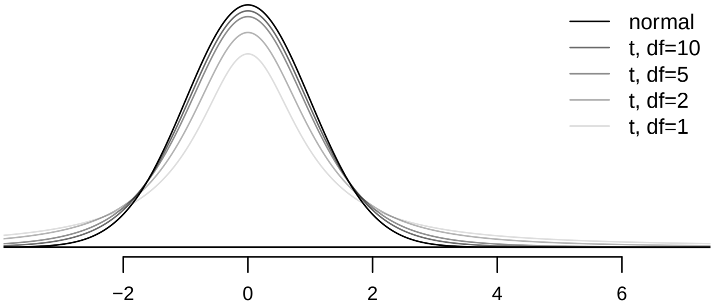
The T distribution

- With large samples, we can assume $\sigma \approx s$.
- With small samples, there is usually uncertainty about σ. We have not accounted for this additional uncertainty until now.
- With small n, sampling distributions (from approximately symmetric populations) follow the T distribution, which is like the Z distribution, but it has fatter tails.
- As n grows, the T distribution approaches the Z distribution.



Using the t table

- If $df = 8$ and $P(T < t) = 0.95$, what is t?
- If $df = 10$ and $P(|T| > t) = 0.005$, what is t?
- If $df = 5$ what is $P(T > 2.76)$?
- If $df = 5$ what is $P(T < -2.76)$?

Example confidence interval of single sample with small n, unknown σ.

A random sample of size $n = 10$ was collected from a population which is believed to be approximately symmetric. The sample has a mean $\bar{x} = 135.7$ and standard deviation $s = 24.6$. Find the confidence interval with a confidence level $\gamma = 0.95$.

- The formulas:

$$\begin{aligned} SE &= \frac{s}{\sqrt{n}} \\ df &= n - 1 \\ P(|T| < t^\star) &= \gamma \\ CI &= \bar{x} \pm t^\star SE \end{aligned}$$

A random sample of size $n = 10$ was collected from a population which is believed to be approximately symmetric. The sample has a mean $\bar{x} = 135.7$ and standard deviation $s = 24.6$. Find the confidence interval with a confidence level $\gamma = 0.95$.

- Calculate the standard error (same way as before).

$$SE = \frac{24.6}{\sqrt{10}} = 7.78$$

- Calculate the degrees of freedom.

$$df = 10 - 1 = 9$$

- Determine $t^\star$ such that $P(|T| < t^\star) = 0.95$. We use the T table.

$$t^\star = 2.26$$

- Calculate the confidence interval.

$$CI = 135.7 \pm (2.26)(7.78)$$

$$CI = (118.1,\ 153.3)$$

Practice: confidence interval, single small sample, unknown σ

A random sample of size $n = 15$ was collected from a population which is believed to be approximately symmetric. The sample has a mean $\bar{x} = 11.1$ and standard deviation $s = 2.3$. Find the confidence interval with a confidence level $\gamma = 0.99$.

Working backwards: confidence intervals

A 90% confidence interval for a population mean, μ, is given as (43.84, 55.92). This confidence interval is based on a simple random sample of 12 observations. Calculate the sample mean and standard deviation. Assume that all conditions necessary for inference are satisfied. Use the T distribution in any calculations.

- The formulas:

$$\begin{aligned} SE &= \frac{s}{\sqrt{n}} \\ df &= n - 1 \\ P(|T| < t^\star) &= \gamma \\ CI &= \bar{x} \pm t^\star SE \end{aligned}$$

The sample mean is the average of the bounds of the CI.

$$\bar{x} = \frac{43.84 + 55.92}{2} = 49.88$$

The margin of error is half the difference between the bounds. It is also the distance from $\bar{x}$ to either bound.

$$ME = \frac{55.92 - 43.84}{2} = 55.92 - 49.88 = 6.04$$

That margin of error is the product of $t^\star$ and SE. We find the $t^\star$ when $P(|T| < t^\star) = 0.9$ and $df = n - 1 = 11$.

$$t^\star = 1.8$$

We can calculate SE.

$$ME = t^\star SE$$

$$6.04 = (1.8)SE$$

$$SE = \frac{6.04}{1.8} = 3.3564258$$

continued on next frame...

$$SE = \frac{6.04}{1.8} = 3.3564258$$

We can now calculate the sample standard deviation.

$$SE = \frac{s}{\sqrt{n}}$$

$$3.3564258 = \frac{s}{\sqrt{12}}$$

$$s = (3.356)\sqrt{12} = 11.627$$

Thus, the sample mean is 49.88 and the sample standard deviation is 11.63.

Practice: working backwards with confidence intervals

A 95% confidence interval for a population mean, μ, is given as (88.09, 109.15). This confidence interval is based on a simple random sample of 9 observations. Calculate the sample mean and standard deviation. Assume that all conditions necessary for inference are satisfied. Use the T distribution in any calculations.

Hypothesis testing with single small sample

You will perform a single-sample t test of the null hypothesis claiming $\mu = 22$. Before collecting the sample, you decide to use a two-tail test with a significance level $\alpha = 0.05$. The sample has the following attributes:

$$\begin{aligned} n &= 11 \\ \bar{x} &= 15.49 \\ s &= 8.03 \end{aligned}$$

What is your conclusion?
We state the hypotheses:

$$H_0: \quad \mu = 22$$

$$H_A: \quad \mu \neq 22$$

We estimate the standard error (same way as with z testing).

$$SE = \frac{s}{\sqrt{n}} = \frac{8.03}{\sqrt{11}} = 2.421$$

continued on next slide...

We calculate the t score (same way as with z testing).

$$t = \frac{15.49 - 22}{2.421} = -2.69$$

For the T table, we use the absolute value of t...

$$t = 2.69$$

We determine the degrees of freedom.

$$df = n - 1 = 10$$

We estimate the p-value from the T table.

$$0.02 < p\text{-value} < 0.04$$

We compare the p-value to α.

$$p\text{-value} < \alpha$$

We make our conclusion: we reject the null.

Practice:

You will perform a single-sample t test of the null hypothesis claiming $\mu = 140$. Before collecting the sample, you decide to use a two-tail test with a significance level $\alpha = 0.1$. The sample has the following attributes:

$$\begin{aligned} n &= 3 \\ \bar{x} &= 193.06 \\ s &= 52.22 \end{aligned}$$

What is your conclusion?