

1. Fill in the blanks: parameters describe populations and we use the symbol  $\mu$  to represent the mean and  $\sigma$  to represent the standard deviation.
2. Fill in the blanks: statistics describe samples and we use the symbol  $\bar{x}$  to represent the mean and  $s$  to represent the standard deviation.

The scores of the worst bowling team in the world are as follows: 64, 58, 78, 82, 79, 50, 79, 102

3. Find the mean, the median, and the mode. Which is the best measure of center? Why?

mean: 74

Median b/c of the outlier.

median: 78.5

mode: 79

4. Find the IQR, the range, and the standard deviation. Which gives the best measure of variability? Why?

IQR:  $78.5 - 61 = 17.5$

Range:  $102 - 50 = 52$

SD: 15.19

Consider the following data sets.

Data set X: 4, 7, 1, 9, 10, 3, 6, 8

Data set Y: 0, 6, 6, 2, 5, 3, 3, 8

1, 3, 4, 6, 7, 8, 9, 10, 20

0, 2, 3, 3, 5, 6, 6, 8

5. Find the Five Number Summary for Data set X and draw a box and whisker plot.

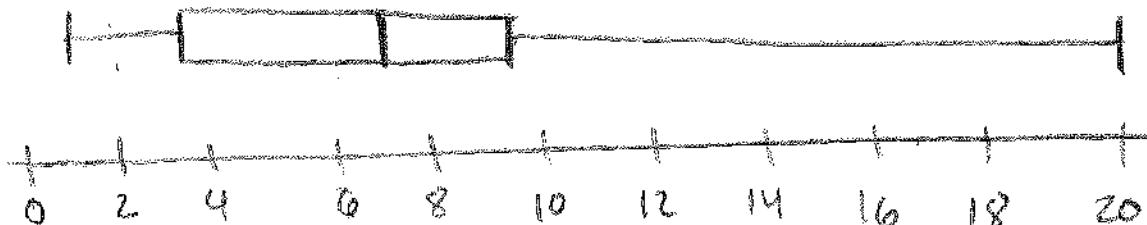
Median = 7

Q1 = 3.5

Q3 = 9.5

Max = 20

Min = 1



6. Find the Five Number Summary for Data set Y and draw a box and whisker plot.

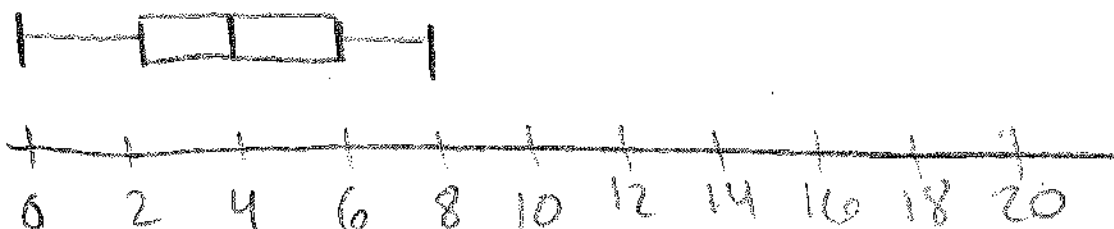
Median = 4

Q1 = 2.5

Q3 = 6

Max = 8

Min = 0



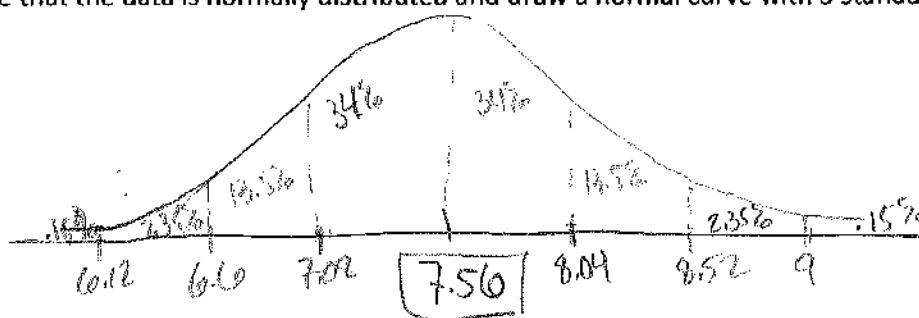
Trina is conducting research on white pine trees for her graduate degree in environmental science. The lengths of the first sample of white pine needles in her study are listed in the table below:

Lengths of White Pine Needles (cm)					
7.4	7.7	7.9	7.7	8.4	7.5
8.1	7.1	7.6	8.6	7.5	6.5
7.6	7.3	7.1	7.7	7.5	6.6
7.5	7.2	7.8	8.5	7.6	7.0
7.6	7.3	8.2	7.7	7.5	7.0

7. Find the mean and the standard deviation of the data set (use your calculator!!). Round to three numbers after the decimal.

$$\bar{x} = 7.56 \quad \sigma = .48$$

8. Assume that the data is normally distributed and draw a normal curve with 3 standard deviations from the mean.



9. Use the empirical rule and the information from problems 7 and 8 to find the following:

- the percent of data between 6.6 and 8.04 cm. 21.5%
- the percent of data is less than 7.56 cm. 50%
- how many pine trees (from the study) had needles more than 7.08 needles? 8%

$$.08(30) = 2.4 \approx 25 \text{ pine trees}$$

10. Use the information on pine needles to find the z-score for:

$$\text{a. } X=6.88 \quad z = \frac{6.88 - 7.56}{.48} = -1.42$$

$$\text{c. } X=8.6 \quad z = \frac{8.6 - 7.56}{.48} = 2.17$$

$$\text{b. } X=8.52 \quad z = \frac{8.52 - 7.56}{.48} = 2$$

$$\text{d. } X=7.45 \quad z = \frac{7.45 - 7.56}{.48} = -.23$$

11. Given the following z-scores, find the data value (X) for a normal distribution with a mean of 45 and a standard deviation of 11:

$$\text{a. } z=1.23 \quad 1.23 = \frac{X-45}{11}$$

$$X = 58.53$$

$$\text{c. } z=-2.8 \quad -2.8 = \frac{X-45}{11}$$

$$X = 14.2$$

$$\text{b. } z=0.87 \quad .87 = \frac{X-45}{11}$$

$$X = 54.57$$

$$\text{d. } z=-3 \quad -3 = \frac{X-45}{11}$$

$$X = 12$$

12. Given a normal distribution with a mean of 15.2 and a standard deviation of 2.2, find the following probabilities:

a.  $P(X > 16) = P(Z > .36)$

c.  $P(X < 21.1) = P(Z < 2.68)$

$z = \frac{16 - 15.2}{2.2} = .36$   
 $1 - .6406 = .3594$

$z = \frac{21.1 - 15.2}{2.2} = 2.68$   
 $.9963$

b.  $P(14 < X < 18.9) = P(-.54 < Z < 1.68) = .9535$

d.  $P(9 < X < 13.4) = P(-2.81 < Z < -.81)$

$z(14) = \frac{14 - 15.2}{2.2} = -.54$   
 $z(18.9) = \frac{18.9 - 15.2}{2.2} = 1.68$   
 $.2946$   
 $.6589$

$z(9) = -2.81$   
 $z(13.4) = -.81$   
 $.2090$   
 $.2025$   
 $.2065$

Use the information from the pine needles problem to answer questions 13-15.

13. What is the probability that a pine needle will have a length greater than 8.55cm?

$P(X > 8.55) = P(Z > 2.06) = 1 - .9803 = .0197$

14. What is the probability that a pine needle's length will fall between 5.8 and 7 cm?

$P(5.8 < X < 7) = P(-1.17 < Z < -1.17) =$  Can't do because of this # whoops!

15. What is the probability that a pine needle's length will be less than 6.83cm?

$P(X < 6.83) = P(Z < -1.52) = .0643$

16. If 300 water bills are randomly selected from a group that has a mean of \$100 and a standard deviation of \$12, about how many water bills would you expect to be less than \$90?

$P(X < 90) = P(Z < -.83) = .2033$

$.2033 (300) = 60.99 \approx 61$

17. A study is done to try to determine the mean number of tweets posted by high school students in one day. In a random sample of 40 students, the mean number of tweets was 6.2 with a standard deviation of 0.9.

a. What is the population being studied?

High School Students

b. What is the variable of interest?

# of tweets in one day

c. Assume the data is normally distributed; construct and interpret a 90% confidence interval for the population mean.

Construct:  $6.2 \pm 1.645 \left( \frac{.9}{\sqrt{40}} \right)$   
 $6.2 \pm .234$   
 $(5.966, 6.434)$

Interpret: I am 90% confident that the average # of tweets posted by h.s. students in one day will be between 5.966 and 6.434 tweets.

18. Of 1012 high school softball players, 83 suffered an injury while playing.

a. In this problem, what is the population and what is the sample?

High School Softball players

b. What is the sample proportion?

$\hat{p} = \frac{83}{1012} = .082$

c. Construct a 99% confidence interval using these statistics.

Construct:  $.082 \pm 2.576 \left( \sqrt{\frac{.082(1-.082)}{1012}} \right)$   
 $.082 \pm .022$   
 $(.06, .104)$

Interpret: I am 99% confident that p, the proportion of h.s. softball players that suffer an injury while playing, will be between .06 and .104.

For each description of sampling, decide if the sampling technique is

- A. Simple Random
- B. Stratified
- C. Cluster
- D. Systematic
- E. Convenience

- D 19. In order to estimate the percentage of defects in a recent manufacturing batch, a quality control manager at Intel selected every 8<sup>th</sup> chip that comes off the assembly line starting with the 3<sup>rd</sup>, until she obtains a sample of 140 chips.
- C 20. In order to determine the average IQ of ninth-grade students, a school psychologist obtains a list of all high schools in the local public school system. She randomly selects five of these schools and administers an IQ test to all ninth-grade students at the selected schools.
- C 21. In an effort to determine customer satisfaction, United Airlines randomly selects 50 flights during a certain week and surveys all passengers on the flights.
- B 22. A member of Congress wishes to determine her constituency's opinion regarding estate taxes. She divides her constituency into three income classes: low-income households, middle-income households, and upper-income households. She then takes a random sample of households from each income class.
- A 23. In an effort to identify whether an advertising campaign has been effective, a marketing firm conducts a nationwide poll by randomly selecting individuals from a list of known users of the product.
- E 24. A radio station asks its listeners to call in their opinion regarding the use of American forces in peacekeeping missions. *→ really - voluntary*
- C 25. A farmer divides his orchards into 50 subsections, randomly selects 4 and samples all of the trees within the 4 subsections in order to approximate the yield of his orchard.
- B 26. A school official divides the student population into five classes: freshman, sophomore, junior, senior and graduate student. The official takes random sample from each class and asks the members' opinions regarding student services.
- E 27. A survey regarding download time on a certain Web site is administered on the Internet by a market research firm to anyone who would like to take it.
- D 28. A lobby has a list of 100 senators of the United States. In order to determine the Senate's position regarding farm subsidies, they decide to talk to every seventh senator on the list starting with the third.

Decide if the following are observational or experimental studies.

29. One study of cell phones and the risk of brain cancer looked at a group of 469 people who have brain cancer. The investigators matched each cancer patient with a person of the same age, sex, and race who did not have brain cancer, then asked about the use of cell phones. Result: "Our data suggest that the use of hand-held cellular phones is not associated with risk of brain cancer." Is this an observational study or an experiment? Justify your answer.

- Observation - did not impose anything on subjects, just compared*
30. An educational software company wants to compare the effectiveness of its computer animation for teaching biology with that of a textbook presentation. The company gives a biology pretest to each of a group of high school juniors, and then divides them into two groups. One group uses the animation, and the other studies the test. The company retests all students and compares the increase in biology test scores in the two groups. Is this an observational study or an experiment? Justify your answer

*Experiment The company decided one group would use the animation, one would not. Something was imposed on the subjects*