**AP Stat Unit 2 Guided Notes**

*All homeworks are due the day of the test. If you finish them early, feel free to turn them in. I highly suggest you do not wait until the night before the test…practice makes perfect!*

**HW #19: page 105 (1, 5, 9, 11, 13, 15) HW #20 page 107 (19–31 odd, 40)**

**HW #21: page 109 (33–38), page 131 (41, 43, 45) HW #22: page 131 (47-53 odd, 56, 58, 59)**

**HW #23 page 132 (54, 60, 68–76)** **HW #24: page 136: Chapter review exercises**

**HW #25: page 138: AP Statistics Practice Test (T2.2: choices are 1, 2, 3, 4, 5; skip T2.6)**

**2.1: Identifying location in a distribution: percentiles and *z*-scores**

**84-85**

What is a percentile? On a test, is a student’s percentile the same as the percent correct?

**Example**: Wins in Major League Baseball

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

 5 9

 6 2455

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

 7 00455589

 8 0345667778

 9 123557

10 3

Calculate and interpret the percentiles for the Colorado Rockies who had 92 wins, the New York Yankees who had 103 wins, and the Cleveland Indians who had 65 wins.

**86–89**

**Example**: State Median Household Incomes

Here is a cumulative relative frequency graph showing the distribution of median household incomes for the 50 states and the District of Columbia.



a) California, with a median household income of $57,445, is at what percentile? Interpret this value.

b) What is the 25th percentile for this distribution? What is another name for this value?

c) Where is the graph the steepest? What does this indicate about the distribution?

Macy, a 3-year-old female is 100 cm tall. Brody, her 12-year-old brother is 158 cm tall. Obviously, Brody is taller than Macy—but who is taller, relatively speaking? That is, relative to other kids of the same ages, who is taller? According to the Centers for Disease Control and Prevention, the heights of three-year-old females have a mean of 94.5 cm and a standard deviation of 4 cm. The mean height for 12-year-olds males is 149 cm with a standard deviation of 8 cm.

**89–91**

How do you calculate and interpret a standardized score (*z*-score)? Do *z*-scores have units? What does the sign of a standardized score tell you?

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Year** | **Player** | **HR** | **Mean** | **SD** |
| 1927 | Babe Ruth | 60 | 7.2 | 9.7 |
| 1961 | Roger Maris | 61 | 18.8 | 13.4 |
| 1998 | Mark McGwire | 70 | 20.7 | 12.7 |
| 2001 | Barry Bonds | 73 | 21.4 | 13.2 |

**Example**: Home run kings

The single-season home run record for major league baseball has been set just three times since Babe Ruth hit 60 home runs in 1927. Roger Maris hit 61 in 1961, Mark McGwire hit 70 in 1998 and Barry Bonds hit 73 in 2001. In an absolute sense, Barry Bonds had the best performance of these four players, because he hit the most home runs in a single season. However, in a relative sense this may not be true. Baseball historians suggest that hitting a home run has been easier in some eras than others. This is due to many factors, including quality of batters, quality of pitchers, hardness of the baseball, dimensions of ballparks, and possible use of performance-enhancing drugs. To make a fair comparison, we should see how these performances rate relative to others hitters during the same year. Calculate the standardized score for each player and compare.

In 2001, Arizona Diamondback Mark Grace’s home run total had a standardized score of *z* = –0.48. Interpret this value and calculate the number of home runs he hit.

**92–97**

What is the effect of adding or subtracting a constant from each observation?

What is the effect of multiplying or dividing each observation by a constant?

In 2010, Taxi Cabs in New York City charged an initial fee of $2.50 plus $2 per mile. In equation form,

*fare* = 2.50 + 2(*miles*). At the end of a month a businessman collects all of his taxi cab receipts and analyzed the distribution of fares. The distribution was skewed to the right with a mean of $15.45 and a standard deviation of $10.20.

a) What are the mean and standard deviation of the lengths of his cab rides in miles?

b) If the businessman standardized all of the fares, what would be the shape, center, and spread of the distribution?

**99–103**

What is a density curve? When would we use a density curve? Why?

 Trace it out on a dot-plot below.

How can you identify the mean and median of a density curve? *Sudoku pg. 83.*

**2.2 Normal Distributions**

**110–111**

According to the CDC, the heights of 12-year-old males are approximately Normally distributed with a mean of 149 cm and a standard deviation of 9 cm. Sketch this distribution, labeling the mean and the points one, two, and three standard deviations from the mean.

Here is a dotplot of Kobe Bryant’s point totals for each of the 82 games in the 2008-2009 regular season. The mean of this distribution is 26.8 with a standard deviation of 8.6 points. In what percentage of games did he score within one standard deviation of his mean? Within two standard deviations?



Points

 *1 SD: 18.2 to 35.4: 57/82 = 69.5%; 2 SD: 9.6 to 44.0: 80/82 = 97.6%*

Here is a dotplot of Tim Lincecum’s strikeout totals for each of the 32 games he pitched in during the 2009 regular season. The mean of this distribution is 8.2 with a standard deviation of 2.8. In what percentage of games were his strikeouts within one standard deviation of his mean? Within two standard deviations?



Strikeouts

**112–114**

What is the 68-95-99.7 rule? When does it apply? Pg. 112 Do you need to know about Chebyshev’s inequality?

Using the earlier example, about what percentage of 12-year-old boys will be over 158 cm tall?

About what percentage of 12-year-old boys will be between 131 and 140 cm tall?

Suppose that a distribution of test scores is approximately Normal and the middle 95% of scores are between 72 and 84. What are the mean and standard of this distribution?

Can you calculate the percent of scores that are above 80? Explain.

**115** *You will get a normal table…I have them laminated for the class to keep in here.*

What is the standard Normal distribution?

Find the proportion of observations from the standard Normal distribution that are:

(a) less than 0.54 (b) greater than –1.12 (c) greater than 3.89

(d) between 0.49 and 1.82. (e) within 1.5 standard deviations of the mean

A distribution of test scores is approximately Normal and Joe scores in the 85th percentile. How many standard deviations above the mean did he score?

In a Normal distribution, *Q*1 is how many SD below the mean?

**Example**: Serving Speed

In the 2008 Wimbledon tennis tournament, Rafael Nadal averaged 115 miles per hour (mph) on his first serves. Assume that the distribution of his first serve speeds is Normal with a mean of 115 mph and a standard deviation of 6 mph.

(a) About what proportion of his first serves would you expect to exceed 120 mph?

***Remember****:*

1. *State distribution*
2. *Identify values of interest*
3. *Show work and answer question*

(b) What percent of Rafael Nadal’s first serves are between 100 and 110 mph?

(c) The fastest 30% of Nadal’s first serves go at least what speed?

(d) What is the *IQR* for the distribution of Nadal’s first serve speeds?

(e) A different player has a standard deviation of 8 mph on his first serves and 20% of his serves go less than 100 mph. If the distribution of his serve speeds is approximately Normal, what is his average first serve speed?

**2.2: Using the Calculator for Normal Calculations**

How do you do Normal calculations on the calculator? What do you need to show on the AP exam?

 ***Finding areas: normalcdf(lower, upper, mean, SD)***

 ***Finding boundaries: invNorm(area to left, mean, SD)***

 ***Mean and SD default to 0, 1 if not entered.***

***Must show three steps: state distribution and identify values of interest, show work, answer.***

Suppose that Zach Greinke of the Kansas City Royals throws his fastball with a mean velocity of 94 miles per hour (mph) and a standard deviation of 2 mph and that the distribution of his fastball speeds is can be modeled by a Normal distribution.

(a) About what proportion of his fastballs will travel over 100 mph?

(b) About what proportion of his fastballs will travel less than 90 mph?

(c) About what proportion of his fastballs will travel between 93 and 95 mph?

(d) What is the 30th percentile of Greinke’s distribution of fastball velocities?

(e) What fastball velocities would be considered low outliers for Zach Greinke?

(f) Suppose that a different pitcher’s fastballs have a mean velocity of 92 mph and 40% of his fastballs go less than 90 mph. What is his standard deviation of his fastball velocities, assuming his distribution of velocities can be modeled by a Normal distribution?

According to [CDC](http://www.cdc.gov/growthcharts/), the heights of 3 year old females are approximately Normally distributed with a mean of 94.5 cm and a standard deviation of 4 cm.

(a) What proportion of 3 year old females are taller than 100 cm?

(b) What proportion of 3 year old females are between 90 and 95 cm?

(c) 80% of 3 year old females are at least how tall?

(d) Suppose that the mean heights for 4 year old females is 102 cm and the third quartile is 105.5 cm. What is the standard deviation, assuming the distribution of heights is approximately Normal?

**124–125**

The measurements listed below describe the useable capacity (in cubic feet) of a sample of 36 side-by-side refrigerators. (Source: *Consumer Reports*, May 2010) Are the data close to Normal?

12.9 13.7 14.1 14.2 14.5 14.5 14.6 14.7 15.1 15.2 15.3 15.3

15.3 15.3 15.5 15.6 15.6 15.8 16.0 16.0 16.2 16.2 16.3 16.4

16.5 16.6 16.6 16.6 16.8 17.0 17.0 17.2 17.4 17.4 17.9 18.4

**126–128**

When looking at a Normal probability plot, how can we determine if a distribution is approximately Normal?

Sketch a Normal probability plot for a distribution that is strongly skewed to the left.