Let $X_1, X_2, \ldots, X_n$ be a random sample from the normal distribution with mean μ and unknown variance σ^2.

Consider testing the simple versus simple hypotheses

$$H_0 : \mu = \mu_0 \qquad H_1 : \mu < \mu_0$$

where μ_0 is fixed and known.

Reject H_0, in favor of H_1, if

$$\overline{X} < \mu_0 + z_{1-\alpha}\frac{\sigma}{\sqrt{n}}$$

unknown!

This is a useless test!

It was based on the fact that

$$\overline{X} \sim N(\mu, \sigma^2/n)$$

and

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

What is we use the sample standard deviation $S = \sqrt{S^2}$ in place of σ?

Let $X_1, X_2, \ldots, X_n$ be a random sample from the normal distribution with mean μ and unknown variance σ^2.

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

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

Let $X_1, X_2, \ldots, X_n$ be a random sample from the normal distribution with mean μ and <u>unknown</u> variance σ^2.

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

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

For the normal distribution, $\overline{X}$ and S^2 are independent.

Thus,

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

Back to the hypothesis test…

Let $X_1, X_2, \ldots, X_n$ be a random sample from the normal distribution with mean μ and <u>unknown</u> variance σ^2.

Consider testing the simple versus simple hypotheses

$$H_0 : \mu = \mu_0 \qquad H_1 : \mu < \mu_0$$

where μ_0 is fixed and known.

$$H_0 : \mu = \mu_0$$
$$H_1 : \mu < \mu_0$$

Step One:

Choose an estimator for μ.

$$\hat{\mu} = \overline{X}$$

Step Two:

Give the "form" of the test.

Reject H_0, in favor of H_1 if $\overline{X} < c$,
where c is to be determined.

Et cetera!

$$H_0 : \mu = \mu_0$$

$$H_1 : \mu < \mu_0$$

Step Three:

Find c.

$$\alpha = \max_{\mu=\mu_0} \mathrm{P}(\text{Type I Error})$$

$$= \max_{\mu=\mu_0} \mathrm{P}(\text{Reject } H_0; \mu)$$

$$= \mathrm{P}(\text{Reject } H_0; \mu_0)$$

$$= \mathrm{P}(\overline{X} < c; \mu_0)$$

$$H_0 : \mu = \mu_0$$
$$H_1 : \mu < \mu_0$$

Step Three:

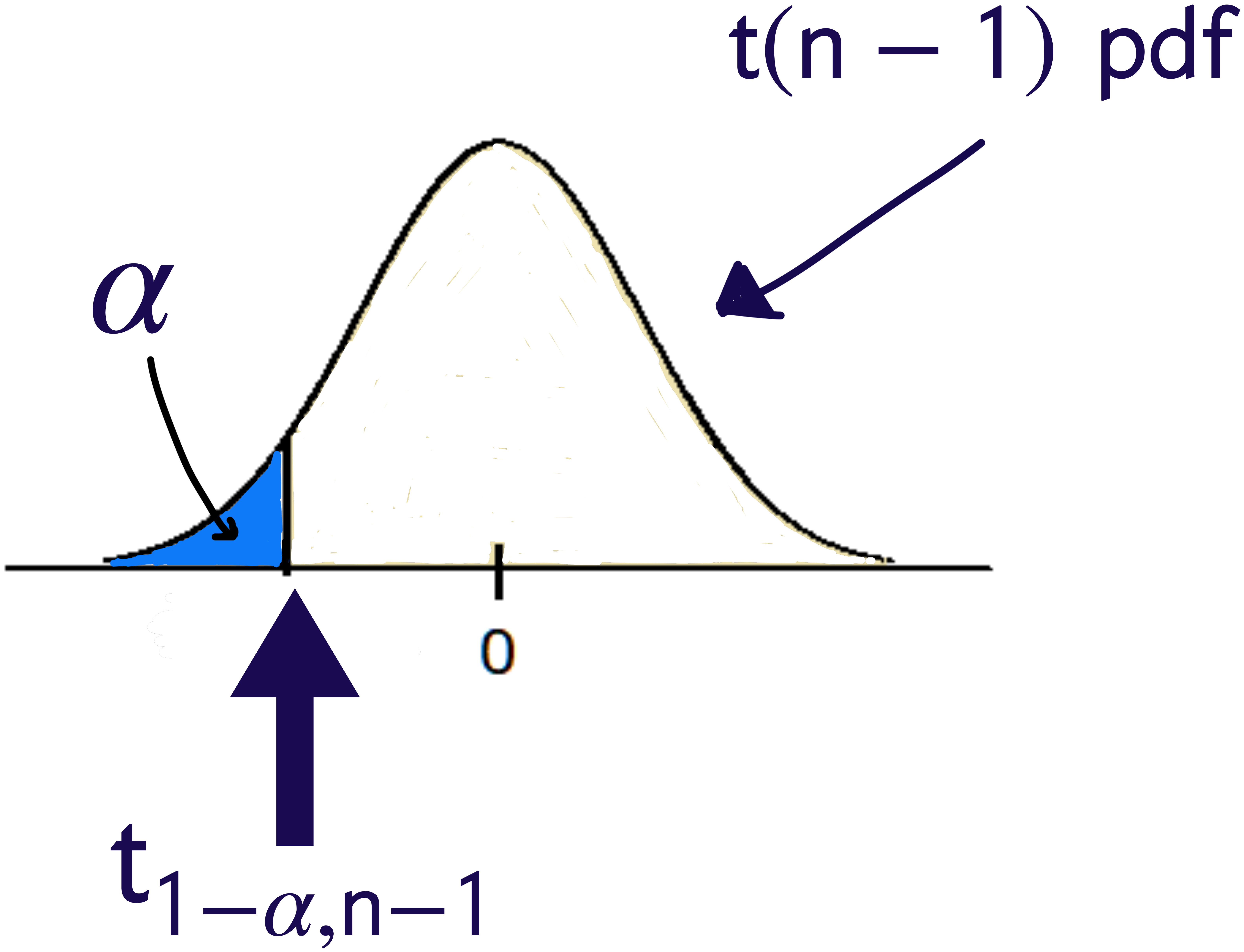
Find c.

$$\alpha = P(\overline{X} < c;\ \mu_0)$$

$$= P\left(\underbrace{\frac{\overline{X} - \mu_0}{S/\sqrt{n}}}_{t(n-1)} < \frac{c - \mu_0}{S/\sqrt{n}};\ \mu_0 \right)$$

$$= P\left(T < \frac{c - \mu_0}{S/\sqrt{n}};\ \mu_0 \right)$$

where $T \sim t(n-1)$


t(n − 1) pdf
α
0
$t_{1-\alpha,n-1}$

Step Three:

$$H_0 : \mu = \mu_0$$
$$H_1 : \mu < \mu_0$$

Find c.

$$\alpha = P\left(T < \frac{c - \mu_0}{S/\sqrt{n}}; \mu_0 \right)$$

$$\Rightarrow \quad \frac{c - \mu_0}{S/\sqrt{n}} = t_{1-\alpha, n-1}$$

$$H_0 : \mu = \mu_0$$
$$H_1 : \mu < \mu_0$$

Step Four:

Conclusion!

Reject H_0, in favor of H_1, if

$$\overline{X} < \mu_0 + t_{1-\alpha,n-1} \frac{S}{\sqrt{n}}$$

Example:

In 2019, the average health care annual premium for a family of 4 in the United States, was reported to be \$6,015.

In a more recent survey, 15 randomly sampled families of 4 reported an average annual health care premium of \$6,033 and a sample variance of \$825.

Can we say that the true average is currently greater than \$6,015 for all families of 4?

Use $\alpha = 0.10$.

Example:

Assume that annual health care premiums are normally distributed.

Let μ be the true average for all families of 4.

Step Zero:

Set up the hypotheses.

$$H_0 : \mu = 6015 \qquad H_1 : \mu > 6015$$

$$H_0 : \mu = 6015$$
$$H_1 : \mu > 6015$$

Step One:

Choose a test statistic.

$$\overline{X}$$

Step Two:

Give the form of the test.

Reject H_0, in favor of H_1, if $\overline{X} > c$
where c is to be determined.

$$H_0 : \mu = 6015$$

$$H_1 : \mu > 6015$$

Step Three:

Find c.

$$\alpha = \max_{\mu=\mu_0} \mathrm{P}(\text{Type I Error})$$

$$= \max_{\mu=6015} \mathrm{P}(\text{Reject } H_0;\ \mu)$$

$$= \mathrm{P}(\text{Reject } H_0;\ \mu = 6015)$$

$$= \mathrm{P}(\overline{X} > c;\ \mu = 6015)$$

$$H_0 : \mu = 6015$$

Step Three:

$$H_1 : \mu > 6015$$

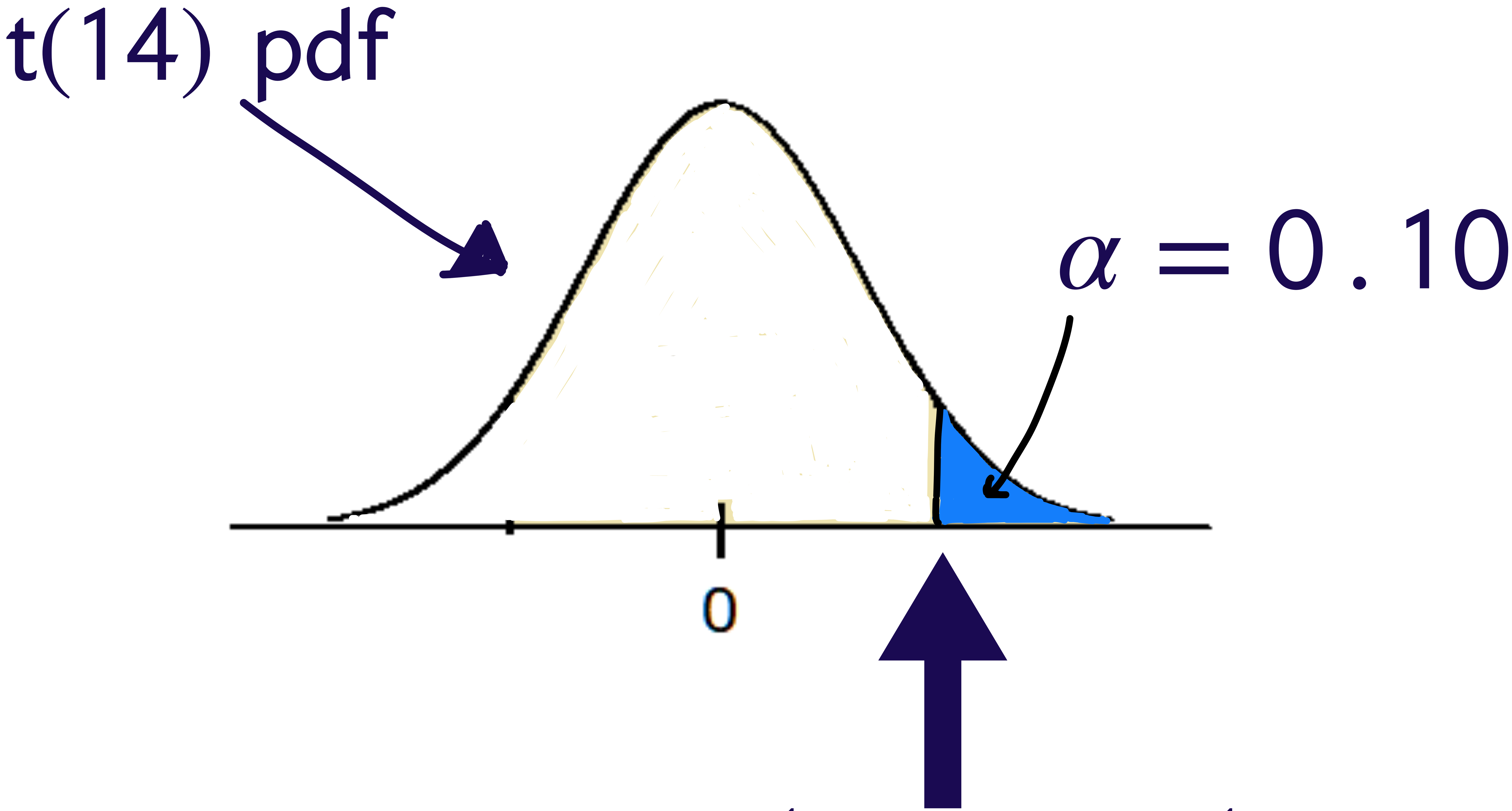
Find c.

$$\alpha = P(\overline{X} > c; \mu = 6015)$$

$$= P\left(\frac{\overline{X} - \mu_0}{S/\sqrt{n}} > \frac{c - 6015}{\sqrt{825}/\sqrt{15}}; \mu = 6015 \right)$$

$$= P\left(T > \frac{c - 6015}{\sqrt{825}/\sqrt{15}} \right)$$

where $t \sim t(14)$



$$t_{\alpha,n-1} = t_{0.10,14}$$
$$= 1.345$$

In R : qt(0.9, 14)

$$H_0 : \mu = 6015$$
$$H_1 : \mu > 6015$$

$$\Rightarrow \quad \frac{c - 6015}{\sqrt{825}/\sqrt{15}} = 1.345$$

$$\Rightarrow \quad c = 6024.98$$

$$H_0 : \mu = 6015$$
$$H_1 : \mu > 6015$$

Step Four:

Conclusion.

Rejection Rule:

Reject H_0, in favor of H_1 if

$$\overline{X} > 6024.98$$

We had $\overline{x} = 6033$ so we reject H_0.

There is sufficient evidence (at level 0.10) in the data to suggest that the true mean annual healthcare premium cost for a family of 4 is greater than $6,015.

$$H_0 : \mu = 6015$$

$$H_1 : \mu > 6015$$

P-Value

$$= P(\overline{X} > 6033; \mu = 6015)$$

$$= P\left(\frac{\overline{X} - \mu}{S/\sqrt{n}} > \frac{6033 - 6015}{\sqrt{825}/\sqrt{15}}; \mu = 6015 \right)$$

$$= P(T > 2.43) \approx 0.015$$

where $T \sim t(14)$

In R : $1 - \text{pt}(2.43, 14)$