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

Derive a hypothesis test of size α for testing

$$H_0 : \mu = \mu_0$$

$$H_1 : \mu \neq \mu_0$$

Derive a hypothesis test of size α for testing

$$H_0 : \mu = \mu_0$$

$$H_1 : \mu \neq \mu_0$$

We will look at the sample mean $\overline{X}$...

... and reject if it is either too high or too low.

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

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 either $\overline{X} < c$ or $\overline{X} > d$ for some c and d to be determined.

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

Step One:

Choose an estimator for μ.

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

Step Two:

Give the "form" of the test.

~~Reject H_0, in favor of H_1 if either $\overline{X} < c$ or $\overline{X} > d$ for some c and d to be determined.~~

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

Step Two:

Easier to make it symmetric!

Reject H_0, in favor of H_1 if either

$$\overline{X} > \mu_0 + c$$

or

$$\overline{X} < \mu_0 - c$$

for some c to be determined.

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

Step Three:

Find c.

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

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

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

$$= \mathrm{P}(\overline{X} < \mu_0 - c \textbf{ or } \overline{X} > \mu_0 + c;\ \mu_0)$$

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

Step Three:

Find c.

$$\alpha = P(\overline{X} < \mu_0 - c \textbf{ or } \overline{X} > \mu_0 + c;\ \mu_0)$$

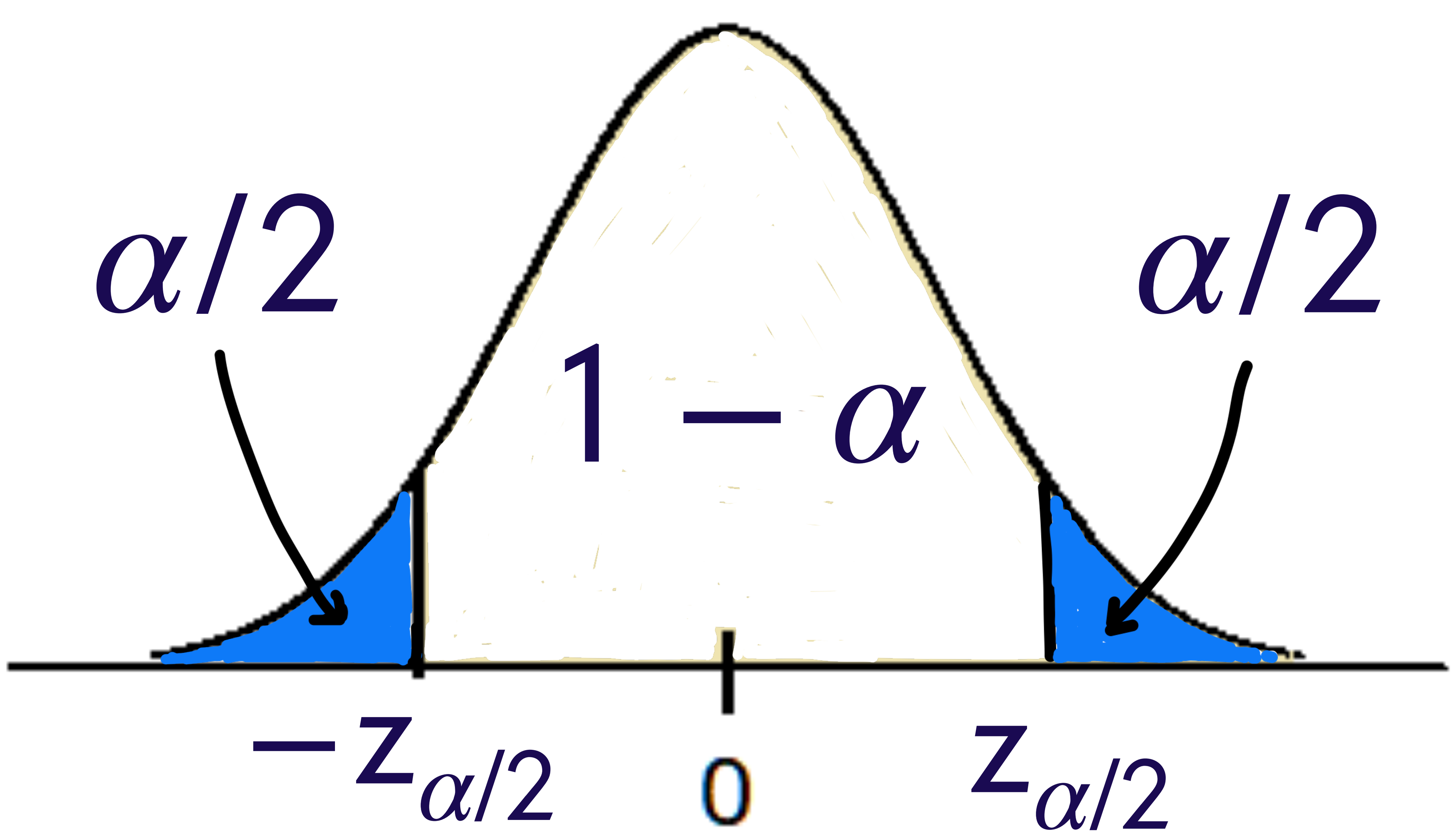
$$= 1 - P(\mu_0 - c \leq \overline{X} \leq \mu_0 + c\,;\mu_0)$$

Subtract μ_0 and divide by $\sigma/\sqrt{n}$.

$$= 1 - P\left(\frac{-c}{\sigma/\sqrt{n}} \leq Z \leq \frac{c}{\sigma/\sqrt{n}}\right)$$

$$\alpha = 1 - P\left(\frac{-c}{\sigma/\sqrt{n}} \le Z \le \frac{c}{\sigma/\sqrt{n}}\right)$$

$$1 - \alpha = P\left(\frac{-c}{\sigma/\sqrt{n}} \le Z \le \frac{c}{\sigma/\sqrt{n}}\right)$$



$$\frac{c}{\sigma/\sqrt{n}} = z_{\alpha/2}$$

$$c = z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

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

Step Four: Conclusion:

Reject H_0, in favor of H_1, if

$$\overline{X} > \mu_0 + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

or

$$\overline{X} < \mu_0 - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

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, 100 randomly sampled families of 4 reported an average annual health care premium of \$6,177.

Can we say that the true average, for all families of 4, is currently different than the sample mean from 2019?

$\sigma = 814$ Use $\alpha = 0.05$.

Assume that annual health care premiums are normally distributed with a standard deviation of \$814.

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

Hypotheses:

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

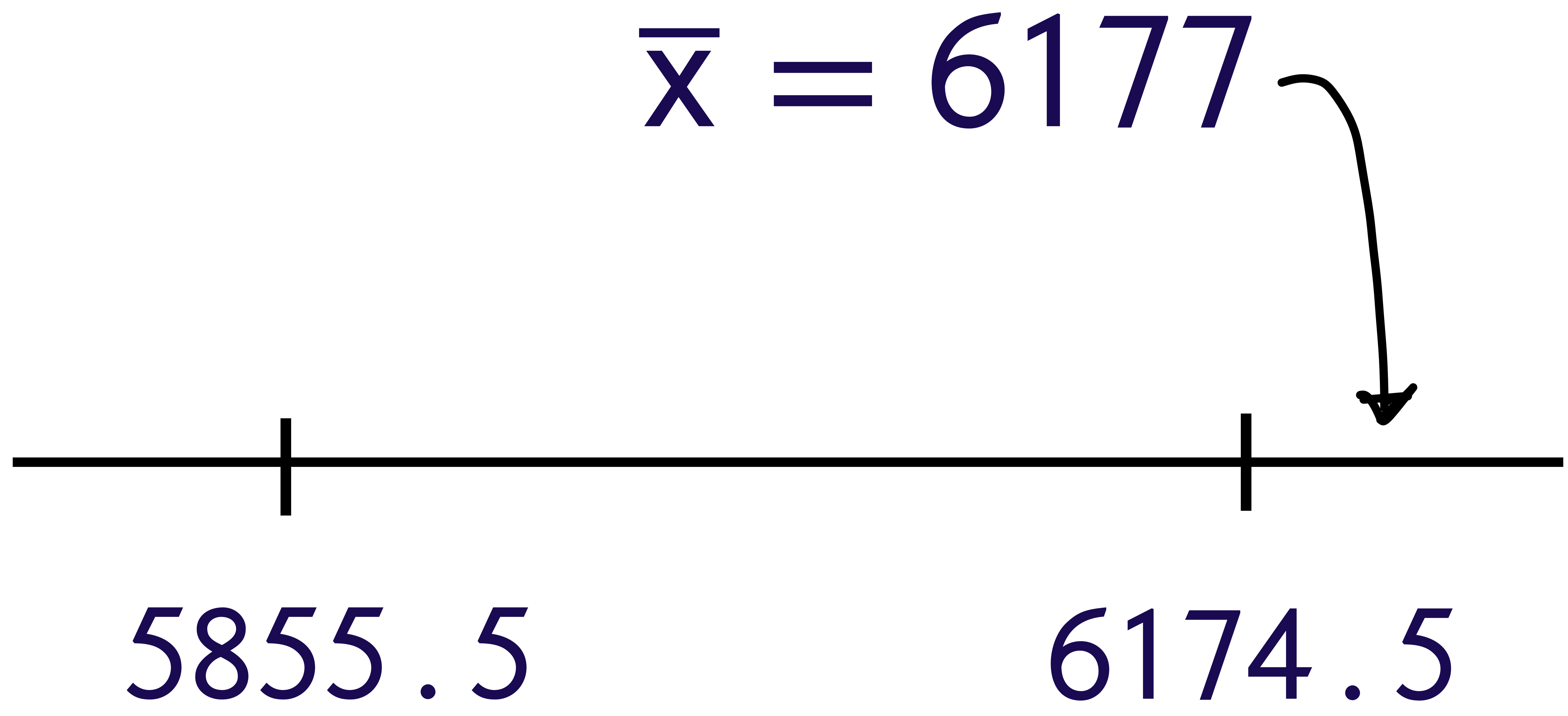
$$\bar{x} = 6177 \quad \sigma = 814 \quad n = 100$$

$$z_{\alpha/2} = z_{0.025} = 1.96$$

In R: qnorm(0.975)

$$6015 + 1.96\frac{814}{\sqrt{100}} = 6174.5$$

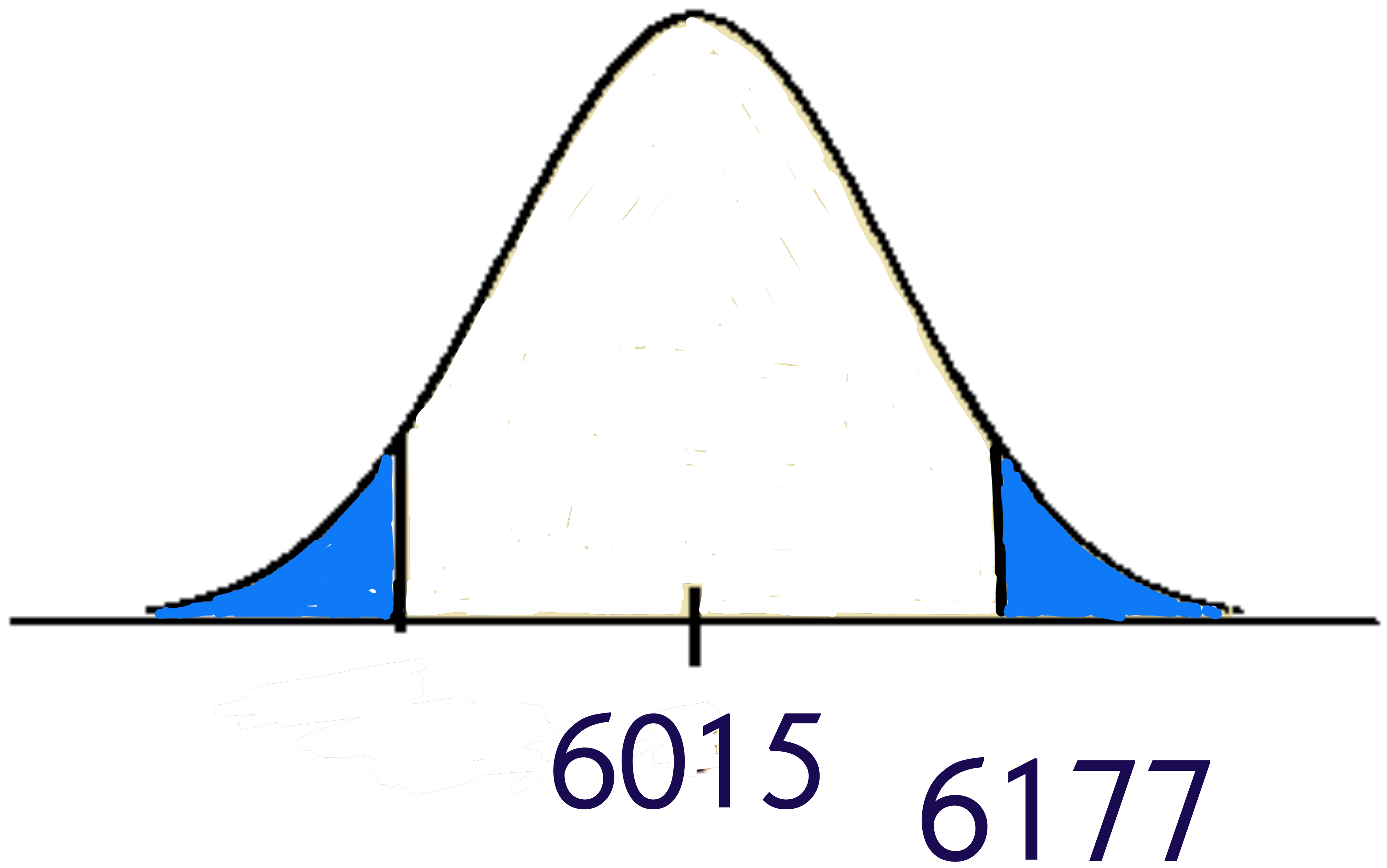
$$6015 - 1.96\frac{814}{\sqrt{100}} = 5855.5$$



We reject H_0, in favor of H_1. The data suggests that the true current average, for all families of 4, is different than it was in 2019.

P-Value:

$$P(\overline{X} > 6174.5 \textbf{ or } \overline{X} < 5855.5;\ \mu_0)$$



P-Value $= 2\,P(\overline{X} > 6177;\ \mu_0 = 6015)$

$$= 2\,P(Z > 1.99)$$

$$= 2(0.023295) = 0.0466$$

This is smaller than 0.05 so we reject H_0 at 0.05 level of significance.

P-Value $= 0.0466$

This is smaller than 0.05 so we reject H_0 at 0.05 level of significance.