JCC Statistics

Final Exam Review Ch 3 & 4

Name:

Date:

1. The following data represents the amount of snowfall (in inches) received at Mount Washington from the winter of 1996-97 and the winter of 2005-06. Treat the data as a sample of size 10.

399	542	347	381	425
299	358	349	620	616

   a. Determine the mean amount of snowfall. $433.6'' = \bar{x}$

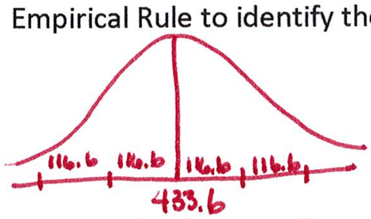
   b. Determine the median amount of snowfall $390''$

   c. Calculate the standard deviation for the amount of snowfall $s = 116.6''$

   d. Calculate the interquartile range (IQR)of the snowfall (show all calculations)

   $542 - 349 = 193$

   e. Assuming that snowfall amounts on Mt Washington are normally distributed, Use the Empirical Rule to identify the interval between which 95% of the data would lie.



   1 sd. = 68%

   2 sd. = 95%

   3 sd = 99%

   $200.4'' - 666.8''$

   f. Use Cebyshev's Inequality to determine the percent of data that lies between snowfall amounts of 550 inches and 565 inches given a mean of 557.5 and a standard deviation of 2.5.

   $\underbrace{550 - 557.5}_{\substack{7.5 \\ 3sd}} \underbrace{- 565}_{\substack{7.5 \\ 3sd}}$

   $K = 3$

   $1 - \frac{1}{K^2} = 1 - \frac{1}{9} = \frac{8}{9} = 89\%$

2. The following data represents the length of time (in minutes) it takes to ride the train from home to work.

Time(minutes)	Frequency (L_2)	x_i (L_1)	$x_i f_i$ (L_3)	$x_i^2 f_i$ (L_4)
40-49	8	45	360	16200
50-59	44	55	2420	133100
60-69	23	65	1495	97175
70-79	16	75	1200	90000
80-89	107	85	9095	773075
	198		14570	1109550

a. Approximate the mean commute time.

$$\mu = \frac{\sum x_i f_i}{\sum f_i} = \frac{14570}{198} = 73.6$$

b. Approximate the standard deviation for the commute times.

Changed

$$\sigma^2 = \frac{\sum x_i^2 f_i - \frac{(\sum x_i f_i)^2}{\sum f_i}}{\sum f_i} = \frac{1109550 - \frac{(14570)^2}{198}}{198} = 188.909$$

$$\sigma = 13.7444$$

$$\sigma = 13.7$$

3. Pascal has taken both the SAT and ACT for admission to college. Pascal scored a 610 on the SAT math and 27 on the ACT math. The SAT has a mean math score of 515 with a standard deviation of 114 and the ACT has a mean math score of 21 with a standard deviation of 5.1 . On which test did Pascal score higher?

SAT: $z = \frac{610-515}{114} = 0.83$

ACT: $z = \frac{27-21}{5.1} = 1.17$

Better on ACT

4. For skewed data, which measure of central tendency would be the best measure of the "center " of the data set? Explain.

Median: it is resistent to extreme values (outliers)

It consistently approximates the center of the data set

5. For the following data set, compute the values for the 5 number summary and on graph paper sketch the box and wiskers plot with outliers (*) .

$H = 33$
$L = 20$
$Q_1 = 21.5$
$M = 25$
$Q_3 = 28$

~~32~~	~~23~~	~~28~~	~~28~~	~~28~~	~~29~~	~~25~~	~~20~~	~~25~~
~~21~~	~~24~~	~~21~~	~~24~~	~~21~~	~~26~~	~~28~~	~~24~~	~~33~~
~~31~~	~~22~~	~~20~~						

20 20 21 21 21 22 23 24 24 25 25 25 26 28 28 28 28 29 31, 32, 33

21.5 M

6. Using the data above, construct a stem and leaf plot.

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

7. What percentile rank does the value of 24 have in the data set above?

~~$\frac{9.5}{21} = 45^{th}$ percentile~~

$X = \frac{\#\text{ of values below}}{n} \times 100\%$ $X = \frac{7}{21} \times 100$

$X = 33\%$

8. Jennifer wishes to create a new hamburger recipe for her restaurant. She decides to combine 2 lbs of ground beef (cost \$2.70/lb), 1 lbs of ground pork (cost \$1.30/lb), and ½ pound of ground turkey (\$1.80/lb). What is the cost per pound of her hamburger mixture?

	Weight	\$	
Gr. Beef	2	2.70	= 5.40
Gr Pork	1	1.30	= 1.30
Gr. turkey	.5	1.80	= .90

$\frac{7.60}{3.5} = 2.17$

9. A study is done to determine the relationship between a driver's age and the number of accidents he or she has over a one-year period. The data are shown in the table below.

Driver's Age, x	16	24	18	17	23	27	32
No. of Accidents, y	3	2	5	2	0	1	1

a. Draw a scatter plot.

b. Compute a correlation coefficient, r . Interpret the meaning of the correlation coefficient.

r= -.61 A moderately strong neg. relationship

c. Determine the regression line equation.

y= -.17x + 5.8

d. Plot the regression line on the scatter plot

e. If there is a significant relationship, predict the number of accidents of a driver who is 28.

y= -.17(28) +5.8 = 1.04 ≈ 1 accident

f. Calculate the residual for a driver who's age is 17 and explain its meaning.

residual = actual - predicted = 2-1 = 1

g. Compute the coefficient of determination (r^2) and explain its meaning.

r^2 = .37

37% of the relationship that exists between age and no. of accidents is explained by the regression equation
63 % is due to other factors