Warm-Up 9/23

Here are scores from a test (out of 40):
18  17  17  28  13  20  22  24  11  18  9
32  16  14  24  18  20  25  22  27  26

What are the mean and standard deviation of the test scores?
\[ \overline{x} = 20.05 \quad s = 5.88 \]

Make a histogram of the original data on the calculator with a class size of 5 starting at 0. Where does the person who scored 20 fall relative to the center of the distribution? What about a person who scored 26?

Matching Boxplots, Histograms, and Summary Statistics

Work with your table partner to match each histogram with a boxplot and the summary statistics that describe the same data set.

<table>
<thead>
<tr>
<th>Histograms</th>
<th>Boxplots</th>
<th>Summary Stats</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>A</td>
<td>5</td>
</tr>
<tr>
<td>II</td>
<td>C</td>
<td>1</td>
</tr>
<tr>
<td>III</td>
<td>D</td>
<td>8</td>
</tr>
<tr>
<td>IV</td>
<td>G</td>
<td>3</td>
</tr>
<tr>
<td>V</td>
<td>E</td>
<td>2</td>
</tr>
<tr>
<td>VI</td>
<td>F</td>
<td>6</td>
</tr>
<tr>
<td>VII</td>
<td>B</td>
<td>4</td>
</tr>
<tr>
<td>VIII</td>
<td>H</td>
<td>7</td>
</tr>
</tbody>
</table>
Measuring Relative Standing: percentiles

At Mackenzie's 2.5 year check up, she was in the 79th percentile in weight, 87th percentile in height, and 92nd percentile for head size. What does this mean?

**Percentile:** The $p$th percentile of a distribution is the value that has $p\%$ of the data less than or equal to it.

$0^{th}$ percentile

$100^{th}$ percentile

Frequency: **number of individuals in each class**

Relative frequency: **percent of data the class contains**

Cumulative frequency: **number of individuals in all classes up to this point**

Relative cumulative frequency: **percent of the data in all classes up to this point**
### Ages of Presidents at Inauguration

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>40-44</td>
<td>2</td>
<td>$2/45 = 4.4%$</td>
<td>2</td>
<td>$2/45 = 4.4%$</td>
</tr>
<tr>
<td>45-49</td>
<td>7</td>
<td>$7/45 = 15.6%$</td>
<td>9</td>
<td>$9/45 = 20%$</td>
</tr>
</tbody>
</table>

**Ogive (oh-JIVE): rel. cumul. freq. graph**

Shown on Ogive Classwork
Plot points at the left endpoint of the next class interval. Start with a point at 0% at the left endpoint of the first class interval.

An ogive tells us what percentage of data falls below a given class.
Presidents' Ages at Inauguration

HW: 2.1(p.99) #2-10 (evens)
Read p.85-88
The graph displays the cumulative relative frequency of the lengths of phone calls made from the mathematics department office at Gabalot High last month.

1. About what percent of calls lasted less than 30 minutes? \(65\%\). Between 15 and 25 minutes? \(20\%\).

2. Estimate Q1, Median, Q3, and the IQR of the distribution.

If the distribution of absences was displayed in a histogram, what would be the best description of the histogram’s shape?

(a) Symmetric
(b) Uniform
(c) Skewed left
(d) Skewed right
(e) Cannot be determined
Recall from yesterday that one way to measure an individual data points position (relative standing) within a distribution is with percentiles.

Based on the data displayed, in what percentile would the student who scored an 86 on the test fall?

Measuring Relative Standing: z-scores

If $x$ is an observation from a distribution that has a known mean and st. dev., the **standardized value** of $x$ is

$$z = \frac{x - \text{mean}}{\text{st. dev.}}$$

A standardized value is what we call a **z-score**.

This is a way to describe how many standard deviations something is away from the mean.
1) Find the z-score for each of the following students, and interpret each value in context:  \( \bar{x} = 80, s = 6.07 \)
   
   a) The student who scored 93
   \[ Z = 2.14 \]
   
   b) The student who scored 72
   \[ Z = -1.32 \]

2) Jenny (who scored 86 on the stats test) earned an 82 on her chemistry test and was disappointed. If the mean of the chemistry test was 76 with a st. dev. of 4, on which test did Jenny perform better relative to the class?

\[ Z_{\text{Stats}} = 0.99 \quad Z_{\text{Chem}} = 1.5 \]

3) Brent is a member of the school’s basketball team. The mean height of the players on the team is 76 inches. Brent’s height of 72 inches translates to a z-score of \(-0.85\) in the team’s height distribution. What is the standard deviation of the team members’ heights?

\[ -0.85 = \frac{72 - 76}{s} \]

\[ s = 4.71 \]
In a population of scores, a raw score of 136 corresponds to a z-score of 1 and a raw score of 122 corresponds to a z-score of -2.5. What is the mean and standard deviation of this population?

\[ I = \frac{136 - \bar{x}}{S} \quad \Rightarrow \quad S = \frac{136 - \bar{x}}{I} \]

\[ -2.5 = \frac{122 - \bar{x}}{S} \quad \Rightarrow \quad S = \frac{122 - \bar{x}}{-2.5} \]

\[ 136 - \bar{x} = \frac{122 - \bar{x}}{-2.5} \]

\[ \bar{x} = 132, \quad S = 4 \]

**HW: 2.1 (p.99) #12,15,25,27**

**Read p.89-91  Due Thursday**
Calculate the relative standing (position) for a temperature of 7 in two ways:

a) as a percentile
\[ \frac{13}{30} = 43.3\% \text{ile} \]

b) as a z-score
\[ z = \frac{x - \text{mean}}{\text{st. dev.}} \]
\[ = \frac{7 - 8.43}{2.27} \]
\[ = -0.63 \]

Changing the Unit of Measurement

\[ z = \frac{x - \text{mean}}{\text{st. dev.}} \]

Linear Transformation: Changes the original variable \( x \) into the new variable \( x_{\text{new}} \) given by the equation: \( x_{\text{new}} = a + bx \)

- Adding \( a \) shifts all values of \( x \) up or down.
- Multiplying by \( b \) changes the size of the unit.
Linear Transformations Classwork

1. Find the measures of center and spread for the original data as well as linear transformations of the original data.

2. Discover the affects that linear transformations have on a set of data.

3. Solve a problem by using a linear transformation.

Teacher raises A school system employs teachers at 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 $1000 raise, what will this do to the mean salary? To the median salary? Explain your answers.

(b) What would a flat $1000 raise do to the extremes and quartiles of the salary distribution? To the standard deviation of teachers’ salaries? Explain your answers.
Teacher raises Refer to Exercise 20. If each teacher receives a 5% raise instead of a flat $1000 raise, the amount of the raise will vary from $1400 to $3000, depending on the present salary.

\[ \times 1.05 \]

(a) What will this do to the mean salary? To the median salary? Explain your answers.

(b) Will a 5% raise increase the IQR? Will it increase the standard deviation? Explain your answers.

Variance \( \times (1.05)^2 \)

Measure up 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).

\[ X_{\text{new}} = (X - 0.2) \times 2.54 \]

\[ \bar{X}_{\text{new}} = (\bar{X}_{\text{old}} - 0.2) \times 2.54 \]

\[ S_{\text{new}} = S_{\text{old}} \times 2.54 \]
HW: 2.1 (p.99) #12,15,25,27
Read p.89-91  Due Thursday

Based on today
No Homework
Read p.92-98

Warm-Up 9/26

Answer the questions on the back of the worksheet given to you about the temperature in Mr. Selvaag's cabin.
Exploratory data analysis of a single quantitative variable:

1. Make a graph: histogram or stemplot
2. Describe the shape, center, spread and any outliers
3. Calculate a numerical summary to help describe the center and spread
4. Describe a regular distribution of a large # of observations with a smooth curve. This can help us describe the location of observations within the distribution.

Vocabulary scores of 7th graders in Gary, Indiana

The curve is a mathematical model: it gives us a picture of the overall pattern but ignores irregularities and outliers.
The area under the curve within any interval is the **proportion** of all observations that fall within that interval.

**Density Curve:**
- Has an area of exactly 1 underneath the curve.
- Always on or above the horizontal axis.

No real data is **exactly** described by a density curve; it is an **approximation** that is easy to use!
Is the graph a valid density curve?

Find the proportion of the observations:

- $0 \leq X \leq 2$
- $X \leq 3$
At which of these points on the curve below do the mean and the median fall?

Two measures of center are marked on the curve. Which of the following is correct?

(a) The median is at the yellow line and the mean is at the red line.
(b) The median is at the red line and the mean is at the yellow line.
(c) The mode is at the red line and the median is at the yellow line.
(d) The mode is at the yellow line and the median is at the red line.
(e) The mode is at the red line and the mean is at the yellow line.

Mean vs. Median of a Density Curve

Median: "equal-areas point"
- The point with half the area under the curve on each side.

Mean: "balance point"
- The point at which the curve would balance if made of solid material.
Mean and Standard Deviation

We need to distinguish between mean and st. dev. of actual observations and the mean and st. dev. of the density curve.

Actual observations: mean = \( \bar{x} \) (x-bar)
st. dev. = \( s \)

Density curve: mean = \( \mu \) (mu)
st. dev. = \( \sigma \) (sigma)

HW: 2.2 (p.128) #33-35,37
Read p.103-107
Warm-Up 9/27

Mr. Starnes uses an unusual grading system in his class. After each test, he transforms the scores to have a mean of 0 and a standard deviation of 1. Mr. Starnes then assigns a grade to each student based on the transformed score. On his most recent test, the class's scores had a mean of 68 and a standard deviation of 15. What transformations should he apply to each test score? Explain.

\[ S_{\text{new}} = b \cdot S \]
\[ \frac{1}{b} = 6 \cdot 15 \]
\[ \frac{1}{15} = b \]

\[ X_{\text{new}} = a + b \cdot \frac{X}{15} \]
\[ 0 = a + \frac{68}{15} \]
\[ -\frac{68}{15} = a \]

\[ X_{\text{new}} = \frac{-68}{15} + \frac{X}{15} \]
\[ = -\frac{68 + x}{15} \]

\[ Z = \frac{x - 68}{15} \]

Against All Odds Video
Normal Curves
Applications of Normal dist. with real data:

ACT scores
Heights
Weights

Many distributions are NOT Normal:

Salaries
Pop of senior citizens

Normal Distributions

Definition: Density curves that are symmetric, single-peaked, and bell-shaped.

Picture

Notation

\[ X \sim N(\mu, \sigma) \]
\[ X \sim N(100, 5) \]
Approx 68% of data fall within 1σ of the mean

Approx 95% of data fall within 2σ of the mean

Approx 99.7% of data fall within 3σ of the mean
Warm-Up 9/30

The distribution of heights of women aged 18-24 is approximately Normal with $\mu = 64.5$ and $\sigma = 2.5$ (measurements in inches).

1. Approximately what proportion of women are between 62 and 67 inches tall?
   \[ 68\% \]

2. Approximately what proportion are between 59.5 and 64.5 inches tall?
   \[ 47.5\% \]

3. Approximately what proportion are less than 62 inches tall?
   \[ 16\% \]
4  Complete Response

Both parts essentially correct

3  Substantial Response

One part essentially correct and the other part partially correct

2  Developing Response

One part essentially correct and the other part incorrect
    OR
    Both parts partially correct

1  Minimal Response

One part partially correct
Normal curve applet

1. Is the Empirical Rule exact?

2. About how many $\sigma$ above or below $\mu$ does the 40th percentile fall? 10th percentile? 75th percentile?

All Normal distributions have the same shape and general characteristics, but have different $\mu$ and $\sigma$. So how can we make comparisons between observations that are in different distributions?

$\mu = 50$
$\sigma = 10$

$\mu = 600$
$\sigma = 30$

To do this we need to standardize our data.
Standardizing important values on a Normal distribution

\[ \sigma = 6 \]

### Standard Normal Distribution

This is the Normal distribution \( N(0,1) \) with a mean of 0 and standard deviation of 1.
### Table A

#### Standard Normal Probabilities

<table>
<thead>
<tr>
<th>$z$</th>
<th>.00</th>
<th>.01</th>
<th>.02</th>
<th>.03</th>
<th>.04</th>
<th>.05</th>
<th>.06</th>
<th>.07</th>
<th>.08</th>
<th>.09</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.0</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
</tr>
<tr>
<td>1.9</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
</tr>
<tr>
<td>1.8</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
</tr>
<tr>
<td>1.7</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
</tr>
<tr>
<td>1.6</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
</tr>
<tr>
<td>1.5</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
</tr>
<tr>
<td>1.4</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
</tr>
<tr>
<td>1.3</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
</tr>
<tr>
<td>1.2</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
</tr>
<tr>
<td>1.1</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
</tr>
<tr>
<td>1.0</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
<td>.0000</td>
</tr>
</tbody>
</table>

### Normal Calculations with z-scores

Classwork
No Homework
Read p.112-117

Fig. 1.0: The Extended Bell Curve.

Warm-Up 10/1

A state's ACT scores are approximated by the Normal distribution $N(23,3)$. Make a sketch of this Normal distribution.

Calculate the z-scores of students who received test scores of 21 and 30. Show the values of 21 and 30 on your Normal curve.

$$Z_{21} = \frac{21-23}{3} = -0.67$$

$$Z_{30} = \frac{30-23}{3} = 2.33$$

What score would place you at approximately the 97.5th percentile?

29
Draw a Normal curve showing the dist. of ACT test scores as well as the z-scores. Plot the students scoring 21 and 30 on the distribution as well.

\[
\frac{1}{\infty} = 0
\]

What proportion of students scored below 21? \[ \textcircled{.2514} \]

What proportion of students scored above 30? \[ \textcircled{-0.991 \pm 0.0099} \]

What proportion of students scored 30 or above? \[ \textcircled{.0099} \]

Use Table A to calculate the value in the distribution.

Two x-values that capture the middle 60%.
\[ \mu = 85, \sigma = 20 \]
More Normal Calculations with z-scores Classwork

HW: Z-Scores, Proportions, and Percentiles worksheet
Read p.112-117
Warm-Up 10/2

In 2004, the male SAT math test scores followed the N(537, 116) distribution \((\mu = 537, \sigma = 116)\). What percent of males scored 750 or better?

\[
Z = \frac{750 - 537}{116} = 1.84
\]

Table:

\[
1 - .9671 = .0329
\]

\[
\text{normalcdf}(750, \infty, 537, 116) = .0332
\]

In the same year, females followed the N(501,110) distribution. What percent of females scored 750 or better?

\[
.0118
\]

---

Given a Value, Find a Proportion

**Step 1:** State the problem in terms of \(x\). Draw a picture of the Normal curve and shade the area of interest.

**Step 2:** Standardize \(x\) (calculate z-score) and label the z-score underneath the \(x\)-value on the curve.

\[
\text{calc} \quad \text{normalcdf}(\text{lower}, \text{upper}, \mu, \sigma)
\]

**Step 3:** Find the area of interest under the curve.

**Step 4:** Write your conclusion in the context of the problem.
Given a Value, Find a Proportion

For teenage boys, the cholesterol level is Normally distributed with \( \mu = 170 \text{ mg/dl} \) and the standard deviation is \( \sigma = 30 \text{ mg/dl} \). Levels above 240 mg/dl require medical attention. What percent of teenage boys have more than 240 mg/dl of cholesterol? Follow the four steps.

\[
P(x > 240) = P\left( z > \frac{240 - 170}{30} \right) = P(z > 2.33)
\]

\[
1 - .9901 = .0099
\]

\[
\text{OR}
\]

\[
\text{normalcdf}(240, \infty, 170, 30) = .0098
\]

0.99% of teenage boys have more than 240 mg/dl of cholesterol.

---

Given a Proportion, Find a Value

**Step 1:** Draw a picture of the curve

**Step 2:** Use the st. N. table (going backwards) to find a \( z \)-score

\[
\text{calc} \quad \text{invNorm}(\text{area}, \mu, \sigma)
\]

**Step 3:** Unstandardize the \( z \)-score to get an \( x \)-value

**Step 4:** Conclusion in context of problem
Given a Proportion, Find a Value

Scores on the SAT Verbal test approximately follow the N(505,110) distribution. How high must a student score in order to place in the top 10%?

\[ P(x > ?) = 0.10 \]

\[ \frac{0.10}{90\%} \rightarrow z = 1.28 \]

\[ 1.28 = \frac{x - 505}{110} \rightarrow x = 645.8 \]

OR

\[ \text{invNorm}(0.9, 505, 110) = 645.9 \]

In order to be in the top 10% on the SAT Verbal test, you need to score at least 645.8.

---

HW: Carrying Out a Complete Normal Calculation Problem worksheet
Read p.118-120
Warm-Up 10/3

The distribution of ACT scores is $N(20.9, 4.8)$. What percent of scores are between 15 and 25?

The heights of American women are approximately normal with a mean of 64.5 inches and a standard deviation of 2.5 inches. How tall is an American woman in the 80th percentile?

Assessing Normality

In order to use the Empirical Rule, z-scores, and Table A (standard Normal distribution table) we need to know that we are dealing with a Normal distribution. So how do we know when a situation is Normal or not?

We need to have methods to check for Normality without assuming that a density curve is Normal.
Assessing Normality

Method 1: Construct a histogram with classes of size $s$

7th grade Gary, IN vocabulary scores

Check Normality with the Empirical Rule

Normal Probability Plots Classwork
Assessing Normality

Method 2: Construct a Normal probability plot
- Plot each data value against its corresponding z-score
  (original data may be plotted on either the horiz. or the vert. axis, but we will usually plot it on the horiz. axis)

Calculator:
- Enter data into a list
- Go to STAT PLOT pag
- Turn on the plot, use the type in the bottom right corner of the options, say where your data is, and highlight Data Axis: X
- Make sure to adjust your window to capture the plot

Check Normality by deciding if the plot is approximately linear.
Is it Normal or skewed?

No Homework
Read p.121-126
Warm-Up 10/4

The 1st and 9th deciles (instead of quartiles) of any distribution are the points that mark off the lowest 10% and the highest 10%. The 1st and 9th deciles of a density curve are therefore the points with area 0.1 and 0.9 to their left under the curve.

1. What are the 1st and 9th deciles of the standard Normal distribution?
   \[ z = -1.28 \quad z = 1.28 \]

2. The heights of women are approximated by \( N(64.5, 2.5) \). What are the 1st and 9th deciles of the heights of women based on this distribution?
   \[ 61.3 \text{ in} \quad 67.7 \text{ in} \]

---

Music Auditions - by U of M Prof Larry Copes

My brother is on the faculty of The Juilliard School. He phoned me one day with a problem.

Most of the school’s faculty members are professional musicians who teach only part-time. So when applicants audition for admission to the school, most of the teachers aren’t able to sit in on all of the auditions. Only two teachers hear every audition with the others sitting in on auditions whenever they are at the school.

At each audition, each teacher who is present rates the applicant’s performance on a scale of 1 to 10. Currently, the ratings for each audition are averaged into a single score for that applicant. My brother was concerned about how that score might be affected by the teachers attending different numbers of auditions and how some teachers tended to give higher overall ratings than others.

He asked me if I could come up with a better plan for determining an overall performance score for each applicant.

How might that be done, and do you think it would be a better system?
The following Normal probability plot shows the distribution of points scored for the 551 players in the 2011–2012 NBA season.

If the distribution of points was displayed in a histogram, what would be the best description of the histogram’s shape?

(a) Approximately Normal
(b) Symmetric but not approximately Normal
(c) Skewed left
(d) Skewed right
(e) Cannot be determined

Go to http://bit.ly/2naTBYA. Select one of the five tabs for a data set that sounds interesting to you and your partner.

1) View the data and decide if it is Normally distributed or not.

2) Then determine the United States' percentile in two ways:
   a) Based on the number of countries with a smaller value
   b) Based on a Normal distribution using the mean and standard deviation from the data set. (Even if the data isn't Normal)

3) Find the data value that would fall at the 40th percentile based on a Normal distribution used in part 2b. Which country is closest to this value?
No Homework
BUT start on the Review
Read Chapter Review
p.134-135

Warm-Up 10/7

Is Michigan Normal? We collected data on the tuition charged by colleges and universities in Michigan. Here are some numerical summaries for the data:

<table>
<thead>
<tr>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>10614</td>
<td>8049</td>
<td>1873</td>
<td>30823</td>
</tr>
</tbody>
</table>

Based on the relationship between the mean, standard deviation, minimum, and maximum, is it reasonable to believe that the distribution of Michigan tuitions is approximately Normal? Explain.
Ch. 2 Things to Remember
- Percentiles
- Ogives
- Z-scores
- Linear transformations
- Characteristics of density curves
- Empirical Rule (68-95-99.7)
- Using x-values and z-scores to find proportions/percentages/area
- Using percentiles to find z-scores and x-values
- Use the four steps to solve a problem involving a Norm. dist.
- Assessing Normality (remember approximately Normal)

No Homework
BUT work on the Review
Read Chapter Review
p.134-135

Deadline for missing Ch. 2 homework tomorrow
Deadline for Ch. 1 Retakes this Friday