Inference about μ: Estimation and Hypothesis Testing

Lecture 16

Reading: Chapter 13

1

Same-Day Term Test Inference

- Mark random sample of tests (benchmarking)
  - Cannot help making an inference about the class
  - Example of a term test in 2018/19: for sample of $n = 12$ papers, $\bar{X} = 58.5$ and $s = 13.43$
    - But we are believers in the law of small numbers so we should use a confidence interval estimate
    - 95% CI estimate to make an inference about the overall class average (μ), yields $LCL = 50.0$ and $UCL = 67.0$
    - But mean for all 445 students was 69.8
    - What went wrong?

2

Review & Preview: Inference about μ

- Sampling distribution $\bar{X}$:
  - $E[\bar{X}] = \mu$
  - $V[\bar{X}] = \frac{\sigma^2}{n}; SD[\bar{X}] = \frac{\sigma}{\sqrt{n}}$
  - CLT: For a random sample drawn from any population the sampling distribution of $\bar{X}$ is approximately Normal for a sufficiently large sample size.
- CI estimation
  - CI = Point Est. $\pm$ ME
    - Margin of error reflects both desired confidence level and sampling error
  - $\bar{X} \pm t_{\alpha/2} \frac{s}{\sqrt{n}}$
- Hypothesis testing
  - H_0: $\mu = \mu_0$
  - H_1: $\mu > \mu_0$ (or < or ≠)
  - P-value; Rejection region

3

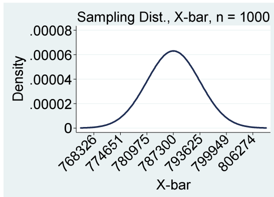
TREB Report: Toronto Housing, 2018

- N = 77,426 residential transactions, μ = \$787,300
- Sampling dist. of $\bar{X}$ if σ is \$200,000 and $n = 1{,}000$?

$E[\bar{X}] = \mu = \$787{,}300$

$SD[\bar{X}] = \frac{\sigma}{\sqrt{n}} = \frac{200{,}000}{\sqrt{1{,}000}}$

$= \$6{,}325$



$\bar{X}$ Normal by CLT ($n = 1{,}000$)

$$P\left(\mu - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \bar{X} < \mu + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha$$

$$P(787{,}300 - 1.96 * 6{,}325 < \bar{X} < 787{,}300 + 1.96 * 6{,}325) = 0.95$$

4

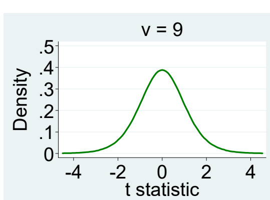
Non-Linearity Causes Trouble

- If $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$ & $Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$ then $Z \sim N(0,1)$
  - But $\frac{\bar{X}-\mu}{s/\sqrt{n}}$ is not distributed standard Normal
    - It is a *non-linear* combination of two random variables: the sample mean and sample s.d.
    - If replace σ with s, cannot use the critical value from the Normal table: $\bar{X} \pm z_{\alpha/2}\frac{s}{\sqrt{n}}$ X
    - 1908 William Gosset often had small samples (of beer): he was making more Type I errors than his chosen α

5

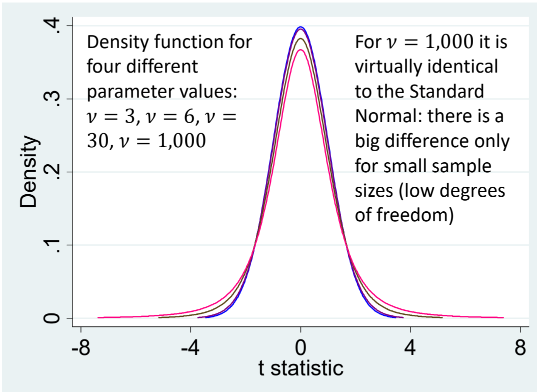
t Distribution

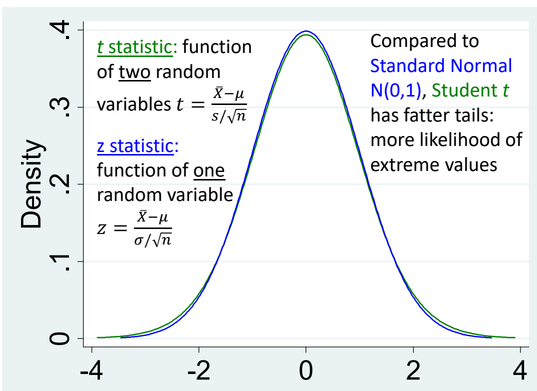
- Name this the t statistic (t ratio): $t = \frac{\bar{X}-\mu}{s/\sqrt{n}}$
  - Complex density function with one parameter ν (nu)
  - $\nu = n - 1$ and is called the degrees of freedom
  - As ν goes to infinity the distribution approaches the Standard Normal



Proof assumes Normal population. BUT robust to departures. For small n, need population roughly symmetric and unimodal ("Nearly Normal Condition"). For large n, CLT kicks in.

6

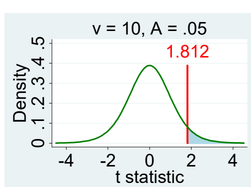


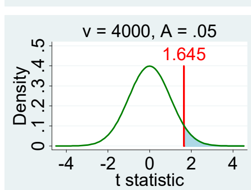


Note: $\nu = 20$ for the t distribution in above graphic

Student *t* Probabilities

- Use probability table
  - See course website (table we use posted next to these slides)
  - Reports t_A such that: $P(t > t_A \mid \nu) = A$ for A = 0.10, 0.05, 0.025, 0.01, 0.005, 0.001, 0.0005
  - When can you use Standard Normal table instead?





CI Estimator of μ

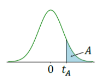
- CI estimator of μ: $\bar{X} \pm t_{\alpha/2} \frac{s}{\sqrt{n}}$ with confidence level $1-\alpha$ yields LCL $\bar{X} - ME$ UCL $\bar{X} + ME$
  - Derivation starts at $P(-t_{\alpha/2} < t < t_{\alpha/2}) = 1 - \alpha$

$$P\left(-t_{\alpha/2} < \frac{\bar{X}-\mu}{s/\sqrt{n}} < t_{\alpha/2}\right) = P\left(-t_{\alpha/2}\frac{s}{\sqrt{n}} < \bar{X}-\mu < t_{\alpha/2}\frac{s}{\sqrt{n}}\right)$$

$$P\left(-\bar{X} - t_{\alpha/2}\frac{s}{\sqrt{n}} < -\mu < -\bar{X} + t_{\alpha/2}\frac{s}{\sqrt{n}}\right) =$$

$$P\left(\bar{X} - t_{\alpha/2}\frac{s}{\sqrt{n}} < \mu < \bar{X} + t_{\alpha/2}\frac{s}{\sqrt{n}}\right) = 1 - \alpha$$

2018/19 Test, Compute 95% CI Est.



Critical Values of Student t Distribution:

ν	$t_{0.10}$	$t_{0.05}$	$t_{0.025}$	$t_{0.01}$	$t_{0.005}$	$t_{0.001}$	$t_{0.0005}$	ν	$t_{0.10}$	$t_{0.05}$	$t_{0.025}$	$t_{0.01}$	$t_{0.005}$	$t_{0.001}$	$t_{0.0005}$
1	3.078	6.314	12.71	31.82	63.66	318.3	636.6	38	1.304	1.686	2.024	2.429	2.712	3.319	3.566
2	1.886	2.920	4.303	6.965	9.925	22.33	31.60	39	1.304	1.685	2.023	2.426	2.708	3.313	3.558
...															
11	1.363	1.796	2.201	2.718	3.106	4.025	4.437	48	1.299	1.677	2.011	2.407	2.682	3.269	3.505
...															
35	1.306	1.690	2.030	2.438	2.724	3.340	3.591	750	1.283	1.647	1.963	2.331	2.582	3.101	3.304
36	1.306	1.688	2.028	2.434	2.719	3.333	3.582	1000	1.282	1.646	1.962	2.330	2.581	3.098	3.300
37	1.305	1.687	2.026	2.431	2.715	3.326	3.574	∞	1.282	1.645	1.960	2.326	2.576	3.090	3.291

Degrees of freedom: ν

How are these useful?

$$\bar{X} \pm t_{\alpha/2}\frac{s}{\sqrt{n}} = 58.5 \pm 2.201\frac{13.43}{\sqrt{12}} = 58.5 \pm 2.201 * 3.877$$

$$= 58.5 \pm 8.5 \quad LCL = 50.0;\ UCL = 67.0$$

Interpretation?

What if accidently use Normal table (i.e. $z_{\alpha/2}$ instead of $t_{\alpha/2}$)?

Check Understanding: $\bar{X} \pm t_{\alpha/2}\frac{s}{\sqrt{n}}$

- Recall the 95% CI estimate [50.0, 67.0]
  - Expected effect on the CI estimate of:
    - Higher confidence level? (e.g. 99%)
    - Benchmark 20 papers instead of 12?
    - Bigger class size?
    - More heterogeneity across students: more perfect papers and more papers with close to 0 marks?
    - 2018/19 test is “curved” by raising scores by 5% (note that is *not* same as raising by 5 percentage points)?

Sparton Resources of Toronto

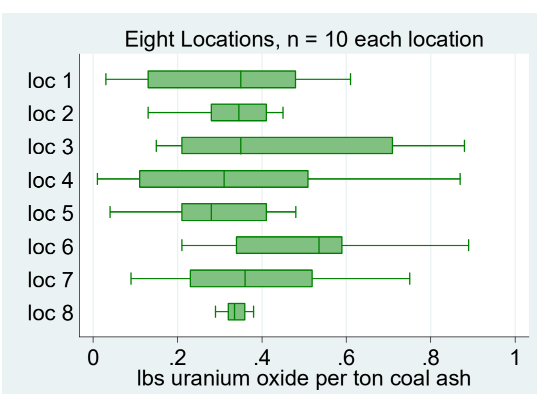
- Mini-case, page 430
  - Scarce uranium ore; required for nuclear power
  - Alternate source: coal ash (waste from creating coal power)
  - Concentration of uranium oxide varies widely depending on properties of the coal
- *To profitably exploit this source requires an average concentration of uranium oxide of at least 0.32 pounds (lbs) per tonne of coal ash*
- Sparton randomly selects 10 batches of ash from each of eight locations:
  1-4 (China), 5-7 (Central Europe), 8 (Africa)

13

Sparton: Raw Data (8 samples)

China				Central Europe			S. Africa
1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
0.32	0.22	0.71	0.33	0.22	0.57	0.41	0.35
0.38	0.28	0.22	0.51	0.21	0.34	0.56	0.31
0.58	0.31	0.78	0.61	0.04	0.59	0.23	0.34
0.61	0.37	0.15	0.11	0.09	0.54	0.09	0.32
0.12	0.39	0.19	0.12	0.25	0.22	0.52	0.33
0.13	0.45	0.88	0.01	0.43	0.89	0.31	0.37
0.48	0.44	0.53	0.07	0.48	0.34	0.18	0.32
0.03	0.13	0.21	0.87	0.39	0.61	0.49	0.36
0.43	0.32	0.33	0.43	0.31	0.53	0.29	0.29
0.17	0.41	0.37	0.29	0.41	0.21	0.75	0.38

14



15

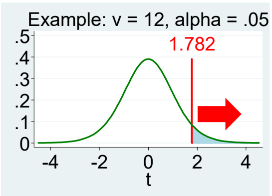
Hypothesis Testing μ

- $H_0: \mu = \mu_0$
- $H_1: \mu > \mu_0$ or $H_1: \mu < \mu_0$ or $H_1: \mu \neq \mu_0$
  - Test statistic: $t = \frac{\bar{X}-\mu_0}{s/\sqrt{n}}$, which is Student t distributed with $\nu = n - 1$
    - Rejection (Critical) Region Approach: Given α, ν, and direction of H_1, find rejection region and check if test statistic t is or is not in rejection region
    - P-value Approach: Using test statistic t, ν, and direction of H_1, compute P-value (area in right, left or both tails)

16

Rejection Region, Right Tailed

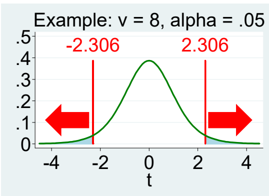
- H_0: $\mu = \mu_0$
- H_1: $\mu > \mu_0$
- Rejection region: (t_α, ∞)
  - Left edge is called the critical value (t^*_α)
    - Depends on degrees of freedom



17

Rejection Region, Two Tailed

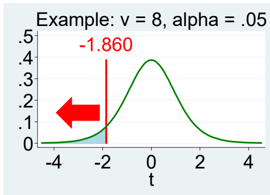
- H_0: $\mu = \mu_0$
- H_1: $\mu \neq \mu_0$
- Rejection region: $(-\infty, -t_{\alpha/2})$ & $(t_{\alpha/2}, \infty)$
  - Edges are called the critical values ($t^*_{\alpha/2}$)
  - Depend on degrees of freedom



18

Rejection Region, Left Tailed

- H_0: $\mu = \mu_0$
- H_1: $\mu < \mu_0$
- Rejection region: $(-\infty, -t_\alpha)$
  - Right edge is called the critical value ($-t^*_\alpha$)
    - Depends on degrees of freedom



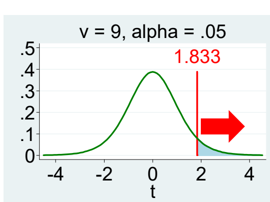
Sparton: Set-up Hypotheses

	n	mean	s.d.
loc 1	10	0.325	0.204
loc 2	10	0.332	0.102
loc 3	10	0.437	0.270
loc 4	10	0.335	0.274
loc 5	10	0.283	0.147
loc 6	10	0.484	0.208
loc 7	10	0.383	0.200
loc 8	10	0.337	0.028

- Hypotheses to test how Location i compares to the 0.32 lbs/tonne profitability threshold?
  - $H_0: \mu_i = 0.32$
    $H_1: \mu_i > 0.32$
  - $H_0: \mu_i = 0.32$
    $H_1: \mu_i < 0.32$
  - $H_0: \mu_i = 0.32$
    $H_1: \mu_i \neq 0.32$

Sparton: Location 8

- Sampled 10 batches of coal ash at Loc. 8
  - Mean conc. of uranium ore is 0.337 lbs/ton
  - S.d. conc. of uranium ore is 0.028 lbs/ton
- H_0: $\mu_8 = 0.32$
- H_1: $\mu_8 > 0.32$

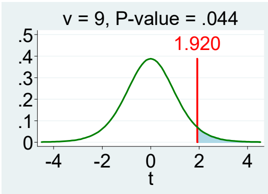


Conclusion?

$$t = \frac{\bar{X}_8 - \mu_0}{\frac{s_8}{\sqrt{n_8}}} = \frac{0.337 - 0.32}{\frac{0.028}{\sqrt{10}}} = 1.92$$

P-value: Location 8

- H_0: $\mu_8 = 0.32$
- H_1: $\mu_8 > 0.32$
- $t = 1.92$
- P-value = $P(t > 1.92 \mid \nu = 9)$
  - With software find exact P-value = 0.044
  - With table find that the P-value is between 0.025 and 0.05

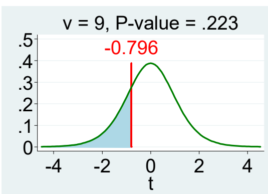


Student t table tells us:
$P(t > 2.262 \mid \nu = 9) = 0.025$
$P(t > 1.833 \mid \nu = 9) = 0.050$

22

Location 5: Confident It's Bad?

- Location 5, n = 10:
  - Mean = 0.283
  - S.d. = 0.147
- How to set-up?
  - $H_0: \mu_5 = 0.32$
    $H_1: \mu_5 > 0.32$
  - $H_0: \mu_5 = 0.32$
    $H_1: \mu_5 < 0.32$
  - $H_0: \mu_5 = 0.32$
    $H_1: \mu_5 \neq 0.32$



$$t = \frac{\bar{X}_5 - \mu_0}{\frac{s_5}{\sqrt{n_5}}} = \frac{0.283 - 0.32}{\frac{0.147}{\sqrt{10}}}$$
$$= -0.796$$

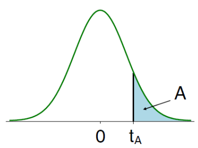
23

Finding P-values with Tables

- You can approximate the P-value when doing hypothesis testing for inference about μ even *without* a computer:
  - With small to fairly large sample sizes ($\nu \leq 1{,}000$) use the Student t table
    - E.g. earlier found P-value between 0.025 and 0.05
    - See also page 422 in your textbook
  - With big sample sizes ($\nu > 1{,}000$) use the Normal table to find P-value (an excellent approximation)

24

Suppose that:
H$_0$: μ = 70
H$_1$: μ > 70
t = 2.147 and ν = 5
What's the P-value?



Critical Values of Student t Distribution:

ν	$t_{0.10}$	$t_{0.05}$	$t_{0.025}$	$t_{0.01}$	$t_{0.005}$	$t_{0.001}$	$t_{0.0005}$
1	3.078	6.314	12.71	31.82	63.66	318.3	636.6
2	1.886	2.920	4.303	6.965	9.925	22.33	31.60
3	1.638	2.353	3.182	4.541	5.841	10.21	12.92
4	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	1.372	1.812	2.228	2.764	3.169	4.144	4.587

Parents' Beliefs About Their Children's Academic Ability: Implications for Educational Investments

Abstract: Schools worldwide distribute information to parents about their children's academic performance. Do frictions prevent parents, particularly low-income parents, from accessing this information to make decisions? A field experiment in Malawi shows that, at baseline, parents' beliefs about their children's academic performance are often inaccurate. Providing parents with clear, digestible performance information causes them to update their beliefs and adjust their investments: they increase the school enrollment of their higher-performing children, decrease the enrollment of lower-performing children, and choose educational inputs that are more closely matched to their children's academic level. Heterogeneity analysis suggests information frictions are worse among the poor.

Source: Rebecca Dizon-Ross, 2019, *American Economic Review*; For a great introduction, watch: "To help students, start by informing parents," *Chicago Booth Review*, March 16, 2018, https://youtu.be/9SM3jSNzxps

Summary Statistics		
	Mean	SD
Academic Performance (Average Achievement Scores)		
Overall score	46.8	17.5
Math score	44.9	20.2
English score	44.2	20.1
Chichewa score	51.2	22.5
(Math – English) score	0.71	19.5
Respondent's Beliefs about Child's Academic Performance		
Believed Overall score	62.4	16.5
Believed Math score	64.7	19.0
Believed English score	55.3	20.9
Believed Chichewa score	66.8	19.4
Believed (Math – English) score	9.48	21.5
Sample size (number of kids)	5,268	

Excerpt Online Appendix Table C.25, Dizon-Ross (2019); 39 randomly selected primary schools in Machinga and Balaka districts in Malawi.

Week 16: inference about *a single mean* (e.g. μ_{AE} where AE means academic English score)

Week 17: inference about *difference in means* (independent samples or paired data)