WELCOME TO SESSION D-5:
Informal Comparative Inference: What is it?

Karen Togliatti  Curriculum and Professional Development Specialist
Process of Statistical Investigation

- Formulate questions that can be answered with data
- Design and use a plan to collect relevant data
- Analyze the data with appropriate methods
- Interpret results and draw valid conclusions from the data that relate to the questions posed.

“The power of statistics lies in making inferences about the world beyond available data.”

-Dr. Katie Makar, University of Queensland
Statistical Inference

“Statistical inference is a process of ‘evidence’ concerning whether or not a set of observations is consistent with a particular hypothesized mechanism that could have produced those observations.”

Harradine, Batanero and Rossman (2011)
Informal Inferential Reasoning

Data-based predictive reasoning that has the following components:

- Making statements or evaluating claims that go beyond the given data (generalizations)
- Explicitly using the data as evidence for generalizations
- Making statements that articulate uncertainty

Makar and Rubin (2009)
CCSSM Standards

Draw informal comparative inferences about two populations.

- **7.SP.3.** Informally assess the degree of visual overlap of two numerical data distributions with similar variabilities, measuring the difference between the centers by expressing it as a multiple of a measure of variability. *For example, the mean height of players on the basketball team is 10 cm greater than the mean height of players on the soccer team, about twice the variability (mean absolute deviation) on either team; on a dot plot, the separation between the two distributions of heights is noticeable.*

- **7.SP.4.** Use measures of center and measures of variability for numerical data from random samples to draw informal comparative inferences about two populations. *For example, decide whether the words in a chapter of a seventh-grade science book are generally longer than the words in a chapter of a fourth-grade science book.*
Informal Comparative Inference

- I notice…describes what is happening in the data in-hand (samples)
- I wonder…stimulates thoughts about what might be happening back in the population
- In comparing two data sets, we want to challenge students to look for “compelling” evidence to support a claim.
- Use of the claim-evidence-reasoning framework
Hand Dominance Activity

- It is estimated approximately 90% of humans are right-handed, with the rest made up of left-handed and ambidextrous individuals.
- Hand preference develops through early childhood.
- Hand dominance plays a role in fine motor tasks such as writing and manual tasks such as throwing accuracy.
- It is argued that reducing the dominance held by a limb increases the performance capacity of athletes.
The statistical question we are investigating is, “Do target hits by middle-school students using their dominant hand tend to be greater than target hits using their non-dominant hand when throwing a foam ball at a target 2 meters away?”
Collect Data

- Discuss our choice of sample.
- Which hand should be used to throw at the target first, the dominant or the non-dominant hand?
- Why might we want to consider having some students throw first with their non-dominant hand while others start with their dominant hand?
- How could we “randomize” which hand throws first?
- How else might we randomize the throws?
Analyze Data

- Create graphic visualizations such as a line plot and a box plot.
- We need to describe the features we see in the data.
- Start with an overall visual comparison.
- Then, use summary statistics and other interesting features of the data to make more detailed comparisons.
- Comparisons should focus within each group and between groups.
Analyze Data

[Images of data analysis and hand dominance]
Analyze Data

I notice…

Dot Plot

Box plot

I wonder…

Descriptive Statistics: Dominant, Non-Dominant

<table>
<thead>
<tr>
<th>Variable</th>
<th>Count</th>
<th>Mean</th>
<th>StDev</th>
<th>Minimum</th>
<th>Q1</th>
<th>Median</th>
<th>Q3</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dominant</td>
<td>25</td>
<td>6.040</td>
<td>2.091</td>
<td>2.000</td>
<td>4.000</td>
<td>6.000</td>
<td>7.500</td>
<td>10.000</td>
</tr>
<tr>
<td>Non-Dominant</td>
<td>25</td>
<td>4.080</td>
<td>1.869</td>
<td>1.000</td>
<td>2.500</td>
<td>4.000</td>
<td>5.500</td>
<td>8.000</td>
</tr>
</tbody>
</table>

Variable Range IQR
Dominant 8.000 3.500
Non-Dominant 7.000 3.000
If there is no overlap, or only a very small overlap make the claim “B tends to be bigger than A” back in the populations.

**The \( \frac{3}{4} - \frac{1}{2} \) Rule**

If the median for one of the samples lies outside the box for the other sample, make the claim “B tends to be bigger than A” back in the populations. (for sample sizes between 20 and 40 in each group)

Distance between medians as proportion of “overall visible spread”

Make the claim “B tends to be bigger than A” back in the populations if distance between medians is greater than about:

1/3 of overall visible spread for sample sizes of around 30
1/5 of overall visible spread for sample sizes of around 100
Interpret Results

☐ What claim can we make comparing throwing accuracy with the dominant hand versus the non-dominant hand for all middle-school students?

☐ What sample data observations can we use as evidence to support the above claim?

☐ Why do the sample statistic(s) or observations that you selected as evidence support the claim?

☐ How certain are you about your claim?
Collect Difference Data

- How might we account for the throwing “ability” of each of the subjects (students) in our study?
- What information would calculating the differences between number of times target hit with dominant hand vs. non-dominant hand provide?
- If we created a distribution of differences, what features might we see on a dot plot? On a boxplot?
Analyze Difference Data

- **What do positive, zero, and negative values indicate?**

- **Do the difference data support the claim?**

- **Can we conclude that the dominant hand is always more accurate than the non-dominant hand?**
Interpret Results

- Do the parallel box plots or difference data provide more “compelling” evidence for your claim?
- Do you think another class or group of students would get the same results? Why or why not?
- What other questions might you ask using the collected data or using additional data?
CCSS Math: Throwing Accuracy

Students build on their previous work with single data distributions to compare two data distributions and address questions about differences between populations. They begin informal work with random sampling to generate data sets and learn about the importance of representative samples for drawing inferences.

- **CCSS.Math.Content.6.SP.A**
  
  Develop understanding of statistical variability.

- **CCSS.Math.Content.6.SP.B**
  
  Summarize and describe distributions.

- **CCSS.Math.Content.7.SP.A.1**
  
  Understand that statistics can be used to gain information about a population by examining a sample of the population.

- **CCSS.Math.Content.7.SP.B.3**
  
  Informally assess the degree of visual overlap of two numerical data distributions with similar variabilities...

- **CCSS.Math.Content.7.SP.B.4**
  
  Use measures of center and measure of variability for numerical data from random samples to draw informal comparative inferences about two populations.
Additional Examples

Body Temperature
Word Memorization
Standing Vertical Jump
Analyzing Body Temperature Activity

Do females have a higher normal body temperature than males? Students will investigate this question by collecting data for their class on body temperatures using disposable student thermometers and will summarize and describe the distributions of gender data using numeric summaries and graphic displays.

http://digitalcommons.imsa.edu/proflearningday/2016/STEM/5/
Word Memorization

Is it easier for students to memorize words with meaning compared to words that are nonsense? Students will investigate this question by being randomly assigned to one of two groups: meaningful words or nonsense words. Students will have one minute to study a list of 20 words, and then will have one minute to recall as many words as possible.

Lesson starter from The Brookhill Institute of Mathematics
Standing Vertical Jump

Is there a difference between middle-school boys’ standing vertical jump heights relative to standing reach and girls’ standing vertical jump heights relative to standing reach?

THANK YOU!

Karen Togliatti
ktogliatti@imsa.edu

Download documents from:
http://digitalcommons.imsa.edu/proflearningday/2017/STEM/8/