Freeport AP Statistics

Chapter 2: Modeling Distributions of Data

2.1 Describing Location in a Distribution

OBJECTIVE(S):

- ***Students will learn to use percentiles to locate individual values within distributions of data.***
- ***Students will learn to interpret a cumulative relative frequency graph.***
- ***Students will learn to find the standardized value (z-score) of an observation and interpret z-scores in context.***
- ***Students will learn to describe the effect of adding, subtracting, multiplying by, or dividing by a constant on the shape, center, and spread of a distribution of data.***
- ***Students will learn to approximately locate the median (equal-areas point) and the mean (balance point) on a density curve.***

***Percentile* -**

*Chapter opener: Activity (page 84)

______________________ _____________________ on a test does not measure the same thing as a _______________________.

Percentiles should be ________________ _________________, so if you get a decimal, you should round your answer to the _________________ ________________. For example, in a class of 13 people the student with the highest score would be at about the _________ percentile (________________).

If two observations have the same value, they will be at the ____________ percentile. To find the percentile, calculate the _____________ of the values in the distribution that are _____________ *both* values.

1. The stemplot below shows the number of wins for each of the 30 Major League Baseball teams in 2009.

Key: **5|9** represents a team with 59 wins.

5	**9**									
6	**2**	**4**	**5**	**5**						
7	**0**	**0**	**4**	**5**	**5**	**5**	**8**	**9**		
8	**0**	**3**	**4**	**5**	**6**	**6**	**7**	**7**	**7**	**8**
9	**1**	**2**	**3**	**5**	**5**	**7**				
10	**3**									

Find the percentiles for the following teams:

a. The Colorado Rockies, who won 92 games.

b. The New York Yankees, who won 103 games.

c. The Kansas City Royals and Cleveland Indians, who both won 65 games.

2. Larry came home very excited after a visit to his doctor. He announced proudly to his wife, “My doctor says my blood pressure is at the 90^{th} percentile among men like me. That means I’m better off than about 90% of similar men.” How should his wife, who is a statistician, respond to Larry’s statement?

3. Peter is a star runner on the track team. In the league championship meet, Peter records a time that would fall at the 80^{th} percentile of all his race times that season. But his performance places him at the 50^{th} percentile in the league championship meet. Explain how this is possible. (Remember that lower times are better in this case!).

Cumulative Relative Frequency Graphs (____________) –

4. **TEXTBOOK p. 100 #10**.
   a. Estimate the 60th percentile of this distribution.

   b. What is the percentile for a lamp that lasted 900hours?

Standardized value (z-score) –

A *z*-score tells us how many ______________________ ______________________ from the ___________ an observation falls, and in what ________________.

5. What do positive *z*-scores tell us?

6. What do negative *z*-scores tell us?

7. Three landmarks of baseball achievement are Ty Cobb's batting average of .420 in 1911, Ted Williams's .406 in 1941, and George Brett's 0.390 in 1980. These batting averages cannot be compared directly because the distribution of major league batting averages has changed over the years. The distributions are quite symmetric, except for outliers such as Cobb, Williams, and Brett. While the mean batting average has been held roughly constant by rule changes and the balance between hitting and pitching, the standard deviation has dropped over time. Here are the facts:

Decade	Mean	Standard Deviation
1910s	0.266	0.0371
1940s	0.267	0.0326
1970s	0.261	0.0317

Compute the standardized batting averages for Cobb, Williams, and Brett to compare how far each stood above his peers.

Effect of Adding (or Subtracting) a Constant –

-

-

Effect of Multiplying (or Dividing) a Constant –

-

-

-

8. A school system employs teacher's salaries between $28,000 and $60,000. The teachers' union and the school board are negotiating the form of next year's increase in the salary schedule.

   a. If every teacher is given a flat $1,000 raise, what will this do to the mean salary? To the median salary? Explain your answers.

   b. What would a flat $1,000 raise do to the extremes and quartiles of the salary distribution? To the standard deviation of teachers' salaries? Explain your answers.

c. If each teacher receives a 5% raise instead of a flat $1,000 raise, the amount of the raise will vary from $1,400 to $3,000, depending on the present salary.
   i. What will this do to the mean salary? To the median salary? Explain your answers.

   ii. Will a 5% increase the IQR? Will it increase the standard deviation? Explain your answers.

9. Clarence measures the diameter of each tennis ball in a bag with a standard ruler. Unfortunately, he uses the ruler incorrectly so that each of his measurements is 0.2 inches too large. Clarence's data had a mean of 3.2 inches and a standard deviation of 0.1 inches. Find the mean and standard deviation of the corrected measurements in centimeters (recall that 1 inch = 2.54 cm).

Freeport AP Statistics

Chapter 2: Modeling Distributions of Data

2.2 Density Curves and Normal Distribution

OBJECTIVE(S):

- Students will learn to use the 68-95-99.7 rule to estimate the percent of observations from a Normal distribution that fall in an interval involving points one, two, or three standard deviations on either side of the mean.
- Students will learn to use the standard Normal distribution to calculate the proportion of values in a specified interval.
- Students will learn to use the standard Normal distribution determine a *z*-score from a percentile.
- Students will learn to make an appropriate graph to determine if a distribution is bell-shaped.
- Students will learn to use their calculator to find the percentile of a value from any Normal distribution and the value that corresponds to a given percentile.
- Students will learn to use the 68-95-99.7 rule to assess Normality of a data set.
- Students will learn to interpret a Normal probability plot.

A *Density Curve* is a curve that:

-

-

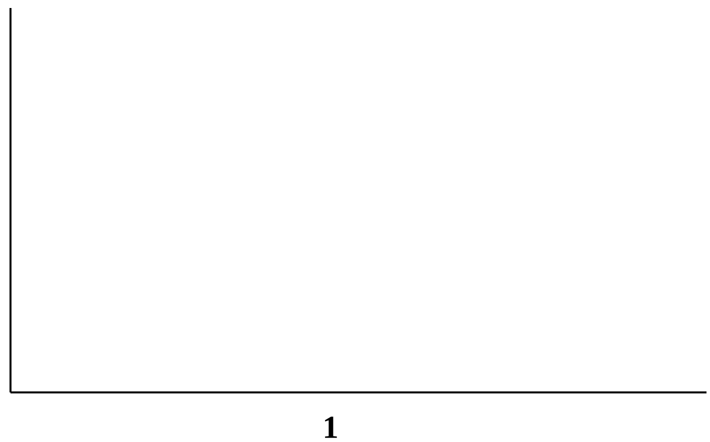
Distinguishing the Median and Mean of a Density Curve

Median –

Mean -

10. Sketch a density curve that might describe a distribution that has a single peak and is skewed to the left.

11. The figure below displays the density curve of a uniform distribution. The curve takes the constant value 1 over the interval from 0 to 1 and is 0 outside the range of values. This means that data described by this distribution take values that are uniformly spread between 0 and 1.



a. Explain why this curve satisfies the two requirements for a density curve.

b. What percent of the observations are greater than 0.8?

c. What percent of the observations lie between 0.25 and 0.75?

d. What is the mean μ and the median of the density curve?

12. **TEXTBOOK p. 129 #40**.

a. Mean = ______ Median = _____

b. Mean = ______ Median = _____

DAY 2

13. What 4 important facts do normal curves illustrate?

-
-
-
-

13. Where do the points of inflection take place on every normal curve?

***<u>Normal Distribution</u>* –**

***<u>Normal Curve</u>* –**

14. Why are normal distributions important in statistics?

a.

b.

c.

<u>The 68-95-99.7 rule ("Empirical Rule")</u>

-
-
-

Standard Normal Distribution -

15. The distribution of weights of 9-ounce bags of a particular brand of potato chips is approximately normal with mean $\mu = 9.12$ ounces and standard deviation $\sigma = 0.05$ ounce. Draw an accurate sketch of the distribution of potato chip bag weights. Be sure to label the mean, as well as the point's one, two, and three standard deviations away from the mean on the horizontal axis.

Use the 68-95-99.7 Rule to answer the following questions.

a. What percent of bags weigh less than 9.02 ounces?

b. Between what weights do the middle 68% of bags fall?

c. What percent of 9-ounce bags of this brand of potato chips weigh between 8.97 and 9.17 ounces?

d. A bag that weighs 9.07 ounces is at what percentile in this distribution?

16. Use the calculator to find the following probabilities:
    a. $z < -2.46$

    b. $z > 2.46$

    c. $0.89 < z < 2.46$

    d. $-2.95 < z < -1.27$

    e. z is between -2.05 and 0.78

    f. z is between -1.11 and -0.32

17. Find the value z from the standard Normal distribution that satisfies each of the following conditions. In each case, sketch a standard Normal curve with your value of z marked on the axis.
    a. The 63rd percentile.

    b. 75% of all observations are greater than z.

18. Scores on the Wechsler Adult Intelligence Scale (a standard IQ test) for the 20 to 34 age group are approximately Normally distributed with $\mu = 110$ and $\sigma = 25$.

    a. At what percentile is an IQ score of 150?

    b. What percent of people aged 20 to 34 have IQs between 125 and 150?

    c. MENSA is an elite organization that admits as member's people who score in the top 2% on IQ tests. What score on the Wechsler Adult Intelligence Scale would an individual have to earn to qualify for MENSA membership?

19. At some fast-food restaurants, customers who want a lid for their drinks get them from a large stack left near straws, napkins, and condiments. The lids are made with a small amount of flexibility so they can be stretched across the mouth of the cup and then snuggly secured. When lids are too small or too large, customers can get very frustrated, especially if they end up spilling their drinks. At one particular restaurant, large drink cups require lids with a "diameter" of between 3.95 and 4.05 inches. The restaurants lid supplier claims that the mean diameter of their large lids is 3.98 inches with a standard deviation of 0.02 inches. Assume that the supplier's claim is true.
    a. What percent of large lids are too small to fit?

    b. What percent of large lids are too big to fit?

    c. Compare your answers to a. and b. Does it make sense for the lid manufacturer to try to make one of these values larger than the other? Why or why not?

(19 cont'd): The supplier is considering two changes to reduce the percent of its large-cup lids that are too small to less than 1%. One strategy is to adjust the mean diameter of its lids. Another option is to alter the production process, thereby decreasing the standard deviation of the lid diameters.

d. If the standard deviation remains at $\sigma = 0.02$ inches, at what value should the supplier set the mean diameter of its large-cup lids to ensure that less than 1% are too small to fit?

e. If the mean diameter stays at $\mu = 3.98$ inches, what value of the standard deviation will result in less than 1% of lids that are too small to fit?

f. Which of the two options in d. and e. do you think is preferable? Justify your answer. (Be sure to consider the effect of these changes on the percent of lids that are too large to fit.)

DAY 3

20. The amount of time Ricardo spends brushing his teeth follows a |Normal distribution with unknown mean and standard deviation. Ricardo spends less than one minute brushing his teeth about 40% of the time. He spends more than two minutes brushing his teeth 2% of the time. Use this information to determine the mean and standard deviation of this distribution.

21. In 1798, the English scientist Henry Cavendish measured the density of the earth several times by careful work with a torsion balance. The variable recorded was the density of the earth as a multiple of the density of water. Here are Cavendish's 29 measurements:

5.50	5.61	4.88	5.07	5.26	5.55	5.36	5.29	5.58	5.65
5.57	5.53	5.62	5.29	5.44	5.34	5.79	5.10	5.27	5.39
5.42	5.47	5.63	5.34	5.46	5.30	5.75	5.68	5.85	

a. Draw a histogram summarizing this distribution.

b. Describe the distribution of density measurements.

c. Calculate the percent of observations that fall within one, two, and three standard deviations of the mean. How do these results compare with the 68-95-99.7 rule?

d. Use your calculator to construct a normal probability plot. Interpret this plot.

e. Having inspected the data from several different perspectives, do you think these data are approximately Normal?

22. The heights of people of the same gender and similar ages follow Normal distributions reasonably closely. Weights, on the other hand, are not Normally distributed. The weights of women aged 20 to 29 have mean 141.7 pounds and median 133.2 pounds. The first and third quartiles are 118.3 pounds and 157.3 pounds. What can you say about the shape of the weight distribution? Why?

SUMMARY OF NORMALCDF v. INVNORM

	normalcdf( )	**InvNorm( )**
Given what type of distribution shape?		
What are you given in the problem?		
What does the problem ask you to find?		