Advanced Statistics
DS2003 (BDS-4A)
Lecture 08

Instructor: Dr. Syed Mohammad Irteza

Assistant Professor, Department of Computer Science, FAST

10 March, 2022

Previous Lecture

- Testing hypotheses using confidence intervals
- Example – Facebook Interest Categories
- Test Statistic
- p-Values
- Making a decision
- College/University Applications:
  - Do the students of Duke follow the recommended number by College Board (8), or do they apply to more universities?

Number of college applications - Making a decision

- p-value = 0.0003
  - If the true average of the number of colleges Duke students applied to is 8, there is only 0.03% chance of observing a random sample of 206 Duke students who on average apply to 9.7 or more schools.
  - This is a pretty low probability for us to think that a sample mean of 9.7 or more schools is likely to happen simply by chance.
- Since p-value is *low* (lower than 5%) we *reject* H_0.
- The data provide convincing evidence that Duke students apply to more than 8 schools on average.
- The difference between the null value of 8 schools and observed sample mean of 9.7 schools is *not due to chance* or sampling variability.

Recap: Hypothesis testing framework

1. Set the hypotheses.
2. Check assumptions and conditions.
3. Calculate a *test statistic* and a p-value.
4. Make a decision, and interpret it in context of the research question.

Recap: Hypothesis testing for a population mean

1. Set the hypotheses

- H_0: μ = null value
- H_A: μ < or > or ≠ null value

2. Calculate the point estimate

3. Check assumptions and conditions

- Independence: random sample/assignment, 10% condition when sampling without replacement
- Normality: nearly normal population or $n \geq 30$, no extreme skew -- or use the *t* distribution (Ch 5)

4. Calculate a *test statistic* and a p-value (draw a picture!)

$$Z = \frac{\bar{x} - \mu}{SE}, \text{ where } SE = \frac{s}{\sqrt{n}}$$

5. Make a decision, and interpret it in context

- If p-value < α, reject H_0, data provide evidence for H_A
- If p-value > α, do not reject H_0, data do not provide evidence for H_A

Choosing a significance level

- Choosing a significance level for a test is important in many contexts, and the traditional level is 0.05. However, it is often helpful to *adjust the significance level based on the application*.
- We may select a level that is *smaller or larger than 0.05* depending on the *consequences of any conclusions reached* from the test.
- If making a *Type 1 Error* is dangerous or especially costly, we should choose a *small significance level* (e.g. 0.01). Under this scenario we want to be very *cautious about rejecting the null hypothesis*, so we demand very strong evidence favoring H_A before we would reject H_0.
- If a *Type 2 Error* is relatively more dangerous or much more costly than a *Type 1 Error*, then we should choose a *higher significance level* (e.g. 0.10). Here we want to be *cautious about failing to reject H_0* when the null is actually false.

Case of Type-1 Error being costly

- Assume we have two drugs for the same condition
- Drug-1 is very affordable, Drug-2 is very expensive
- ***H_0***: *Both drugs are equally effective*
- ***H_A***: *Drug-2 is more effective than Drug-1*
- *Type-1 Error → deciding Drug-2 is more effective when the reality is that it is no better than Drug-1*
- This is not desirable from the patient's perspective
- Best to choose a small significance level (e.g., 0.01, or 1%)

Case of Type-2 Error being costly

- Assume we have two drugs known to be equally effective for a certain condition. Both are equally affordable
- Some suspicion exists that Drug-2 causes a serious side-effect in some patients, whereas Drug-1 has been used for decades with no reported side-effects
- ***H_0***: *Incidence of the side-effect in both drugs is the same*
- ***H_A***: *the incidence of the side-effect in Drug-2 is greater than in Drug-1*
- *Falsely rejecting H_0 when its actually true (Type-1 Error) wouldn't have great consequences for the consumer, but*
- *A Type-2 Error → i.e., failing to reject H0, when in fact the H_A is true)*
- This may have serious consequences for public health
- Best to choose a larger significance level (e.g., 0.10, or 10%)

Inference for a single proportion: Example

- In a telephone poll of 1000 American adults, 440 said they approve of the way the President is handling his job
- $\hat{p} = \frac{440}{1000} = 0.44$ (sample proportion)→ estimate of p
- Construct a 95% confidence interval for p
- $\hat{p} \ \pm Z_{\propto/2} \ * SE(\hat{p})$ → term after ± is called the *margin of error*
- $SE(\hat{p}) = \sqrt{\frac{0.44*(1-0.44)}{1000}} = 0.0157$
- 0.44 ± 1.96*0.0157 = 0.44 ± 0.031 → (0.409, 0.471)

What do we infer from this?

- We can be 95% confident that if *all* adult Americans were contacted in this manner, between 40.9% and 47.1% would say they approve of the way the President is doing his job.

Same Matter, Hypothesis Test

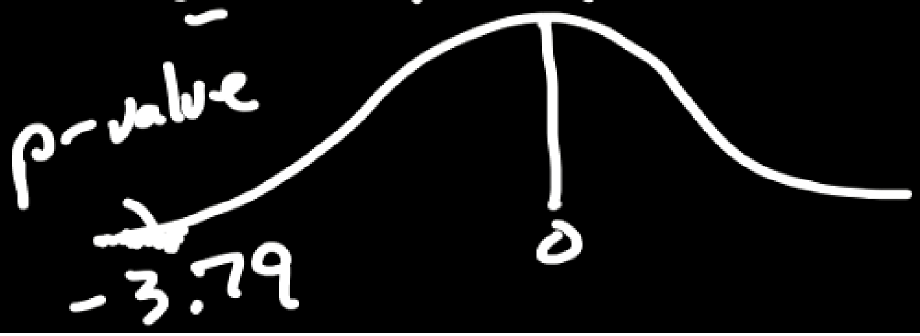
- In a telephone poll of 1000 American adults, 440 said they approve of the way the President is handling his job
- $\hat{p} = \frac{440}{1000} = 0.44$ (sample proportion)→ estimate of p
- Test the null hypothesis that the true proportion is 0.5 (i.e., 50%), against the alternative hypothesis that it is less than 0.5

$$H_0\colon p = 0.5, \quad \boxed{H_a\colon p < 0.5}$$

$$Z = \frac{\hat{p} - p_0}{SE_0(\hat{p})} \qquad SE_0(\hat{p}) = \sqrt{\frac{p_0(1-p_0)}{n}}$$

$$Z = \frac{0.44 - 0.50}{0.0158} = -3.79$$

$$SE_0(\hat{p}) = \sqrt{\frac{0.5(1-0.5)}{1000}} = 0.0158$$



$$\text{p-value} = 7 \times 10^{-5} = 0.00007$$

Look at this video for the working:

- https://www.youtube.com/watch?v=fVrSVYbrRzc

Back to experimental design...

The GSS found that 571 out of 670 (85%) of Americans answered the question on experimental design correctly. Estimate (using a 95% confidence interval) the proportion of all Americans who have good intuition about experimental design?

Back to experimental design...

The GSS found that 571 out of 670 (85%) of Americans answered the question on experimental design correctly. Estimate (using a 95% confidence interval) the proportion of all Americans who have good intuition about experimental design?

Given: $n = 670$, $\hat{p} = 0.85$. First check conditions.

Back to experimental design...

The GSS found that 571 out of 670 (85%) of Americans answered the question on experimental design correctly. Estimate (using a 95% confidence interval) the proportion of all Americans who have good intuition about experimental design?

Given: $n = 670$, $\hat{p} = 0.85$. First check conditions.

1. *Independence*: The sample is random, and 670 < 10% of all Americans, therefore we can assume that one respondent's response is independent of another.

Back to experimental design...

The GSS found that 571 out of 670 (85%) of Americans answered the question on experimental design correctly. Estimate (using a 95% confidence interval) the proportion of all Americans who have good intuition about experimental design?

Given: $n = 670$, $\hat{p} = 0.85$. First check conditions.

1. *Independence*: The sample is random, and 670 < 10% of all Americans, therefore we can assume that one respondent's response is independent of another.
2. *Success-failure*: 571 people answered correctly (successes) and 99 answered incorrectly (failures), both are greater than 10.

Practice

We are given that n = 670, $\hat{p}$ = 0.85, we also just learned that the standard error of the sample proportion is

$$SE_{\hat{p}} = \sqrt{\frac{p\,(1-p)}{n}}$$

Which of the below is the correct calculation of the 95\% confidence interval?

(a) $0.85 \pm 1.96 \times \sqrt{\frac{0.85 \times 0.15}{670}}$

(b) $0.85 \pm 1.65 \times \sqrt{\frac{0.85 \times 0.15}{670}}$

(c) $0.85 \pm 1.96 \times \frac{0.85 \times 0.15}{\sqrt{670}}$

(d) $571 \pm 1.96 \times \sqrt{\frac{571 \times 99}{670}}$

Practice

We are given that $n = 670$, $\hat{p} = 0.85$, we also just learned that the standard error of the sample proportion is

$$SE_{\hat{p}} = \sqrt{\frac{p\,(1-p)}{n}}$$

Which of the below is the correct calculation of the 95\% confidence interval?

(a) $0.85 \pm 1.96 \times \sqrt{\frac{0.85 \times 0.15}{670}} \rightarrow (0.82, 0.88)$

(b) $0.85 \pm 1.65 \times \sqrt{\frac{0.85 \times 0.15}{670}}$

(c) $0.85 \pm 1.96 \times \frac{0.85 \times 0.15}{\sqrt{670}}$

(d) $571 \pm 1.96 \times \sqrt{\frac{571 \times 99}{670}}$

Choosing a sample size

How many people should you sample in order to cut the margin of error of a 95% confidence interval down to 1%.

Choosing a sample size

How many people should you sample in order to cut the margin of error of a 95% confidence interval down to 1%.

$$ME = z^{\star} \times SE$$

Choosing a sample size

How many people should you sample in order to cut the margin of error of a 95% confidence interval down to 1%.

$$ME = z^{\star} \times SE$$

$$0.01 \quad \geq \quad 1.96 \times \sqrt{\frac{0.85 \times 0.15}{n}} \rightarrow \textit{Use estimate for } \hat{p} \textit{ from previous study}$$

Choosing a sample size

How many people should you sample in order to cut the margin of error of a 95% confidence interval down to 1%.

$$ME = z^{\star} \times SE$$

$$0.01 \geq 1.96 \times \sqrt{\frac{0.85 \times 0.15}{n}} \rightarrow \textit{Use estimate for } \hat{p} \textit{ from previous study}$$

$$0.01^2 \geq 1.96^2 \times \frac{0.85 \times 0.15}{n}$$

Choosing a sample size

How many people should you sample in order to cut the margin of error of a 95% confidence interval down to 1%.

$$ME = z^\star \times SE$$

$$0.01 \geq 1.96 \times \sqrt{\frac{0.85 \times 0.15}{n}} \rightarrow \text{Use estimate for } \hat{p} \text{ from previous study}$$

$$0.01^2 \geq 1.96^2 \times \frac{0.85 \times 0.15}{n}$$

$$n \geq \frac{1.96^2 \times 0.85 \times 0.15}{0.01^2}$$

Choosing a sample size

How many people should you sample in order to cut the margin of error of a 95% confidence interval down to 1%.

$$ME = z^{\star} \times SE$$

$$0.01 \geq 1.96 \times \sqrt{\frac{0.85 \times 0.15}{n}} \rightarrow \textit{Use estimate for } \hat{p} \textit{ from previous study}$$

$$0.01^2 \geq 1.96^2 \times \frac{0.85 \times 0.15}{n}$$

$$n \geq \frac{1.96^2 \times 0.85 \times 0.15}{0.01^2}$$

$$n \geq 4898.04 \rightarrow n \textit{ should be at least } 4{,}899$$

What if there isn't a previous study?

... use $\hat{p}$ = 0.5

why?

What if there isn't a previous study?

... use $\hat{p}$ = 0.5

why?

- if you don't know any better, 50-50 is a good guess

What if there isn't a previous study?

... use $\hat{p}$ = 0.5

why?

- if you don't know any better, 50-50 is a good guess
- $\hat{p}$ = 0.5 gives the most conservative estimate -- highest possible sample size

Sources

- openintro.org/os (Chapter 6)
- Type I and II Errors (utexas.edu)
- https://www.youtube.com/watch?v=fVrSVYbrRzc