Normal Distributions

Chapter 2

AP Statistics

2nd six weeks (2013-2014) AP Stats

MONDAY	TUESDAY	WEDNESDAY	THURSDAY	FRIDAY
Oct 7	8	9	10	11
Standard Normal Calculations HW: Worksheet	Assessing Normality **Quiz** HW: 2.26,2.27 WS MC HW 1-10	Chapter 2 Review HW: MC 11-14 Ch 2 Review #2	Chapter 2 Review HW: AP Problem Set	**Chapter 2 Test**
14	15	16	17	18
Student Holiday	Scatterplots HW: 3.2-3.4, 3.16, 3.19,3.20, 3.21	Correlation HW: 3.28,3.30,3.31, 3.34,3.37	**Quiz 3.1** HW: AP Problem Set	Linear Regression HW: 3.38 (use data on page 127),3.40, 3.41, 3.52
21	22	23	24	25
LSRL and r-squared Interpret Slope HW: 3.45, 3.71 ws - Slope	Residuals HW: 3.46-3.48, 3.61	**Quiz 3.3** HW: Log WS	Modeling non-linear Data HW: WS	Modeling Non-Linear Data (Natural Logs) HW: WS #2
28	29	30	31	November 1
Quiz 4.1 HW: AP Problem Set	Power Models HW: WS	Chapter 3/4 Review HW: Chapter 3 Review	Chapter 3/4 Review HW: AP Problem Set	**Chapter 3/4 Test** Read Chapter 5 Sampling Designs
4	5	6	7	8
Relations in Categorical Data HW: 4.51-4.53, 4.58	Interpreting Correlation and Regression HW: 4.33,4.37,4.38, 4.42,4.44,4.76	Intro to Random Digit Table HW: 5.1, 5.2, 5.7, 5.8, 5.10	6 Weeks Cumulative Test	**Half Day** Go over HW. Assign Project

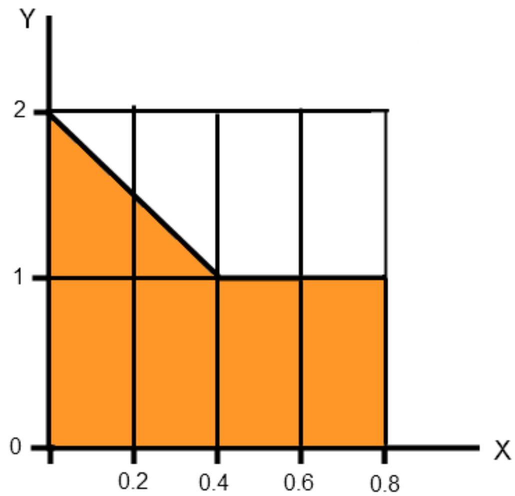
**All assignments are subject to last minute changes!!!

Name ______________________________ Period ______________

Irregular Density Curve Practice

Refer to the curve given and answer the questions which follow.

1.



a) Verify the above is a density curve (i.e. show the total area is equal to 1).

For each of the following use areas under this density curve to find the proportion of observations within the given interval.

b) $0.6 \leq X \leq 0.8$

c) $0 \leq X \leq 0.4$

d) $0 \leq X \leq 0.2$

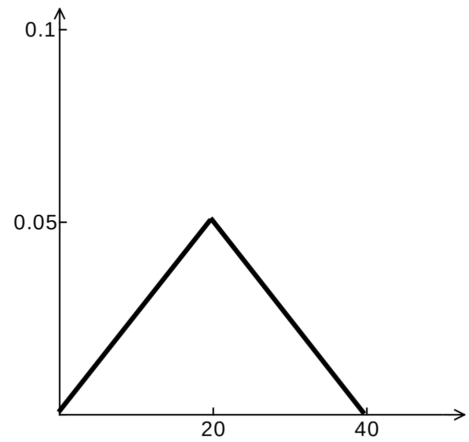
2. If a density curve has a **uniform distribution** over the interval (on x) from 3 to 12, what value, on the y-axis, would mark the height of this density curve? Use the template below to draw the curve.

a) What proportion of observations is greater than 7?

b) What proportion of observations is between 4 and 10?

c) What proportion of observations is equal to 5?

3. Given the following curve:



a) Verify that the total area under the curve is equal to 1 (that it is a valid density curve).

b) What percent of observations occur when $x < 20$? When $x \leq 20$?

c) What percent of observations occur when $x < 10$? When $x > 30$?

d) What percent of observations occur when $10 < x < 30$?

4. A certain density curve consists of a straight-line segment that begins at the origin (0,0) and has a slope of 1.

a) Sketch the density curve. What is the right endpoint of the line segment?

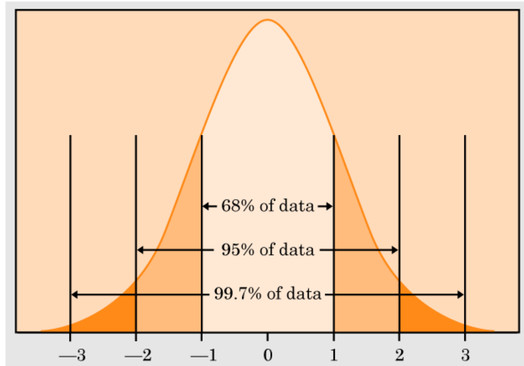
b) Determine the median, Q1 and Q3.

c) Where is the mean relative to the median?

d) What percent of the observations lie below 0.5? Above 1.5?

Name ___________________________________ Period ________________

Notes for Empirical Rule



Example: An IQ test is normally distributed with mean 100 and standard deviation 15. Just by knowing the mean and standard deviation, we can answer the following questions:

(1) Between what two values do approximately 68% of IQ scores lie?

(2) Between what two values do approximately 95% of IQ scores lie?

(3) Between what two values do approximately 99.7% of IQ scores lie?

(4) What percent of IQ scores are less than 115?

(5) What percent of IQ scores are greater than 130?

(6) What percent of IQ scores are less than 85?

In the picture above, the values on the horizontal axis represent the number of standard deviations above or below the mean. To get the actual values, we have to add or subtract.

In this example, the mean is 100 and the standard deviation is 15.

So, one standard deviation above the mean is 100 + 15 = 115.

Similarly, one standard deviation below the mean is 100 – 15 = 85.

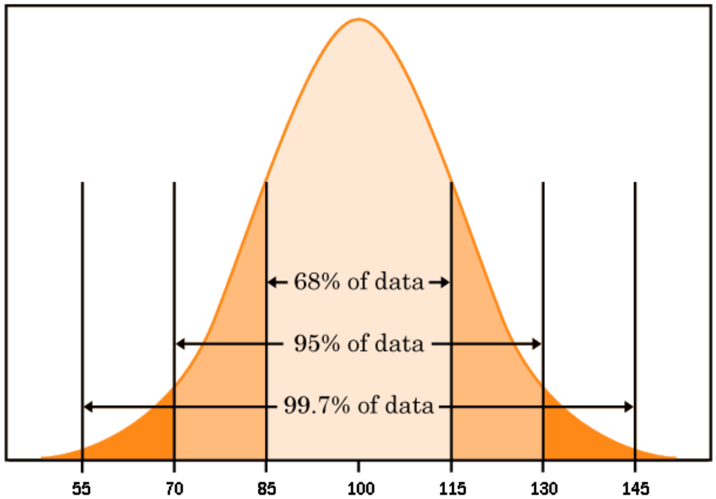
Two standard deviations above the mean is
$100 + 2(15) = 100 + 30 = 130$.

Two standard deviations below the mean is
$100 - 2(15) = 100 - 30 = 70$.

Three standard deviations above the mean is
$100 + 3(15) = 100 + 45 = 145$.

Three standard deviations below the mean is
$100 - 3(15) = 100 - 45 = 55$.

So, if we re-label the horizontal axis with the actual values instead of the number of standard deviations, we get this picture of the curve.



To answer the questions listed earlier:

(1) Approximately 68% of IQ scores fall between 85 and 115.

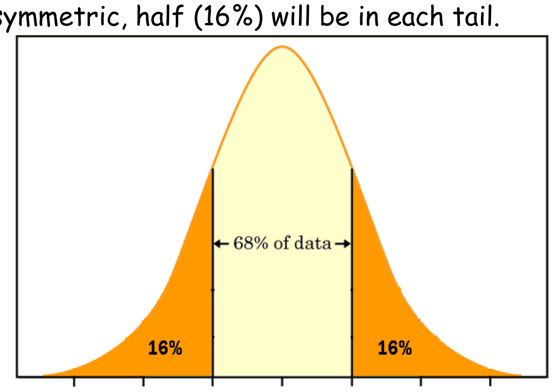
(2) Approximately 95% of IQ scores fall between 70 and 130.

(3) Approximately 99.7% of IQ scores fall between 55 and 145.

(4) To get the percentage of IQ scores less than 115:

> We know that about 68% of IQ scores are between 85 and 115. Now, if we can figure out the percentage less than 85, we can add those two percentages together to get the percentage less than 115.

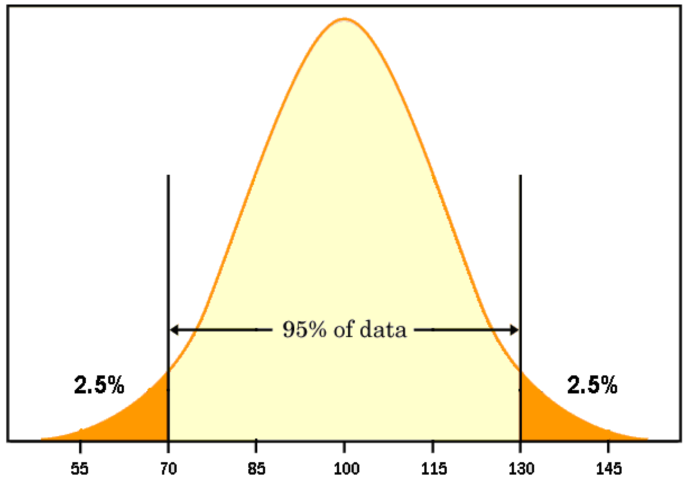
If 68% of IQ scores are between 85 and 115, then the other 32% must be either less than 85 or greater than 115. Because the curve is symmetric, half (16%) will be in each tail.



Thus, the percentage of IQ scores less than 115 is about

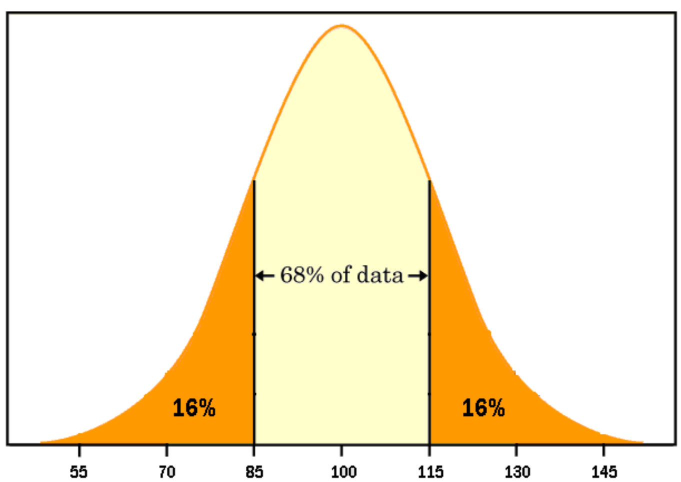
68% + 16% = 84%.

(5) To determine the percentage of scores above 130, we follow a similar procedure. We know that about 95% of IQ scores are between 70 and 130. This means that the other 5% are either less than 70 or greater than 130. Since the curve is symmetric, 2.5% are less than 70, and the other 2.5% are greater than 130.



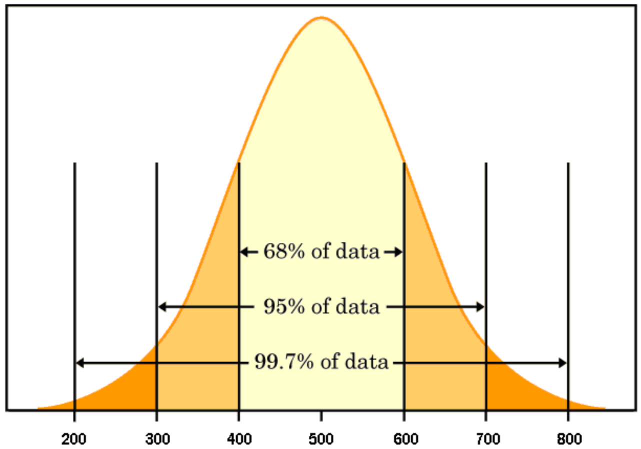
(6) To determine the percentage of IQ scores less than 85:

We know about 68% of the data is between 85 and 115. So the other 32% is either less than 85 or greater than 115. By symmetry, 16% of the data is in the left tail (less than 85), and 16% is in the right tail (greater than 115).



Another Example: On the math portion of the SAT, the scores are roughly normally distributed with mean 500 and standard deviation 100.

Since the mean is 500 and standard deviation is 100, we have:



(a) What percentage of SAT scores are above 400?

We know that about 68% of SAT scores are between 400 and 600. So the other 32% are either less than 400 or greater than 600. By symmetry, 16% are less than 400, and the other 16% are greater than 600. So 68% + 16% = 84% of SAT scores are above 400.

(b) What percentage of SAT scores are above 500?

Since 500 is the median, 50% of SAT scores are above it.

(c) What percentage of SAT scores are below 700?

We know that 95% of SAT scores are between 300 and 700 (within two standard deviations of the mean). The other 5% are either less than 300 or greater than 700. By symmetry, 2.5% are less than 300, and the other 2.5% are greater than 700. So the percentage of SAT scores below 700 is 95% + 2.5% = 97.5%.

(d) Between what two values do approximately 95% of SAT scores lie?

95% of SAT scores will be within two standard deviations of the mean, or between 500 – 2(100) = 500 – 200 = 300 and
500 + 2(100) = 500 + 200 = 700.

(e) What is the score of a person who scores higher than 84% of people who take the SAT?

Since 500 is the median, 50% of SAT scores are below it. Further, since 68% of the scores are between 400 and 600, by symmetry, we have that 34% are between 400 and 500, while 34% are between 500 and 600. Since 50% of SAT scores are below 500, and 34% are between 500 and 600, 84% of SAT scores are below 600.

Empirical Rule Worksheet

1. Given an approximately normal distribution what percentage of all values are within 1 standard deviation from the mean?

2. Given an approximately normal distribution what percentage of all values are within 2 standard deviations from the mean?

3. Given an approximately normal distribution what percentage of all values are within 3 standard deviations from the mean?

4. Given an approximately normal distribution with a mean of 175 and a standard deviation of 37,

   a) Draw a normal curve and label 1, 2, and 3 standard deviations on both sides on the mean.

   b) What percent of values are within the interval (138, 212)?
   c) What percent of values are within the interval (101, 249)?
   d) What percent of values are within the interval (64, 286)?
   e) What percent of values outside the interval (138, 212)?
   f) What percent of values are outside the interval (101, 249)?
   g) What percent of values are outside the interval (64, 286)?

5. Given an approximately normal distribution with a mean of 122 and a standard deviation of 22,

   a) Draw a normal curve and label 1, 2, and 3 standard deviations on both sides on the mean.

   b) What interval contains 68% of all values?
   c) What interval contains 95% of all values?
   d) What interval contains 99.7% of all values?
   e) What percent of values are above 122?
   f) What percent of values are below 78?

6. Given an approximately normal distribution with a mean of 159 and a standard deviation of 17,

   a) Draw a normal curve and label 1, 2, and 3 standard deviations on both sides on the mean.

   b) What percent of values are within the interval (142, 176)?
   c) What percent of values are within the interval (125, 193)?
   d) What interval contains 99.7% of all values?
   e) What percent of values are above 176?
   f) What percent of values are below 125?

7. Assume that the heights of college women are normally distributed, with mean 65 in. and standard deviation 2.5 in.

   a. What percentage of women are taller than 65 in.?
   b. What percentage of women are shorter than 65 in.?
   c. What percentage of women are between 62.5 in. and 67.5 in.?
   d. What percentage of women are between 60 in. and 70 in.?
   e. What percentage of women are between 60 and 67.5 in?
   f. What percentage of women are shorter than 70 in.?
   g. Approximately how tall would a woman need to be in order to be located in the 50th percentile? 97th percentile?

8. The incubation time for Rhode Island Red chicks is normally distributed with mean 21 days and standard deviation approximately 1 day. If 1000 eggs are being incubated, how many chicks do we expect will hatch?

a. In 19 to 23 days?
b. In 21 days or fewer?
c. In 19 days or fewer
d. In 18 to 24 days?
e. In 22 days or more?
f. In 18 to 21 days?
g. In 19 to 24 days?
h. In 18 days or more?

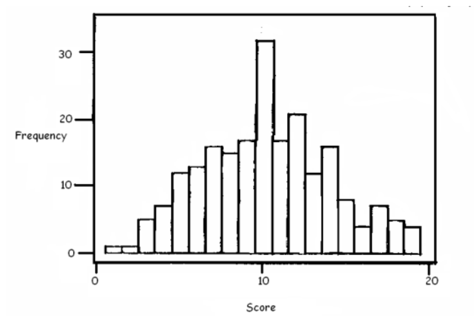
Normal Distributions – Packet #1

Activity #1:

The Department of Mathematics and Computer Science of Dickinson College gives an exam each fall to freshmen that intend to take calculus; scores on the exam are used to determine into which level of calculus a student should be placed. The exam consists of 20 multiple-choice questions. Scores for the 213 students who took the exam in 1992 are tallied in the following table.

Score	1	2	3	4	5	6	7	8	9	10
Count	1	1	5	7	12	13	16	15	17	32
Score	11	12	13	14	15	16	17	18	19	20
Count	17	21	12	16	8	4	7	5	4	0

The mean score on this exam is $\bar{x} = 10.221$. The standard deviation of the exam scores is s = 3.859. A histogram of this distribution follows:



a) Does this distribution appear to be roughly symmetric and mound-shaped?

b) Consider the question of how many scores fall within one standard deviation of the mean (denoted by $\bar{x} \pm s$). Determine the upper endpoint of this interval by adding the value of the standard deviation to that of the mean. Then determine the interval's lower endpoint by subtracting the value of the standard deviation from that of the mean.

c) What proportion of the 213 scores fall within one standard deviation of the mean?

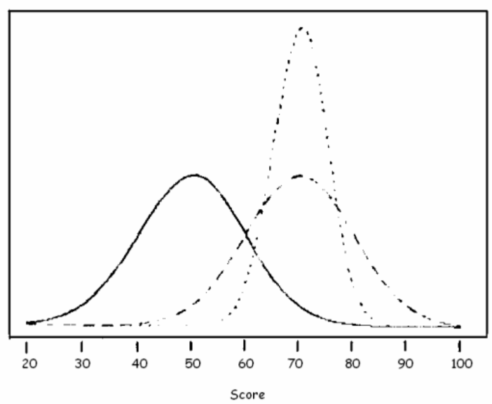
d) Determine how many of the 213 scores fall within two standard deviations of the mean, which turns out to be between 2.503 and 17.939. What proportion is this?

e) Determine how many of the 213 scores fall within three standard deviations of the mean, which turns out to be between –1.356 and 21.798. What proportion is this?

f) How do the values determined in c) to e) compare to the values learned for the empirical rule?

Activity #2:

The sketch below contains three normal curves; think of them as approximating the distribution of exam scores for three different classes. One (call it A) has a mean of 70 and a standard deviation of 5; another (call it B) has a mean of 70 and a standard deviation of 10; the third (call it C) has a mean of 50 and a standard deviation of 10. Identify which is which by labeling each curve with its appropriate letter.



Activity #3:

Suppose that a college admissions office needs to compare scores of students who take the SAT with those who take the ACT. Suppose that among the college's applicants who take the SAT, scores have a mean of 896 and a standard deviation of 174. Further suppose that among the college's applicants who take the ACT, scores have a mean of 20.6 and a standard deviation of 5.2.

a) If applicant Bobby scored 1080 on the SAT, how many points above the SAT mean did he score?

b) If applicant Kathy scored 28 on the ACT, how many points above the ACT mean did she score?

c) Is it sensible to conclude that since your answer to (a) is greater than your answer to b), Bobby outperformed Kathy on the admissions test? Explain.

d) Determine how many standard deviations above the mean Bobby scored by dividing your answer to (a) by the standard deviation of the SAT scores. Do the same for Kathy's score.

What you have found is the z-score, or standardized score. It should only be used when working with mound-shaped distributions. The z-score indicates how many standard deviations above (or below) the mean a particular value falls. The formula to find a z-score is $z = \frac{x-\mu}{\sigma}$.

e) Which applicant has the higher z-score for his or her admissions test score?

f) Explain in your own words which applicant performed better on his or her admissions test.

g) Calculate the z-score for applicant Peter, who scored 740 on the SAT, and for applicant Kelly, who scored 19 on the ACT.

h) Which of Peter and Kelly has the higher z-score?

i) Under what conditions does a z-score turn out to be negative?

Activity #4:
The probability of a randomly selected observation falling in a certain interval is equivalent to the proportion of the population's observations falling in that interval. Since the total area under the curve of a normal distribution is 1, this probability can be calculated by finding the area under the normal curve for that interval.

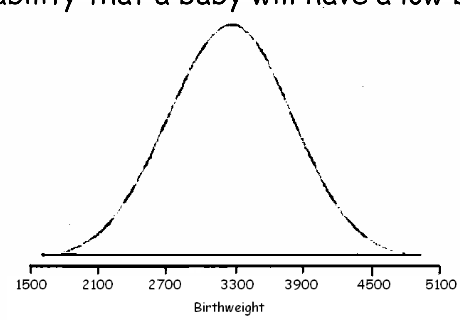
To find area under the normal curve, one can use either technology or tables. Table A in the front of your book (the Standard Normal Probability table) reports the area to the left of a given z-score under the normal curve. Enter the table at the row corresponding to the first two digits in the z-score. Then move to the column corresponding to the hundredths digit.

a) Use Table A to look up the area to the left of z = -0.45 under the normal curve.

We use Z to denote the standard normal distribution. The notation P(a<Z<b) denotes the probability lying between the values a and b, calculated as the area under the standard normal curve in that region. The notation $P(Z \geq z)$ denotes the area to the right of a particular value z, while $P(Z \leq z)$ refers to the area to the left of a particular z value. Your table gives you the area to the left of a particular z value. To find the area to the right, subtract the value in the table from 1.

Birth weights of babies in the United States can be modeled by a normal distribution with mean 3250 grams and standard deviation 550 grams. Those weighing less than 2500 grams are considered to be of low birth weight.

b) A sketch of this normal distribution appears below. Shade in the region whose area corresponds to the probability that a baby will have a low birth weight.



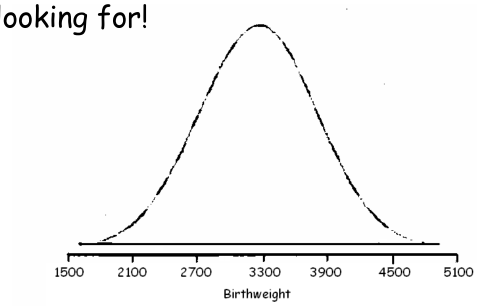
c) Based on this shaded region (remembering that the total area under the normal curve is 1), make an educated guess as to the proportion of babies born with a low birth weight.

d) Calculate the z-score for a birth weight of 2500 grams.

e) Look this z-score up in Table A to determine the proportion of babies born with a low birth weight. In other words, find P(Z<z), where z represents the z-score calculated in d).

f)

g) What proportion of babies would the normal distribution predict as weighing more than 10 pounds (4536 grams) at birth? You should always start with a sketch of the area that you are looking for!



h) Determine the probability that a randomly selected baby weighs between 3000 and 4000 grams at birth (again, start with a sketch!)

i) Data from the *National Vital Statistics Report* indicate that there were 3,880,894 births in the US in 1997. A total of 291,154 babies were of low birth weight, while 2,552,852 babies weighed between 3000 and 4000 grams. Calculate the observed proportions in each of these two groups, and comment on how well the normal calculations in (e) and (g) approximate these values.

j) How little would a baby have to weigh to be among the lightest 2.5% of all newborns? [Hint: Start with a sketch and read your table "in reverse". Find the appropriate z-score, then "unconvert" the z-score back to the birth weight scale.]

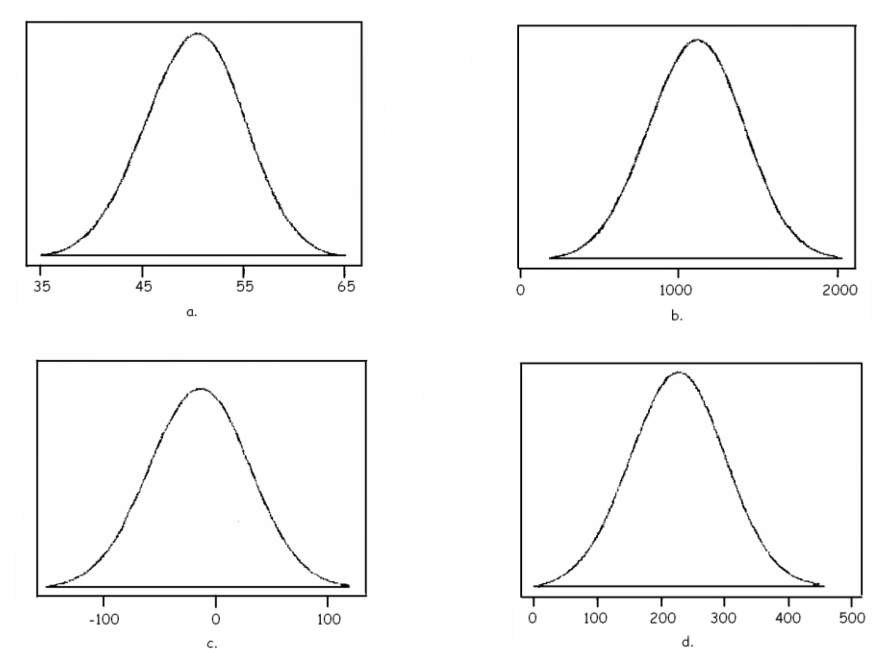
k) How much would a baby have to weigh to be among the heaviest 10% of all newborns?

Normal Distribution Practice Problems – Packet #2

1. Suppose the distribution of GPAs at Jefferson High School has a mean of 2.7 and a standard deviation of 0.37. The GPAs at Washington High School has a mean of 2.8 and a standard deviation of 0.33.
   a. Ted, a student at Washington High School, has a GPA of 3.25, and Frank, at Jefferson High School, has a GPA of 3.17. Calculate the z-score for Ted and Frank and comment on which of them has the higher GPA relative to his peers.
   b. What GPA would Ted need to have the same z-score as Frank?
   c. Torsten, another student at Jefferson High School, has a GPA of 3.07. Assuming that these GPAs follow a mound-shaped distribution, approximately what proportion of Jefferson High School students have a larger GPA? (Use the empirical rule to answer this question.)
   d. What GPA would you need to have to be in the top 10% of the class at each high school?

2. Suppose the average height of women collegiate volleyball players is 5'9", with a standard deviation of 2.1". Assume that heights among these players follow a mound-shaped distribution.
   a. According to the empirical rule, about 95% of women collegiate volleyball players have heights between what two values?
   b. What does the empirical rule say about the proportion of players who are between 62.7 inches and 75.3 inches?
   c. Reasoning from the empirical rule, what is the tallest we would expect a woman collegiate volleyball player to be?

3. For each of the following normal curves, identify (as accurately as you can from the graph) the mean μ and standard deviation σ of the distribution.



4. Data from the *National Vital Statistics Report* reveal that the distribution of the duration of human pregnancies (i.e., the number of days between conception and birth) is approximately normal with mean $\mu = 270$ and standard deviation $\sigma = 15$. Use this normal model to determine the probability that a given pregnancy comes to term in:
    a. less than 244 days (which is about 8 months).
    b. more than 275 days (which is about 9 months).
    c. over 300 days.
    d. between 260 and 280 days.
    e. Data from the *National Vital Statistics Report* reveal that of 3,880,894 births in the US in 1997, the number of pregnancies that resulted in a preterm delivery, defined as 36 or fewer weeks since conception, was 436,600. Compare this to the prediction that would be obtained from the model.

5. Suppose that you are deciding whether to take Professor Fisher's class or Professor Savage's next semester. You happen to know that each professor gives A's to those scoring above 90 on the final exam and F's to those scoring below 60. You also happen to know that the distribution of scores on Professor Fisher's final is approximately normal with mean 74 and standard deviation 7 and that the distribution of scores of Professor Savage's final is approximately normal with mean 78 and standard deviation 18.
    a. Produce a sketch of both teachers' grade distributions, on the same scale.
    b. Which professor gives the higher proportion of A's? Show the appropriate calculations to support your answer.
    c. Which professor gives the higher proportion of F's? Show the appropriate calculations to support your answer.
    d. Suppose that Professor DeGroot has a policy of giving A's to the top 10% of the scores on his final, regardless of the actual scores. If the distribution of scores on his final turns out to be normal with mean 69 and standard deviation 9, how high does your score have to be to earn an A?

6. Suppose that the IQ scores of students at a certain college follow a normal distribution with mean 115 and standard deviation 12.
    a. Draw a sketch of this distribution. Be sure to label the horizontal axis.
    b. Shade in the area corresponding to the proportion of students with an IQ below 100. Based on this shaded region, make an educated guess as to this proportion.
    c. Use the normal model to determine the proportion of students with an IQ score below 100.
    d. Find the proportion of these undergraduates having IQs greater than 130.
    e. Find the proportion of these undergraduates having IQs between 110 and 130.
    f. With his IQ of 75, what would the percentile of Forrest Gump's IQ be?
    g. Determine how high one's IQ must be in order to be in the top 1% of all IQs at this college.

7. Suppose that Professors Wells and Zeddes have final exam scores that are approximately normally distributed with mean 75. The standard deviation of Wells' scores is 10, and that of Zeddes' scores is 5.
    a. With which professor is a score of 90 more impressive? Support your answer with appropriate probability calculations and with a sketch.
    b. With which professor is a score of 60 more discouraging? Again support your answer with appropriate probability calculations and with a sketch.

8. Suppose that the wrapper of a certain candy bar lists its weight as 2.13 ounces. Naturally, the weights of individual bars vary somewhat. Suppose that the weights of these candy bars vary according to a normal distribution with mean μ = 2.2 ounces and standard deviation σ = 0.04 ounces.
    a. What proportion of candy bars weigh less than the advertised weight?
    b. What proportion of candy bars weigh more than 2.25 ounces?
    c. What proportion of candy bars weigh between 2.2 and 2.3 ounces?
    d. If the manufacturer wants to adjust the production process so that only 1 candy bar in 1000 weighs less than the advertised weight, what should the mean of the actual weights be (assuming that the standard deviation of the weights remains 0.04 ounces)?

9. Sample data from the National Center for Health Statistics reveal that weights of American men aged 20 – 29 have a mean of about 175 pounds and a standard deviation of about 35 pounds. For women the mean is about 140 pounds and the standard deviation is about 30 pounds.
    a. If these distributions are roughly normal, what percentage of men would you expect to weigh less than 150 pounds? Less than 200 pounds? Less than 250 pounds?
    b. Answer a. for women.
    c. Sample data from the National Center for Health Statistics reveal that the observed percentages in these ranges are 29.0%, 82.1%, and 96.2% for men, compared to 70.4%, 92.5%, and 99.0% for women. How well does the normal model predict these percentages?

10. A person with too much time on his hands collected 1000 pennies that came into his possession in 1999 and calculated the age (as of 1999) of each. The distribution has mean 12.264 years and standard deviation 9.613 years. Knowing these summary statistics but without seeing the distribution, can you comment on whether the normal distribution is likely to provide a reasonable model for these penny ages? Explain.

11. Use the table of standard normal probabilities to determine the proportion of the normal curve that falls within:
    a. one standard deviation of its mean (in other words, between z-scores of –1 and 1).
    b. two standard deviations from the mean
    c. three standard deviations from the mean.
    d. Compare these values to the values obtained from the empirical rule.

AP Statistics **Standard Normal Calculations** **Name:** ____________________
Ch. 2 **Worksheet #2** **Period:** ______________

1. The EPA fuel economy estimates for automobile models tested recently predicted a Normal model with a mean of 24.8 mpg and a standard deviation of 6.2 mpg.

    a) In what interval would you expect the central 68% of autos to be found?

    b) About what percent of autos should get less than 31 mpg?

    c) What percent of cars should get between 31 and 37 mpg?

    d) What percent of cars should get more than 20 mpg?

    e) Describe the gas mileage of the worst 20% of all cars?

2. Some IQ tests are standardized to a Normal model with a mean of 100 and a standard deviation of 16.

    a) What score would begin the interval for the top 16% of all scores? You may use the Empirical Rule to answer this.

    b) The top 10% of all scores represent the label of "genius". What is the range of scores for anyone who qualifies as a genius?

    c) What proportion of test takers score a 130 or higher?

3. Assume the cholesterol levels of Adult American women can be described by a Normal model with a mean of 188 mg/dL and a standard deviation of 24.

   a) What percent of adult women do you expect to have cholesterol levels over 200 mg/dL?

   b) What percent of adult women do you expect to have cholesterol levels between 150 and 170 mg/dL?

   c) Estimate the interquartile range of the cholesterol levels.

   d) Above what value are the highest 15% of women's cholesterol levels?

4. While only 5% of babies have learned to walk by the age of 10 months, 75% have earned to walk by 13 months of age. If the age at which babies develop the ability to walk can be described by a Normal model, find the mean and standard deviation of the Normal model.

5. Wildlife biologists believe that the weights of adult trout can be described by a Normal model. Fishermen report that 22% of all trout caught were thrown back because they weigh below the 2-pound minimum, and only 6% weighed over 5 pounds. What are the mean and standard deviation of the model?

Assessing Normality
Normal Probability Plots

Activity #1
Type the following data set into L_1 of your calculator:

154	109	137	115	152	140	154	178	200
103	126	126	137	165	165	129	200	148

Draw a normal probability plot to assess normality. Confirm symmetry by drawing a boxplot.

Activity #2
Simulate 50 administrations of an IQ test. This test is normally distributed with a mean of 100 and a standard deviation of 15, or N(100, 15). Use randnorm (100, 15, 50) -> L_2. Make and draw a normal probability plot to confirm that a normal distribution is the best model for IQ scores.

Activity #3
Type the following data set into L_3 of your calculator:

32	31	29	10	30	33	22	25	32	20
30	20	24	24	31	30	15	32	23	23

Is this data set best modeled by using a normal distribution? Create and draw a normal probability plot to decide. Explain your results.

Name ______________________________ Period ______________

Normal Distribution Multiple Choice HW Assignment

1. Which of the following are true statements?
   I. The area under a normal curve is always equal to 1, no matter what the mean and standard deviation are.
   II. The smaller the standard deviation of a normal curve, the higher and narrower the graph.
   III. Normal curves with different means are centered around different numbers.
      (A) I and II
      (B) I and III
      (C) II and III
      (D) I, II, and III
      (E) None of the above gives the complete set of true responses.

2. Which of the following are true statements?
   I. The area under the standard normal curve between 0 and 2 is twice the area between 0 and 1.
   II. The area under the standard normal curve between 0 and 2 is half the area between –2 and 2.
   III. For the standard normal curve, the interquartile range is approximately 3.
      (A) I and II
      (B) I and III
      (C) II only
      (D) I, II and III
      (E) None of the above gives the complete set of true responses.

3. Populations P1 and P2 are normally distributed and have identical means. However, the standard deviation of P1 is twice the standard deviation of P2. What can be said about the percentage of observations falling within two standard deviations of the mean for each population?
   (A) The percentage for P1 is twice the percentage for P2.
   (B) The percentage for P1 is greater, but not twice as great, as the percentage for P2.
   (C) The percentage for P2 is twice the percentage for P1.
   (D) The percentage for P2 is greater, but not twice as great, as the percentage for P1.
   (E) The percentages are identical.

4. Which of the following are true statements?
   I. In all normal distributions, the mean and median are equal.
   II. All bell-shaped curves are normal distributions for some choice of μ and σ.
   III. Virtually all the area under a normal curve is within three standard deviations of the mean, no matter what the particular mean and standard deviation are.
      (A) I and II
      (B) I and III
      (C) II and III
      (D) I, II and III
      (E) None of the above gives the complete set of true responses.

5. A trucking firm determines that its fleet of trucks averages a mean of 12.4 miles per gallon with a standard deviation of 1.2 miles per gallon on cross-country hauls. What is the probability that one of the trucks averages fewer than 10 miles per gallon?
   (A) 0.0082
   (B) 0.0228
   (C) 0.4772
   (D) 0.5228
   (E) 0.9772

6. A factory dumps an average of 2.43 tons of pollutants into a river every week. If the standard deviation is 0.88 tons, what is the probability that in a week more than 3 tons are dumped?
   (A) 0.2578
   (B) 0.2843
   (C) 0.6500
   (D) 0.7157
   (E) 0.7422

7. An electronic product takes an average of 3.4 hours to move through an assembly line. If the standard deviation is 0.5 hour, what is the probability that an item will take between 3 and 4 hours?
   (A) 0.2119
   (B) 0.2295
   (C) 0.3270
   (D) 0.3811
   (E) 0.6730

8. The mean score on a college placement exam is 500 with a standard deviation of 100. Ninety-five percent of the test takers score above what?
   (A) 260
   (B) 336
   (C) 405
   (D) 414
   (E) 664

9. The average noise level in a restaurant is 30 decibels with a standard deviation of 4 decibels. Ninety-nine percent of the time it is below what value?
   (A) 20.7
   (B) 32.0
   (C) 33.4
   (D) 37.8
   (E) 39.3

10. The mean income per household in a certain state is \$9500 with a standard deviation of \$1750. The middle 95% of incomes are between what two values?
   (A) \$5422 and \$13578
   (B) \$6070 and \$12930
   (C) \$6621 and \$12379
   (D) \$7260 and \$11740
   (E) \$8049 and \$10951

11. Jay Olshansky from the University of Chicago was quoted in *Chance News* as arguing that for the average life expectancy to reach 100, 18% of people would have to live to 120. What standard deviation is he assuming for this statement to make sense (assuming life expectancies are normally distributed)?
    (A) 21.7
    (B) 24.4
    (C) 25.2
    (D) 35.0
    (E) 111.1

12. Cucumbers grown on a certain farm have weights with a standard deviation of 2 ounces. What is the mean weight if 85% of the cucumbers weigh less than 16 ounces?
    (A) 13.92
    (B) 14.30
    (C) 14.40
    (D) 14.88
    (E) 15.70

13. If 75% of all families spend more than $75 weekly for food, while 15% spend more than $150, what is the mean weekly expenditure and what is the standard deviation?
    (A) $\mu = 83.33$, $\sigma = 12.44$
    (B) $\mu = 56.26$, $\sigma = 11.85$
    (C) $\mu = 118.52$, $\sigma = 56.26$
    (D) $\mu = 104.39$, $\sigma = 43.86$
    (E) $\mu = 139.45$, $\sigma = 83.33$

14. A coffee machine can be adjusted to deliver any fixed number of ounces of coffee. If the machine has a standard deviation in delivery equal to 0.4 ounce, what should be the mean setting so that an 8-ounce cup will overflow only 0.5% of the time?
    (A) 6.97 ounces
    (B) 7.22 ounces
    (C) 7.34 ounces
    (D) 7.80 ounces
    (E) 9.03 ounces

Warm-up

Helen recently took the SAT Math test. The distribution of scores for this test is normally distributed with $\mu = 500$ and $\sigma = 100$.

If Helen scores somewhere within the distribution's IQR, what range of values will Helen's SAT Math score fall within?

Name: ___________________________________ Date:__________ AP Statistics

Chapter 2 Review #2

1. Based on data concerning the performance of machines owned by the Zorro Company, it is known that the length of time a machine can go without repair is approximately normally distributed with μ = 350 hours and σ = 55 hours. What is the probability that the machine can go without repair for:

    a) less than 424 hours?
    b) between 198 and 302 hours?
    c) between 267 and 389 hours?
    d) more than 412 hours?
    e) What is the range of scores that we would expect 68% of the repair times to center around?

2. An educational testing service has designed a new test of mechanical aptitude. Scores on this test are normally distributed with μ = 400 and σ = 60.

    a) What score would you need to be in the top 15%?
    b) What score represents the 45th percentile?
    c) If 200 students at McCallum took the test, how many would you expect to score below 300?

3. A student scores 75 on a test with μ = 60 and σ = 11. What would be the equivalent score in a distribution with μ = 20 and σ = 20?

4. Thirty students were asked at random to pick a number from zero to twenty (inclusive). Here are the results:

7	8	13	5	7	10	13	7	10	2	1	11	12
17	4	11	7	6	18	14	17	11	9	1	12	10
12	2	15	12									

    Assess the normality.

5. The following is the list reflects the number of children of the presidents of the United States up through George W. Bush.

0	0	0	0	0	0	1	1	2	2	2	2	2
2	2	2	2	3	3	3	3	3	3	4	4	4
4	4	4	4	5	5	5	6	6	6	6	6	6
7	8	10	14									

    Assess the normality.

6. A researcher notes that two populations of lab mice – one consisting of mice with white fur, and one of the mice with grey fur – have the same mean weight, and both have approximately normal distributions. However, the population of white mice has a larger standard deviation than the population of grey mice. If the weights of both of these populations were plotted, how would the curves compare to each other?

7. Given the following statement: the percent of observations that are smaller than **z** in a standard normal distribution is 8%. What z-score make the statement true?

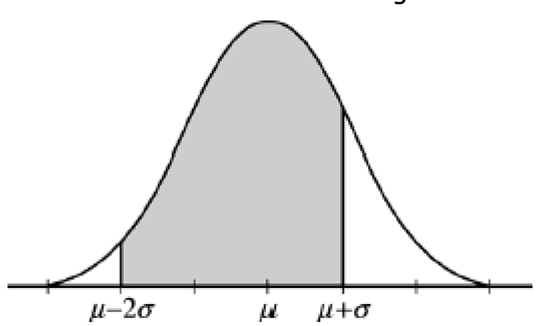
8. In a certain southwestern city the air pollution index averages 62.5 during the year with a standard deviation of 18.0. Assuming that the empirical rule is appropriate, the index falls within what interval about 95% of the time?

9. Population P1 and P2 are both normally distributed. The standard deviation of P1 is 5 with a mean of 23 while the standard deviation of P2 is 10 with a mean of 17. What can be said about the percentage of observations falling within one standard deviation of the mean for each population?

10. A distribution of test scores is not symmetric. What is the best estimate of the z-score of the third quartile?

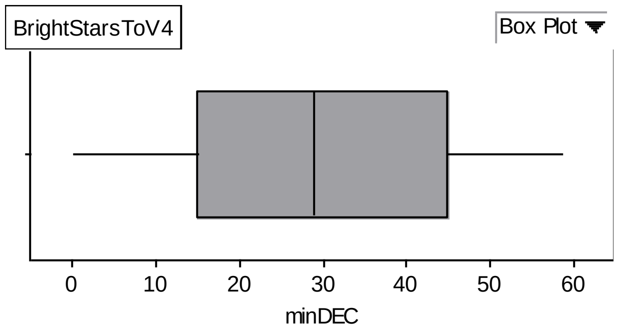
11. In a certain large population, 40% of households have a total annual income of at least $70,000. A simple random sample of 4 of these households is selected. What is the probability that 2 or fewer of the households in the survey have an annual income of at least $70,000.

12. A certain type of remote-control car has a fully charged battery at the time of purchase. The distribution of running times of cars of this type, before they require recharging of the battery for the first time after its period of initial use, is approximately normal with a mean of 80 minutes and a standard deviation of 2.5 minutes. The shaded area in the figure below represents what probability?

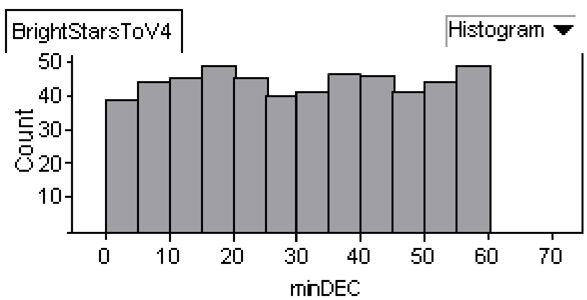


13. Declination is the star's angle north or south of Earth's celestial equator. The following boxplot and histogram display the minimum declination angle for the 513 brightest stars.

a) Assess the normality of the data for the boxplot below. Write no more than a sentence or two.



b) Assess the normality of the data for the histogram below. Write no more than a sentence or two.



c) Based on you assessments for a) and b), would you expect a straight-line relationship to be found in a normal probability plot from the same data? Explain fully your choice.

14. The length of human pregnancies from conception to birth varies according to a distribution that is approximately normal with mean 266 days and standard deviation 16 days. How long would a pregnancy last to be in the 85^{th} percentile?

15. Big Town Fisheries recently stocked a new lake in a city park with 2,000 fish of various sizes. The distribution of the lengths of these fish is approximately normal. Suppose the standard deviation is 0.3 inches while the mean length of the fish is 8 inches.

    a. Compute the probability that a random fish will have a length less then 7.75 inches.

    b. If you catch 18 fish, what is the probability that half of them will be less than 7.75 inches?

    c. If you catch 18 fish, with is the probability that between 7 and 12 will be less than 7.75 inches?

    d. You decide to keep a fish only if the fish is least 8.5 inches long. What is the probability that you keep your first fish on the 3rd catch?

16. Suppose we have examined several graphical displays of a data set. Describe how one could use the mean and median of a data set to **help** decide if the data set is considered normal.

17. The length of human pregnancies from conception to birth varies according to a distribution that is approximately normal with mean 266 days and standard deviation 16 days. Draw and label a distribution. Note: **some** of the following questions will allow the use of the Empirical Rule.

a. What percent of pregnancies last between 250 and 282 days?

b. We know roughly 99.7% of all pregnancies fall between how many days?

c. A pregnancy located in the 16^{th} percentile would last how long?

d. If 15 women are surveyed, what is the probability that 5 of them have a pregnancy that last between 250 and 266 days?

e. If a woman is claiming her pregnancy is lasting longer than 90% of pregnancies, how long might her pregnancy last?