PSY 201: Statistics in Psychology

Lecture 22

Estimation of population mean

Less than 5% of published psychological research should be wrong (and why that probably isn't true).

Greg Francis

Purdue University

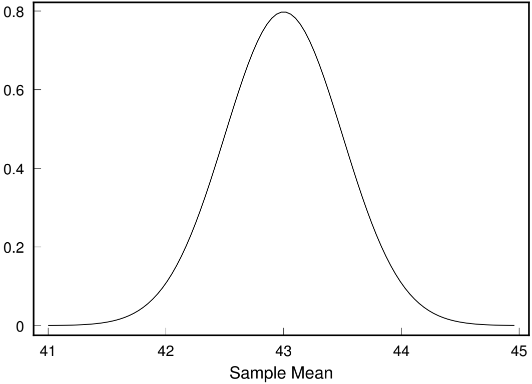
Fall 2023

LAST TIME

- construct an interval around an observed statistic, $\overline{X}$

$$\text{CI} = \text{statistic} \pm (\text{critical value})\,(\text{standard error of the statistic})$$

$$\text{CI} = \overline{X} \pm (t_{cv})(s_{\overline{X}})$$



CONFIDENCE

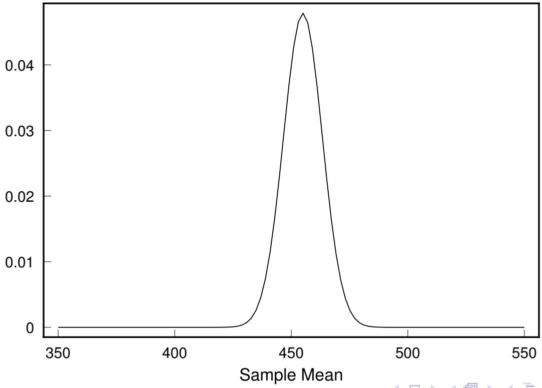
- we **never** say that a specific 95% confidence interval contains μ with probability 0.95
- either the interval contains μ or it does not
- we **can** say that the procedure of producing CI's produce intervals that contain μ with probability 0.95
- we do talk about the **confidence** that an interval includes μ
- we would say that the confidence interval contains μ with confidence of 0.95
- the confidence is in the **procedure** of calculating CIs

HYPOTHESIS TESTING

- remember SAT data:

$$H_0 : \mu = 455$$
$$H_a : \mu \neq 455$$

- calculated sampling distribution
  - for $\alpha = 0.05$, $\overline{X} = 535$, $s_{\overline{X}} = 8.33$
  - $t = 9.6$, $p \approx 0$, we rejected H_0



CONFIDENCE INTERVAL

- given our data, we could also compute confidence intervals around $\overline{X} = 535$
- $t_{cv} = \pm 1.96$

$$\text{CI} = \overline{X} \pm (t_{cv})(s_{\overline{X}})$$

$$\text{CI}_{95} = 535 \pm (1.96)(8.33)$$

$$\text{CI}_{95} = (518.67, 551.33)$$

COMPARISON

- note: the rejected $H_0 : \mu = 455$ is consistent with the CI
- 455 is **not** in the 95% confidence interval (518.67, 551.33)
- CI contains only tenable values of μ, given the sampled data

CI AND HYPOTHESIS TESTS

- CIs ask: which values of μ would it be reasonable for me to get the value of $\overline{X}$ that I found?
- Hypothesis tests ask: is the value of $\overline{X}$ I found consistent with a hypothesized value of μ?
- "reasonable" and "consistent" are defined relative to Type I error (α), and confidence (1-α)

HYPOTHESIS TESTING

- constructing a CI is like testing a large number of non-directional hypotheses simultaneously:

$$H_0 : \mu = 435$$

$$H_0 : \mu = 22$$

$$H_0 : \mu = 522$$

$$H_0 : \mu = 549$$

$$H_0 : \mu = 563$$

- anything in the CI (518.67, 551.33) would be not be rejected, anything not in the CI would be rejected

EXAMPLE

- On the papers going around the room, write down the number of math-based courses you have taken at college (include physics, engineering, and computer science, if it was largely math-based)
- Now go around the room and sample this information from 6 other people
- Calculate the mean and standard deviation for your sample (use the on-line calculator of the textbook)

$$\overline{X} = \frac{\sum X_i}{n}$$

$$s = \sqrt{\frac{\sum_i X_i^2 - [(\sum_i X_i)^2/n]}{n-1}}$$

HYPOTHESIS TEST

- (1) State the hypothesis and set the criterion:

$$H_0 : \mu = 3$$

$$H_a : \mu \neq 3$$

- $\alpha = 0.05$
- (2) Compute test statistic:
  - Calculate standard error of the mean

$$s_{\overline{X}} = \frac{s}{\sqrt{n}} = \frac{s}{\sqrt{6}} =$$

  - Compute the t-value

$$t = \frac{\overline{X} - 3}{s_{\overline{X}}} =$$

- (3) Compute the p-value:
  - use the t Distribution calculator with $df = n - 1 = 5$ to compute

$$p =$$

- (4) Make your decision:

CONFIDENCE INTERVAL

$$\text{CI} = \overline{X} \pm (t_{cv})(s_{\overline{X}})$$

- Calculate standard error of the mean

$$s_{\overline{X}} = \frac{s}{\sqrt{n}} = \frac{s}{\sqrt{6}} =$$

- with $n = 6$, $df = 5$, so the Inverse t Distribution calculator gives $t_{cv} = 2.571$
- plug everything into your CI formula

COMPARISON

- Who rejected H_0?
- Who have the value 3 outside their CI?
- Should be similar!
- Now repeat everything for $H_0 : \mu = 4$, using the textbook's online calculator
- notice what is required in the new calculations!

STATISTICAL PRECISION

- consider the equation for confidence intervals

$$\text{CI} = \overline{X} \pm (t_{cv})(s_{\overline{X}})$$

- where
  - $\overline{X}$ is the sample mean
  - t_{cv} is the critical value using the appropriate t distribution for the desired level of confidence
  - $s_{\overline{X}}$ is the estimated standard error of the mean

$$s_{\overline{X}} = \frac{s}{\sqrt{n}}$$

- smaller t_{cv} or $s_{\overline{X}}$ produce narrower widths

STATISTICAL PRECISION

- since

$$s_{\overline{X}} = \frac{s}{\sqrt{n}}$$

- increasing the sample size n produces narrow widths of CI
- narrower widths imply greater precision about where μ is located
- increasing n also modifies t_{cv} by changing degrees of freedom

$$\text{df} = n - 1$$

- larger df leads to smaller t_{cv}
  (see the Inverse t Calculator)

STATISTICAL PRECISION

- we can also change t_{cv} by changing the level of confidence
- larger level of confidence, implies smaller α, which implies larger t_{cv}, which implies larger width of CI
- makes sense, we become more confident the interval includes μ by broadening the interval
- of course, then we are less sure about the value μ

PUBLISHED DATA

- most researchers in the behavioral sciences use $\alpha = 0.05$
- this means that they make a Type I error only 5% of the time (or less)
- no way to completely avoid making mistakes
- this makes it quite likely that **some** of the data in published journals is wrong
- it is important in science to double (and triple) check everything
- if a bit of data is tremendously important, better replicate the experimental finding

PUBLISHING CHALLENGES

- Researchers often use statistical significance as a way of identifying what findings should be published
- If only findings with $p < .05$ are published, then journals can be filled with findings where H_0 is actually true
- even if H_0 is true, around 5% of samples will produce a significant p value
- If non-significant findings are not published, then it becomes hard to interpret the findings that actually are published (publication bias)

SAMPLING CHALLENGES

- Suppose you run a study with $n = 20$ subjects and get $p = .07$. This does not meet the $\alpha = 0.05$ criterion.
- It is tempting to add an additional 10 subjects (for a total of $n = 30$) and do the analysis again
- This is a problem because you have given yourself an extra chance to get a significant outcome. Your Type I error is bigger than the $\alpha = 0.05$ that you intended.
- Cannot add subjects to an experiment and re-analyze. Nor can you stop data collection when you get a significant result (data peeking, optional stopping).
- The sampling distribution is only valid for a **fixed** sample size. In the above cases, the sample size is not fixed.
- To avoid these problems, you have to plan your experiment carefully in advance (power).

PRECISION FOCUS

- One way to avoid these issues is to run your study to focus on measuring things "well enough".
- You might want to keep gathering data until the width of a 95% confidence interval is "small enough"
- Then you could test the H_0
- Of course, you have to come up with some definition of small enough

CONCLUSIONS

- estimation
- confidence intervals
- relationship with hypothesis testing
- statistical precision
- challenges

NEXT TIME

- more hypothesis testing
- tests for a proportion

Can you read my mind: Part II?