Math 153 - Test 1
September 14, 2017

Name key

Score _____

Show all work to receive full credit. Supply explanations where necessary. You may use your calculator for all statistical computations.

---

1. (2 points) In a poll of 2065 American men over the age of fifty, 929 had drunk alcohol within the last 24 hours. Identify the population and the sample.

Pop: All American men over 50

Sample: 2065 men chosen to participate

2. (6 points) Determine whether the given value is a statistic or a parameter.

   (a) In a recent Gallup poll, 39% of respondents enjoy spending money more than saving money.

Sample

39% is a statistic.

   (b) There are 50 state capitols in the United States.

All of U.S.

50 is a parameter.

   (c) The mean atomic weight of all elements in the periodic table is 134.355 atomic mass units.

All

134.355 is a parameter.

3. (3 points) In a recent week, a motorcycle dealer sold 12 motorcycles with 2 wheels and 1 motorcycle with 3 wheels. A customer asked for the average number of wheels, and the dealer computed the following:

$$\frac{12(2)+1(3)}{13} \approx 2.08.$$

Is 2.08 an appropriate measure of "average"? Briefly explain your reasoning.

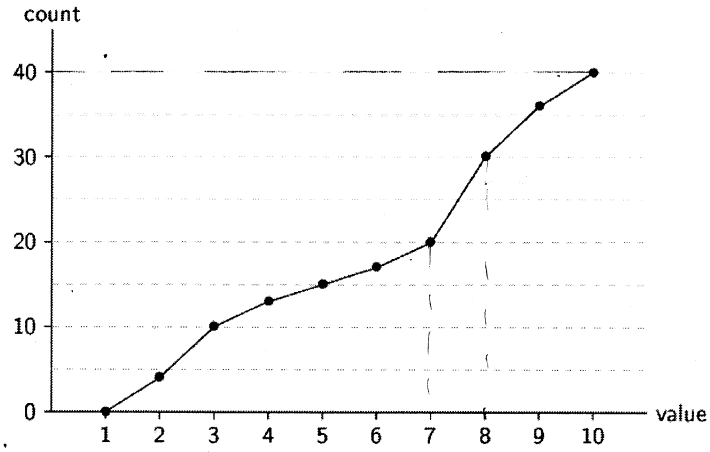
In this context, "average" probably means most frequently sold. The mode is a better measure of "average".

You could also argue that the number of wheels is more of a label, a kind of motorcycle, rather than a count of something. In this case, 2.08 has no meaning.

4. (2 points) A small-time, local farmer surveyed several other farmers by asking the following question: *Because large factory farms are bad for the environment, should they be banned?* What is wrong with this survey question?

THE QUESTION IS A LOADED QUESTION.
IT CONTAINS AN UNJUSTIFIED PRESUMPTION.

5. (8 points) The graph below is an **ogive** showing the **cumulative frequencies** of certain recorded values.



(a) What is the total number of values in the sample?

40

(b) Over what range of values did the frequency increase the most?

STEEPEST SLOPE FROM 7 TO 8

(c) About how many times did the value of 3 occur in the sample?

$10 - 4 = 6$

(d) What is the relative frequency of the value 3?

$\frac{6}{40} = 15\%$

6. (16 points) The following frequency distribution shows the white blood cell counts of **males** in a certain sample of size 40.

White Blood Cells	Relative Frequency	
3.0–4.9	20%	8
5.0–6.9	37.5%	15
7.0–8.9	27.5%	11 ←
9.0–10.9	12.5%	5
11.0–12.9	2.5%	1

(a) What is the class width?

$5.0 - 3.0 = \boxed{2.0}$

(b) What are the class boundaries associated with the first class listed above?

2.95 & 4.95

(c) If the relative frequency distribution was changed to a cumulative frequency distribution, what count would be associated with the class "≤ 8.9"?

$8 + 15 + 11 = \boxed{34}$

(d) What are the class midpoints?

$\frac{3.0 + 4.95}{2} = 3.95,\ 5.95,\ 7.95,\ 9.95,\ 11.95$

(e) Use class midpoints to estimate the mean white blood cell count.

$$\frac{8(3.95) + 15(5.95) + 11(7.95) + 5(9.95) + 1(11.95)}{40} = \frac{270}{40} = \boxed{6.75}$$

(f) Use class midpoints to estimate the median white blood cell count.

MEDIAN IS $\frac{20^{TH} + 21^{ST}}{2} = \frac{5.95 + 5.95}{2} = \boxed{5.95}$

(g) Do the counts in the sample appear to be normally distributed? If so, explain why you think so. If not, describe the type of distribution.

No, SKEWED RIGHT

(h) Compare the mean and median you computed above. How does your comparison support your answer in part (g)?

YES, MEAN > MEDIAN IS TYPICAL OF DISTRIBUTIONS SKEWED RIGHT.

7. (6 points) A collection of test scores have mean 71.5 and standard deviation 3.8. Is this a small or large standard deviation? What are the cut-offs for unusually low and high test scores?

Low:
$71.5 - 2(3.8)$
$= 63.9$

High:
$71.5 + 2(3.8)$
$= 79.1$

8. (6 points) For each of the following situations, tell which type of graph would best display the data. Choose from *dot plot*, *bar graph*, *time-series graph*, *scatterplot*, *pie chart*, *ogive*, *histogram*, *Pareto*, or *stem-and-leaf plot*. You may get partial credit if you offer brief explanations.

(a) After collecting data from over 1000 people, Stephanie is constructing a graph that shows the seven most common causes for quitting a job.

PARETO - BAR GRAPH IN DESCENDING ORDER

(b) The author of a geography textbook wants to show a graph displaying the portions of the earth's land surface taken up by the seven continents.

PIE CHART - PORTIONS OF WHOLE

(c) Health researchers weighed 500 fourth-grade children and summarized their data in a frequency distribution. They want to make a graph illustrating their frequency distribution.

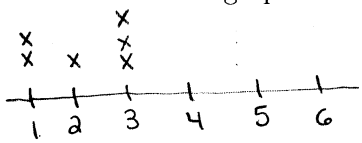
HISTOGRAM

(d) A stock chart shows the value of a particular stock throughout the day. What type of graph is a stock chart?

TIME-SERIES GRAPH

(e) Jon rolled a six-sided die 25 times. He would like to make a graph showing the numbers that he rolled.

DOT PLOT



(f) At the end of the semester, a professor made a list containing his students' ages and their corresponding final exam scores. He formed ordered pairs and plotted the data.

SCATTERPLOT

9. (8 points) Listed below are the numbers of manatee deaths caused each year from 2000 to 2013 by collisions with watercraft.

78 81 95 73 69 79 92 73 90 97 83 88 82 73

Compute the mean and standard deviation. Based on the numbers, is it unusual to have 99 annual deaths? What about 65 deaths?

CALCULATOR: $\bar{X} = 82.357$

$s = 8.906$

$\bar{X} - 2s \approx 64.5 < 65 \Rightarrow$ 65 IS NOT UNUSUAL

$\bar{X} + 2s \approx 100.2 > 99 \Rightarrow$ 99 IS NOT UNUSUAL

10. (3 points) Recent research indicates there is a strong correlation between the extinction of the woolly mammoths and the disappearance of a certain type of flower eaten by the mammoths. Does this suggest that mammoth extinction was caused by the flower disappearance? Explain.

NO, THIS IS AN OBSERVATIONAL STUDY.

CORRELATION DOES NOT IMPLY CAUSATION.

11. (8 points) Determine the level of measurement. Choose from nominal, ordinal, interval, or ratio. You may get partial credit if you offer brief explanations.

(a) Weights of adult women

RATIO

(b) Athletes' jersey numbers

NOMINAL

(c) Years in which total solar eclipses occurred

INTERVAL

(d) Names of books written by Nikolai Gogol

NOMINAL

12. (9 points) A sample of PSC students is obtained as described. Identify the type of sampling (random, systematic, convenience, stratified, cluster). You may get partial credit if you offer brief explanations.

(a) Students are selected as they walk in the main door.

CONVENIENCE

(b) Students are separated into groups according to age, then 20 students are selected from each age group.

STRATIFIED

(c) A complete list of students is compiled and every 150th name is selected.

SYSTEMATIC

(d) Students are separated into groups according to last initial. Ten letters are chosen at random and all students with that initial are selected.

CLUSTER

(e) Student ID numbers are selected at random by using a computer.

RANDOM

(f) Students are grouped according to which high school they attended. Two students are selected from each high school.

STRATIFIED

13. (6 points) A sample of Chips Ahoy cookies had a mean of 24.0 chocolate chips with a standard deviation of 2.6 chips. A sample of cans of regular Coke had a mean weight of 0.81682 lb with a standard deviation of 0.00751 lb.

(a) Compute the coefficients of variation for the two products. Which product had greater variation?

CHIPS:

$CV = \frac{2.6}{24} \approx 10.8\%$ *

COKE:

$\frac{0.00751}{0.81682} \approx 0.9\%$

* → CHIPS HAVE MUCH GREATER VARIATION.

(b) Looking at your result above, did you expect that product to have the greater variation? Why?

YES, IT IS PROBABLY EASIER AND MORE IMPORTANT TO PRECISELY CONTROL THE AMOUNT OF COKE POURED INTO A CAN RATHER THAN CHIPS ENDING UP IN A COOKIE.

14. (6 points) Organize the following data into a stem-and-leaf plot. Are the data approximately normally distributed? Explain.

2.1 3.2 4.8 1.2 2.8 3.2 3.6 3.7 3.8 4.9
4.7 3.2 4.8 5.1 1.0 2.7 5.0 2.1 2.4 1.2

1	0 2 2
2	1 1 4 7 8
3	2 2 2 6 7 8
4	7 8 8 9
5	0 1

2 | 8 means 2.8

THE DATA ARE ROUGHLY NORMAL --- SMOOTHLY INCREASING TO PEAK IN MIDDLE, THEN SMOOTHLY DECREASING. ROUGH SYMMETRIC BELL SHAPE.

15. (3 points) A survey conducted by LA Fitness asked members to report the amount of time they work out each day. What is wrong with such a survey? How could the data collection be improved?

VOLUNTARY RESPONSE SURVEY -- ONLY PEOPLE WITH STRONG FEELING TEND TO PARTICIPATE.

TO IMPROVE DATA COLLECTION, ACTUALLY OBSERVE AND RECORD WORKOUT TIMES.

16. (5 points) A professor separated her students' lab reports into two piles—those of the passing students and those of the failing students. Twenty-two students passed, and their average score was 78.5. Seven students failed, and their average score was 61.3. What was the average score of all the students?

$$\frac{22(78.5) + 7(61.3)}{22+7} = \frac{2156.1}{29} \approx \underline{74.35}$$

17. (3 points) Sketch a histogram showing a distribution in which the mean is much greater than the median.

SKEWED RIGHT!