t-Distribution

Suppose that we have a random sample of size n from a Normally distributed population with mean μ and standard deviation σ. The sample mean $\bar{X}$ is then Normally distributed with mean μ and standard deviation $\sigma/\sqrt{n}$.

When σ is **not** known, we estimate it with the sample standard deviation s, and then we estimate the standard deviation of $\bar{X}$ by $s/\sqrt{n}$. – $\bar{X} \sim t(\mu, \frac{s^2}{n})$.

For a sample of size n, the sample standard deviation s is $s = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n-1}}$ or $S = \sqrt{\frac{\sum(X_i - \bar{X})^2}{n-1}}$.

The **degrees of freedom** are the number of values that are free to vary after a sample statistic has been computed. The degrees of freedom ($\boldsymbol{df}$) is found by subtracting 1 from the sample size, i.e., $\boldsymbol{df = n - 1}$.

When s is used, especially when the sample size is small, critical values greater than the values for $z_{\alpha/2}$ are used in confidence intervals in order to keep the interval at any given level. These values are taken from the *Student t-distribution*, most often called the ***t-distribution***.

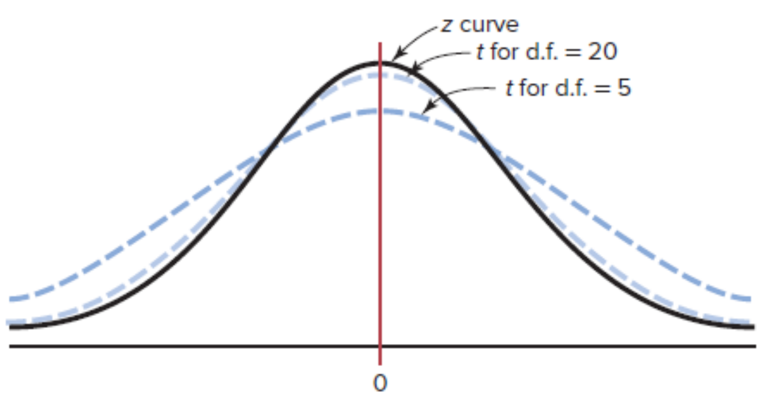
The population from which samples are selected must be normally distributed or the sample size must be 30 or more by CLT.

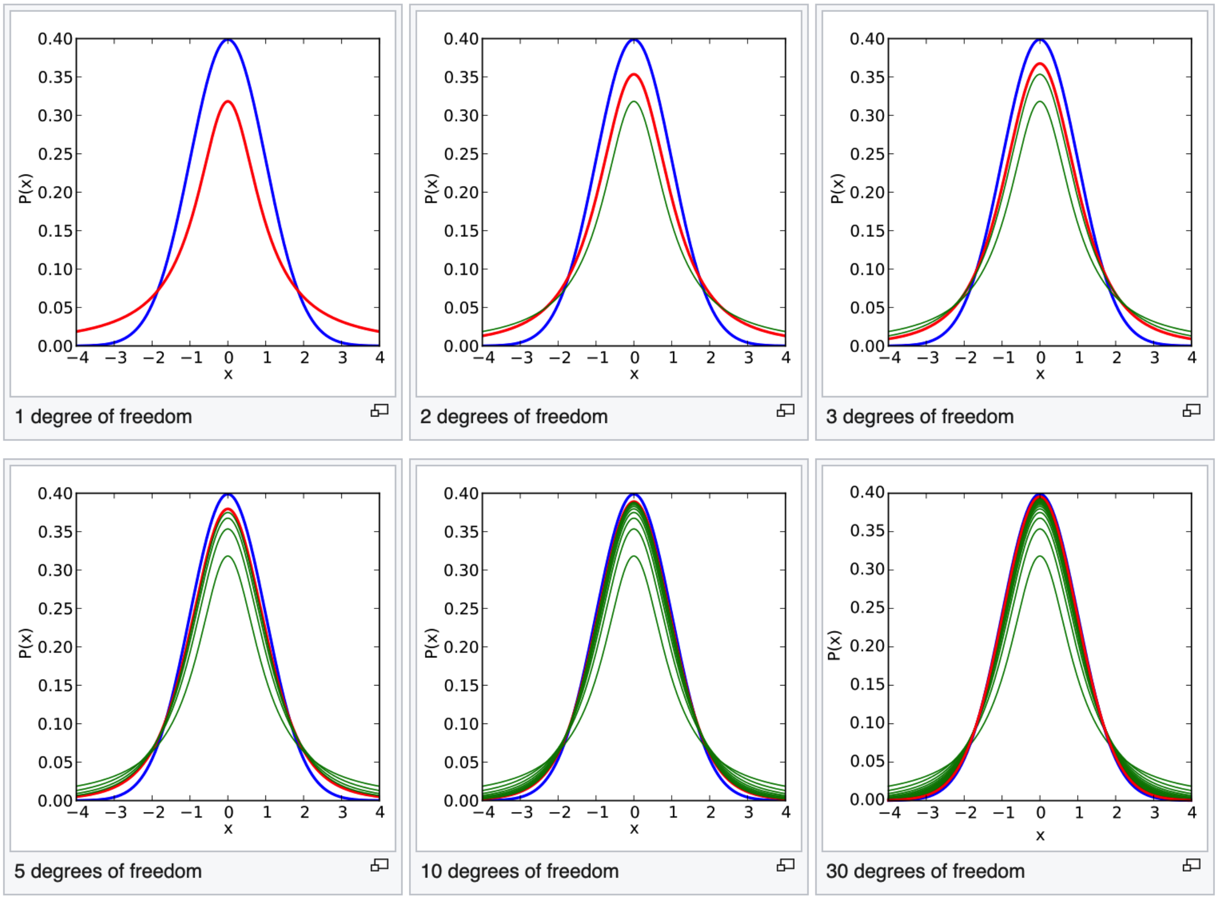
- When σ is known, the sampling distribution is $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$.
- When σ is not known and estimated from the sample standard deviation s, the sampling distribution follows a t distribution $\bar{X} \sim t(\mu, \frac{s^2}{n})$ with degrees of freedom $n - 1$.

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$ is called the **one-sample *t* statistic.**

When n is very large, s is a very good estimate of σ, and the corresponding t distributions are very close to the normal distribution. The t distributions become wider for smaller sample sizes, reflecting the lack of precision in estimating σ from s.

The t distribution is a family of curves based on degrees of freedom, which is related to sample size. As the sample size increases, the t distribution approaches the standard normal distribution.





1 degree of freedom | 2 degrees of freedom | 3 degrees of freedom | 5 degrees of freedom | 10 degrees of freedom | 30 degrees of freedom

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

■ **Note** that the t distribution

1. Is symmetric about ~~zero~~ μ like the standard normal.
2. Has somewhat heavier tails than the standard normal.
3. Depends on a single parameter called the degrees of freedom (df).

■ **Note:** $t_{\alpha,df}$ = the (1- α) percentile of a t-distribution with degrees of freedom (df) such that
$P(t < t_{\alpha,df}) = 1 - \alpha$

■ **Calculator:** 2nd → Vars → #4: invT

$$t_{\alpha,df} = \textbf{invT}(1 - \alpha, df)$$

Example 1: Find:

1. $t_{0.05,20}$ 95th percentile $t = 1.72$ ← invT(0.95, 20)

2. $t_{0.025,24}$ 97.5th percentile $t = 2.06$ ← invT(0.975, 24)

100(1-α)% Confidence Interval for the Mean When σ Is Unknown

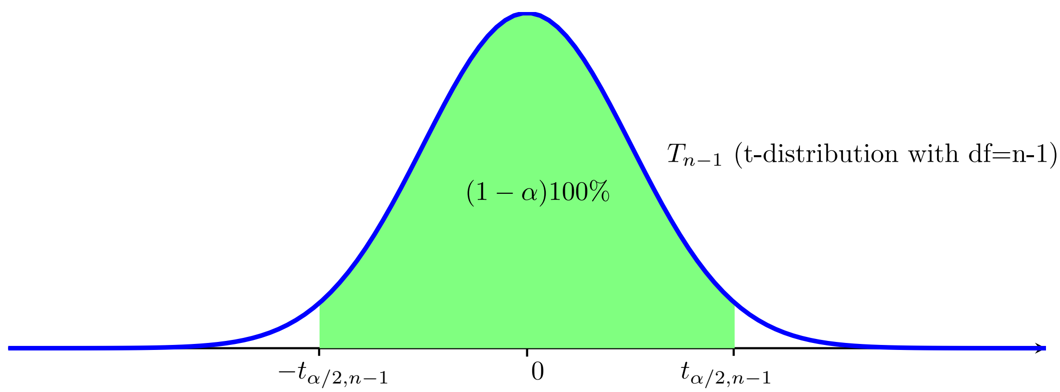
$$\bar{x} - t_{\alpha/2,n-1}\,\frac{s}{\sqrt{n}} < \mu < \bar{x} + t_{\alpha/2,n-1}\,\frac{s}{\sqrt{n}}$$

E

$$\bar{x} \pm t_{\alpha/2,n-1}\,\frac{s}{\sqrt{n}}$$

$$\left(\bar{x} - t_{\alpha/2,n-1}\,\frac{s}{\sqrt{n}},\ \bar{x} + t_{\alpha/2,n-1}\,\frac{s}{\sqrt{n}}\right)$$

The **critical value $t_{\alpha/2,n-1}$** is the $\left(1-\frac{\alpha}{2}\right)$ percentile of a t-distribution with $n-1$ degrees of freedom and the **margin of error** is $t_{\alpha/2,n-1}\,\frac{s}{\sqrt{n}}$.



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

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

■ Calculator

Finding a ***t* Confidence Interval** for the Mean (**Data**)

1. Enter the data into L_1.
2. Press **STAT** and move the cursor to TESTS.
3. Press **8** for TInterval.
4. Move the cursor to Data and press **ENTER**.
5. Type in the appropriate values.
6. Move the cursor to Calculate and press **ENTER**.

STAT → edit to edit lists

STAT → Calc → 1Var Stat

Finding a *t* **Confidence Interval** for the Mean (**Statistics**)

1. Press **STAT** and move the cursor to TESTS.
2. Press **8** for TInterval.
3. Move the cursor to Stats and press **ENTER**.
4. Type in the appropriate values.
5. Move the cursor to Calculate and press **ENTER**.

We assume that a population is normally distributed for a sample size less than 30.

lots is yet CLT

Example 2. Founded in 1998, Telephia provides a wide variety of information on cellular phone use. In 2006, Telephia reported that, on average, United States (U.S.) subscribers with third generation technology (3G) phones spent an average of 8.3 hours per month listening to full-track music on their cell phones. Determine a 95% confidence interval for the U.S. average and draw the following random sample of size 8 from the U.S. population of 3G subscribers:

7 9 1 6 13 10 3 5

$n = 8,\ df = 7$

$$\bar{X} = \frac{7+9+1+6+13+10+3+5}{8} = 6.75$$

$$S = \sqrt{\frac{\sum (X_i - \bar{X})^2}{n-1}} = \sqrt{\frac{(7-6.5)^2 \ldots + (10-6.5)^2 \ldots}{7}} = 3.88$$

$\alpha = 0.05$

$\frac{\alpha}{2} = 0.025$

$t_{0.025, 7} = 2.364$

TInterval $(\bar{x} = 6.75,\ s = 3.88,\ n = 8)$

$$\Rightarrow 3.5062 < \mu < 9.9938$$

Example 3. You randomly choose 16 unfurnished one-bedroom apartments from a large number of advertisements in your local newspaper. You calculate that their mean monthly rent is \$613 and their standard deviation is \$96. Construct a 95% confidence interval for the mean monthly rent of all advertised one-bedroom apartments.

$n = 16,\ df = 15$

$\bar{X} = 613$

$s = 96$

$\alpha = 0.05 \Rightarrow \frac{\alpha}{2} = 0.025$

$t_{0.025, 15} = 2.13$

$$\bar{X} - t_{\frac{\alpha}{2}, df} \frac{S}{\sqrt{n}} < \mu < \bar{X} + t_{\frac{\alpha}{2}, df} \frac{S}{\sqrt{n}}$$

$$613 - 2.13 \cdot \frac{96}{\sqrt{16}} < \mu < 613 + 2.13 \cdot \frac{96}{\sqrt{16}}$$

$$561.85 < \mu < 664.15$$

$n=16,\ df=15$

Example 4. A sample of 16 wolf dens in the southwestern US found a mean of 5.6 pups with s=1.8. Find a 99% confidence interval for the population mean if the number of pups is normal.

$\alpha = 0.01$

$\frac{\alpha}{2} = 0.005$

$t_{0.005,15} = 2.946$

T Interval
$\bar{x} = 5.6$
$s = 1.8$
$n = 16$

$\Rightarrow \boxed{4.274 < \mu < 6.926}$

$n=10,\ df=9$

Example 5. A random sample of 10 children found that their average growth for the first year was 9.8 inches. Assume the variable is normally distributed and the sample standard deviation is 0.96 inch. Find the 95% confidence interval of the population mean for growth during the first year.

$\alpha = 0.05$

$\frac{\alpha}{2} = 0.025$

$n=10,\ df=9$

$t_{0.025,9} = 2.262$

T Interval
$\bar{x} = 9.8$
$s = 0.96$
$n = 10$

$\Rightarrow \boxed{9.1133 < \mu < 10.487}$