Math 153 - Test 1
February 8, 2018

Name key

Score _____

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

---

1. (3 points) A survey question concerning Wi-Fi hotspot security was posted on the online edition of *USA Today*. 721 subjects responded. Is this a good way to conduct a survey? Explain.

   NO, THIS IS A VOLUNTARY RESPONSE SURVEY. THEY TEND TO BE BIASED BECAUSE PEOPLE WITH STRONG OPINIONS PARTICIPATE.

2. (3 points) Yesterday, in a hearing before Illinois legislators, Illinois Public Health Director Dr. Nirav Shah said, "Your assumption that there is a cover-up at play is 150 thousand percent unfounded." What do you think about this statement? Explain.

   100% UNFOUNDED MEANS COMPLETELY UNFOUNDED.

   150,000% UNFOUNDED IS COMPLETE NONSENSE.

3. (3 points) In the first decade of the 2000's, there was a strong correlation between the number of cell phones being used and the number of pirates roaming the world's oceans. Would this seem to indicate that rising cell phone usage could be attributed to pirates? Explain.

   NO, CORRELATION DOES NOT IMPLY CAUSATION.

4. (3 points) In a study of batteries conducted by ALDI, researchers found that ALDI brand batteries outperformed Energizer batteries. What do you think about this study?

   I AM SKEPTICAL BECAUSE IT IS A SELF-INTEREST STUDY.

5. (3 points) What is wrong with the following survey question: *Are you in favor of a border wall that will keep out criminals?*

   IT IS A LOADED QUESTION. ALMOST EVERYONE WOULD WANT TO KEEP CRIMINALS AWAY, BUT THAT IS NOT NECESSARILY RELATED TO A BORDER WALL.

6. (8 points) Determine whether the data are discrete or continuous.

   (a) Decelerations measured in car crash tests

   CONTINUOUS

   (b) Sizes of shoes

   DISCRETE

   (c) Heights of past U.S. presidents

   CONTINUOUS

   (d) Numbers of wheels on vehicles

   DISCRETE

7. (4 points) Determine whether each number is a parameter or a statistic.

   (a) There were 2223 passengers aboard the *Titanic* when it sank.

   PARAMETER

   (b) The median age of all U.S. presidents at inauguration is 55.5 years.

   PARAMETER

8. (4 points) Determine the most appropriate level of measurement. Choose from nominal, ordinal, interval, or ratio.

   (a) Colors of M&M's in a bag of candy

   NOMINAL

   (b) Depths of earthquakes measured in kilometers

   RATIO

9. (15 points) The frequency distribution below summarizes the magnitudes of the world's stronger earthquakes over the last week.

Magnitude	Frequency
4.5–4.8	48
4.9–5.2	28
5.3–5.6	9
5.7–6.0	3
6.1–6.4	2

$48+28+9+3+2 = 90$

(a) What are the class boundaries associated with the last class listed above?

6.05 & 6.45

(b) What is the class width?

$4.9-4.5 = \underline{0.4}$

(c) What are the class midpoints?

$\frac{4.5+4.8}{2} = 4.65,\ 5.05,\ 5.45,\ 5.85,\ 6.25$

(d) If the frequency distribution above was changed to a cumulative frequency distribution, how many earthquakes would be associated with magnitudes of "≤ 5.6"?

$48+28+9 = \underline{85}$

(e) Use class midpoints to estimate the mean (weighted) magnitude.

$$\frac{48(4.65)+28(5.05)+9(5.45)+3(5.85)+2(6.25)}{90} = \underline{4.93}$$

(f) Use class midpoints to estimate the median (weighted).

90 VALUES $\Rightarrow$ $\text{MED} = \frac{45^{TH}+46^{TH}}{2} = \underline{4.65}$

(g) Based on the table above, how would you describe the distribution of data.

SKEWED RIGHT.

10. (12 points) Use *frequency polygon*, *dot plot*, *bar graph*, *time-series graph*, *scatterplot*, *pie chart*, *ogive*, *histogram*, *stem-and-leaf plot*, or *Pareto chart* to answer each question. You may get partial credit if you offer brief explanations.

(a) What type of graph shows cumulative frequencies of data separated into classes?

Ogive

(b) What type of graph should be used to show how portions of a whole are divided among categories?

Pie chart

(c) What type of bar graph has the bars arranged in descending order according to their heights?

Pareto chart

(d) A six-sided die is rolled 15 times. What type of graph would be appropriate for showing how many times each outcome was obtained?

Dot plot

(e) What type of graph would be used to display the temperature throughout the day?

Time-series graph

(f) What type of graph simply displays a collection of plotted ordered pairs?

Scatterplot

11. (4 points) Some teens and adults were asked how much cash they were carrying. 12 teens were carrying a mean amount of \$22.15, and 20 adults were carrying a mean amount of \$12.80. What was the mean amount carried by the entire group of people?

$$\bar{X} = \frac{12(22.15) + 20(12.80)}{12 + 20} = \frac{521.80}{32}$$

$$\approx 16.31$$

12. (10 points) What type of sampling is described in each situation. Choose the most appropriate choice from random, systematic, convenience, stratified, or cluster.

(a) A researcher takes a sample of lake water, every 5 meters along a line as she crosses the lake.

Systematic

(b) A student collected sample data by randomly selecting 20 different pages from a dictionary and counting every word that was defined on each of those pages.

Cluster

(c) On the day of a recent election, a newspaper organized an exit poll in which specific polling stations were randomly selected and all voters were surveyed as they left the premises.

Cluster

(d) News stations often obtain opinions by interviewing neighbors of the people who are the focus of the news story.

Convenience

(e) Air quality was evaluated along I-80 by taking an air sample at every 10th mile marker.

Systematic

13. (14 points) In the following display, 4 | 5 means 45.

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

(a) What is the name of this type of display?

STEM AND LEAF PLOT

(b) Are the data values shown above approximately normally distributed? Briefly explain.

APPROXIMATELY NORMAL, YES

THEY ARE PEAKED IN THE 50's ARE SMOOTHLY DECREASE EVENLY ON BOTH SIDES.

(c) Compute the mean, median, and mode(s).

$\bar{X} = \frac{31+34+\cdots+70}{15} = \frac{785}{15} = \underline{52.\overline{3}}$

MED = 8TH VALUE = $\underline{52}$

MODES ARE 50 & 68 (BOTH OCCUR TWICE)

(d) Compute the range.

$70 - 31 = \underline{39}$

(e) Briefly explain how the values of the mean and median support your conclusion in part (b).

IN A DISTRIBUTION THAT IS NORMAL,

MED = MEAN.

IN THIS CASE, THE MEAN & MEDIAN ARE VERY CLOSE.

14. (8 points) Sammy Sosa was a major league baseball player from 1989 to 2007. His numbers of yearly regular season home runs are shown below in the order in which they occurred.

4	15	10	8	33	25	36	40	36
66	63	50	64	49	40	35	14	21

(a) Use your calculator to compute the mean and sample standard deviation.

$\bar{x} = 33.8\overline{3} \approx 33.8$

$s \approx 19.5$

(b) Use your results from above to determine what numbers of home runs would be unusually small or unusually large.

$33.8 - 2(19.5) = -5.2 \rightarrow$ No number of home runs is unusually small

$33.8 + 2(19.5) = 72.8$

$\rightarrow$ 73 or more is unusually large.

15. (5 points) For Yellowstone's Old Faithful geyser, the mean time between eruptions is 1.55 hr with a standard deviation of 0.11 hr. For Yellowstone's Lone Star geyser, the mean is 3.00 hr with a standard deviation of 0.16 hr. Compute the coefficient of variation (CV) for each geyser. Which geyser's eruption cycle has more variation?

OF...

$CV = \frac{0.11}{1.55} \approx 7.1\%$

LS...

$CV = \frac{0.16}{3.00} \approx 5.3\%$

Old Faithful has more variation.

16. (3 points) Eight males and seven females are selected at random. Is this a simple random sample of size 15? Explain.

No, because some samples of size 15 (e.g. 15 males) are impossible. Therefore not all samples of 15 are equally likely.