ECON 3818

Chapter 17

Kyle Butts

27 September 2021

Chapter 17: Tests for Significance

Outline

Things to cover in Chapter 17:

- Reason we test for significance

- Stating a hypothesis

- p-value and statistical significance

- Tests for a population mean

Statistical Inference

Two most common types of statistical inference:

Confidence interval

- when the goal is to estimate a population parameter

Tests of significance

- asses the evidence provided by data about a claim concerning the population

Tests of Significance

Statistical Test: Formal procedure for comparing observed data with a claim (hypothesis) whose truth we want to assess

Basic idea: an outcome that would rarely happen if a certain claim is true, is good evidence that that claim is not true

- *Example:* say someone claims that the average points scored by a team in the NBA over the season is 50 points. If the 2020 season has a sample mean of 80 points scored, that's pretty good evidence against the claim.

Example

Suppose that the return on Spotify stock is normally distributed with an unknown mean and a variance $\sigma^2 = 1$.

You measure gains and losses for 10 trading days and collect the following sample:

$$1.6, 0.4, -2.5, 1.5, -1.1, 1.3, -0.1, -0.3, 1.2$$

The average rate of return is $\bar{X} = 0.3$, but there are quite a few days with losses.

Example

In hypothesis testing, we proceed under the assumption our claim is true and then try to disprove it.

For example we could claim $\mu = 0$, that is the average Spotify return is 0%.

Since the data are drawn from a normal distribution, the sample mean is distributed $\bar{X} \sim N(0, \frac{1}{10})$.

Example

In hypothesis testing, we proceed under the assumption our claim is true and then try to disprove it.

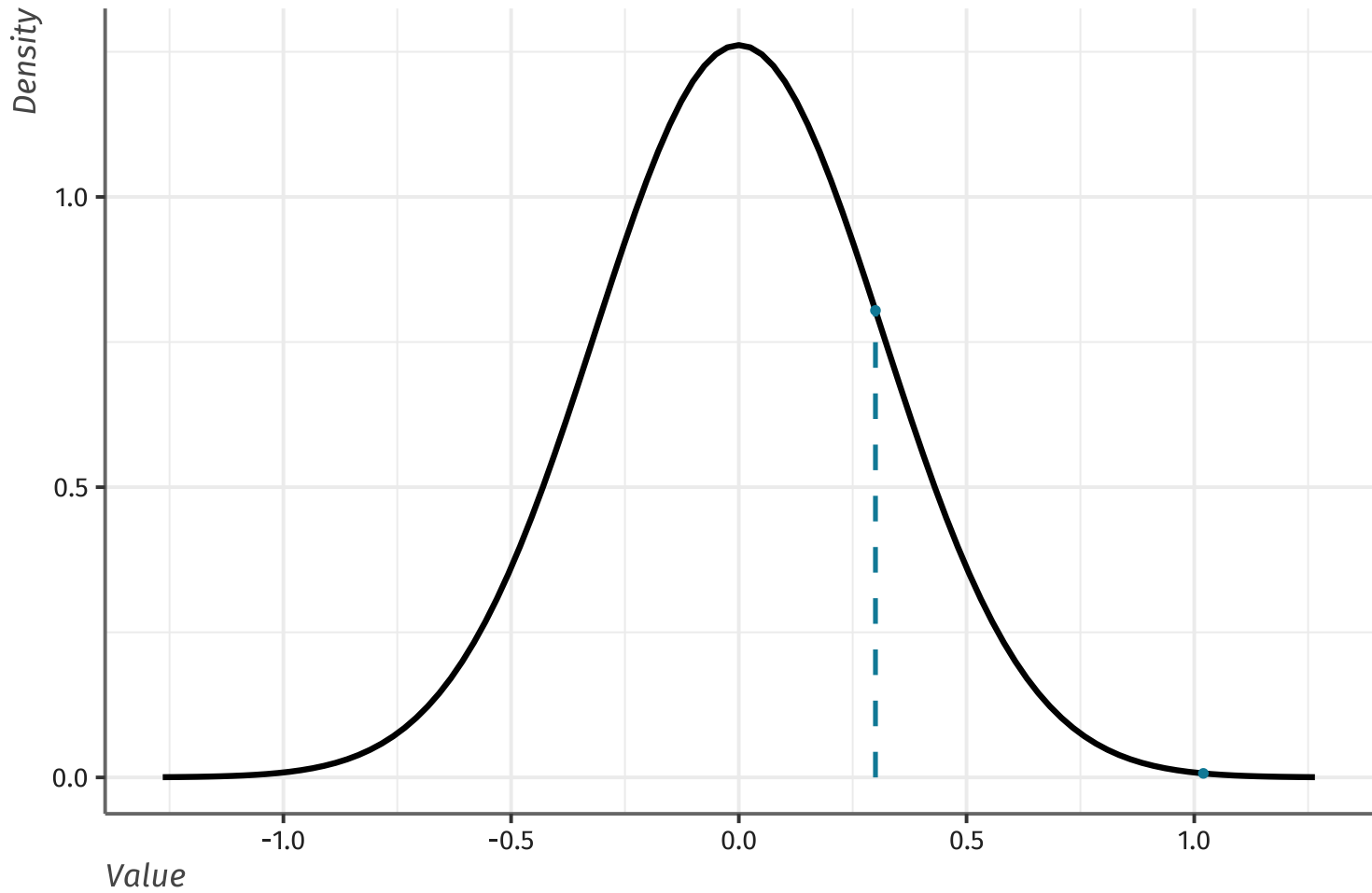For example we could claim $\mu = 0$, that is the average Spotify return is 0%.

Since the data are drawn from a normal distribution, the sample mean is distributed $\bar{X} \sim N(0, \frac{1}{10})$.

Important:

Note that we take the hypothesis that $\mu = 0$ as true!!

Now we can use the sampling distribution to comment on how unlikely this claim is.

Example

Stating a Hypothesis

It is very important to state the hypothesis clearly. A hypothesis has two parts:

Null hypothesis, denoted H_0, is the claim we are trying to disprove

- test is designed to test the strength of evidence against the null hypothesis

Alternative hypothesis, denoted H_1, is the claim we are trying to find evidence for*

- antithesis of the null

* Note that we are **NOT** trying to prove the alternative hyptothesis is correct

Alternative Hypothesis

There are two types of alternative hypotheses:

One-sided alternative

- states the true parameter is larger *or* smaller (not both)

Two-sided alternative

- states the true parameter is *different* from the null

Example

Back to the Spotify example:

Lets say we're claiming there is no return on average. We would like to see if we can disprove this hypothesis in favor of a claim that there is a positive return

$$H_0 : \mu = 0$$

$$H_1 : \mu > 0$$

Clicker Question: What type of hypothesis test is this?

   a. alternative

   b. null

   c. one-sided

   d. two-sided

Continuing Example

So in the Spotify example, we had 10 observations with a sample mean, $\bar{X} = 0.3$. We also somehow know that $\sigma^2 = 1$.
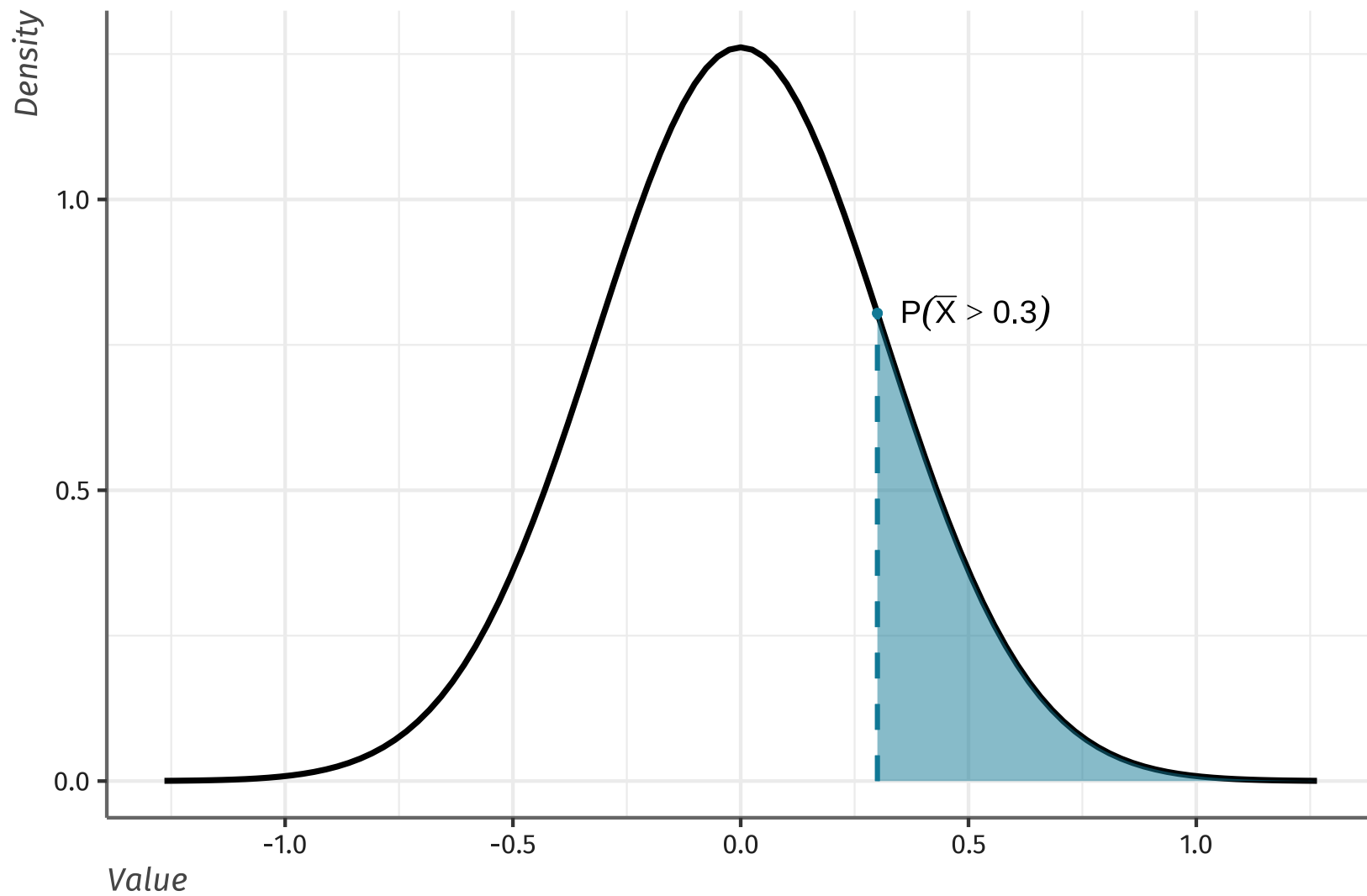
If we want to test the hypothesis

$$H_0 : \mu = 0$$

$$H_1 : \mu > 0$$

We want to calculate the probability we got an $\bar{X}$ as extreme as 0.3 if the true population $\mu = 0$. Is it rare that we got 0.3 or is that quite common from the sampling distribution?

This means we want to calculate $P(\bar{X} > 0.3)$ if $\bar{X} \sim N(0, \frac{1}{10})$

Continuing Example

We want to calculate $P(\bar{X} > 0.3)$ if $\bar{X} \sim N(0, \frac{1}{10})$

$$P(\bar{X} > 0.3) = P\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} > \frac{0.3 - 0}{1/\sqrt{10}}\right) = P(Z > 0.95) = 0.1711$$

This says that if the *true* population mean is $\mu = 0$, then 17% of the time we would expect to see a sample mean greater than 0.3, 17% of the time.

- This is not so rare (~ 1 out of 5), so this isn't strong evidence against the null. It doesn't prove it, but it does not rule it out.

p-values

Suppose you calculate a test statistic $\hat{\theta}$ (e.g. mean or standard deviation) from sample data, and formulate a null hypothesis $H_0 : \theta = \theta_0$ about the true value of θ

- **p-value** is the probability that the test statistic takes a value as or *more extreme* than the observed value $\hat{\theta}$, **under the assumption that H_0 is true***

In general, the test statistic we calculate is $\bar{X}$ and formulate a hypothesis about μ

- In this context, the p-value is the probability of getting an $\bar{X}$ as or *more extreme*, if the true population mean is equal to μ

* Can not emphasize this enough. Under the assumption that H_0 is true.

18 / 55

Calculating p-values

The way a p-value is calculated changes depending on the type of alternative hypothesis we have:

$$H_0 : \mu = \mu_0$$

Left-tailed Test:

$$H_1 : \mu > \mu_0 \implies p = P(\bar{X} \leq \bar{X}_{obs} \mid \mu = \mu_0)$$

Right-tailed Test:

$$H_1 : \mu < \mu_0 \implies p = P(\bar{X} \geq \bar{X}_{obs} \mid \mu = \mu_0)$$

Two-tailed Test:

$$H_1 : \mu \neq \mu_0 \implies p = 2 \cdot P(\bar{X} \geq |\bar{X}_{obs}| \mid \mu = \mu_0)$$

where $\bar{X}$ is a random draw from our sample distribution and $\bar{X}_{obs}$ is our **observed** value for the sample statistic.

Example -- One-sided

Lets say we have a sample of 30 observations and calculate a sample mean of 1020. We also know that the standard deviation is $\sigma = 40$. We want to test the following hypothesis:

$$H_0 : \mu = 1000$$

$$H_1 : \mu > 1000$$

- Write out p-value:

$$p = P(\bar{X} \geq 1020 \mid \mu = 1000)$$

- Solve for p-value:

$$P(\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \geq \frac{1020 - 1000}{40/\sqrt{30}}) = P(Z > 2.74) = 0.0032$$

So $p = 0.003$, the probability we get a sample mean as large as 1020 if the true mean is 1000 is .3%, super rare!

Example -- Two-sided

Say we have nine observations, and calculate a sample mean of $\bar{X} = 0.5$. We know the underlying distribution has a standard deviation $\sigma = 1$. We want to test the following hypothesis:

$$H_0 : \mu = 0$$

$$H_1 : \mu \neq 0$$

- Write out p-value:

$$p = 2 \cdot P(\bar{X} \geq 0.5 \mid \mu = 0)$$

- Solve for p-value:

$$2 \cdot P\left(\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \geq \frac{0.5 - 0}{1/\sqrt{9}}\right) = 2 \cdot P(Z > 1.5) = 2 \cdot (0.067) = 0.134$$

So $p = 0.134$, meaning there is a 13.4% probability of getting a sample mean *as extreme as* 0.5 if the true mean is 0.

Breakout Group

Say we have a sample of 25 observations and we calculate a sample mean of 11,500. We know that the standard deviation is $\sigma = 5000$. We want to test the following hypothesis:

$$H_0 : \mu = 12,675$$

$$H_1 : \mu < 12,675$$

Calculate the p-value:

a. 0.407

b. 0.814

c. 0.119

d. 0.238

Breakout Group

Say we have a sample of 25 observations and we calculate a sample mean of 11,500. We know that the standard deviation is $\sigma = 5000$. We want to test the following hypothesis:

$$H_0 : \mu = 12,675$$

$$H_1 : \mu < 12,675$$

Hypothesis Testing

A restaurant owner is determining whether to expand her restaurant to another location. She counts the number of customers per weekend and calculates an average of 538 customers over 50 weekends. She's heard that restaurants of her size shouldn't expand unless they get more than 510 customers each weekend. Conduct a hypothesis at the $\alpha = 0.05$ level and determine whether or not she can feel confident she gets at least 510 customers each weekend. Assume $\sigma = 120$.

- State Null and Alternative hypotheses

$$H_0 : \mu = 510$$
$$H_1 : \mu \geq 510$$

Hypothesis Testing

A restaurant owner is determining whether to expand her restaurant to another location. She counts the number of customers per weekend and calculates an average of 538 customers over 50 weekends. She's heard that restaurants of her size shouldn't expand unless they get more than 510 customers each weekend. Conduct a hypothesis at the $\alpha = 0.05$ level and determine whether or not she can feel confident she gets at least 510 customers each weekend. Assume $\sigma = 120$.

- Calculate p-value

$$P(\bar{X} \geq 538 | \mu = 510) = 0.049$$

If true population mean is 510, there is a 4.95% probability of calculating a sample mean as large as 538

Interpreting p-Values

Small p-values are evidence against the null, because they say **the observed value of the sample statistic if very unlikely to occur if the null hypothesis, H_0, is true.**

- Larger p-values fail to give evidence against the null because the observed sample statistic could have resulted by chance if the null hypothesis, H_0 is true.

p-values do not tell you the probability that H_0 is true!!*

p-Values and Statistical Significance

To review, our decision process using p-values is

- p-value is sufficiently small $\rightarrow$ reject H_0 $\rightarrow$ accept H_1
- p-value is too large $\rightarrow$ fail to reject H_0 $\rightarrow$ do not accept H_1

We *never* say "accept H_0". All we can say is if there is enough evidence to say it is false, we cannot know if it is true.

We say our results are **statistically significant** if we reject H_0*

---

* This saying is short for "statistically significantly different from zero" (or H_0 more generally).

p-Values and Statistical Significance

How small does p need to be in order to be "sufficiently small"? When can we claim our results are statistically significant?

Suppose you calculate a p-value, and specify a **level of significance**, α. We say our results are **statistically significant** if:

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

We will discuss how to think about choosing α. Some of the most common levels of significance are .01, .05, and 0.1.

Group Example

A restaurant owner is determining whether to expand her restaurant to another location. She counts the number of customers per weekend and calculates an average of 538 customers over 50 weekends. She's heard that restaurants of her size shouldn't expand unless they get more than 510 customers each weekend. Conduct a hypothesis at the $\alpha = 0.05$ level and determine whether or not she can feel confident she gets at least 510 customers each weekend. Assume $\sigma = 120$.

1. State Null and Alternative hypotheses

2. Calculate p-value

3. Is our estimate statistically siginificant?

Group Example

1. State Null and Alternative hypotheses

   ○ $H_0 : \mu = 510; \quad H_1 : \mu \geq 510$

2. Calculate p-value

   ○ $P(\bar{X} \geq |\mu = 510) = 0.049$

3. Is our estimate statistically siginificant?

   ○ $0.049 < 0.05 \implies p - value < \alpha \implies \text{Reject } H_0$

Reject null hypothesis in favor of alternative, that the true population mean number of customers is greater than 510

Hypothesis Testing Algorithm

A hypothesis test consists of 4 parts:

- Null hypothesis, H_0

- Alternative hypothesis, H_1

- Test statistic, $\hat{\theta}$, most commonly $\bar{X}$

- Level of significances α

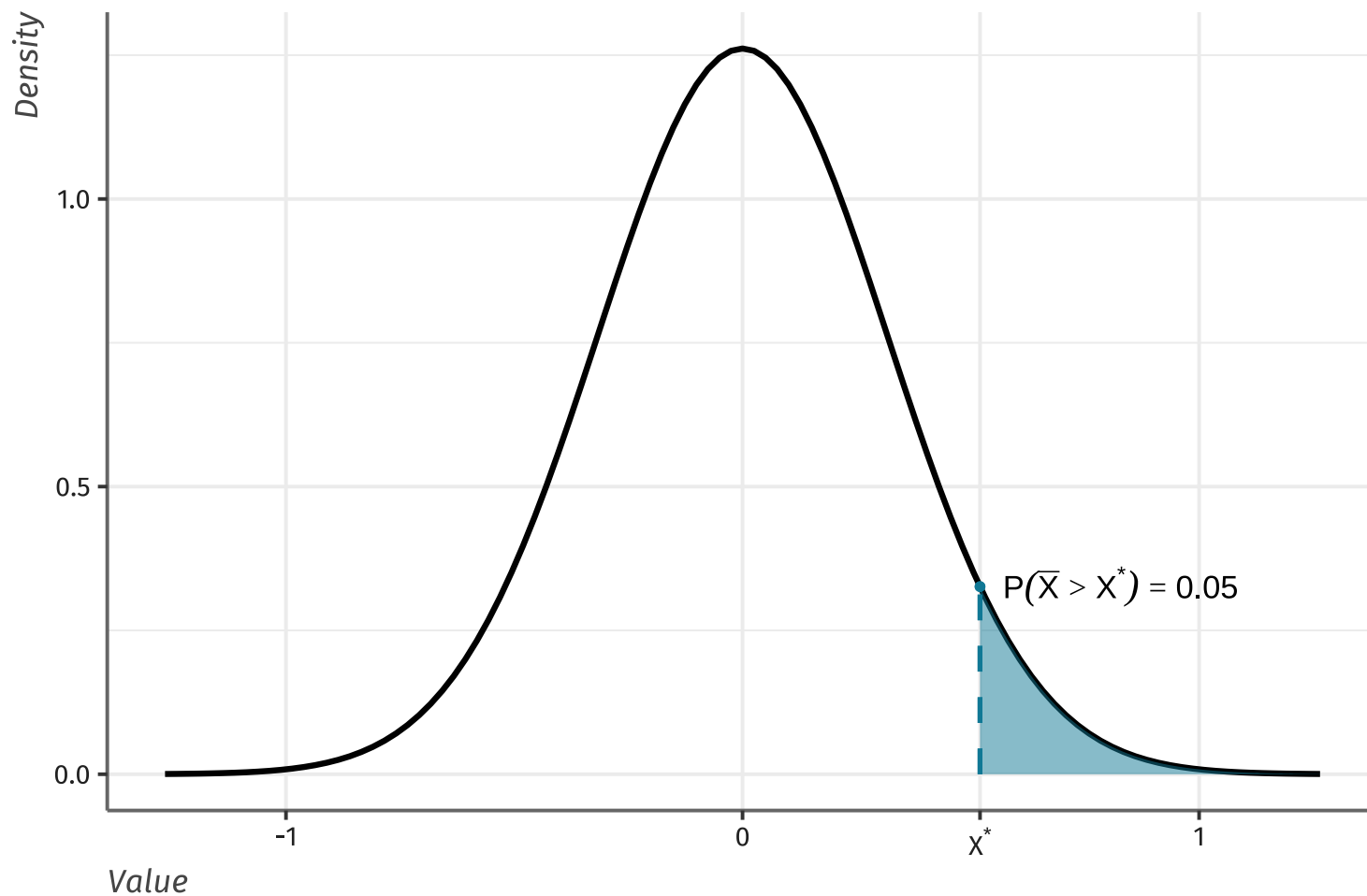With these, we reach the conclusion:

- reject H_0 if $p \leq \alpha$

- fail to reject H_0 if $p > \alpha$
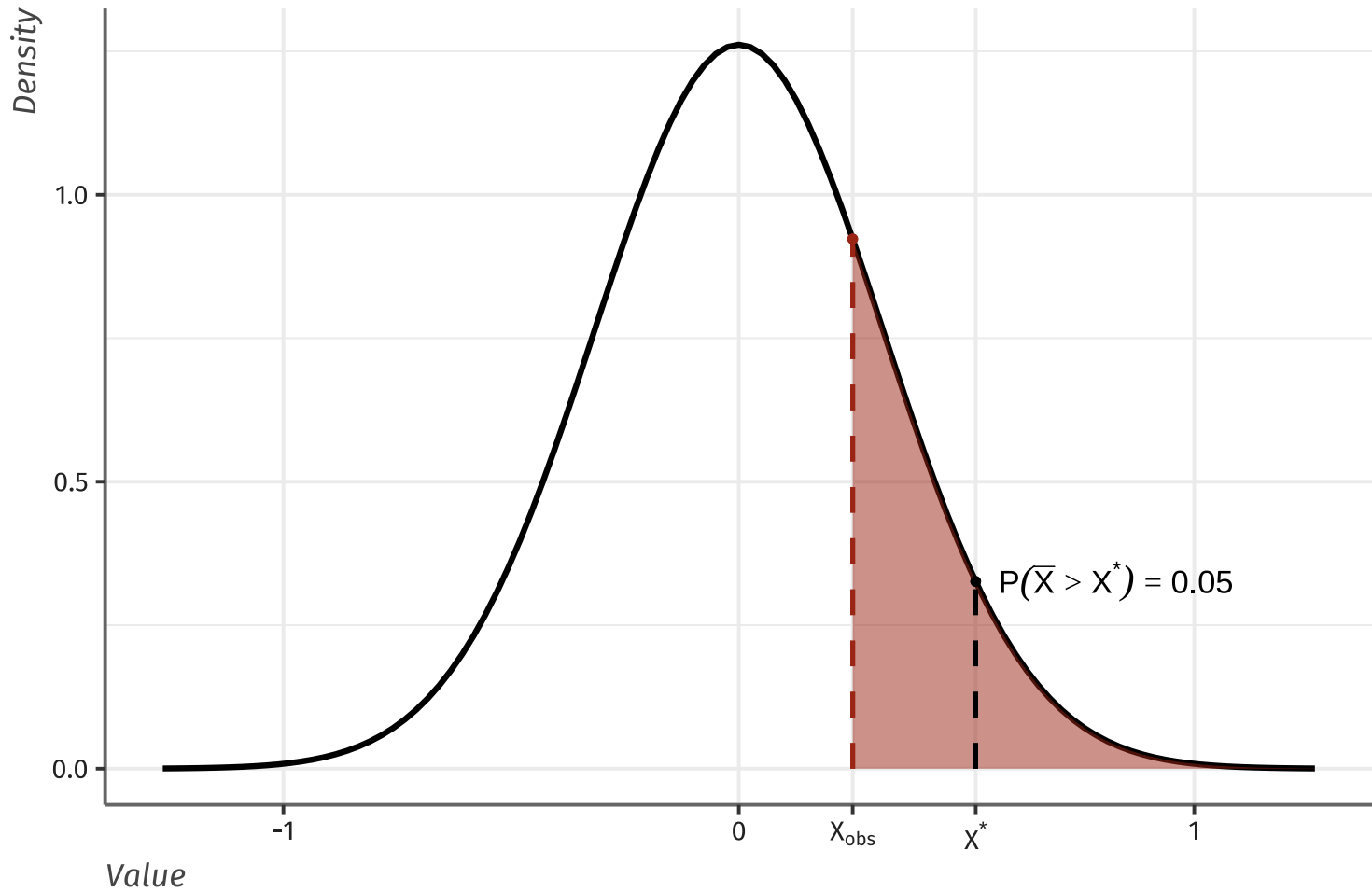
Rejection Region

Imagine getting many many observed sample means $(\bar{X}_{1,obs}, \bar{X}_{2,obs}, \ldots, \bar{X}_{100,obs})$. We don't want to calculate 100 p-values.

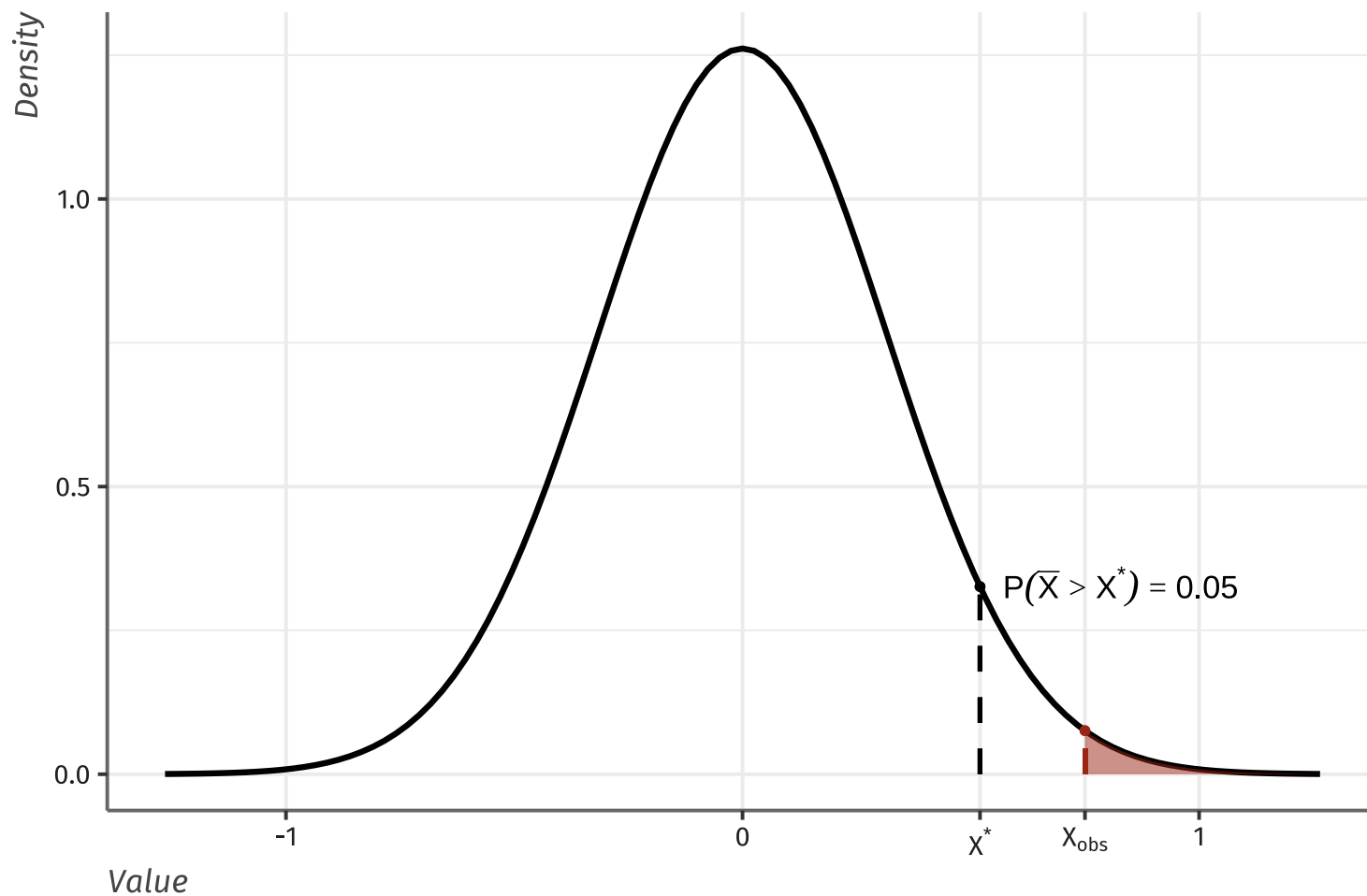It would be better if we could figure out for what range of values that would give us p-value less than α.

Rejection Region

Rejection Region

Rejection Region

Rejection Region

We call the range of values for our *observed* test statistic $\bar{X}_{obs}$ which we would reject the null the rejection region

Calculate rejection region using

- the population mean, μ under the null hypothesis

- the standard deviation, σ

- level of significance, α and Z-score associated with it

The rejection region tells us how extreme the sample mean needs to be for us to reject the null hypothesis $H_0 : \mu = \mu_0$

Calculate Rejection Region -- One-Sided

Let's say we want to test a hypothesis about the number of children in a family. We have a sample of 25 observations, and we know $\sigma = 0.8$. We decide to test the following hypothesis:

$$H_0 : \mu = 2$$

$$H_1 : \mu < 2$$

If we are testing at the 5% significance level, calculate the rejection region.

- The rejection region is the values of $\bar{X}$ that are far enough away from μ that we would reject H_0

We want to find $\bar{X}^*$ that solves

$$R = P(\bar{X} \leq \bar{X}^* \mid \mu_0 = 2) = 0.05$$

Calculating Rejection Region -- One-Sided

$$P(\bar{X} \leq \bar{X}^* | \mu_0 = 2) = 0.05 = P(Z \leq \frac{\bar{X}^* - \mu}{\sigma/\sqrt{n}})$$

Plugging in μ and σ, we have:

$$P(Z \leq \frac{\bar{X}^* - 2}{0.8/5}) = 0.05 = P(Z \leq \frac{\bar{X}^* - 2}{.16}) = 0.05$$

We know that $P(Z \leq -1.645) = 0.05$, therefore:

$$\frac{\bar{X}^* - 2}{0.16} = -1.645 \implies \bar{X}^* = 1.74$$

This means our rejection region is all $\bar{X} \leq 1.74$. In words, we will reject $H_0 : \mu = 2$ at the $\alpha = 0.05$ level if we calculate a sample mean of 1.74 or smaller.

Breakout Group

A study is conducted to determine the average income of 18-25 year olds in San Antonio. They are trying to test whether the average is greater than \$21,000. The study is able to collect 16 different salaries. We somehow know that the population standard deviation is $\sigma =$ \$2000. What sample averages would we reject the null hypothesis at the $\alpha = 0.05$ level?

$$H_0 : \mu = \$21,000$$

$$H_1 : \mu > \$21,000$$

a. Reject if $\bar{X} < \$21,822.5$
b. Reject if $\bar{X} > \$21,822.5$
c. Reject if $\bar{X} < \$20,177.5$
d. Reject if $\bar{X} > \$20,177.5$

Breakout Group

A study is conducted to determine the average income of 18-25 year olds in San Antonio. They are trying to test whether the average is greater than \$21,000. The study is able to collect 16 different salaries. We somehow know that the population standard deviation is $\sigma =$ \$2000. What sample averages would we reject the null hypothesis at the $\alpha = 0.05$ level?

$$H_0 : \mu = \$21,000$$

$$H_1 : \mu > \$21,000$$

Calculating Rejection Region -- Two-Sided

Solving for a rejection region changes a little bit when we consider a two-tailed test.

Example: Suppose we have a sample of 25 observations, we know $\sigma = 3.5$. We want to test the following hypothesis at the $\alpha = 0.05$ level.

$$H_0 : \mu = 13.5$$

$$H_1 : \mu \neq 13.5$$

Our rejection region is calculated by finding two different $\bar{X}^*$'s.

$$P(\bar{X} \geq \bar{X}_U^*) = \frac{0.05}{2}$$

$$P(\bar{X} \leq \bar{X}_L^*) = \frac{0.05}{2}$$

Calculating Rejection Region -- Two-Sided

$$P(\bar{X} \geq \bar{X}_U^*) = \frac{0.05}{2} = 0.025 = P(Z > \frac{\bar{X}_U^* - 13.5}{0.64})$$

- We know $P(Z > 1.96) = 0.025$.

- Therefore:

$$\frac{\bar{X}_U^* - 13.5}{0.64} = 1.96 \implies \bar{X}_U^* = 14.75$$

Breakout Group

For the same problem, solve for $\bar{X}_L^*$.

Remember it solves: $P(\bar{X} \leq \bar{X}_L^*) = \frac{0.05}{2}$

   a. -14.75

   b. 12.246

   c. 11.54

Solving Hypothesis Testing Problem

p-value method

- calculate p-value, varies based off type of alternative hypothesis
- p-value: probability of getting as extreme of a sample statistic, if the null hypothesis is true

$$\text{If } p < \alpha \implies \text{ reject } H_0$$

Rejection region method

- calculate rejection region based off level of significance
- compare sample statistic with rejection region

$$\text{If sample statistic is in rejection region, reject } H_0$$