Suppose that $X_1, X_2, \ldots, X_n$ is a random sample from any distribution with mean μ and variance σ^2.

Suppose that n is “large”. (n>30)

Suppose that μ and σ^2 are both unknown.

By the Central Limit Theorem we know that the sample mean $\overline{X}$ is approximately normally distributed.

Let's find an approximate $100(1-\alpha)\%$ confidence interval for μ.

$$\text{CLT} \quad \Rightarrow \quad \overline{X} \overset{\text{asymp}}{\sim} N(\mu, \sigma^2/n)$$

$$\Rightarrow \quad \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \to N(0,1)$$

$$\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \overset{\text{approx}}{\sim} N(0,1) \qquad \text{for large n}$$

$$-z_{\alpha/2} < \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} < z_{\alpha/2}$$

Solve for μ “in the middle”.

But this involves σ, which is unknown!

Consider instead using the sample variance.

$$S^2 = \frac{\sum_{i=1}^{n} (X_i - \overline{X})^2}{n-1} \xrightarrow{P} \sigma^2$$

S^2 is approximately σ^2 in some sense.

So,
$$\frac{\overline{X}-\mu}{S/\sqrt{n}} \overset{\text{approx}}{\sim} N(0,1)$$

We can get an approximate $100(1-\alpha)\,\%$ confidence interval for μ is given by solving

$$-z_{\alpha/2} < \frac{\overline{X}-\mu}{S/\sqrt{n}} < z_{\alpha/2}$$

for μ "in the middle".

- $X_1, X_2, \ldots, X_n$ is a random sample from any distribution with mean μ and variance σ^2.
- n is "large". (n>30)
- μ and σ^2 are both unknown.

An approximate $100(1-\alpha)\%$ confidence interval for μ is given by

$$\boxed{\overline{X} \pm z_{\alpha/2}\frac{S}{\sqrt{n}}}$$

Sample Variance

$$S^2 = \frac{\sum_{i=1}^{n} (X_i - \overline{X})^2}{n - 1}$$

$$= \frac{\sum_{i=1}^{n} X_i^2 - \frac{\left(\sum_{i=1}^{n} X_i\right)^2}{n}}{n - 1}$$

Suppose that $X_1, X_2, \ldots, X_n$ is a random sample from any distribution with mean μ and variance σ^2.

Suppose that n is "small". ($n \leq 30$)

Suppose that μ and σ^2 are both unknown.

A $100(1-\alpha)\%$ CI or approximate CI for μ?

No!

Suppose that $X_1, X_2, \ldots, X_n$ is a random sample from the normal distribution with mean μ and variance σ^2.

Suppose that n is "small". ($n \leq 30$)

Suppose that μ and σ^2 are both unknown.

A $100(1-\alpha)\,\%$ CI or approximate CI for μ?

- $\overline{X}$ has a normal distribution with mean μ and variance σ^2/n.
- σ^2 and hence σ^2/n, are unknown.
- Want to use S^2 in place of σ^2.
- Small sample means the approximation is not good!
- What should we do?!?

What is the distribution of $\dfrac{\overline{X} - \mu}{S/\sqrt{n}}$ **?**

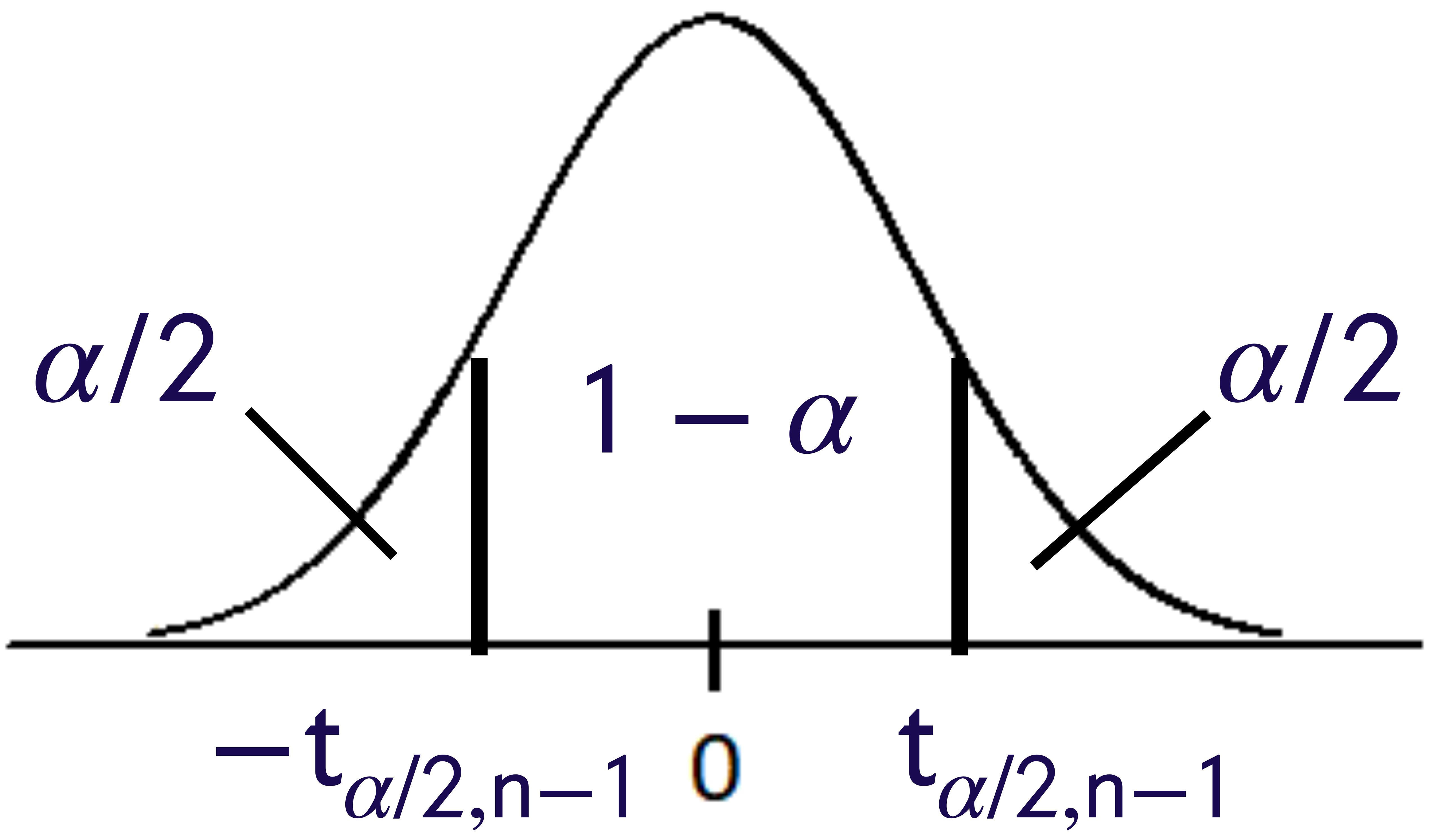
$$\frac{\overline{X} - \mu}{S/\sqrt{n}} = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \cdot \frac{\sigma}{S}$$

$$= \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \Bigg/ \sqrt{\frac{S^2}{\sigma^2}}$$

$$= \underbrace{\frac{\overline{X} - \mu}{\sigma/\sqrt{n}}}_{N(0,1)} \Bigg/ \sqrt{\frac{\underbrace{\frac{(n-1)S^2}{\sigma^2}}_{\chi^2(n-1)}}{n-1}}$$

What is the distribution of $\dfrac{\overline{X} - \mu}{S/\sqrt{n}}$ **?**

$$\frac{\overline{X} - \mu}{S/\sqrt{n}} = \frac{Z}{\sqrt{W/(n-1)}} \sim t(n-1)$$

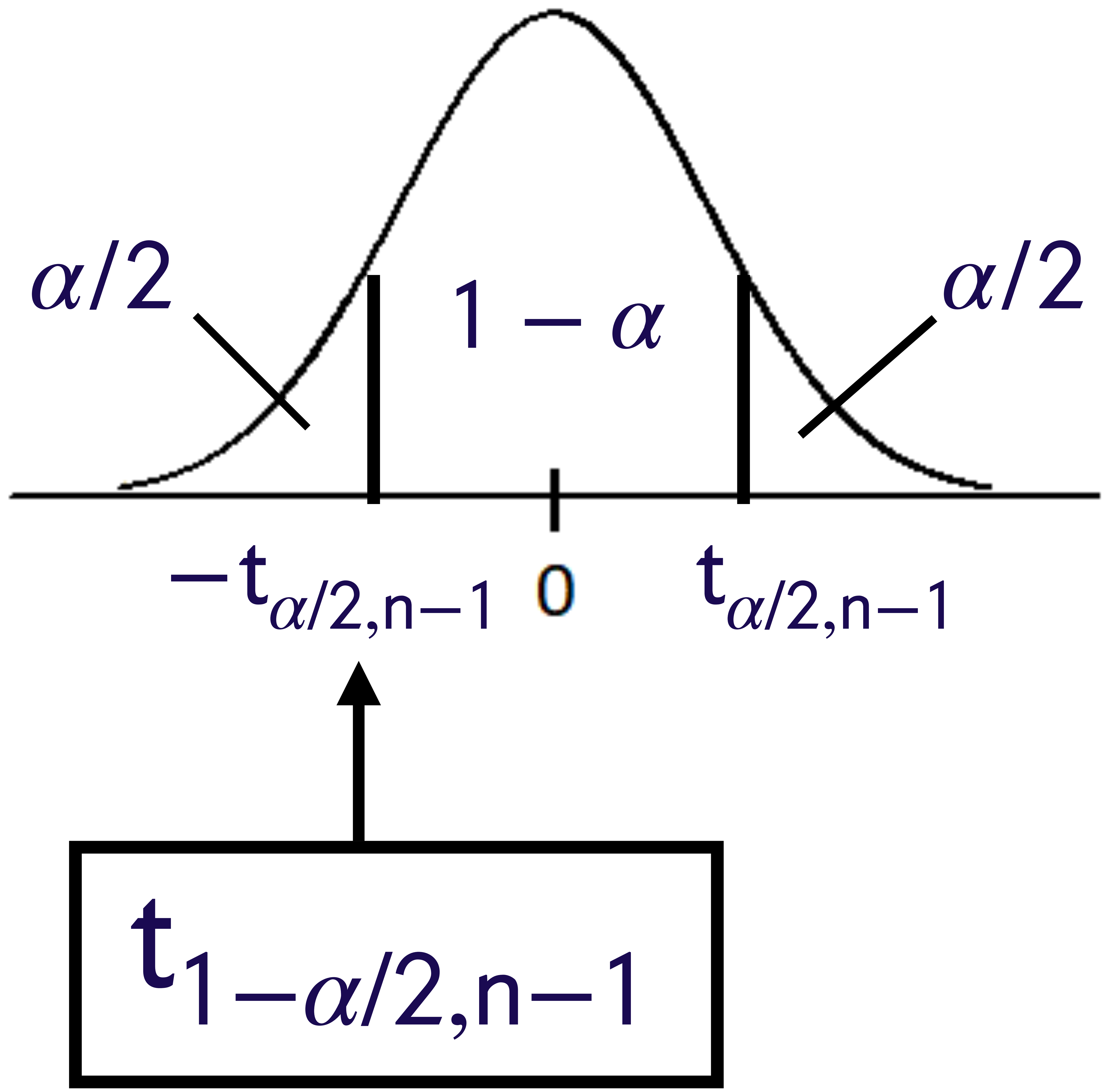


$\boxed{\begin{array}{c} t(n-1) \\ \textbf{pdf} \end{array}}$

$$-t_{\alpha/2,n-1} < \frac{\overline{X} - \mu}{S/\sqrt{n}} < t_{\alpha/2,n-1}$$

$$\Downarrow$$

$$\boxed{\overline{X} \pm t_{\alpha/2,n-1} \frac{S}{\sqrt{n}}}$$



α/2
1 − α
α/2
−t α/2,n−1
0
t α/2,n−1
t 1−α/2,n−1

Suppose that $X_1, X_2, \ldots, X_n$ is a random sample from the normal distribution with mean μ and variance σ^2.

Suppose that n is "small". ($n \leq 30$)

Suppose that μ and σ^2 are both unknown.

A $100(1-\alpha)\%$ CI or approximate CI for μ?
Is given by

$$\boxed{\overline{X} \pm t_{\alpha/2,n-1} \frac{S}{\sqrt{n}}}$$

Example

A small study is being conducted to test a new sensor for a continuous glucose monitoring system. Based on previous studies, the lifetime of the sensors, in days, is expected to be normally distributed.

A random sample of 20 patients were fitted with the new sensor. On average, it took 187 days for the sensors to wear out . The sample variance of the sensor lifetimes was 16.2 days.

Find a 95% confidence interval for the true sensor mean lifetime.

Example

$$n = 20, \ \bar{x} = 187, \ s^2 = 16.2, \ \alpha = 0.05$$

In R: qt(0.975,19) = 2.093024

$$\bar{x} \pm t_{\alpha/2, n-1} \frac{s}{\sqrt{n}} \qquad 187 \pm 2.093 \frac{\sqrt{16.2}}{\sqrt{20}}$$

The 95% confidence interval for μ is

(185.11, 188.88).