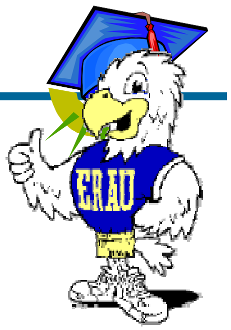


EMBRY-RIDDLE
Aeronautical University
WORLDWIDE

Some Lessons Learned in Developing an Introductory Online Undergraduate Statistics Course

Dr. Eugene L. Round

Improving Statistics Teaching

EMBRY-RIDDLE Aeronautical University WORLDWIDE

- GAISE Guidelines
- Randomization and bootstrapping techniques
- Student projects in an introductory statistics class
- Developing a nine-week introductory undergraduate statistics course that incorporates these ideas
- AND MOST IMPORTANT, your input on techniques used at other schools as well as comments on our process will be sought in an effort to continue to improve course delivery.

GAISE - Guidelines for Assessment and Instruction in Statistics Education

1. Emphasize statistical literacy and develop statistical thinking.
2. Use real data
3. Stress conceptual understanding rather than mere knowledge of procedures.
4. Foster active learning in the classroom
5. Use technology for developing concepts and analyzing data.
6. Use assessment to improve and evaluate student learning.

1. Emphasize statistical literacy and develop statistical thinking.

- Use technology and show students how to use technology to:
  - Manage and explore data
  - Perform inference
  - Check conditions that underlie inference procedures.
- Include open-ended problems and projects.
- Give students practice in choosing techniques.

2. Use real data

3. Stress conceptual understanding rather than mere knowledge of procedures.

- Primary goal – Discover Concepts, not Cover Methods
- Pare down content to focus on core concepts
- Perform routine computations using technology to allow greater emphasis on interpretation of results.

4. Foster active learning in the classroom

- Precede computer simulations with physical explorations.
- Don’t use activities that lead students step-by-step through a list of procedures.
- Include assessment as an important component of an activity.

5. Use technology for developing concepts and analyzing data.

- Implement computer intensive methods.
- Allow students to focus on interpretation of results and testing of conditions rather than on computations.
- Help students visualize concepts.
- Explore “What if” questions.
- HOWEVER, don’t replace a set of rote “by hand” procedures with a set of rote computer procedures.

6. Use assessment to improve and evaluate student learning.

- Students will value what you assess, so assessment should be aligned with learning goals.
- Use projects and open-ended investigative tasks.

Goals for our course

1. Integrate GAISE guidelines into a nine-week course – classroom and online.
2. Reduce the number of techniques covered to concentrate on understanding.
3. Use Homework and Testing software that would give students immediate feedback.
4. Use Technology to explore data, reduce computations, and give students time to explain results.
5. Include a project.
6. Use computer intensive methods to get at concepts

1. Reduce Coverage

- ❑ Emphasize concepts that students traditionally have trouble with
  - Distributions, confidence interval, hypothesis testing, p-value
- ❑ Cover only the Basics of Simple Linear Regression – Strength of Relationships
  - Scatter plots, Correlation, Simple Linear Regression Line
- ❑ Cover only the Basics of Probability
  - Randomness, Independence, Conditional Probability, Normal Probability Model

1. Reduce Coverage

- Included (p-value method):
  - One sample test for proportion
  - Two sample test for proportion
  - One sample t-test for mean
  - Two sample t-test for mean paired data
  - Two-sample t-test for mean, independent samples, w/o equal variation assumption
- Not Included:
  - Two-sample t-test for mean assuming equal variation
  - ANOVA
  - Chi-Square
  - Linear Regression

Technology

2. Use Homework and Testing software that gives immediate feedback - MyStatLab

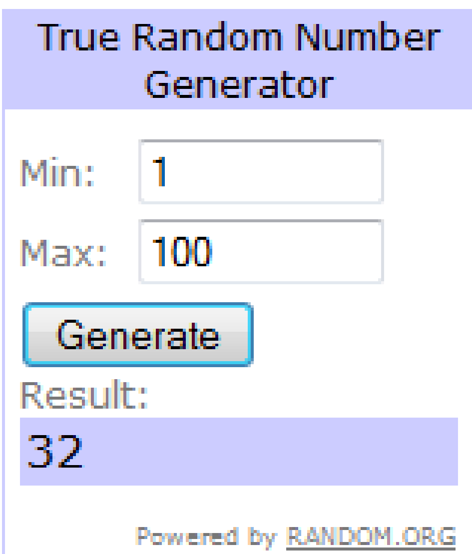
3. Use Technology - StatCrunch

4. Include a Project

- http://faculty.ung.edu/DJSpence/NSF/materials.html
- Incorporated a t-test project
  - Made large data sets available for students to use or they can find their own data.
  - Detailed scoring rubric given to students
  - Must turn in project prior to taking final exam
- Excellent article at: http://www.amstat.org/publications/jse/v21n3/bailey.pdf

5. Use Computer Intensive Methods

- StatCrunch incorporates the capability to do Bootstrapping and Randomization methods
- Bootstrapping Illustration



- Randomization Illustration

Bootstrapping- Reaction Times

- Dropper holds a meter stick vertically with one hand near the top of the stick with 0-centimeter mark is at the bottom.
- Catcher positions his or her thumb and index finger about 5 cm apart on opposite sides of the meter stick at the bottom.
- Dropper drops the meter stick without warning, and Catcher catches it.
- Location of the middle of the thumb of the catcher is recorded (reaction distance)

Bootstrapping

- Assume the first student is right handed. Then the reaction distance for the right hand was 8.0 centimeters and for the left hand was 8.5 centimeters.
- File shows the reaction distances of all 40 students.

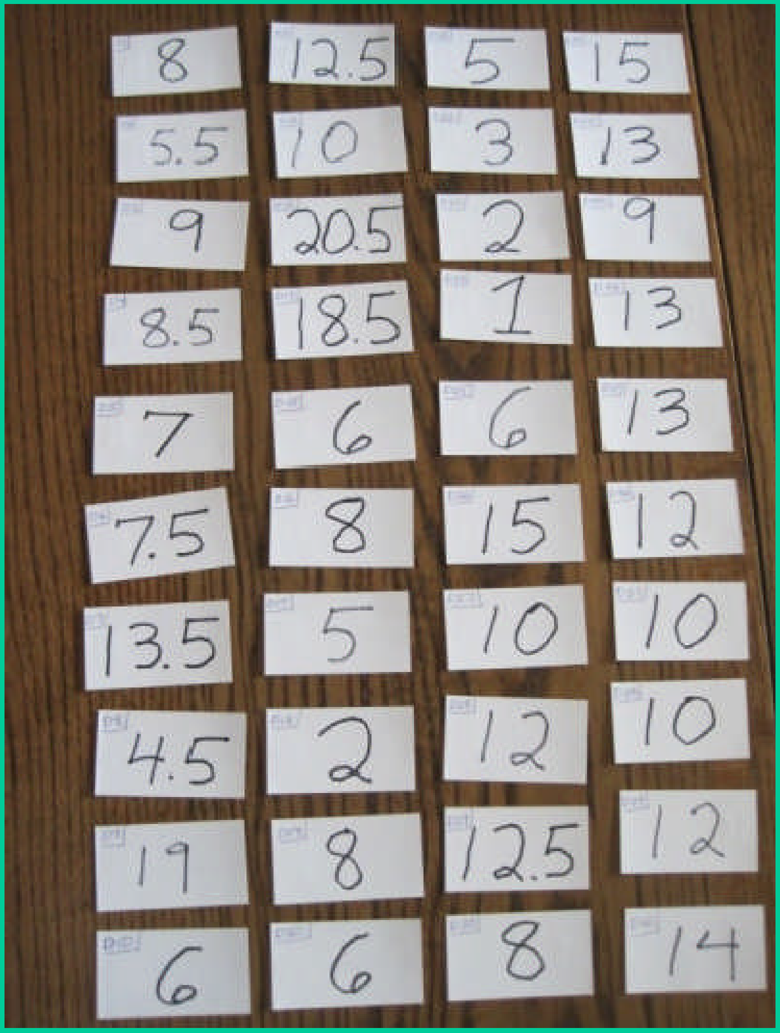
Row	Dom	Non
1	8	8.5
2	5.5	7.5
3	9	14.5
4	8.5	11
5	7	6
6	7.5	15
7	13.5	15.5
8	4.5	5
9	19	22
10	6	17.5
11	12.5	19
12	10	7.5
13	20.5	20
14	18.5	15
15	6	5
16	8	10.5
17	5	7
18	2	3
19	8	9
20	6	6
21	5	5
22	3	4
23	2	12
24	1	8

Bootstrapping

- In a bootstrap sample, we take the original values from the sample of 40 and select another sample of size 40 by selecting 40 values with replacement from the original sample.
- Bootstrapping allows us to quickly take a large number of samples from our population and plot the distribution of sample statistics, such as the sample mean.

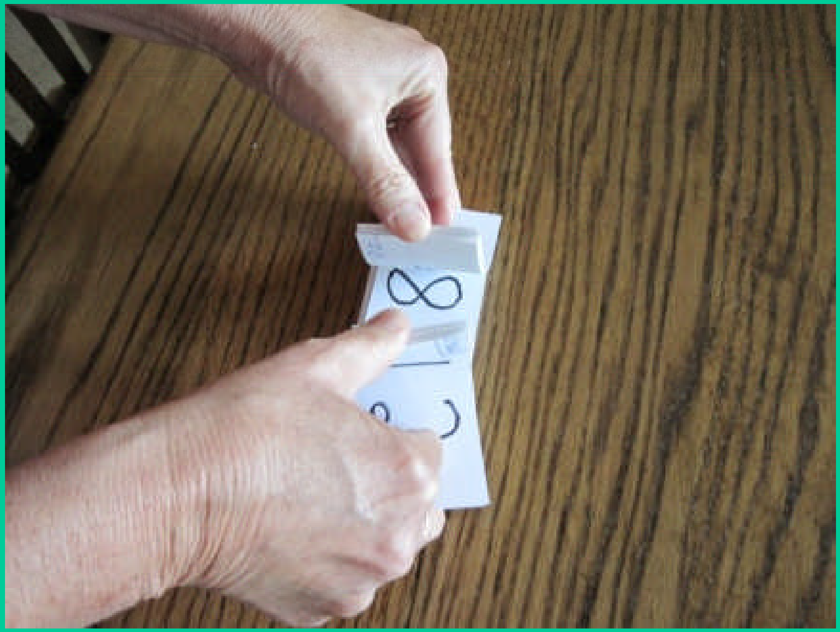
Bootstrapping-In Class Activity

We can simulate this resampling technique with a "hands-on" example. Suppose we have 40 cards with the original dominant hand reaction distances written on them as shown.

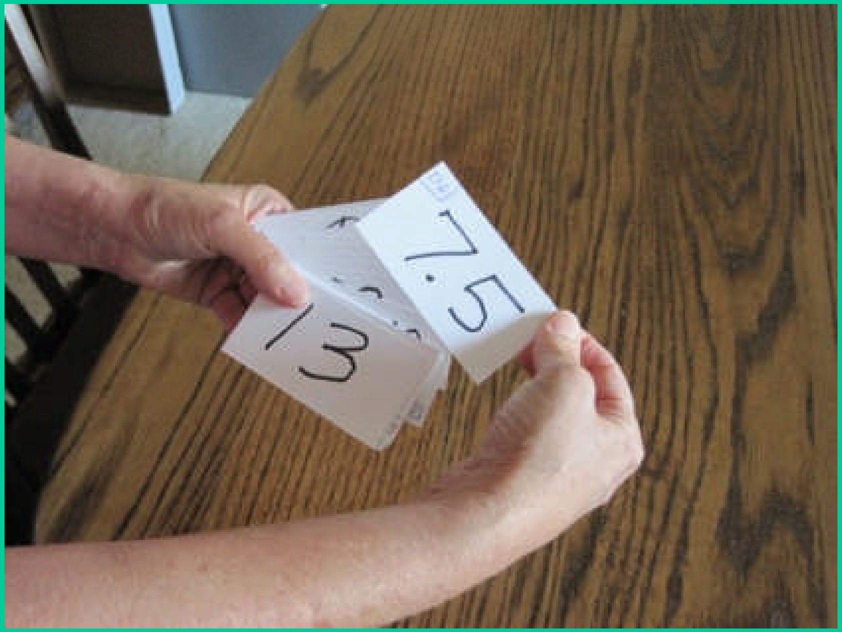


Bootstrapping-In Class Activity

We want to construct a bootstrap sample of size 40. First we would shuffle the cards.



Then we would select one card from the shuffled deck.



Bootstrapping-In Class Activity

EMBRY-RIDDLE Aeronautical University WORLDWIDE

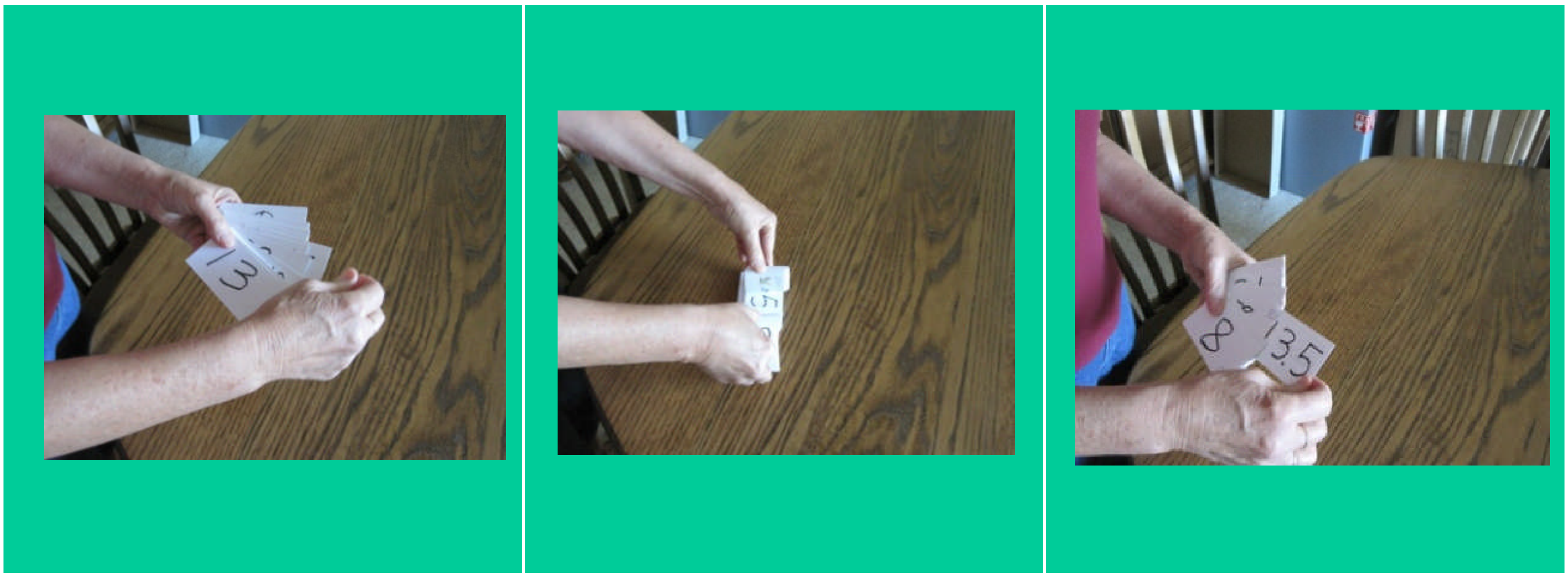
We would record that number as part of our bootstrap sample on the spreadsheet as shown.

	Original Data	ReSample
	Dom	Dom
D1	8	7.5
D2	5.5	
D3	9	
D4	8.5	
D5	7	
D6	7.5	
D7	13.5	
D8	4.5	
D9	19	

Bootstrapping-In Class Activity

EMBRY-RIDDLE Aeronautical University WORLDWIDE

We would put that card back into the deck, shuffle again, and select another card.



Bootstrapping-In Class Activity

EMBRY-RIDDLE Aeronautical University WORLDWIDE

We selected 13.5, so we would record it on our spreadsheet.

	Original Data	ReSample
	Dom	Dom
D1	8	7.5
D2	5.5	13.5
D3	9	
D4	8.5	
D5	7	
D6	7.5	
D7	13.5	
D8	4.5	
D9	19	

Bootstrapping-In Class Activity

EMBRY-RIDDLE Aeronautical University WORLDWIDE

- Continue the process until we had a new sample of size 40.
- If each student in a classroom course with 25 students went through this process, we would have 25 bootstrap samples that we could use in our analysis. We could plot the sample means and look at the distribution, for example.
- Time consuming process! Automate using StatCrunch

Bootstrapping with StatCrunch

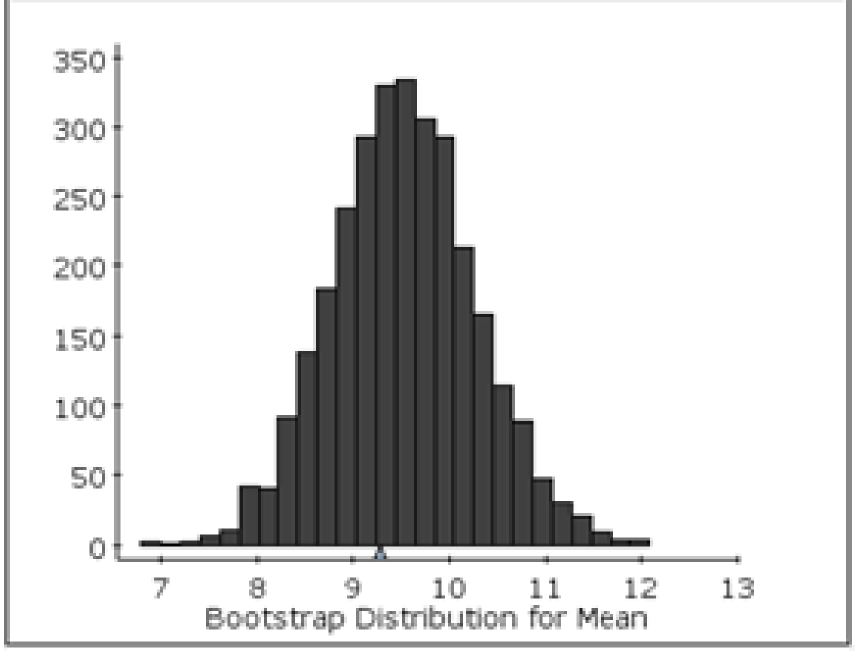
- StatCrunch Demonstration
  - File Hands Data
  - www.statcrunch.com

Bootstrapping with StatCrunch

- This display from StatCrunch shows the distribution of the 3006 sample means.
- Students often have difficulty understanding what a “distribution” is.
- The graphic is a picture of the distribution of sample means.

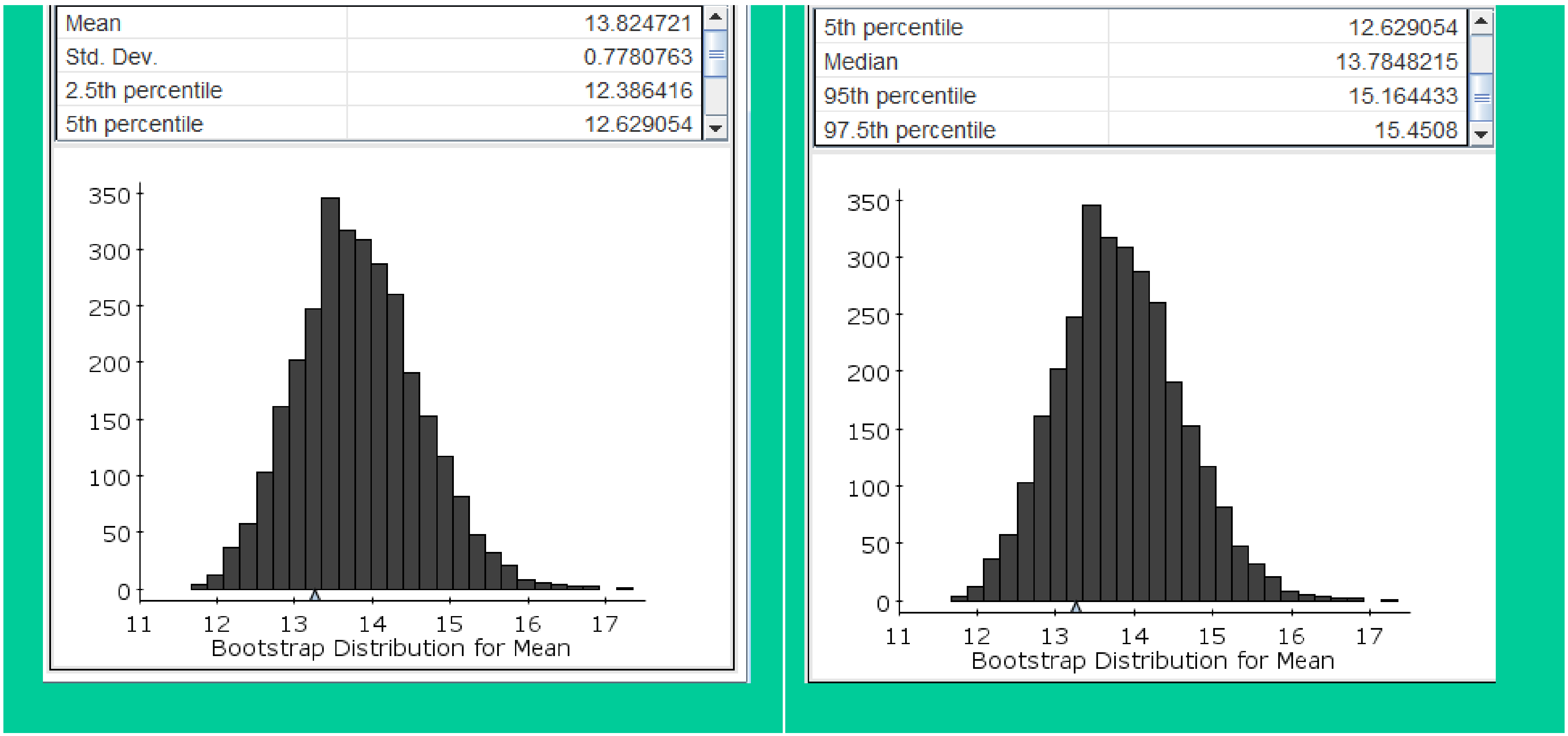
Bootstrap results for 3006 resamples:

Mean	9.525516
Std. Dev.	0.7385718
2.5th percentile	8.075
5th percentile	8.3375



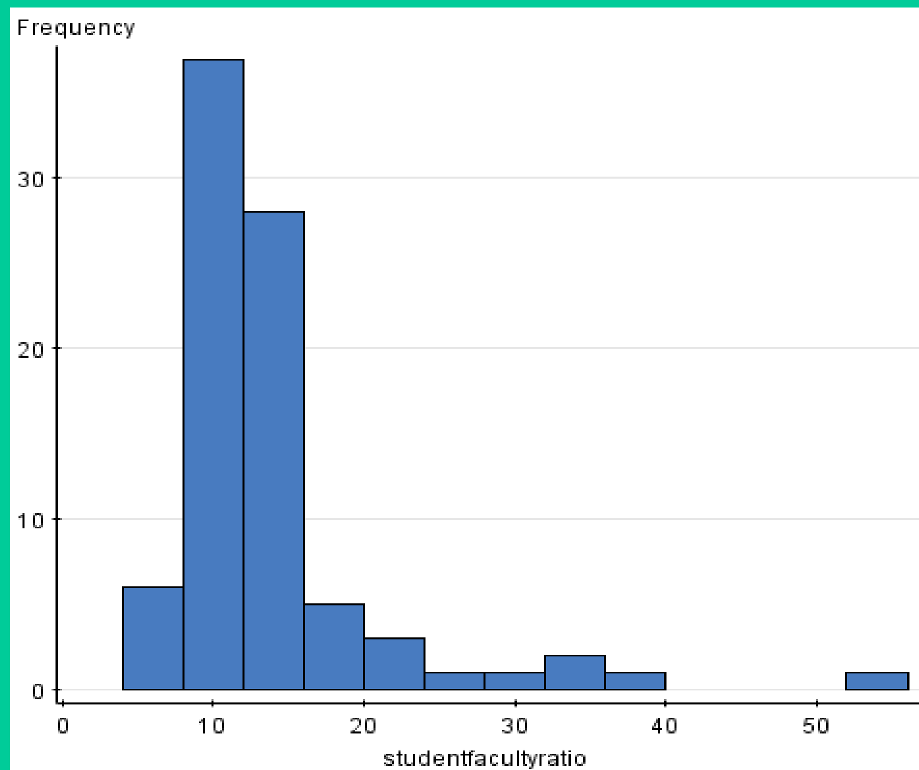
How would you describe the shape and spread?

The screen shots below show the results from 3000 resamples of student/faculty ratio data. They give the 2.5th , 5th , 95th, and 97.5th percentiles. We can use those to construct bootstrap confidence intervals and discuss the meaning of a CI.



Another Bootstrap Example

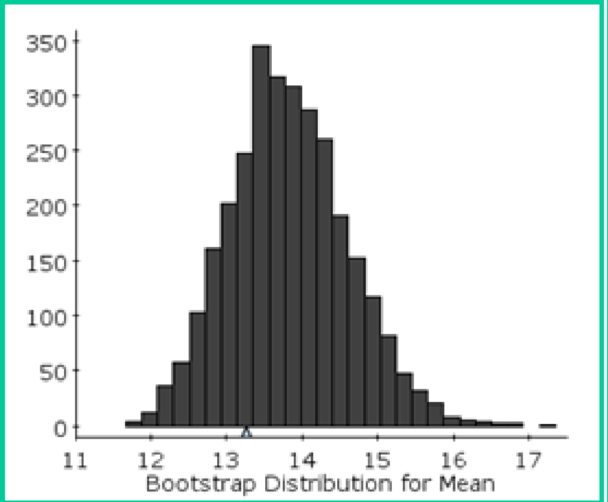
Here is a histogram for the student-faculty ratio data. How would you describe the shape of the distribution?



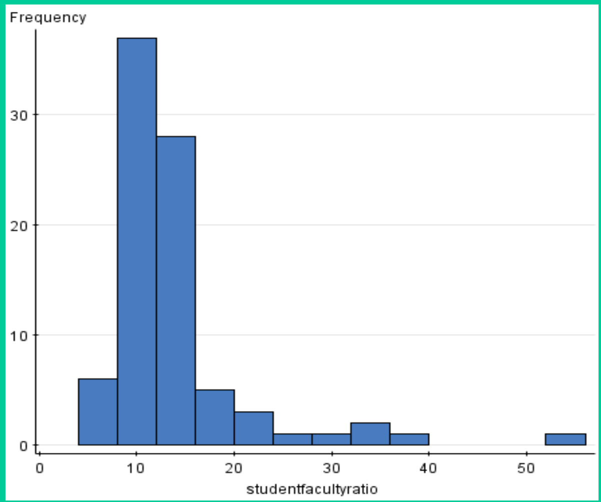
Another Bootstrap Example

What is the shape of the distribution of sample means from the 3000 bootstrap samples?

Distribution of Sample Means



Original Distribution



Randomization

- Use randomization techniques to lay groundwork for hypothesis testing.
- The data in the file represent the amounts spent on clothing during a one month period for random samples of 14 college women and 19 college men.
- Our hypotheses are:

  H_0: μ(women) = μ(men)

  H_a: μ(women) ≠ μ(men)

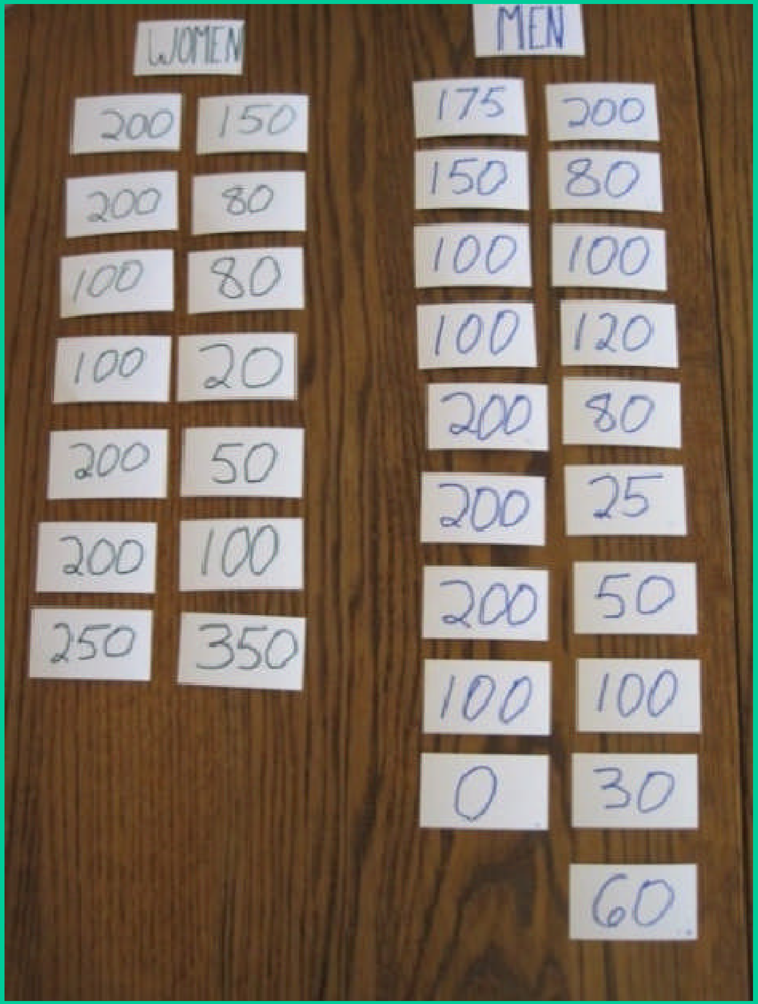
Row	$clothesperw>	$clothesperw>
1	200	175
2	200	150
3	100	100
4	100	100
5	200	200
6	200	200
7	250	200
8	150	100
9	80	0
10	80	200
11	20	80
12	50	100
13	100	120
14	350	80
15		25

Randomization

- In conducting the hypothesis test, we assume that H_0 is true.
- A key idea in the randomization technique is that if the null is true and there is no difference in the average monthly amounts spent on clothing by women and by men, then any of the data points could come from a man or a woman.

Randomization-In Class Activity

In a classroom course, we would first simulate this in a hands-on activity using cards as follows. We have two sets of cards, one showing the data for women and another showing the data for men.



Randomization-In Class Activity

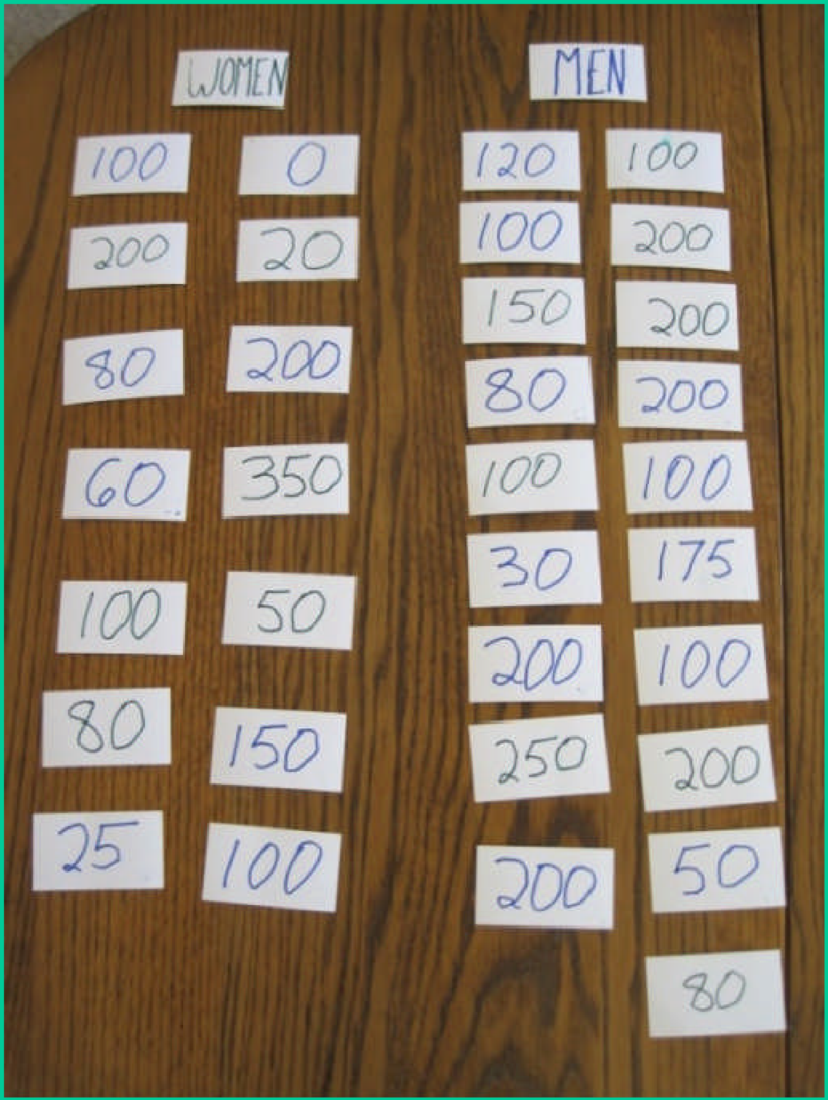
EMBRY-RIDDLE Aeronautical University WORLDWIDE

If any of the data values in the picture could come from either women or men, then we can put all the cards together in one pile, reshuffle them, and then form new samples of 14 values for the women and 19 values for the men.

Putting all the cards together	Reshuffling the cards	Forming new samples
		

Randomization-In Class Activity

The new sets of data make up our new randomized distribution. In a class with 25 students, we would have 25 "randomizations" and we could look at the distributions of sample means.



Randomization using StatCrunch

- This a somewhat time-consuming process. Its purpose is to give students a feel for what is going on in randomization.
- We will turn to technology and use StatCrunch to do this now.
- Use file Clothes Data

www.statcrunch.com

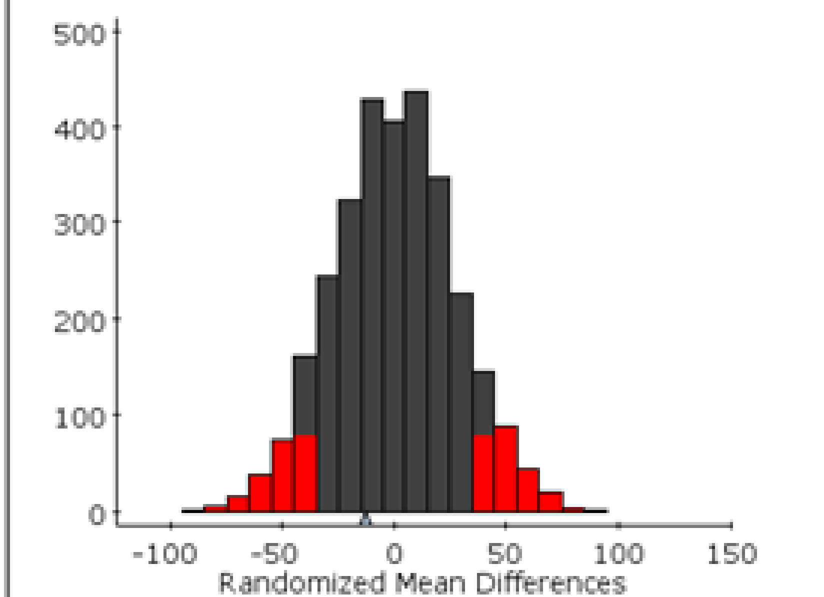
Randomization

EMBRY-RIDDLE Aeronautical University WORLDWIDE

- In the display, there is a region for -39.62406 or below and +39.62406 and above.
- As we do the simulation, we will be interested in the number and proportion of results that fall outside of those values. Those would be the results that give "more extreme" values than what we got in the original samples.

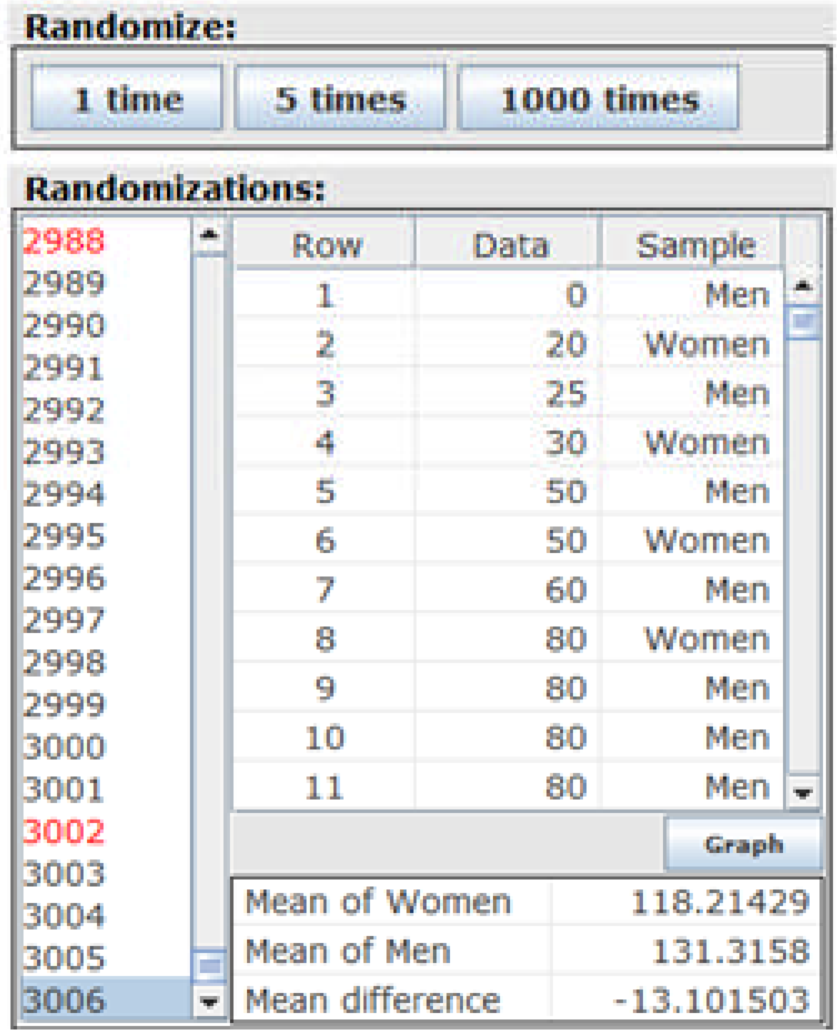
Results for 3006 randomizations:

Region	Count	Proportion
-39.62406 or below	210	0.06986028
39.62406 or above	234	0.077844314
Total	444	0.14770459



Randomization

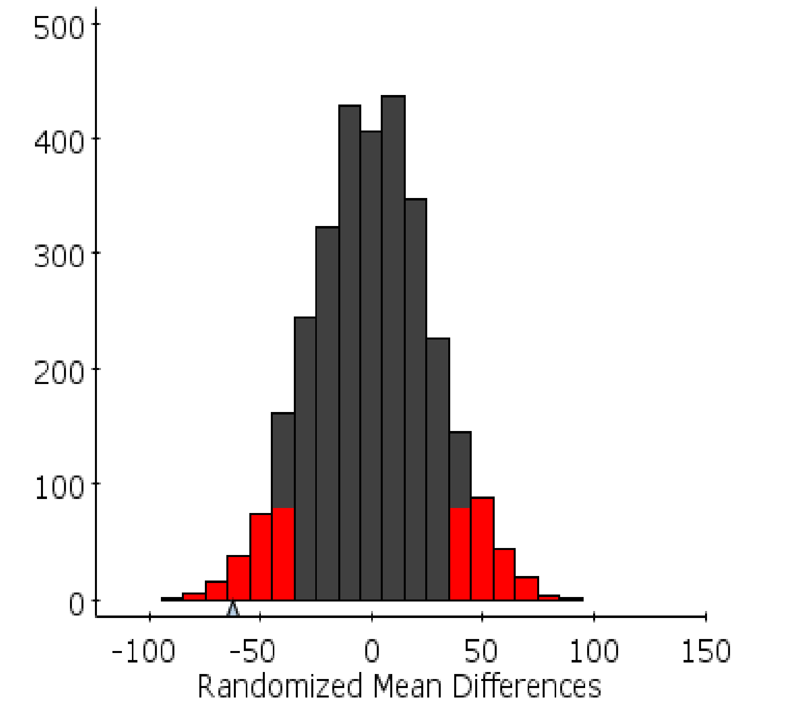
Notice that on the left side of the display most of the samples are shown in black, but some are in red. Those in red are the ones in which the difference in sample means was more extreme than what we had in our original data, that is, less than -39.62406 or greater than +39.62406.



The right side of the display shows that 210 or 6.99 percent of the samples had a difference in sample means less than what we got in our original sample and 234 or 7.78 percent had a difference in sample means greater than what we got in our original sample. A total of 444 or 14.77% of the 3006 samples resulted in a difference in sample means more extreme than what we got in the original data. Those results are also shown in red in the histogram.

Results for 3006 randomizations:

Region	Count	Proportion
-39.62406 or below	210	0.06986028
39.62406 or above	234	0.077844314
Total	444	0.14770459



Randomization

EMBRY-RIDDLE Aeronautical University WORLDWIDE

- To summarize—the question we want to answer is whether or not the results of our original sample show an “unusual” occurrence—something that would not happen just by chance but could be the result of a real difference in spending habits of college women and men.
- We assumed there is no difference in spending and ran 3006 simulations based on that assumption.
- Approximately 15% of the simulations resulted in differences in mean spending more extreme than what we got in our original sample. Stated another way, our simulation showed that even if there is no difference in mean spending between the two genders, we will get a difference at least as great as what we got in our original data approximately 15% of the time.

Randomization

- Do our results support the null hypothesis of no difference in mean spending or the alternate hypothesis that there is a real difference?
- The answer depends on whether or not we think an event that happens 15% of the time is unusual.
- We would probably say that something that can happen 15% of the time is not too unusual so there is not sufficient evidence to reject the null hypothesis.
- We would conclude that there isn’t a significant difference in mean monthly amounts spent on clothes.
- Even though the original data showed a difference, that difference isn’t large enough for us to conclude there is a real difference.

Goals for our course

EMBRY-RIDDLE Aeronautical University WORLDWIDE

- Integrate these guidelines into a nine-week course – classroom and online.
- Reduce the number of techniques covered to concentrate on understanding.
- Use Homework and Testing software that would give students immediate feedback.
- Use Technology to explore data, reduce computations, and give students time to explain results.
- Include a project.
- Use computer intensive methods to get at concepts

Questions?
Comments?

EMBRY-RIDDLE
Aeronautical University
WORLDWIDE

Some Lessons Learned in Developing an Introductory Online Undergraduate Statistics Course

Dr. Eugene L. Round