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, hypothesis tests
  - Repeated questions
  - Human progress

Statistical thinking

- Example: Sex discrimination?

<table>
<thead>
<tr>
<th></th>
<th>Men</th>
<th>Women</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accepted</td>
<td>533</td>
<td>113</td>
</tr>
<tr>
<td>Denied</td>
<td>667</td>
<td>356</td>
</tr>
<tr>
<td>Total</td>
<td>1300</td>
<td>469</td>
</tr>
</tbody>
</table>

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

1. Statistical thinking

<table>
<thead>
<tr>
<th></th>
<th>Men</th>
<th>Women</th>
</tr>
</thead>
<tbody>
<tr>
<td>Program A</td>
<td>32</td>
<td>51</td>
</tr>
<tr>
<td>Program B</td>
<td>55</td>
<td>367</td>
</tr>
</tbody>
</table>

- Program A
  - Men: 51/825 ≈ .061
  - Women: 367/1088 ≈ .334
- Program F
  - Men: 22/373 ≈ .059
  - Women: 24/341 ≈ .070

https://askgoodquestions.blog
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

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

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

- 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

Example: 1970 Draft Lottery
Any reason to doubt randomness?

Calculate median draft number for each month
Statistical thinking

- Calculate median draft number for each month
  - Jan 211
  - Feb 201
  - Mar 256
  - Apr 225
  - May 226
  - Jun 207.5
  - Jul 188
  - Aug 145
  - Sep 168
  - Oct 201
  - Nov 131.5
  - Dec 100
- Do you see a pattern/trend?

Statistical thinking

- Example: 1970 Draft Lottery

Statistical thinking

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

Example: Lung capacity and smoking

- $t = 7.15$
- $p$-value = 0.0000

What's going on here???

Confounding variable

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

Age is associated with both smoking status and lung capacity

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

Undergraduate statistics enrollments
- 2015 CBMS Survey
  - Look at this growth! Especially at TYCs
  - Also 222K took AP Statistics exam in 2018

Descriptive statistics
- I suspect that when I flew from California to Wisconsin, 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 Wisconsin)

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

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?

How does this graph indicate sampling bias?

Would closing eyes and pointing be unbiased?

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.

$t = -4.17; p-value \approx .0000$

Very strong evidence of anchoring effect

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?

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

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

- 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!

- 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!

Intervals and tests

- Survey of 47,000 U.S. households in 2006 found that 32.4% had a pet cat
- Does this provide very strong evidence that the population proportion with a cat is different from one-third?
  - Yes! Test stat \( z = -4.29 \), p-value \( \approx 0.00002 \)
- Does this provide strong evidence that the population proportion with a cat is very different from one-third?
  - No! 99.9% CI: (.317 → .331)

- Is the problem that 5% of all 95% CIs fail to capture actual value of population parameter?
  - Not
- 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)
  - What are the observational units in this study?
  - What are the variables in this study? Which type is which variable? Which variable plays which role?
  - Did this study make use of random sampling, random assignment, both, or neither?
  - Is this an observational study or an experiment?

- Example: Anchoring (Mandela’s age)
  - Summarize your conclusion from the (approximate) p-value.
  - Estimate magnitude of effect with confidence interval.
  - Is it reasonable to draw a cause-and-effect conclusion? Explain why or why not.
  - 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

Ignorance about 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

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: $z = \frac{0.05 - 0.33}{\sqrt{0.05(0.95)/1005}} = -19.1$ (!!!)
- Overwhelming evidence that Americans do worse than random guessing on 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?

- 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