Math 206 - Test 2

March 14, 2012

Name __key__

Score _____

Show all work. Supply explanations where necessary.

1. (4 points) The pH level of the water in a volcanic hot spring was measured daily for 15 days. The data are shown in the diagram below.

$$
\begin{array}{c|cccccc}
6 & 0 & 1 \\
5 & 5 & 7 & 7 & 8 \\
4 & 2 & 3 & 3 & 8 & 8 & 8 \\
3 & 5 & 8 & 8 \\
\end{array}
$$

4|2 represents 4.2

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

STEM-AND-LEAF PLOT

(b) What are some of the advantages and disadvantages of this type of display?

ADVANTAGES: ALL DATA IS SHOWN. DATA IS IN NUMERICAL ORDER. TURNING PLOT SIDEWAYS RESEMBLES A BAR GRAPH OR HISTOGRAM. SHOWS DISTRIBUTION OF DATA.

DISADVANTAGES: NOT GOOD FOR LARGE COLLECTIONS OF DATA.

(c) Could a line graph be an appropriate type of display for this data? Explain.

ABSOLUTELY! LINE GRAPHS ARE GOOD FOR SHOWING TRENDS IN TIME. THIS DATA IS MEASURED DAILY, SO AS LONG AS WE KNOW WHICH DATA WERE COLLECTED ON WHICH DAYS, A LINE GRAPH WOULD BE GREAT.

(d) What was the median pH level?

4.8 --- DATA ELEMENT #8

2. (3 points) An experiment has two outcomes, O_1 and O_2, with probabilities 4/7 and 3/7, respectively. Explain how this experiment could be simulated.

PLACE 4 BLUE MARBLES AND 3 RED MARBLES IN A JAR. SELECT ONE MARBLE AT RANDOM. CHOOSING A BLUE MARBLE CORRESPONDS TO O_1 (PROB 4/7) AND CHOOSING A RED MARBLE CORRESPONDS TO O_2 (PROB 3/7).

1

37, 45, 45, 52, 79, 98, 98, 98

3. (3 points) Consider the following collection of test scores.

$$98 \quad 52 \quad 98 \quad 45$$
$$37 \quad 45 \quad 98 \quad 79$$

(a) Without using your calculator's statistical features, find the mean, median, and mode of the test scores.

$$\bar{X} = \frac{98 + 52 + 98 + 45 + 37 + 45 + 98 + 79}{8} = \frac{552}{8} = 69$$

$$MED = \frac{52 + 79}{2} = 65.5$$

$$MODE = 98$$

(b) Which is not an appropriate measure of the center of the data set and why?

THE MODE IS NOT APPROPRIATE BECAUSE IT

DOESN'T REPRESENT THE "CENTER" OF THE DATA.

4. (6 points) The boxplot shown below describes a certain collection of data. Find approximate values for the median, first and third quartiles, and the interquartile range. Based on your approximations, what would be the cutoff values for outliers?

$$1 \quad 2 \quad 3 \quad 4 \quad 5 \quad 6 \quad 7 \quad 8 \quad 9 \quad 10$$
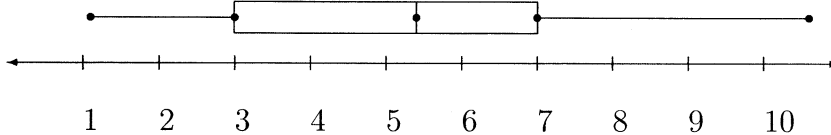
MEDIAN $\approx$ 5.4

$Q_1 \approx 3$

$Q_3 \approx 7$

IQR $\approx 7 - 3 = 4$

CUTOFFS:

$$3 - 1.5(4) = -3$$

$$7 + 1.5(4) = 13$$

2

5. (3 points) Following are the ages of 30 children who participated in Krug School's science fair.

$$
\begin{array}{cccccccccc}
10 & 10 & 11 & 10 & 13 & 8 & 10 & 13 & 14 & 9 \\
14 & 13 & 10 & 14 & 11 & 9 & 13 & 10 & 11 & 12 \\
11 & 12 & 14 & 13 & 12 & 8 & 13 & 14 & 9 & 14
\end{array}
$$

(a) Sketch a line plot (dot plot) that displays the data.

SEE ATTACHED GRAPH PAPER.

(b) Use your calculator to compute the standard deviation.

$$\sigma = 1.91$$

6. (3 points) A contestant on a game show must select a suitcase containing money or be paid by the host to stop playing. There are 10 suitcases: six contain $10, two contain $100, one contains $1000, and one contains $10000. How much should the host pay the contestant to stop playing? (Hint: Think about expected value.)

THE HOST SHOULD BE WILLING TO PAY THE EXPECTED

VALUE (AT LEAST):

EXPECTED VALUE =

$$\frac{6}{10}(\$10) + \frac{2}{10}(\$100) + \frac{1}{10}(\$1000) + \frac{1}{10}(\$10000)$$

$$= 6 + 20 + 100 + 1000$$

$$= \$1126$$

3

7. (5 points) Fred has ten projects that are due next week. He has already decided that he will not do any of them on Saturday or Sunday. So he has ten projects that he will randomly schedule on the five remaining days. (Each project will be randomly assigned a day.)

(a) Explain how Fred could use a random-digit table to assign his projects to days.

THIS IS LIKE THE MONTANA DUCK HUNTER PROBLEM. PICK 10 DIGITS --- ONE FOR EACH PROJECT. IF THE DIGIT IS 0-1, THE PROJECT IS DONE MONDAY, 2-3 TUES, 4-5 WED, 6-7 THURS, 8-9 FRI.

(b) Use your random-digit table to do eight trials.

① 77705 28891 = Th, Th, Th, M, W, T, F, F, F, M
② 12106 56281 = M, T, M, M, Th, W, Th, T, F, M
③ 86222 66116 = F, Th, T, T, T, Th, Th, M, M, Th
④ 39626 06080 = T, F, Th, T, Th, M, Th, M, F, M

⑤ 05505 45420
⑥ 44016 79662
⑦ 92069 27628
⑧ 50002 32540

(c) Based on your trials, what is the probability that Fred will have at least one day off (in addition to Saturday and Sunday).

③ 1 DAY OFF   ⑤ 2 DAYS OFF
④ 1 DAY OFF   ⑦ 1 DAY OFF
                ⑧ 2 DAYS OFF

3 FAILURES
5 SUCCESSES

$\Rightarrow$ PROB $\approx \dfrac{5}{8}$

8. (3 points) A paper company made an annual revenue of $275 million: $26 million from the sales of toliet paper, $38 million from the sales of paper cups, $20 million from the sales of paper plates, $32 million from the sales of tissues, $120 million from the sales of napkins, and $39 million from the sales of miscellaneous other paper products.

(a) Explain why a circle graph is particularly appropriate for displaying this data.

A CIRCLE GRAPH IS PERFECT FOR SHOWING PORTIONS OF A WHOLE.

(b) Recall that there are 360 degrees in a circle. Suppose you constructed the circle graph corresponding to the data above. What should be the degree measure of the angle in the portion associated with napkins?

$$\frac{120}{275} \times 360° = 157.\overline{09}°$$

9. (4 points) For each of the following situations, tell which type of graph would best display the data. Choose from *line plot, bar graph, histogram, line graph, stem-and-leaf plot*, or *circle graph*. Give a brief explanation to support your answer.

(a) A company has expenditures that can be grouped into seven broad categories. Company executives would like to make a graph showing how the expenditures are divided among the seven categories.

THE EXECUTIVES WANT TO SHOW PARTS OF A WHOLE. THEY SHOULD USE A CIRCLE GRAPH.

(b) Emma is writing a report about Greenland. In her report she would like to include a graph that shows the values of the five highest grossing exported products in 2010.

EMMA WILL PROBABLY SHOW DOLLAR AMOUNTS ASSOCIATED WITH THE 5 INDIVIDUAL PRODUCTS. SHE SHOULD USE A BAR GRAPH.

(c) The National Center for Health Statistics keeps detailed records on the births and deaths of US residents. What type of graph would best display the birth weights of US babies in 2010?

BIRTH WEIGHTS ARE CONTINUOUS DATA. HISTOGRAM WOULD BE BEST.

(d) Mr. Smith has been keeping track of the daily price of IBM common stock. He would like to make a graph showing how the prices have changed over the last three weeks.

A LINE GRAPH IS BEST FOR SHOWING TRENDS IN TIME.

10. (3 points) The class mean on a reading test was 27.5 out of 40 possible points. The 19 girls in the class scored a total of 532 points. If there were 11 boys in the class, what was the mean of the boys' scores?

$$\text{Sum of all scores} = (19 + 11) \times 27.5 = 825$$

$$\text{Sum of Boys scores} = 825 - 532 = 293$$

$$\text{Mean of Boys scores} = \frac{293}{11} = 26.\overline{63}$$

11. (2 points) Willis scored 95 on a test with mean 82 and standard deviation 8.4. Gloria scored 193 on a test with mean 156 and standard deviation 22. Compute their z-scores. Who had a better score?

Willis: $z = \dfrac{95 - 82}{8.4} \approx 1.55$ ← His score is 1.55 standard deviations above mean

Gloria: $z = \dfrac{193 - 156}{22} \approx 1.68$ ← Her score is 1.68 standard deviations above mean

Gloria scored better.

12. (3 points) Joe had scores of 90, 95, 85, 90, and 20 on his tests. Without computing them, which measure of center (mean, median, mode) do you think Joe would want his teacher to use for his "average" test score? Explain.

The mode will be best for Joe because

it is completely unaffected by the

low score of 20.

Coincidentally, the median and the mode

are both the same.

Definitely not the mean, because it will be

pulled down by the low

score.

6

13. (8 points) Following are the weights, in pounds, of 15-day-old pigs raised on two different feeds.

Feed A: 7.2, 9.8, 12.3, 13.1, 13.3, 8.4, 10, 11.6, 9.9, 11.5, 5.8, 9.8, 12.2, 12.3

Feed B: 6.8, 8.5, 9.5, 9.2, 10.1, 11.3, 11.7, 7.9, 7.5, 13.3, 6.2, 5.8

(a) For each data set (separately), compute the median, the first and third quartiles, the IQR, and the outlier cutoffs.

★

FEED A: 5.8, 7.2, 8.4, 9.8, 9.8, 9.9, 10, 11.5, 11.6, 12.2, 12.3, 12.3, 13.1, 13.3

MEDIAN = $\frac{10 + 11.5}{2}$ = 10.75

Q_1 = 9.8

Q_3 = 12.3

IQR = 12.3 - 9.8 = 2.5

CUTOFFS: $Q_1 - 1.5 \times IQR$ = 6.05

$Q_3 + 1.5 \times IQR$ = 16.05

5.8 IS THE ONLY OUTLIER IN FEED A group.

FEED B: 5.8, 6.2, 6.8, 7.5, 7.9, 8.5, 9.2, 9.5, 10.1, 11.3, 11.7, 13.3

MEDIAN = $\frac{8.5 + 9.2}{2}$ = 8.85

Q_1 = $\frac{6.8 + 7.5}{2}$ = 7.15

Q_3 = $\frac{10.1 + 11.3}{2}$ = 10.7

IQR = 10.7 - 7.15 = 3.55

CUTOFFS: $Q_1 - 1.5 \times IQR$ = 1.825

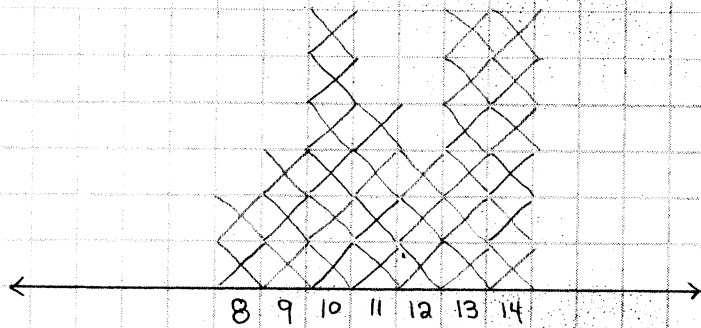$Q_3 + 1.5 \times IQR$ = 16.025

No OUTLIERS IN FEED B group.

(b) On graph paper, sketch the boxplot for each feed. Sketch one boxplot above your number line and one boxplot below.

SEE ATTACHED SHEET.

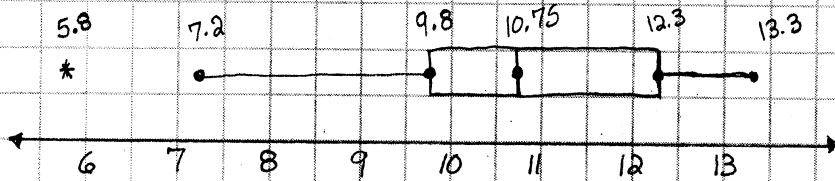(c) Compare the boxplots. Which feed is better and why do you think so?

BOTH COLLECTIONS OF WEIGHTS HAVE THE SAME RANGE, BUT THE SPREAD IN THE FEED A DATA IS SMALLER. FEED A MORE CONSISTENTLY PRODUCES GREATER WEIGHTS. IN FACT, THE UPPER HALF OF THE FEED A DATA LIES IN THE UPPER EXTREME OF THE FEED B DATA.

7

#5a

8 9 10 11 12 13 14

#13b

FEED A

5.8   7.2   9.8   10.75   12.3   13.3
*

6   7   8   9   10   11   12   13

FEED B

5.8   7.15   8.85   10.7   13.3