Teaching Statistics with Developmental Math

Mary Parker

Austin Community College

www.austincc.edu/mparker/talks

mparker@austincc.edu

Outline

- What are we doing now?
  - For whom?
  - How did we get to this?
- What's the formative question?
- What are our main challenges?
- What products can we share with you?
  - Website and applets

What are we doing now?

- Have students engage with
  - thinking about concepts
  - making connections,
  - different strategies of thinking

  as soon as possible in their classes with us.
- Not so much emphasis on traditional developmental math topics.

For whom?

- For students going to Math for Liberal Arts and Elementary Statistics
- Not for STEM students

How did we get to this?

- Statway from the Carnegie Foundation for the Advancement of Teaching Pathways
- New Math Pathways from the Charles A Dana Center
- Our own course, “Developing Mathematical Thinking” which we started in 2010.

(Texas has a statewide mandate that by Fall 2015 all community colleges provide pathways like these.)

What’s the formative question?

What do the “A students” already know,

but the “C students” don’t seem to get?

Re-interpreted as:

What do the “A students” already know (or with minimal instruction,) but the “C students” don’t seem to get (even with some instruction)?

What are our main challenges?

- Difficult to get students to take developmental math pathways courses if they aren’t **required** to take developmental math.
- Difficult to get advisers to put developmental math students into these courses since it it’s completely clear that the student won’t change their career plan to be one in a STEM field.
- Also, some of the students who aren’t required to take developmental math are weak on some of the crucial skills needed.

As we refine the materials we’re using in these developmental pathways courses,

also provide materials for students to use outside of class

to reinforce these same “habits of mind” and skills around which our new developmental classes are built.

And then make them available to students before and during their college-level classes.

Website

Find the URL from   http://www.austincc.edu/mparker/talks/

Or just email me: mparker@austincc.edu

Prepare for Elementary Statistics

What is statistics?

Statistics is the science of getting information from data.

Statistics is not mainly computing numbers. Of course, you must compute some numbers and make some graphs. Mostly, however, you must interpret the numbers and the graphs, usually in the context of a real-world problem. Students in these classes vary quite a lot in how much attention they have paid to discussions of getting information from data. This website is a first step in providing all the students in the class with some common background as they begin the course.

What do you need in your "mental toolkit" to be prepared for this class?

Answer: You need

- some skills from your previous math classes
- some ideas about how to discuss data

How to think about statistics questions, Part 1 • Producing data • Visualizing data • Summarizing data How to think about statistics questions, Part 2	What algebra and geometry skills are needed? • Numerical / algebraic toolkit

Estimate

Activity:

Consider the following study and decide which of the four possible conclusions stated is the best conclusion.

It's October 2014 and we're interested in what proportion of the voters in our state will vote for the Democratic candidate for Governor next month. A leading polling organization has done a scientifically-designed poll of 1000 registered voters in the state and found that 48.1% of the sample of registered voters in the state favor the Democratic candidate.

Which of the following best describes a reasonable prediction about the outcome of the election next month? (Notice that I am not saying that only one of these is correct. This is a question to think about, not one where I expect you to fully understand all the possible answers.)

1. We predict that exactly 48.1% of the voters will vote for the Democratic candidate for Governor next month.
2. We don't have any idea what percent of the voters will vote for the Democratic candidate for Governor next month.
3. If there aren't any major public relations upheavals in the next month, we believe that the percentage of the voters who vote for the Democratic candidate for Governor next month will be fairly close to 48.1%.
4. If there aren't any major public relations upheavals in the next month, we have 95% confidence that the percentage of voters who vote for the Democratic candidate for Governor next month will be 48.1% plus or minus 3%.

Discussion:

Navigation

- Prepare for Elementary Statistics
  - Producing data
  - Visualizing data
  - Summarizing data
  - Estimate
  - Test a claim
  - Math Toolkit
  - Contact information
- Recent site activity

Prepare for Elementary Statistics >

Summarizing data

Activities:

Part 1:

Suppose you're taking a class with five tests and the entire grade of the class is based on performance on those tests. Here are grades of two students:

Student A: 85, 85, 19, 85, 85
Student B: 66, 76, 79, 62, 70

What letter grade do you think each student will make? Is that letter grade a good summary of each of these students' performanc (There is not just one correct answer to this second question - this is one for you to think about.)

Part 2:

Do you suppose it would be useful to measure not only the averages (center) of each of their distributions of grades, but also the "spread" of them? If you want to do that, one way is to look at the range of the data. The range is the maximum score minus th minimum score. Find the range for each of the student's test scores. What do these ranges tell you?

Part 3:

According to the 2000 US Census, the average number of children per family was 1.86 (for families with any children.) Why do you suppose they chose to report it with a decimal number? Clearly, all the numbers that went into that average were whole numbers: 1 child, 2 children, etc.

Discussion:

Part 1. I expect that you computed the average grade for each of those two students and found that both averages were 71.8. In mos classes, that would be a letter grade of C. I expect that most students would say that a C is a good summary of the grades for Stude B. But some of you might have had some hesitation about saying that a C is a good summary for Student A. (I expect that each of you, if you were Student A, would want to try to convince the teacher that a B is a more appropriate grade.)

Part 2. For Student A, the range is 85 - 19 = 66 points. For Student B, the range is 79 - 62 = 17 points.
These two summary statistics illustrate that the overall set of grades for Student B is much more consistent than it is for Student A. (Which was glaringly obvious just by looking at the grades!)
Using the two summary statistics for each student of the average and the range allows us to give a numerical summary that illustrates very substantial difference in the overall distribution of the grades.

The point of these two activities is that we must learn about interpreting summary statistics and this helps motivate why we will learn about different types of averages and different types of measures of spread of a distribution.

Part 3. When you compute the average, if you round it off to the appropriate level for an individual value, you will lose information that could be useful. So it is more correct to say "the average is 1.86 children per family" than to say "the average is about 2 children per family." One important use for an average is to compute a total. For instance, if, in a particular state, it was know that there were about 5 million families, then you could multiply and estimate that there are about 9.3 million children in the state. That's quite differe from multiplying by the rounded number, which would give you an estimate of 10 million children in the state. The people thinking about how to get the resources needed to educate the children would prefer to have as precise an estimate as is reasonable from the data available, so they'd want to be told 1.86 children per family rather than 2 children per family.

Part 3. When you compute the average, if you round it off to the appropriate level for an individual value, you will lose information that could be useful. So it is more correct to say "the average is 1.86 children per family" than to say "the average is about 2 children per family." One important use for an average is to compute a total. For instance, if, in a particular state, it was know that there were about 5 million families, then you could multiply and estimate that there are about 9.3 million children in the state. That's quite different from multiplying by the rounded number, which would give you an estimate of 10 million children in the state. The people thinking about how to get the resources needed to educate the children would prefer to have as precise an estimate as is reasonable from the data available, so they'd want to be told 1.86 children per family rather than 2 children per family.

Major Ideas:

While it is important to learn to compute various summary statistics for data, it is even more important to learn to interpret them, and to realize what they aren't telling you as well as realizing what they are telling you.

Looking ahead:

In Elementary Statistics, you will learn to compute more than one measure of center and more than one measure of spread and you will see how the features of each of these lead to ideas about what situations are appropriate to use each of these.

Shoe size data: Answering the three questions.

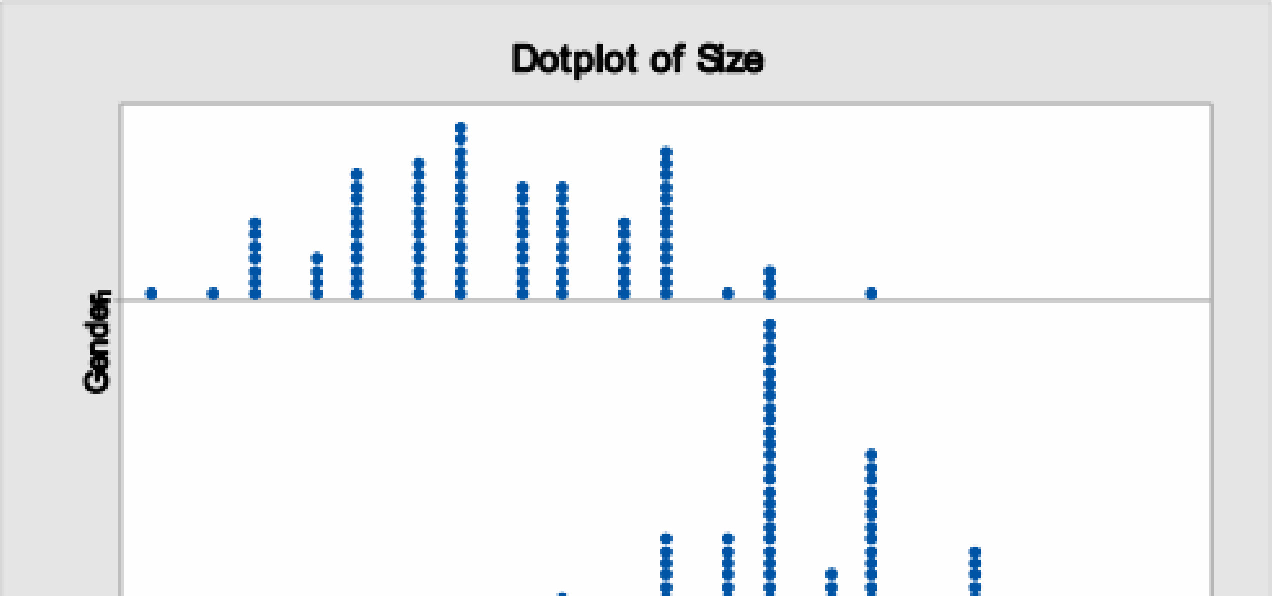
1. For women, what are typical shoe sizes? What are unusual shoe sizes?

2. For men, what are typical heights? What are unusual heights?

3. For the students in this sample, if we use height to predict shoe size, are the predictions pretty good?

1. For women, what are typical shoe sizes? What are unusual shoe sizes?

Since the dataset has both males and females, and I only want the graph for the females, I could delete all the data for males before making the graph.
Or I could use the software to make the comparative graphs and only look at the graph for females.
Since it is tedious to delete all of the data for males, so I choose to make the comparative graphs and only discuss the part that is for the females.



1. For women, what are typical shoe sizes? What are unusual shoe sizes?
2. For men, what are typical heights? What are unusual heights?
3. For all the students in the sample, if we use height to predict shoe size, are the predictions pretty good?

Discussion of these questions: View Download
Other comments about graphs using these data: View Download

References:
Read the story: http://www.amstat.org/publications/jse/v20n3/mclaren/documentation.doc
Look at the data: http://www.amstat.org/publications/jse/v20n3/mclaren/shoesize.xls
Article: http://www.amstat.org/publications/jse/v20n3/mclaren.pdf

Major ideas:

There may be multiple ways to make a correct and useful graph to answer a question.

Different types of data (one variable or multiple variables, categorical data or quantitative data) require different kinds of graphs.

Some of the questions we answer using graphs do not have precise answers. (Examples: What are "typical" data values? What are "unusual" data values?)

Link to our applets - subset for students BEFORE enrolling in statistics.

More background

The first two graphs I made in this document View Downloa… …ade on the second page of this document View Download are histograms. Hav… …ever seen such graphs being made? Would you like to see that? Look at this website with two …mple applets showing how a histogram and a dotplot are made: http://www.tlok.org/visualize/stat-prelim/

Histograms: sensitivity to binning settings

Paste data values below and **Display**

```
107.53 87.38 101.09 104.47 87.56 41.76
72.4 85.76 59.74 64.31 73.1 71.2 89.85
78.72 96.48 63.8 72.81 91.78 73.53
98.01 75.37 71.47 32.33 92.87 41.42
56.24 45.37 99.2 72.49 63.47 110.3
72.25 66.11 39.93 118.22 81.33 75.2
```

Or generate random values:

25 values **100 values** **500 values**

10 Bins to use

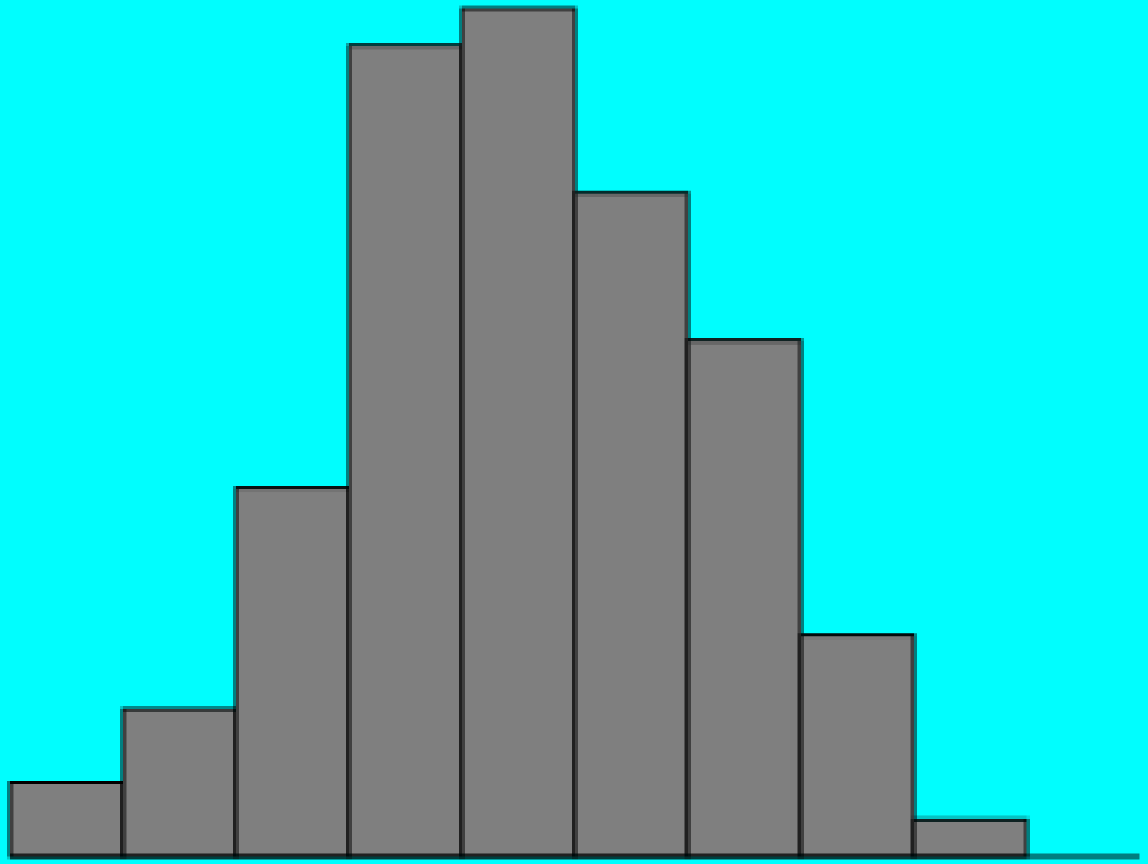
16.15 Lowest value

126.65 Highest value

Shift bins this fraction of their width:

○ -½ ○ -¼ ◉ none ○ +¼ ○ +½

100 values ranging from 24.91 to 118.22 tabulated in 10
between 16.15 and 137.15 (each bin is 12.1 wide); no lab



Histograms: sensitivity to binning settings

Paste data values below and **Display**

```
107.53 87.38 101.09 104.47 87.56 41.76
72.4 85.76 59.74 64.31 73.1 71.2 89.85
78.72 96.48 63.8 72.81 91.78 73.53
98.01 75.37 71.47 32.33 92.87 41.42
56.24 45.37 99.2 72.49 63.47 110.3
72.25 66.11 39.93 118.22 81.33 75.2
```

Or generate random values:

25 values **100 values** **00 values**

15 Bins to use

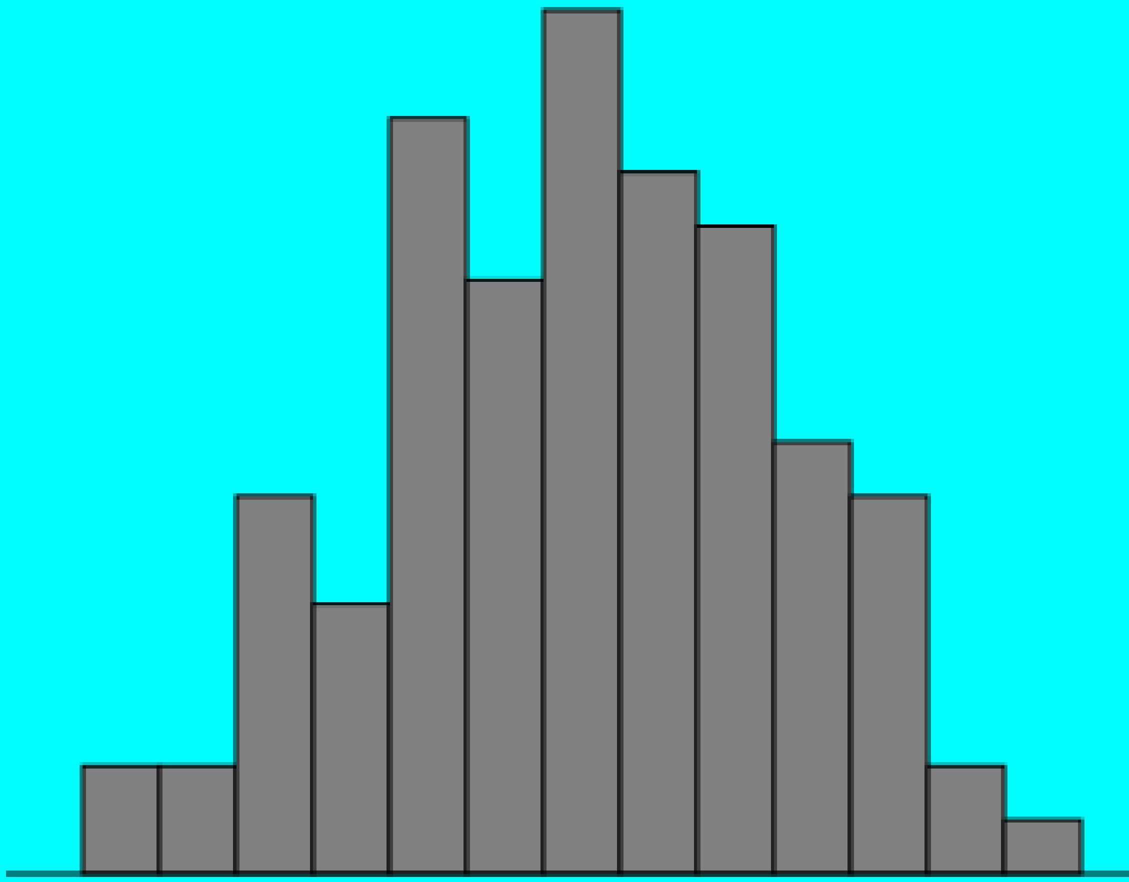
16.15 Lowest value

126.65 Highest value

Shift bins this fraction of their width:

-½ -¼ none +¼ +½

100 values ranging from 24.91 to 118.22 tabulated in 1
between 16.15 and 127.15 (each bin is 7.4 wide); no lo

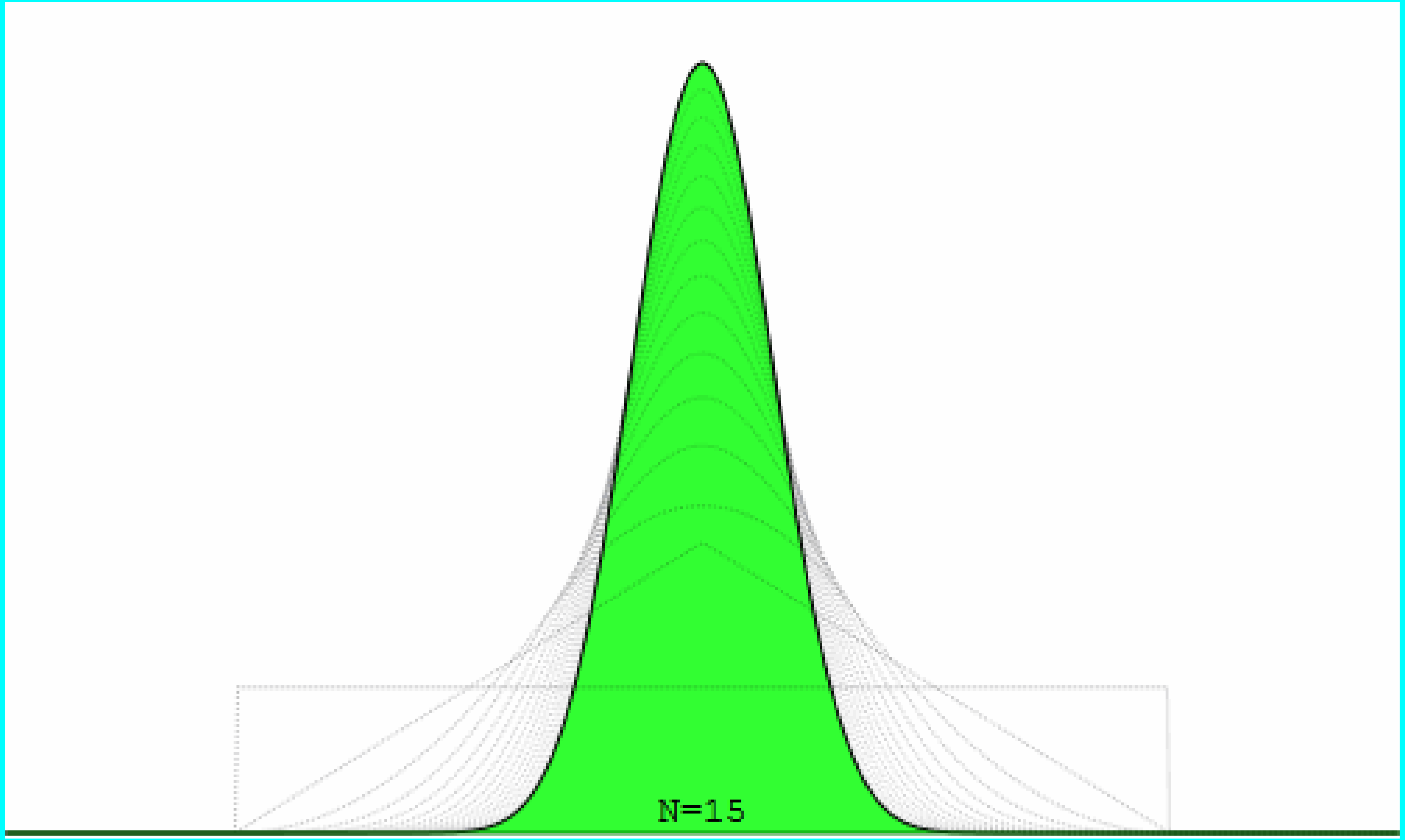


◉ Uniform ○ Normal ○ Exponential ○ Binomial ○ Notch ○ Overlap ○ Triangle ○ Triangles ○ Comb

Sample sizes: ○ 5 ○ 10 ◉ 15 ○ 20 ○ 25 ○ 30 ○ 35 ○ 40 ○ 50 ○ 75 ○ 100 or list: (e.g., 2,5,10 or 1:50)

Display Height: 300 Width: 500

Uniform distribution: probability distributions of the sample means for increasing sample sizes



austincc/mparker/talks/jmm2015

Teaching Statistics with Developmental Math

When we teach statistics with more class time and more activities, we find that we can integrate in most of the developmental math that students need and, at the same time, get them into a much better frame of mind about learning the concepts well and deeply. Student outcomes compare favorably with a comparison group of university Elementary Statistics students. Early versions of the activities we use are publicly available, along with detailed instructor notes. Teachers of all elementary statistics courses may find useful material in these. These are the foundation of the Pathways project of the Carnegie Foundation for the Advancement of Teaching and the New Math Pathways project of the Charles A Dana Center.

Slides for talk:

Website: Prepare for Elementary Statistics

Visualize" Applets: For prospective students | Jan. 2015 version of all the applets | | Feb 2015 version for ACC's Elem. Stat classes | May 2015 version |