STAT 100 Lecture 19: Student's t Distribution

Nate Strawn

November 14

Last Time...

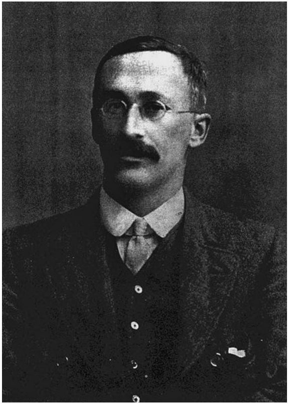
1. Inferences about population proportion

Today's Agenda

1. Student's t-Distribution
2. Confidence Intervals for μ when X is Normal
3. Examples.

Historical Note

In 1908, William Sealy Gosset published “The probable error of a mean”[1] under the alias Student. He did so because he was contracted under Arthur Guiness and Son, which prohibited its researchers from publishing.



[1] Student [William Sealy Gosset]. "The probable error of a mean". Biometrika 6 (1): 1-25. 1908.

Sampling Distributions Derived from $N(\mu, \sigma)$

Let X_1, X_2,...,X_n be drawn from $N(\mu, \sigma)$. We have learned that

$$Z = \frac{\overline{X} - \mu}{\sigma / \sqrt{n}}$$

is then distributed like the Standard Normal. But, it is rarely the case that we actually know σ. Instead, we form the statistic

$$T = \frac{\overline{X} - \mu}{S / \sqrt{n}},$$

where $S = \sqrt{\sum (X_i - \overline{X})^2 / (n-1)}$ is the sample standard deviation. How is this T distributed?

Definition

If $X_1, \ldots, X_n$ is a random sample from $N(\mu, \sigma)$,

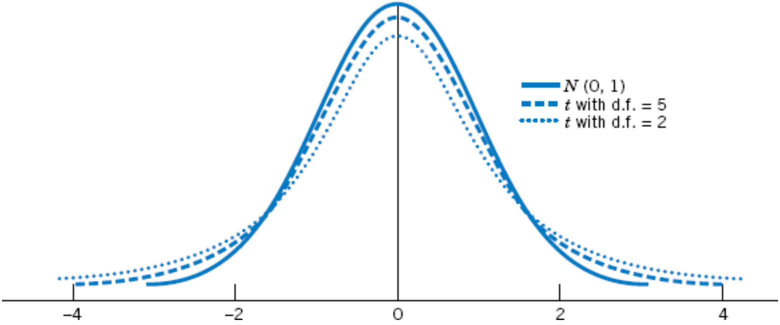
$$\overline{X} = \frac{1}{n} \sum X_i, \text{ and } S^2 = \frac{\sum (X_i - \overline{X})^2}{n-1},$$

then the distribution of

$$T = \frac{\overline{X} - \mu}{S/\sqrt{n}}$$

is called Student's t-distribution with $n-1$ degrees of freedom, and we write $d.f. = n - 1$.

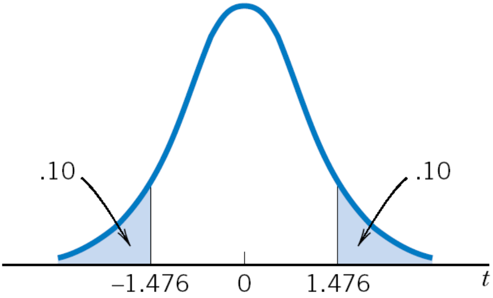
Comparision of $N(0, 1)$ and t



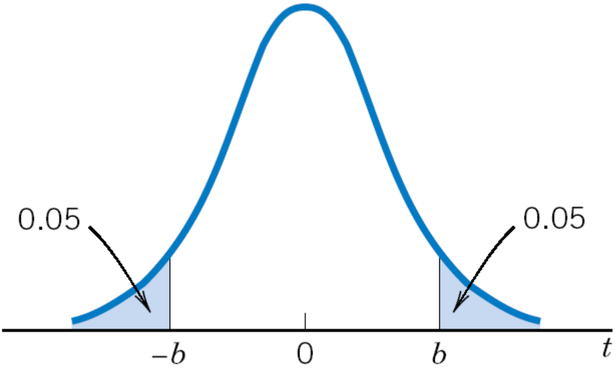
Below, we have found the value $t_{0.10} = 1.476$ for the t-distribution with $d.f. = 5$.

Percentage Points of t distributions

d.f. \ α	$\cdots$	.10
.	.	.
.	.	.
.	.	.
5	$\cdots$	1.476



Example: Find t_α if $\alpha = 0.05$ and $d.f. = 25$



Example: Calculating Probabilities

Example

- *Find* $P(T < 1.33)$ *if* $d.f. = 18$.
- *Find* $P(-0.687 < T < 2.613)$ *if* $d.f. = 20$.
- *Asses the probability that* $|T| > 2$ *if* $d.f. = 13$.

Confidence Intervals for μ when X is Normal

Have a normal population X with mean μ and standard deviation σ. We know that $\overline{X} \sim N(\mu, \sigma/\sqrt{n})$, and thus

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1).$$

A $100(1-\alpha)\%$ confidence interval for μ is bounded by $\overline{x} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$.

In reality, we only have S, so we get that

$$T = \frac{\overline{X} - \mu}{S/\sqrt{n}}$$

has the t-distribution with $n-1$ degrees of freedom. We can still get a confidence interval for μ if we replace $z_{\alpha/2}$with $t_{\alpha/2}$!

Confidence Intervals for μ when X is Normal

A $100(1 - \alpha)\%$ Confidence Interval for a Normal Population Mean

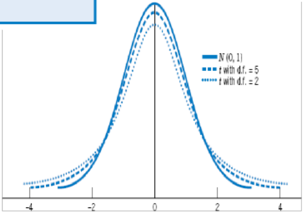
$$\left(\overline{X} - t_{\alpha/2}\frac{S}{\sqrt{n}}, \quad \overline{X} + t_{\alpha/2}\frac{S}{\sqrt{n}}\right)$$

where $t_{\alpha/2}$ is the upper $\alpha/2$ point of the t distribution with d.f. $= n - 1$.

We can now construct confidence intervals for any n!

Before, we only constructed confidence intervals for large n.

for large n, T is like Z



Large Sample Confidence Interval for μ

When n is large, a $100(1 - \alpha)\%$ confidence interval for μ is given by

$$\left(\overline{X} - z_{\alpha/2}\frac{S}{\sqrt{n}}, \quad \overline{X} + z_{\alpha/2}\frac{S}{\sqrt{n}}\right)$$

where S is the sample standard deviation.

Example: Confidence Interval for μ

Example

Suppose X is normally distributed, and one acquires $n = 25$ samples with sample mean $\overline{x} = 50$ and sample standad devation $S = 5$.

1. *Construct a 99% confidence interval for μ.*
2. *If another sample of $n = 25$ samples is taken, does the length of the interval change?*

For Next Time

- Read Section 8.5 from Johnson and Bhattacharyya
- Group Problems:

Group	1	2	3	4	5
Problem	9.1	9.3	9.5	9.7	9.9
Group	6	7	8	9	10
Problem	9.1	9.3	9.5	9.7	9.9