Small-Sample CI's for Normal Pop. Mean

Engineering Statistics
Section 7.3

Josh Engwer

TTU

04 April 2016

PART I:

GOSSET'S t DISTRIBUTION

William Sealy Gosset (1876-1937)

Gosset's employer made him publish under the pseudonym "Student".

Gosset's t Distribution (AKA Student's t Distribution)

Definition

Notation	$T \sim t_\nu$
Parameter(s)	$\nu \equiv$ # Degrees of Freedom $(\nu = 1, 2, 3, 4, \cdots)$
Support	Supp$(T) = (-\infty, \infty)$
pdf	$f_T(t;\nu) := \frac{\Gamma((\nu+1)/2)}{\sqrt{\pi\nu}\cdot\Gamma(\nu/2)} \cdot \frac{1}{[1+(t^2/\nu)]^{(\nu+1)/2}}$
cdf	$\Phi_t(t;\nu) = \frac{\Gamma((\nu+1)/2)}{\sqrt{\pi\nu}\cdot\Gamma(\nu/2)} \int_{-\infty}^{t} \frac{1}{[1+(\tau^2/\nu)]^{(\nu+1)/2}}\, d\tau$
Mean	$\mathbb{E}[T] = +\infty$, for $\nu = 1$ $\mathbb{E}[T] = 0$, for $\nu > 1$
Variance	$\mathbb{V}[T] = +\infty$, for $\nu = 1, 2$ $\mathbb{V}[T] = \nu/(\nu - 2)$, for $\nu > 2$
Model(s)	(Used exclusively for Statistical Inference)

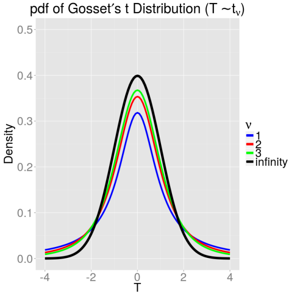
ν is the lowercase Greek letter "nu" τ is the lowercase Greek letter "tau"

Gosset's t Distribution (AKA Student's t Distribution)

Proposition

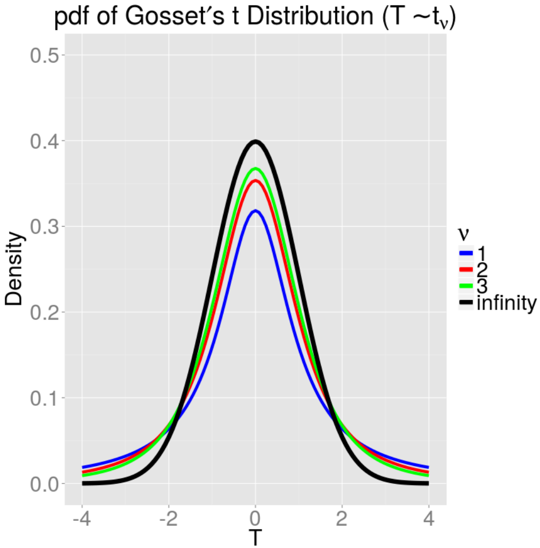
Properties of the t_ν distribution:

- t_ν *is symmetric, bell-shaped and centered at zero.*
- t_ν *is more spread out than the standard normal curve.*
- *The spread of the t_ν curve decreases as ν increases.*
- *As $\nu \to \infty$, the t_ν curves approaches the standard normal curve.*



The black curve is the **Standard Normal curve**.

Plots of t Distributions (A Closer Look)



The black curve is the **Standard Normal curve**.

t Critical Values

A key component to some CI's is the t **critical value**:

Definition

$t^*_{\nu,\alpha/2}$ is called a t **critical value** of the t distribution with ν df's such that its upper-tail probability is exactly its subscript value $\alpha/2$: (Here, $T \sim t_\nu$)

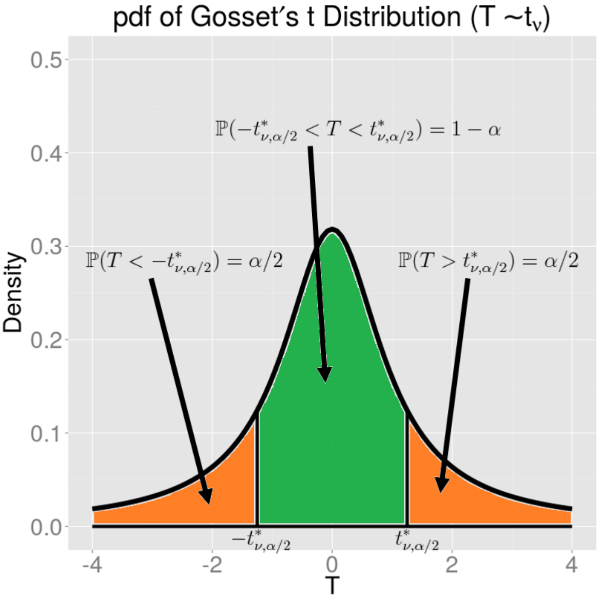
$$\mathbb{P}(T > t^*_{\nu,\alpha/2}) = \alpha/2$$

IMPORTANT: Do not confuse t critical value $t^*_{\nu,\alpha/2}$ with the t percentile $t_{\nu,\alpha/2}$:

$$\mathbb{P}(T \leq t_{\nu,\alpha/2}) = \alpha/2$$

Finally, notice that $t^*_{\nu,\alpha/2}$ is always **positive**.

t Critical Values



t Critical Value Table

ν	**90% CI** ($\alpha/2 = 0.05$)	**95% CI** ($\alpha/2 = 0.025$)	**99% CI** ($\alpha/2 = 0.005$)
1	6.314	12.706	63.657
2	2.920	4.303	9.925
3	2.353	3.182	5.841
4	2.132	2.776	4.604
5	2.015	2.571	4.032
6	1.943	2.447	3.707
7	1.895	2.365	3.499
8	1.860	2.306	3.355
9	1.833	2.262	3.250
10	1.812	2.228	3.169
11	1.796	2.201	3.106
12	1.782	2.179	3.055
13	1.771	2.160	3.012
14	1.761	2.145	2.977
15	1.753	2.131	2.947
16	1.746	2.120	2.921

PART II:

SMALL-SAMPLE CI'S FOR NORMAL POPULATION MEAN μ

A Statistic related to the t Distribution

Theorem

Let $X_1, \ldots, X_n$ be a random sample from a Normal(μ, σ^2) population. Then:

$$\frac{\overline{X} - \mu}{S/\sqrt{n}} \sim t_{n-1}$$

PROOF: (It's complicated...)

Small-Sample CI for Normal Pop. Mean μ (Motivation)

Given a normal population with unknown mean μ and std dev σ.
Let $\mathbf{X} := (X_1, \ldots, X_n)$ be a random sample from the population.

Then, construct the $100(1-\alpha)\%$ CI for parameter μ:

1. Produce a suitable **pivot**: Let $Q(\mathbf{X};\mu) = \dfrac{\overline{X}-\mu}{S/\sqrt{n}}$
2. Then the pivot is a t distribution with $\nu = (n-1)$ df's: $Q(\mathbf{X};\mu) \sim t_{n-1}$
3. Find constants $a < b$ such that $\mathbb{P}(a < Q(\mathbf{X};\mu) < b) = 1-\alpha$

   Since the t_{n-1} pdf is symmetric, $a = -t^*_{n-1,\alpha/2}$ and $b = t^*_{n-1,\alpha/2}$
4. Manipulate the inequalities to isolate parameter μ:

$$-t^*_{n-1,\alpha/2} < \frac{\overline{X}-\mu}{S/\sqrt{n}} < t^*_{n-1,\alpha/2} \implies \overline{X}-t^*_{n-1,\alpha/2}\cdot\frac{S}{\sqrt{n}} < \mu < \overline{X}+t^*_{n-1,\alpha/2}\cdot\frac{S}{\sqrt{n}}$$

5. Take a size n sample $\mathbf{x} := (x_1, \ldots, x_n)$ from the population.
6. Replace point estimators $\overline{X}$ & S with $\overline{x}$ & s computed from sample:

$$\overline{x} - t^*_{n-1,\alpha/2}\cdot\frac{s}{\sqrt{n}} < \mu < \overline{x} + t^*_{n-1,\alpha/2}\cdot\frac{s}{\sqrt{n}}$$

Small-Sample CI for Normal Population Mean μ

Proposition

Given a normal population with unknown mean μ and std dev σ.
Let $x_1, \ldots, x_n$ be a small sample taken from the population.

*Then the $100(1-\alpha)\%$ **small-sample CI for** μ is*

$$\left(\overline{x} - t^*_{n-1,\alpha/2} \cdot \frac{s}{\sqrt{n}},\ \ \overline{x} + t^*_{n-1,\alpha/2} \cdot \frac{s}{\sqrt{n}}\right)$$

— *OR WRITTEN MORE COMPACTLY* —

$$\overline{x} \pm t^*_{n-1,\alpha/2} \cdot \frac{s}{\sqrt{n}}$$

Textbook Logistics for Section 7.3

- Difference(s) in Notation:

CONCEPT	TEXTBOOK NOTATION	SLIDES/OUTLINE NOTATION
Probability of Event	$P(A)$	$\mathbb{P}(A)$
z Critical Value	$z_{\alpha/2}$	$z^*_{\alpha/2}$
t Critical Value	$t_{\alpha/2,\nu}$	$t^*_{\nu,\alpha/2}$

- Ignore any mention of **one-sided CI's**
- Ignore "A Prediction Interval for a Single Future Value" (pg 299-301)
  - Prediction Intervals (PI's) are useful in some applications.
  - Since there's enough work to be done with CI's, PI's will not be covered.
- Ignore "Tolerance Intervals" section (pg 300-301)
- Ignore "Intervals Based on Nonnormal Population Distributions" (pg 302)
  - Bootstrap CI's are very effective for nonnormal populations.

Fin.