# Statistics – Week 2

# Diverse Goals of the Grant

- Any grant like this will have varied, complex, and potentially (too many?) ambitious goals.

# Diverse Goals of the Grant

- Deepening content knowledge is an important goal for this grant
  - This might, at times, feel like the least connected to your daily work
  - This is also makes for a challenging goal to explore with a group with varied grade level interests

# Diverse Goals of the Grant

- One thing we know from research is that an important part of improving 'Mathematics Knowledge for Teaching' (MKT) is to personally engage in mathematical thinking tasks

- The goal with many of the probability and statistics activities was to do just that.

# Diverse Goals of the Grant

- Wrestling with the CCSSM and the 'story-line' of content is important
- Analyzing, and writing formative assessment probes is also important
  - Writing high-quality formative assessment probes is extremely difficult
  - It requires deep content knowledge of the subject
  - It requires recognition of prior knowledge and future trajectory
  - Is complicated by the fact that the learning of many topics is not well understood

# My Plans/Goals: as stated on day 1

- 1. A main focus of the summer PDs, specifically the morning statistics activities, is content knowledge.
- 2. Our grant is targeted at high school readiness, so I will focus on 6-12 content
- 3. Engage in 6-12 stats lessons, not just to 'learn' content – but to investigate content.
- 4. Attempt to give you experience with some statistics activities you may have experience with, and others you have not.
- 5. My intention is to expose you to 'statistical thinking' as much as possible…de-emphasizing computations, and encouraging statistical exploration

# Today's Agenda

- Reflect on Last Week's activities with the GAISE framework

- Engage in one (at most two) short statistics activities
  - Dice Sampling Activity
  - Michael and You Activity

# GAISE

- I brought virtually no attention to the 'GAISE framework' last week (other than a very quick mention of it on day 1)
- I did this on purpose

# GAISE

- I believed that if I detailed the ‘levels’ of thinking framework for statistics, we would:
  - Spend more time analyzing the tasks
  - And spend less time engaging with the mathematics of the tasks.
- I felt this was antithetical to the goal of the grant for the first week.

# GAISE

- Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report:
  - A Pre-K -12 Curriculum Framework

# GAISE

- Framework for Development of Statistical Thinking
- Levels 'A' 'B' and 'C'
- Applies to '4 components to Statistical Problem Solving':
  - ***I. Formulate Questions***
  - ***II. Collect Data***
  - ***III. Analyze Data***
  - ***IV. Interpret Results***

# GAISE

## – *I. Formulate Questions*

| Process Component | Level A | Level B | Level C |
|---|---|---|---|
| **I. Formulate Question** | **Beginning awareness of the *statistics question distinction***<br>Teachers pose questions of interest<br>Questions restricted to the classroom | **Increased awareness of the *statistics question distinction***<br>Students begin to pose their own questions of interest<br>Questions not restricted to the classroom | **Students can make the *statistics question distinction***<br>Students pose their own questions of interest<br>Questions seek generalization |

**The formulation of a statistics question requires an understanding of the difference between a question that anticipates a deterministic answer and a question that anticipates an answer based on data that vary.**

# Formulate Questions: Example(s)

- *Word Length Example*
  - Level A: How long are the words on this page?
  - Level B: Are the words in a chapter of a fifth-grade book longer than the words in a chapter of a third- grade book?
  - Level C: Do fifth-grade books use longer words than third-grade books?

# Formulate Questions: Example(s)

- *Height and Arm Span Example*
  - Level A: In our class, are the heights and arm spans of students approximately the same?
  - Level B: Is the relationship between arm span and height for the students in our class the same as the relationship between arm span and height for the students in another class?
  - Level C: Is height a useful predictor of arm span for the students in our school?

# GAISE

## – *II. Collect Data*

| Process Component | Level A | Level B | Level C |
|---|---|---|---|
| **II. Collect Data** | Do not yet *design for differences*<br>Census of classroom<br>Simple experiment | Beginning awareness of *design for differences*<br>Sample surveys; begin to use random selection<br>Comparative experiment; begin to use random allocation | Students make *design for differences*<br>Sampling designs with random selection<br>Experimental designs with randomization |

**Data collection designs must acknowledge variability in data, and frequently are intended to reduce variability. Random sampling is intended to reduce the differences between sample and population. The sample size influences the effect of sampling variability (error).**

# Collect Data: Example(s)

- *Word Length Example*
  - Level A: How long are the words on this page?
    - The length of every word on the page is determined and recorded.
  - Level B: Are the words in a chapter of a fifth-grade book longer than the words in a chapter of a third- grade book?
    - A simple random sample of words from each chapter is used.
  - Level C: Do fifth-grade books use longer words than third-grade books?
    - Different sampling designs are considered and compared, and some are used. For example, rather than selecting a simple random sample of words, a simple random sample of pages from the book is selected and all the words on the chosen pages are used for the sample.

# GAISE

## – *III. Analyze Data*

| Process Component | Level A | Level B | Level C |
|---|---|---|---|
| **III. Analyze Data** | *Use* particular properties of *distributions* in the context of a specific example<br>Display variability within a group<br>Compare individual to individual<br>Compare individual to group<br>Beginning awareness of group to group<br>Observe association between two variables | Learn to *use* particular properties of *distributions* as tools of analysis<br>Quantify variability within a group<br>Compare group to group in displays<br>Acknowledge sampling error<br>Some quantification of association; simple models for association | Understand and *use distributions* in analysis as a global concept<br>Measure variability within a group; measure variability between groups<br>Compare group to group using displays and measures of variability<br>Describe and quantify sampling error<br>Quantification of association; fitting of models for association |

**The main purpose of statistical analysis is to give an accounting of the variability in the data. One key idea is the confidence level, which tells us how often estimates produced by the method employed will produce correct results. This analysis is based on the distribution of estimates from repeated random sampling.**

# Analyze Data: Example(s)

- *Popular Music Example*
  - Level A: What type of music is most popular among students in our class?
    - A bar graph is used to display the number of students who choose each music category.
  - Level B: How do the favorite types of music compare among different classes?
    - For each class, a bar graph is used to display the percent of students who choose each music category. The same scales are used for both graphs so that they can easily be compared.
  - Level C: What type of music is most popular among students in our school?
    - A bar graph is used to display the percent of students who choose each music category. Because a random sample is used, an estimate of the margin of error is given.

# Analyze Data: Example(s)

- *Height and Arm Span Example*
  - Level A: In our class, are the heights and arm spans of students approximately the same?
    - The difference between height and arm span is deter- mined for each individual. An X-Y plot (scatterplot) is constructed with X = height, Y = arm span. The line Y = X is drawn on this graph.
  - Level B: Is the relationship between arm span and height for the students in our class the same as the relationship between arm span and height for the students in another class?
    - For each class, an X-Y plot is constructed with X = height, Y = arm span. An “eye ball” line is drawn on each graph to describe the relationship between height and arm span. The equation of this line is determined. An elementary measure of association is computed.
  - Level C: Is height a useful predictor of arm span for the students in our school?
    - The least squares regression line is determined and assessed for use as a prediction model.

# GAISE

## – *IV. Interpret Results*

| Process Component | Level A | Level B | Level C |
|---|---|---|---|
| **IV. Interpret Results** | Students do not look *beyond the data*<br><br>No generalization beyond the classroom<br><br>Note difference between two individuals with different conditions<br><br>Observe association in displays | Students acknowledge that *looking beyond the data* is feasible<br><br>Acknowledge that a sample may or may not be representative of the larger population<br><br>Note the difference between two groups with different conditions<br><br>Aware of distinction between observational study and experiment<br><br>Note differences in strength of association<br><br>Basic interepretation of models for association<br><br>Aware of the distinction between association and cause and effect | Students are able to *look beyond the data* in some contexts<br><br>Generalize from sample to population<br><br>Aware of the effect of randomization on the results of experiments<br><br>Understand the difference between observational studies and experiments<br><br>Interpret measures of strength of association<br><br>Interpret models of association<br><br>Distinguish between conclusions from association studies and experiments |

Statistical interpretations are made in the presence of variability and must allow for it. Looking beyond the data to make generalizations must allow for variability in the data.

# Interpret Results: Example(s)

- *Word Length Example*
  - Level A: How long are the words on this page?
    - The dotplot of all word lengths is examined and summarized. In particular, students will note the longest and shortest word lengths, the most common and least common lengths, and the length in the middle.
  - Level B: Are the words in a chapter of a fifth-grade book longer than the words in a chapter of a third- grade book?
    - Students interpret a comparison of the distribution of a sample of word lengths from the fifth-grade book with the distribution of word lengths from the third- grade book using a boxplot to represent each of these. The students also acknowledge that samples are being used that may or may not be representative of the complete chapters.
  - The boxplot for a sample of word lengths from the fifth-grade book is placed beside the boxplot of the sample from the third-grade book.
    - Level C: Do fifth-grade books use longer words than third-grade books?
    - The interpretation at Level C includes the interpretation at Level B, but also must consider generalizing from the books included in the study to a larger population of books.

# GAISE Framework

- The *Framework* uses three developmental Levels:
  - A, B, and C.
  - Although these three levels may (at times) parallel grade levels, they are based on development in statistical literacy, not age.
  - Thus, a middle-school student who has had no prior experience with statistics will need to begin with Level A concepts and activities before moving to Level B.
  - This holds true for a secondary student as well. If a student hasn't had Level A and B experiences prior to high school, then it is not very appropriate for that student to jump into Level C expectations.
  - The learning is more teacher-driven at Level A, but becomes student-driven at Levels B and C.

# Last Week's Activities

- We will focus on each of the activities that was set up more like a traditional 'lesson' rather than just a problem.

- **Help 911**
- **How Long is 30 Seconds?**
  - **Boxplot Generalizations**
- **How Random is the iPod Shuffle?**
- **Did I trap the Median?**
- **How Wet is the Earth?**
- **The Case of the Careless ZooKeeper**

# Your Task

- With the details and examples provided, do your best with your group members to:
  - For each activity as implemented: Determine (*roughly*) the level (A, B, or C) for each part of the statistical process.
  - Consider the question: Could the activity be adapted to be bumped up (or down) a level? If so, what might that look like?

# Help 911

- Formulate Question:
- Collect Data:
- Analyze Data:
- Interpret Results:

# How Long is 30 Seconds? + Boxplot Extensions

- Formulate Question:
- Collect Data:
- Analyze Data:
- Interpret Results:

# Random iPod

- Formulate Question:
- Collect Data:
- Analyze Data:
- Interpret Results:

# Did I trap the Median?

- Formulate Question:
- Collect Data:
- Analyze Data:
- Interpret Results:

# How Wet is the Earth?

- Formulate Question:
- Collect Data:
- Analyze Data:
- Interpret Results:

# The Careless ZooKeeper

- Formulate Question:
- Collect Data:
- Analyze Data:
- Interpret Results:

# Using Dice to Introduce Sampling Distributions

- For this activity, we are going to explore a simple question:  what proportion of rolls of a six-sided die will result in a 6?

- We probably already know the answer to this question.

- In fact – I'll tell you it should be 1/6 (or roughly P(6) = 0.167) if the die is a fair die.

# Using Dice to Introduce Sampling Distributions

- What we are really going to investigate is variability in this short activity.

- Step 1: go to https://www.random.org/dice/ and determine the proportion of rolls that result in a 6 with 10, 20 and 60 rolls. (log this proportion as a decimal – for example if 3 of 20 rolls are a six, use p = .15 rather than 15%)

- Step 2: log results on the class chart

- **Step 3: Construct histograms showing the distribution for each (10, 20, and 60 rolls)**

- **Step 4: Determine the mean and standard deviation of each (10, 20, and 60 rolls)**

$$\sigma_{\widehat{p}} = \sqrt{\frac{p(1-p)}{n}}$$

# Human Histogram

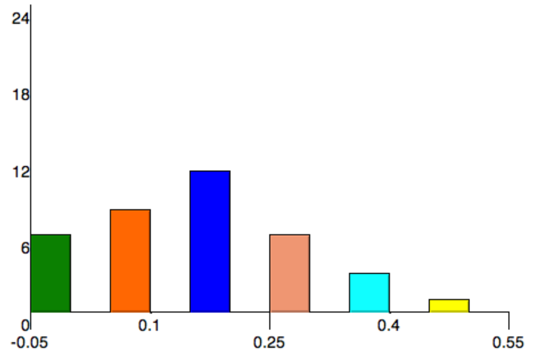
- I thought it would be fun (though possibly impractical) to make a human histogram for each of our three sample sizes (n=10, n=20, and n=60)

# Histogram groups

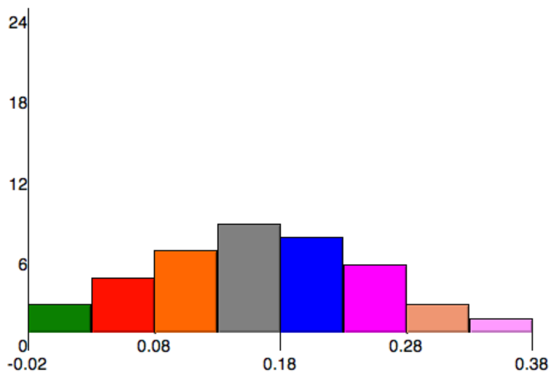
- 0 <= x <= 0.05
- 0 < x <= 0.10
- 0 < x <= 0.15
- 0 < x <= 0.20
- 0 < x <= 0.25
- 0 < x <= 0.30
- 0 < x <= 0.35
- 0 < x <= 0.40

- Each time we make a histogram (for n = 10, 20 and 60)
- Move to the interval containing your sample proportion

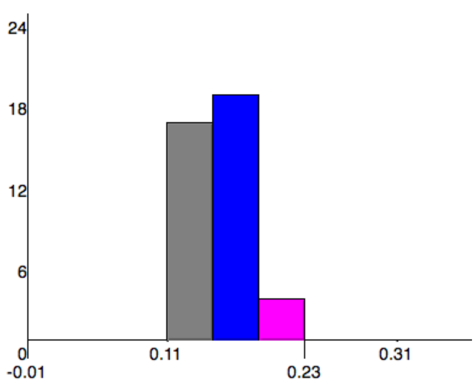
# Using Dice to Introduce Sampling Distributions




N = 35
Mean = .186
Standard deviation = .129




N = 35
Mean = .16
Standard deviation = .084




N = 35
Mean = .167
Standard deviation =  .044

# Discussion

- For which sample size is the standard deviation the largest and for which sample size is the standard deviation the smallest? Why do you suppose this happens?

$$\sigma_{\widehat{p}} = \sqrt{\frac{p(1-p)}{n}}$$

# Discussion

- Chris talked about the ‘Law of Large Numbers’ from probability theory
- Wikipedia definition:
  - According to the law, the average of the results obtained from a large number of trials should be close to the expected value, and will tend to become closer as more trials are performed.

# Discussion

- The Law of Large Numbers is similar in to another property relevant to this activity called the ‘Central Limit Theorem’
  - Roughly, the central limit theorem states that the distribution of the sum (or average) of a large number of independent, identically distributed variables will be approximately normal, regardless of the underlying distribution.