Set 23 - The T Distribution

March 21, 2023 1:29 PM

Stat 260 Lecture Notes

Set 23 - The T Distribution

Properties of T Random Variables:

standardized version

- There are infinitely many T random variables, each identified by the parameter ν, called the *degrees of freedom* (d.f.). The parameter ν is always a positive integer. The notation T_ν indicates that T is a random variable with ν degrees of freedom.

ν nu

$T_7 \to$ T distribution df = 7

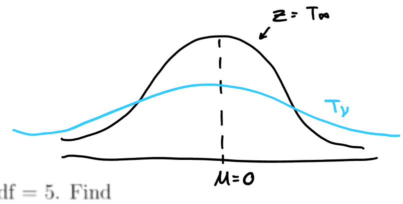
- The T random variable is continuous.
- The graph of the pdf for a T random variable is a symmetric, bell-shaped curve centered at $\mu = 0$.
- The parameter ν dictates the shape of the T pdf. As ν increases, the variance of the T_ν random variable decreases. (So a higher degrees of freedom results in a more compact bell curve.)
- As ν increases, the T pdf curve approaches the standard normal Z distribution.
  Actually, $T_\infty = Z$. In general, the pdf of T_ν is flatter and wider than the pdf of Z.

↑ df, more compact shape (approaches perfect normal bell curve)

↓ df, more spread out

Normal is perfect bell cure

T is flatter/wider bell curve



T distribution tables

↳ kind of backwards

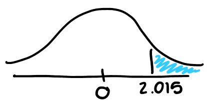
→ getting > probabilities

Example 1: Say df = 5. Find

(a) $P(T_5 > 2.015)$ = 0.05

(b) $P(T_5 < 2.015)$ = 0.95

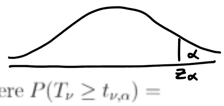
complement of 0.05



1

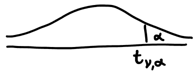
Notation: $t_{\nu,\alpha}$

Before we used z_α as the point where $P(Z \geq z_\alpha) = \alpha$.

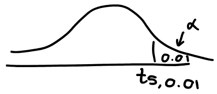


Here we use $t_{\nu,\alpha}$, or just t_α as shorthand, as the point where $P(T_\nu \geq t_{\nu,\alpha}) = \alpha$.

Use the critical value on the T table for this.



Example 2: Say df = 5. Find



df = ν = 5

(a) $t_{5,0.01}$ = 3.365

(b) $t_{5,0.20}$ = 0.920

(c) $t_{5,0.05}$ = 2.015

Example 3: Look at $t_{\nu,0.05}$.

ν	5	10	20	120	∞
$t_{\nu,0.05}$	2.015	1.812	1.725	1.658	1.645

$T_{\infty,0.05} = Z_{0.05}$

Will accept Z table and T table values

Example 4: Say $\nu = 21$. Find $P(T_{21} > 1.90)$.

On the table: $P(T_{21} > 1.721) = 0.05$

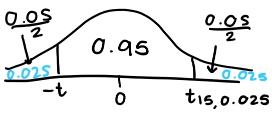
$P(T_{21} > 2.080) = 0.025$

We can't find exact probability, but we can say

$$\boxed{0.025 < P(T_{21} > 1.90) < 0.05}$$

For T distribution, give a range of probabilities for our answer.

Example 5: Find the point t such that $P(-t < T_{15} < t) = 0.95$.



$t_{15,0.025} = 2.131$

We have $P(T_{15} > 2.131) = 0.025$

$P(T_{15} < -2.131) = 0.025$

So if we want $P(T_\nu < -\#)$ we find $P(T_\nu > +\#)$

and $P(T_\nu < -\#) = P(T_\nu > +\#$

Confidence Intervals:

Recall: s is a good point estimate of σ.

Also recall: We have that $\overline{X}$ is normally (or approximately normally) distributed whenever:

summary

- $X_1, X_2, \ldots, X_n$ from a normal distribution and we know σ, or
- $X_1, X_2, \ldots, X_n$ from any distribution and n is big ($n \geq 30$) and we know σ, or
- $X_1, X_2, \ldots, X_n$ from any distribution and n is big ($n \geq 30$) and we don't know σ (so we use the estimate s instead)

And we standardize as $Z = \dfrac{\overline{X} - \mu}{\sigma/\sqrt{n}}$ or $Z = \dfrac{\overline{X} - \mu}{s/\sqrt{n}}$.

no restriction on n

Rule: If $X_1, X_2, \ldots, X_n$ is a random sample of size n from a normal distribution with mean μ and unknown standard deviation (so we would use the estimate s). Then $\dfrac{\overline{X} - \mu}{s/\sqrt{n}}$ follows a T distribution with $n - 1$ degrees of freedom.

overlaps with 3rd point

Note: We can use this new rule if n is large or small. However, in Stat 260 we will follow the convention that we use the normal distribution here when $n \geq 30$ (as we saw in Set 21) and the T distribution whenever $n < 30$. Use this convention on tests and assignments. Our textbook, and other places on the internet, may follow the rule that they always use the T distribution whenever σ is unknown and the estimate s is being used instead.

When $\dfrac{\overline{X}-\mu}{s/\sqrt{n}}$ follows a T distribution we can find a $(1-\alpha)\%$ CI for μ with:

$$[L, U] = \left[\overline{x} - t_{\nu,\alpha/2}\cdot\frac{s}{\sqrt{n}},\ \ \overline{x} + t_{\nu,\alpha/2}\cdot\frac{s}{\sqrt{n}}\right]$$

$$= \overline{x} \pm t_{\nu,\alpha/2}\cdot\frac{s}{\sqrt{n}}$$

estimate ± (c.v.) (e.s.e.)

pattern on formula sheet

degrees of freedom

where $\nu = n-1$.

Example 6: Suppose we have the following random sample of observations which come from a normal distribution:

↳ Xs are normal

56, 44, 62, 36, 53, 39, 50

$\bar{x} = 48.57$

$s = 9.38$

sx button on calculator

Compute a 95% confidence interval for the population mean μ.

n = 7 (n is small)

Xs are normal

do we know σ? No, use s from sample ⇒ $\bar{X}$ is T distributed

Use Z or T?

$\bar{x} \pm t_{\nu,\alpha/2}\cdot\frac{s}{\sqrt{n}} = 48.57 \pm 2.447\cdot\frac{9.38}{\sqrt{7}} = [39.89, 57.25]$

$t_{\nu,\frac{\alpha}{2}} = t_{6,0.025} = 2.447$

$df = \nu = n-1 = 7-1 = 6$

When do we use Z versus T? → for $\bar{x}$