Teaching Introductory Statistics: GAISEing into the Future (www.amstat.org/education/gaise)

Allan J. Rossman
Dept of Statistics
Cal Poly – San Luis Obispo
arossman@calpoly.edu

GAISE

- Guidelines for Assessment and Instruction in Statistics Education
- Recommendations for teaching introductory statistics at college level
  - Comparable PreK-12 guidelines
- Developed, endorsed by American Statistical Association
  - 2005 committee chaired by Joan Garfield
  - 2016 committee chaired by Michelle Everson, Megan Mocko

Before GAISE: Cobb (1992) Report

I. Emphasize statistical thinking
   a) Need for data
   b) Importance of data production
   c) Omnipresence of variability
   d) Quantification and explanation of variability
II. More data and concepts; less theory and fewer recipes
III. Foster active learning

Since GAISE (2005):

- More students studying statistics
- More exposure to statistics in grades 6 – 12
- Huge increase in available data
- Emergence of discipline of data science
- More and better technology tools
- Alternative learning environments
- Calls for revision to "consensus curriculum"
- Newer ways for teaching logic of inference

GAISE recommendations

1. Teach statistical thinking.
2. Focus on conceptual understanding.
3. Integrate real data with a context and purpose.
4. Foster active learning.
5. Use technology to explore concepts and analyze data.
6. Use assessments to improve and evaluate student learning.

1. Teach statistical thinking

- Example: Sex discrimination?

	Men	Women
Accepted	533	113
Denied	665	336
Total	1198	449

- Men: 533/1198 ≈ .445
- Women: 113/449 ≈ .252

1. Teach statistical thinking

- Does this provide evidence of discrimination against women?

1. Teach statistical thinking

	Men		Women	
	Accepted	Denied	Accepted	Denied
Program A	511	314	89	19
Program F	22	351	24	317
Total	533	665	113	336

- Program A
  - Men: 511/825 ≈ .619
  - Women: 89/108 ≈ .824
- Program F:
  - Men: 22/373 ≈ .059
  - Women: 24/341 ≈ .070

1. Teach statistical thinking

- Consider where the data came from
  - Observational vs. experimental
- Engage in proportional reasoning
  - Take sample sizes into account
- Determine scope of conclusions
  - Random sampling, random assignment, both, neither
- Think about alternative explanations
  - Confounding variables

1. Teach statistical thinking

- Example: Cancer pamphlets

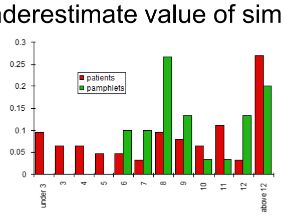
Researchers in Philadelphia investigated whether pamphlets containing information for cancer patients are written at a level that the cancer patients can comprehend

Patients' reading levels	< 3	3	4	5	6	7	8	9	10	11	12	> 12	Total
Count (number of patients)	6	4	4	3	3	2	6	5	4	7	2	17	63

Pamphlets' readability levels	6	7	8	9	10	11	12	13	14	15	16	Total
Count (number of pamphlets)	3	3	8	4	1	1	4	2	1	2	1	30

1. Teach statistical thinking

- Think in terms of distributions of data
- Be sure to address motivating question
- Do not underestimate value of simple graphs



2. Focus on conceptual understanding

- Example: Variability/SD

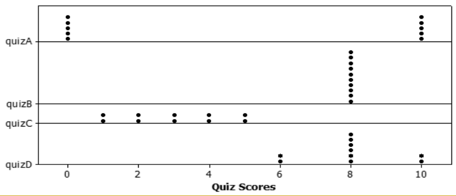
Suppose that Abby records the ages of customers at The Avenue (on-campus snack bar) from 11am-2pm today, while Mary records ages of customers at McDonald's (near freeway).

Who will have the larger standard deviation of customer ages: Abby or Mary? Explain.

2. Focus on conceptual understanding

- Example: Variability/SD

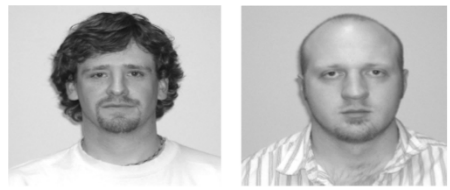
Arrange these distributions of quiz scores in order from largest SD to smallest SD.



3. Integrate real data

- Example: Facial prototyping

Do people tend to associate certain names with aces? (Lea, Thomas, Lamkin, & Bell, 2007)



Who is on the left: Bob or Tim?

3. Integrate real data

Example: Facial prototyping

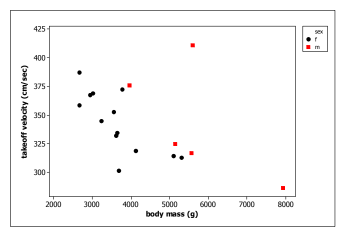
- What are two possible explanations for our observed sample result?
- Which explanation can we investigate/model? How?
- How often would such an extreme sample result occur by chance alone (if there were no facial prototyping)?
- Use coin and then technology to investigate

3. Integrate real data

The Journal of Experimental Biology 205, 3877–3889 (2002)
Printed in Great Britain © The Company of Biologists Limited 2002
JEB4511

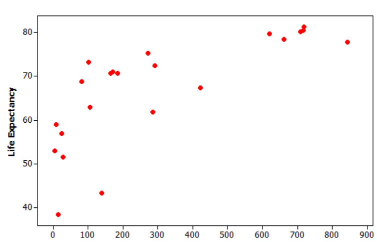
The relationship between maximum jumping performance and hind limb morphology/physiology in domestic cats (*Felis silvestris catus*)

Michelle A. Harris[1,*] and Karen Steudel[2]



4. Foster active learning

- Example: Televisions and life expectancy



4. Foster active learning

- Example: Televisions and life expectancy
  - Is the number of televisions (per thousand people) in a country associated with the country's life expectancy?
  - Would you conclude that sending more TVs to Angola would cause Angolans to live longer?
  - Can you suggest a confounding variable?
  - Is it reasonable to draw a cause-and-effect conclusion after observing a strong association between two variables?

4. Foster active learning

- Example: AIDS testing

ELISA test used to screen blood for the AIDS virus
- Sensitivity: Pr(+ | AIDS) = .977
- Specificity: Pr(- | no AIDS) = .926
- Base rate: Pr(AIDS) = .005

Determine Pr(AIDS | +)
- Initial guess?
- Construct table for hypothetical population

4. Foster active learning

	Positive	Negative	Total
AIDS			
No AIDS			
Total			1,000,000

4. Foster active learning

	Positive	Negative	Total
AIDS			5,000
No AIDS			995,000
Total			1,000,000

4. Foster active learning

	Positive	Negative	Total
AIDS	4885	115	5,000
No AIDS			995,000
Total			1,000,000

4. Foster active learning

	Positive	Negative	Total
AIDS	4885	115	5,000
No AIDS	73,630	921,370	995,000
Total			1,000,000

4. Foster active learning

	Positive	Negative	Total
AIDS	4885	115	5,000
No AIDS	73,630	921,370	995,000
Total	78,515	921,485	1,000,000

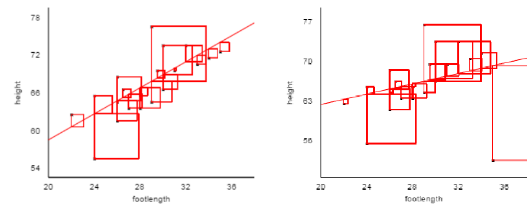
4. Foster active learning

	Positive	Negative	Total
AIDS	4885	115	5,000
No AIDS	73,630	921,370	995,000
Total	78,515	921,485	1,000,000

Pr(AIDS | +) = 4885 / 78,515 ≈ .062

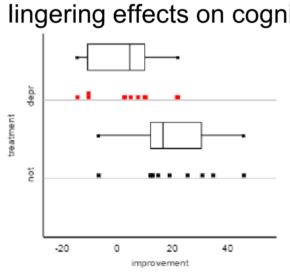
5. Use technology to explore concepts

- Example: Effect of outlier on least squares line



5. Use technology to explore concepts

- Example: Does sleep deprivation produce lingering effects on cognitive functioning?

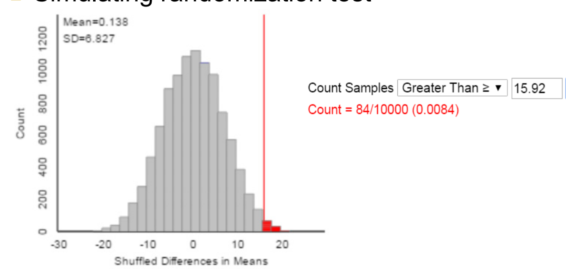


Summary Statistics:

	n	Mean	SD
depr	11	3.90	12.17
not	10	19.82	14.73
pooled	21	11.48	13.44

5. Use technology to explore concepts

- Simulating randomization test



6. Use assessments to improve learning

- Example: Sleep deprivation

a) Is this an observational study or an experiment?

b) What are the observational units in this study?

c) What are the variables in this study? Which type is which variable? Which variable plays which role?

d) Did this study make use of random sampling, random assignment, both, or neither?

6. Use assessments to improve learning

- Example: Sleep deprivation (cont)

e) What was the assumption behind the simulation: harmful effect, no effect, or helpful effect of sleep deprivation?

f) What is the variable in the graph of the simulation results?

g) What are the observational units in the graph of the simulation results?

6. Use assessments to improve learning

- Example: Sleep deprivation (cont)

h) Why did we count the simulation results equal to 15.92 or higher?

i) Interpret the (approximate) p-value: probability of what, assuming what?

j) Summarize your conclusion from the (approximate) p-value.

k) Estimate the magnitude of the effect with a confidence interval.

6. Use assessments to improve learning

- Example: Sleep deprivation (cont)

l) Is it reasonable to draw a cause-and-effect conclusion? Explain why or why not.

m) Is it reasonable to generalize the results to all people? Explain why or why not.

6. Use assessments to improve learning

- Example: Uniform colors

457 matches in 2004 Olympics: 248 won by wrestler in red, 209 by wrestler in blue

a) Test whether one uniform color (red, blue) tends to win Olympic wrestling matches significantly more than the other color.

b) How would test statistic, p-value, confidence interval, conclusion change if you analyzed the proportion of wins for *blue* rather than red?

6. Use assessments to improve learning

- Example (adapted from Jay Lehman):

a) Which would be larger – the mean weight of 10 randomly selected people or the mean weight of 1000 randomly selected cats? Explain briefly.

b) Which would be larger – the standard deviation of the weights of 10 randomly selected people or the standard deviation of the weights of 1000 randomly selected cats? Explain briefly.

6. Use assessments to improve learning

- Example (easy to grade):

For each of the following, indicate whether it can SOMETIMES or NEVER take a negative value.

a) Standard deviation

b) Correlation coefficient

c) Slope coefficient

d) Inter-quartile range

e) *p*-value

6. Use assessments to improve learning

- Example (harder to grade):

The purpose of a confidence interval is to estimate the unknown value of a population parameter with an interval of values determined from a sample.

Convince me that you understand this by describing an example of a situation (not shown in class or the textbook) in which you might do this. Be sure to clearly identify the variable, population, parameter, sample, and statistic.

Brief tangent

- What's the key to being a successful singer?
  - Sing Good Songs

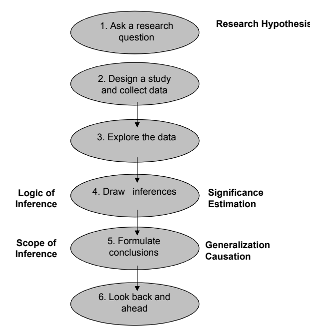
My similarly succinct advice

- What's the key to effective teaching?
  - Ask Good Questions

New emphases in GAISE revision

- Teach statistical thinking
  - Teach statistics as investigative process of problem-solving and decision-making
  - Give students experience with multivariable thinking

1a. Investigative process



1a. Investigative process

1b. Multivariable thinking

- Some examples
  - Labeled scatterplots
  - Matrix plots
  - 2×2×2 tables
  - Confounding variables
  - Simpson's paradox
- Not multiple regression

New emphases (cont)

- Many of the previous examples involve aspects of multivariable thinking, investigative process
- Many more examples, activities to come in workshop immediately following this presentation
  - Roxy Peck: multivariable thinking
  - Michael Posner: investigative process
  - Rob Gould: data science

Goals for introductory students

1. Become *critical consumers.*
2. Be able to apply *investigative process*.
3. Produce and interpret results of *graphical displays and numerical summaries*.
4. Recognize and explain fundamental role of *variability*.
5. Recognize and explain central role of *randomness* in designing studies and drawing conclusions.

Goals for introductory students (cont)

6. Gain experience with *statistical models*, including multivariable ones.
7. Demonstrate understanding of, and ability to apply, *statistical inference* in variety of settings.
8. Interpret and draw conclusions from standard output of *statistical software*.
9. Demonstrate awareness of *ethical issues* associated with sound statistical practice.

Topics that might be omitted

- Probability theory
  - Counting rules
  - Rules for probabilities of unions and intersections
- Constructing plots by hand
- Basic statistics learned in grades 6 – 12
- Drills with probability distribution tables
- Advanced training with statistical software packages

GAISE Appendices

A. Evolution of Introductory Statistics
B. Multivariable Thinking
C. Activities, Datasets, and Projects
D. Examples of Using Technology
E. Examples of Assessment Items
F. Learning Environments

Applicability of GAISE?

- To all kinds of introductory statistics courses
  - Statistical literacy vs. methods
  - All types of student majors
  - Class sizes
  - Learning environments: face-to-face, online, hybrid
  - Institution types: universities, colleges, two-year colleges, high schools
- Beyond introductory courses