

STAT 157 – Test 1: Version A

November 15, 2017

Name: Solution Set

Instructor: Patricia Wrean

Total: 30 points

1. (5 points) A Learning Skills instructor is interested in understanding the study habits of Camosun students. To do this, she surveyed 250 Camosun students and recorded the amount of time, rounded to the nearest hour, that each student said that they had studied in the last week.

(a) What are the experimental units?

the Camosun students

①

(b) What is the population?

all Camosun students

①

(c) Name a variable being measured.

study time

①

(d) For your answer to part (c), is it

(i) qualitative

(ii) quantitative and discrete

(iii) quantitative and continuous

because time is rounded to the nearest hour

①

(e) Is this univariate, bivariate, or multivariate data?

①

2. (2 points) State whether the following studies are experimental or observational by circling the correct choice.

(a) A study took random sample of adults and asked them about their bedtime habits. The data showed that people who drank a cup of tea before bedtime were more likely to go to sleep earlier than those who didn't drink tea. Expt / Observ

(b) A study randomly assigned volunteers to one of two groups. The first group was directed to use social media sites as they usually do, while the second group was blocked from social media sites. The researchers looked at which group tended to be happier. Expt / Observ

3. (3 points) A medical researcher wishes to study the patient outcomes for certain operations at BC hospitals within the last year. For the following situations, identify the survey method used to pick these operations.

(a) The researcher randomly selects 10 hospitals in BC and studies all operations that occurred at those hospitals within the last year.

cluster

(b) The researcher makes a list of all of the operations that took place in BC within the past year and randomly chooses a certain number of operations from that list.

simple random

(c) For each hospital in BC, the researcher chooses a random selection of operations that occurred ~~there~~ within the last year.

at that
hospital

stratified

4. (2 points) Consider the following sample data set, which has a standard deviation of 2.24.

4, 6, 7, 8, 10

(a) What is the standard deviation of the set: 8, 12, 14, 16, 20?

$$s = 2(2.24) = 4.48$$

because if the values double,
so does the spread

(b) What is the standard deviation of the set 14, 16, 17, 18, 20?

$$s = 2.24$$

because the set is only shifted and
has same spread as before

5. (2 points) Every year, Camosun wishes to poll its students to find out whether the library hours are convenient for studying. To do this, the registrar makes a list of all students currently registered at Camosun and puts the names in alphabetical order. He then picks a name randomly from the first 100 on the list, and then takes every 100th name thereafter. This process is repeated on a yearly basis.

Do you think that this will give a representative sample of Camosun students? Briefly explain.

yes, this is 1-in-100 systematic sampling, which is
a valid method

6. (6 points) A small warehouse employs a supervisor at \$1200 a week, an inventory manager at \$700 per week, eight stock workers at \$400 per week, and two drivers at \$500 a week.

set: 4, 4, 4, 4, 4, 4, 4, 4, 4, 5, 5, 7, 12 (in \$100 per week)

- (a) Find the mean and median wage.

$$\text{mean} = \frac{\text{sum}}{\text{number}} = \frac{8 \cdot 400 + 2 \cdot 500 + 700 + 1200}{12} = 508.\bar{3} = \$508.33$$

$$\text{median} = \$400$$

- (b) How many employees earn more than the mean wage?

2

- (c) Calculate the range of wages at this warehouse.

$$\text{range} = \$1200 - \$400 = \$800$$

(-1/2) if wrote 400 - 1200

- (d) Which measure of centre best describes a typical wage at this company? Explain briefly.

the median is better because the mean is pulled in the direction of the outliers at the high end of the data set

7. (2 points) The following graph is taken from the Erickson Times. It attempts to show the number of medals won by various countries during the Summer Olympic Games. Give the main reason that this graph is badly designed.



← this number is almost twice 999, but there are not twice as many medal pictures shown!

so the scale is wrong:

the number of medal pictures/icons in the graph does not accurately represent the medal count

8. (4 points) The accounts of a random sample of Twitter users were examined to determine how many tweets the users had made on the previous day, with the following results. The standard deviation of this sample data is 5.85.

$$0, 2, 8, 2, 1, 19, 1, 1, 3, 0 \quad \text{mean} = \frac{\text{sum}}{\text{number}} = \frac{37}{10} = 3.7$$

- (a) Calculate the z -score for the highest and lowest values in this set.

$$z_{\text{low}} = \frac{x - \bar{x}}{s} = \frac{0 - 3.7}{5.85} = -0.632479 \approx -0.63$$

$$z_{\text{high}} = \frac{x - \bar{x}}{s} = \frac{19 - 3.7}{5.85} = 2.6538 \approx 2.62$$

- (b) State whether the values of the highest and lowest scores are likely or unlikely.

z_{low} is likely, so the lowest score is likely

z_{high} is unlikely, so the highest score is unlikely

9. (4 points) A survey of a number of the leading brands of cereal shows that the content of potassium per serving is bell-shaped, with a mean of 95 milligrams and a standard deviation of 2 milligrams.

- (a) Find the interval in which at least 75% of the data points will fall.

Tchebysheff - at least 75% within 2 std dev of mean

$$\bar{x} + 2s = 95 + 2(2) = 99$$

$$\bar{x} - 2s = 95 - 2(2) = 91$$

so

$$\boxed{91 - 99 \text{ mg}}$$

- (b) Find the interval in which about 68% of the measurements will fall.

Empirical - ~68% within 1 std dev of mean

so

$$\boxed{93 - 97 \text{ mg}}$$