds?

Check your knowledge:

Test your knowledge with the following multiple-choice questions.

For each question, select the correct answer. Explain your decision.

Question 1

$\overline{X}$ is an example of a:

a) population parameter
b) sample statistic
c) population variance
d) mode
e) None of the above answers is correct.

Question 2

The sum of the percent frequencies for all classes will always equal

a) one
b) the number of classes
c) the number of items in the study
d) 100
e) None of the above answers is correct.

Question 3

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

a) variance
b) interquartile range
c) range
d) coefficient of variation
e) None of the above answers is correct.

Question 4

The sample mean is:

a) always equal to the mean of the population
b) always smaller than the mean of the population
c) computed by summing the data values and dividing the sum by (n - 1)
d) computed by summing all the data values and dividing the sum by the number of items
e) None of the above answers is correct.

Question 5

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

a) mean
b) median
c) variance
d) mode
e) None of the above answers is correct.

Question 6

Which of the following value is used as equivalent of the median?

a) the smallest value
b) the largest value
c) the range
d) the 50th percentile
e) the mean
f) the mode

Question 7

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

a) range
b) mode
c) mean
d) median
e) None of the above answers is correct.

Question 8

A statistics professor asked students in a class their ages. On the basis of this information, the professor states that the average age of all the students in the university is 21 years. This is an example of

a) a census
b) descriptive statistics
c) an experiment
d) statistical inference
e) None of the above answers is correct.

Question 9

A tabular summary of a set of data showing the percentage of the total number of items in several classes is a

a) Absolute frequency distribution
b) relative frequency distribution
c) cumulative frequency distribution
d) None of the above answers is correct.

Question 10

A financial analyst's sample of six companies' book value were

£25, £7, £22, £33, £18, £15.

The sample mean and sample standard deviation are (approximately):

a) 20 and 79.2 respectively
b) 20 and 8.9 respectively.
c) 120 and 79.2 respectively.
d) 20 and 8.2 respectively.
e) 120 and 8.9 respectively

Question 11

The standard deviation of a sample of 100 observations equals 64. The variance of the sample equals:

a) 8
b) 10
c) 6,400
d) 4,096
e) None of the above answers is correct.

Question 12

In general, which of the following statements is FALSE?

a) The sample mean is more sensitive to extreme values than the median.
b) The sample range is more sensitive to extreme values than the standard deviation.
c) The sample standard deviation is a measure of spread around the sample mean.
d) The sample standard deviation is a measure of central tendency around the median.
e) If a distribution is symmetric, then the mean will be equal to the median.

Question 13

A sample of 99 distances has a mean of 24 meters and a median of 24.5 meters. Unfortunately, it has just been discovered that an observation which was erroneously recorded as "30" actually had a value of "35". If we make this correction to the data, then:

a) the mean remains the same, but the median is increased.
b) the mean and median remain the same.
c) the median remains the same, but the mean is increased.
d) the mean and median are both increased.
e) we do not know how the mean and median are affected without further calculations; but the variance is increased.

Question 14

Which of the following is not a measure of dispersion?

a) the range
b) the 50th percentile
c) the standard deviation
d) the variance

Question 15

If the variance of a data set is correctly computed with the formula using n - 1 in the denominator, which of the following is true?

a) the data set is a sample.
b) the data set is a population.
c) the data set could be either a sample or a population.
d) the data set is from a census.
e) None of the above answers is correct.

Question 16

Measures of dispersion are influenced by extreme values

a) True
b) False

Question 17

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

a) the range
b) the interquartile range
c) both a and b
d) the standard deviation
e) None of the above answers is correct.

Question 18

A researcher has collected the following sample data. The mean of the sample is 5.

3 5 12 3 2

Which of the following statement is correct?

a) The mean and the median are the same value.
b) The mean is greater than the median.
c) The median is greater than the mean.

Question 19

A researcher has collected the following sample data. The mean of the sample is 5.

3 5 12 3 2

The variance is:

a) 80
b) 4.062
c) 13.2
d) 16.5
e) None of the above answers is correct.