Teaching Introductory Statistics: Ask Good Questions

Allan J. Rossman
Department of Statistics
Cal Poly – San Luis Obispo
arossman@calpoly.edu
https://askgoodquestions.blog
@allanjrossman

Frank Sinatra

- What's the secret to success as a singer?
  - "Sing Good Songs"

My similarly succinct suggestion

- What's the secret to success as a teacher?
  - "Ask Good Questions"

I'm tempted to stop right now

- My take-home message
  - Ask Good Questions
- Quiz at end of presentation
  - Q: What was my point?
  - A: Ask Good Questions
- I have started writing a weekly blog
  - https://askgoodquestions.blog

What kinds/purposes of questions?

- Guide students to develop their understanding and skills
  - Formative assessment
  - Learning activities
- Assess what students have learned
  - Summative assessment
  - Quiz/exam questions

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.

www.amstat.org/education/gaise/

Examples of questions for teaching …

- Statistical thinking
- Descriptive statistics
- Sampling and experimentation
- Simulation-based inference
- Confidence intervals
- More on inference (human progress)

Statistical thinking

- Example: Sex discrimination?

	Men	Women
Accepted	533	113
Denied	665	336
Total	1198	449

- Men: 533/1198 ≈ .445 were accepted
- Women: 113/449 ≈ .252 were accepted
- Does this provide evidence of discrimination against women?

1. 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

Statistical thinking

- Describe and explain the oddity, using the data given
- Most men apply to the easy program to get in; most women apply to the hard program to get in
- Multivariable thinking
  - Simpson's paradox

Statistical thinking

Follow-up assessment:

- Two softball players Amy (A) and Barb (B)
- Create an example in which
  - A has a higher proportion of hits than B in June
  - A has a higher proportion of hits than B in July
  - B has a higher proportion of hits than A when June and July are combined

Statistical thinking

- June: Amy gets 9 hits in 10 attempts (90%), Barb gets 80 hits in 100 attempts (80%)
- July: Amy gets 20 hits in 100 attempts (20%), Barb gets 1 hit in 10 attempts (10%)
- Combined: Amy gets 29 hits in 110 attempts (26.4%), Barb gets 81 hits in 110 attempts (73.6%)

Statistical thinking

- Example: Cancer pamphlets

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

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

Statistical thinking

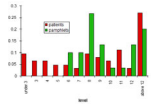
- Explain why you cannot calculate the mean reading level of a patient
- Calculate both medians.
- Do the medians indicate that the patients and the pamphlets are well matched?
- What percentage of patients have reading level below that of simplest pamphlet?

Statistical thinking

- Measures of center do not tell the whole story
- Think in terms of *distributions* of data
- Be sure to address motivating question

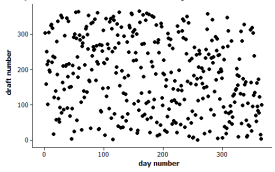
Statistical thinking

- Measures of center do not tell the whole story
- Think in terms of *distributions* of data
- Be sure to address motivating question
- Do not underestimate value of simple graphs



Statistical thinking

- Example: 1970 Draft Lottery



- Any reason to doubt randomness?

Statistical thinking

- Calculate median draft number for each month

Statistical thinking

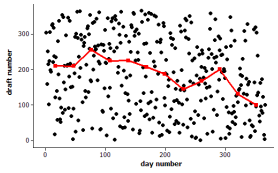
- Calculate median draft number for each month

Jan 211	Jul 188
Feb 201	Aug 145
Mar 256	Sep 168
Apr 225	Oct 201
May 226	Nov 131.5
Jun 207.5	Dec 100

- Do you see a pattern/trend?

Statistical thinking

- Example: 1970 Draft Lottery



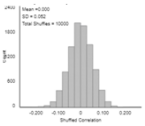
- Any reason to doubt randomness?

Statistical thinking

- Medians help to see pattern ("signal" amidst "noise")
- Randomness can be hard to achieve

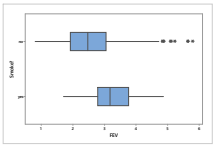
Statistical thinking

- How often would such an extreme outcome occur with a truly random lottery?
- What statistic might you use to determine "extreme"-ness?



Statistical thinking

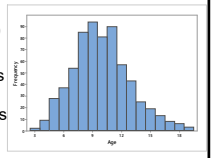
- Example: Lung capacity and smoking
- $t = 7.15$
- p-value ≈ .0000



- What's going on here???

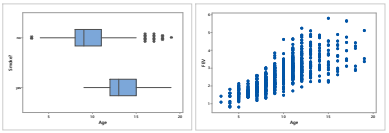
Statistical thinking

- Confounding variable
- These data are from children aged 3 – 19
- How does age explain why smokers have significantly larger lung capacities than non-smokers?



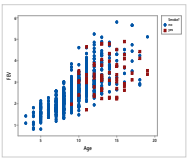
Statistical thinking

- Age is associated with both smoking status and lung capacity



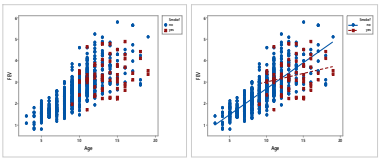
Statistical thinking

- Let's look at all three variables together



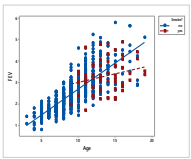
Statistical thinking

- Let's look at all three variables together



Statistical thinking

- After controlling for age, smokers have smaller lung capacity than non-smokers (12 or older)
- Rate of increase in lung capacity per year of age is smaller for smokers than for non-smokers



Descriptive statistics

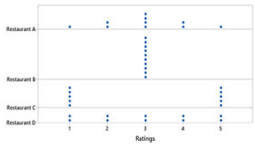
- I suspect that when I drove from California to Arizona, the average IQ dropped in both states!
  - Is this possible?
  - What would have to be true (in principle) for this to happen?
- Yes, if (my IQ > average IQ in California) and (my IQ < average IQ in Arizona)

Descriptive statistics

- 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.

Descriptive statistics

- Put the four restaurants in order from smallest SD of ratings to largest SD of ratings



Descriptive statistics

- Example (adapted from Jay Lehmann):

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 1000 randomly selected people or the standard deviation of the weights of 10 randomly selected cats? Explain briefly.

Descriptive statistics

- SAT scores follow a bell-shaped distribution with mean 1050 and SD 200
- ACT scores follow a bell-shaped distribution with mean 21 and SD 6
- Bob:1250 on SAT; Kathy: 29 on ACT
- Who did better, relative to their peers – Bob or Kathy? Justify your answer.

Descriptive statistics

- Exam scores have mean 70, SD 8
- Arturo's score is 75
- Bella's score is 1.5 SDs above Arturo's score.
- What is Bella's score on the exam? Show your work.

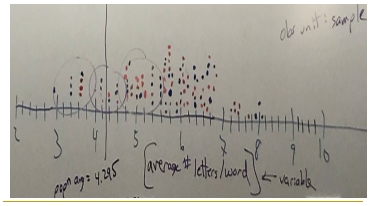
Sampling and experimentation

- Example: Gettysburg Address
- Select a sample of 10 words from the population of 268 words in the Gettysburg Address. (Just circle 10 words.)
- Record the length (# of letters) of each word.
- Calculate the average length for your sample.
- Produce graph of sample averages.

Sampling and experimentation

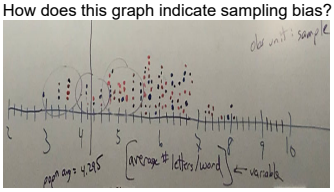
Is this a reasonable sampling method for estimating the average word length in the population?

Sampling and experimentation



Sampling and experimentation

- How does this graph indicate sampling bias?



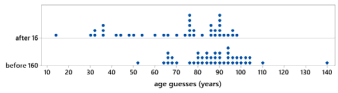
- Would closing eyes and pointing be unbiased?

Sampling and experimentation

- Example: Mandela's age
  - Group 1: Nelson Mandela was the first president of South Africa following apartheid.
    - Do you think he was older or younger than 16 years old when he died?
    - Make a guess for how old he was when he died.
  - Group 2: Nelson Mandela was the first president of South Africa following apartheid.
    - Do you think he was older or younger than 160 years old when he died?
    - Make a guess for how old he was when he died.

Sampling and experimentation

- Example: Mandela's age



- t = -4.17; p-value ≈ .0000
- Very strong evidence of anchoring effect

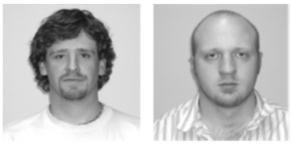
Sampling and experimentation

- Random sampling and random assignment are different tools with different purposes
- Random sampling
  - Goal: Select representative sample
  - Benefit: Generalize results
- Random assignment
  - Goal: Produce similar groups
  - Benefit: Draw cause-effect conclusions

Simulation-based inference

- Example: Facial prototyping

Do people tend to associate names with faces?



Who is on the left: Bob or Tim?

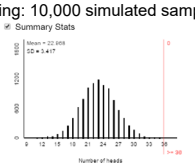
Simulation-based inference

Example: Facial prototyping (cont)

- 36 of 46 students put Tim on the left
  - 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)?
  - Have students flip coins to investigate

Simulation-based inference

- Facial prototyping: 10,000 simulated samples of 46 students



- Very strong evidence: people have a genuine tendency to put Tim on left

Confidence intervals

- Suppose that an alien lands on earth and sets out to estimate the proportion of human beings who are female
- The alien took a good statistics course on its home planet and knows to take a sample, produce a confidence interval
- Sample: the 2019 U.S. Senate, which has 25 women (the most ever!) and 75 men

Confidence intervals

- Calculate the alien's 95% CI
  - (.165 → .335)
- Interpret the CI for the alien
  - The alien is 95% confident that between 16.5% and 33.5% of all humans are female.
- Is this interval consistent with your experience as a long-time resident of planet Earth?
  - Duh!

Confidence intervals

- Is the problem that 5% of all 95% CIs fail to capture actual value of population parameter?
  - No!
- Then what went wrong???
  - Biased sampling method!
- If the alien was only interested in estimating the proportion of 2019 U.S. Senators who are female, would this 95% CI make sense?
  - No, exactly 25% of current senators are female!

Repeated questions

- Example: Anchoring (Mandela's age)

a) What are the observational units in this study?
b) What are the variables in this study? Which type is which variable? Which variable plays which role?
c) Did this study make use of random sampling, random assignment, both, or neither?
d) Is this an observational study or an experiment?

Repeated questions

- Example: Anchoring (Mandela's age)

e) Summarize your conclusion from the (approximate) p-value.
f) Estimate magnitude of effect with confidence interval.
g) Is it reasonable to draw a cause-and-effect conclusion? Explain why or why not.
h) Is it reasonable to generalize the results to all people? Explain why or why not.

A question for you

- In the last 20 years, the proportion of the world population living in extreme poverty has …?
  - Almost doubled
  - Remained more or less the same
  - Almost halved

Human progress

"Over the past twenty years, the proportion of the global population living in extreme poverty has halved. This is absolutely revolutionary. I consider it to be the most important change that has happened in the world in my lifetime." – Hans Rosling

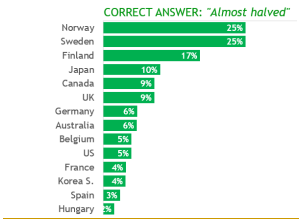
Human progress

"Here is a shocker: *The world has made spectacular progress in every single measure of human well-being.*

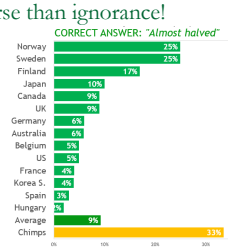
Here is a second shocker: *Almost no one knows about it.*"

– Steven Pinker

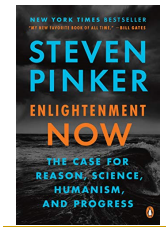
Ignorance about human progress

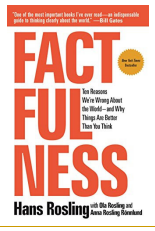


Worse than ignorance!



Human progress





Human progress (more on inference)

- U.S. survey: 5% of 1005 answered correctly
- Do the sample data provide strong evidence that less than one-third of all Americans would answer correctly?
- Test statistic: $z = \frac{.05-.3333}{\sqrt{\frac{.3333 \times .6667}{1005}}} \approx -19.1$ (!!!)
- Overwhelming evidence that Americans do worse than random guessing on this question

Human progress (more on inference)

- U.S. survey: 59% of 1005 answered "doubled"
- Estimate the proportion of all U.S. adults who would answer "doubled" with 95% confidence
  .59 ± .030: (.560, .620)
- Interpret this interval.
  - We are 95% confident that between 56% and 62% of all U.S. adults would give the most wrong answer to this question.

Some final questions for you

- Have you ever attended a presentation titled
  - Ask Bad Questions?
- Have you ever attended a presentation titled
  - Don't Ask Questions?
- Of course not! So, …
  - Why am I wasting your time offering such obvious advice as "Ask Good Questions"??

Yeah, why am I wasting your time?

- I respectfully suggest that the next time you or I feel compelled to invest some of our valuable time, thought, energy, and creativity on …
  - Preparing crystal-clear lectures
  - Writing lucid paragraphs of exposition
  - Developing software illustrations
  - Crafting beautiful presentations

What should we do?

What should we do?

- Resist this temptation!

So, what do I suggest instead?

- Instead we should invest these precious commodities (time, thought, energy, creativity) on
  - Developing engaging classroom activities
  - Preparing thought-provoking assignments
  - Writing lab or project expectations
  - Crafting effective assessment items
- In other words, we should be sure to …

What's the point again?

Ask Good Questions!

Thanks very much!

- Please follow my weekly blog with many more examples and advice:

https://askgoodquestions.blog

- Please contact me with questions and comments:

arossman@calpoly.edu