hypothesis testing

- hypothesis testing via CI
- formal hypothesis testing using p-values
- one and two-sided hypothesis tests

recap: hypothesis testing framework

- start with a *null hypothesis* (H_0) that represents the present situation.
- set an *alternative hypothesis* (H_A) that represents our research question, i.e. what we're testing for.
- conduct a hypothesis test under the assumption that the null hypothesis is true, either via simulation or theoretical methods-methods that rely on the CLT
  - If the test results suggest that the data do not provide convincing evidence for the alternative hypothesis, we stick with the null hypothesis.
  - If they do, then we reject the null hypothesis in favor of the alternative.

hypothesis

null - H_0	Often either a skeptical perspective or a claim to be tested	$=$
alternative - H_A	Represents an alternative claim under consideration and is often represented by a range of possible parameter values.	$<, >, \neq$

The skeptic will not abandon the H_0 unless the evidence in favor of the H_A is so strong that she rejects H_0 in favor of H_A

Practice

A random sample of 50 college students were asked how many exclusive relationships they have been in so far. This sample yielded a mean of 3.2 and a standard deviation of 1.74. Estimate the true average number of exclusive relationships using this sample.

$\bar{x}$=3.2 s = 1.74

The approximate 95% confidence interval is defined as

point estimate ± 2 x SE

$$SE = \frac{s}{\sqrt{n}} = \frac{1.74}{\sqrt{50}} \approx 0.25$$

$\bar{x}$ ± 2 x SE → 3.2 ± 2 x 0.25

→ (3.2 - 0.5, 3.2 + 0.5)

→ (2.7, 3.7)

Earlier we calculated a 95% confidence interval for the average number of exclusive relationships college students have been in to be (2.7, 3.7). Based on this confidence interval, do these data support the hypothesis that college students on average have been in more than 3 exclusive relationships?

H_0: $\mu = 3$ College students have been in 3 exclusive relationships, on average

H_A: $\mu > 3$ College students have been in more than 3 exclusiverelationships, on average

always about pop. parameters,
never about sample statistics

decision based on the p-value

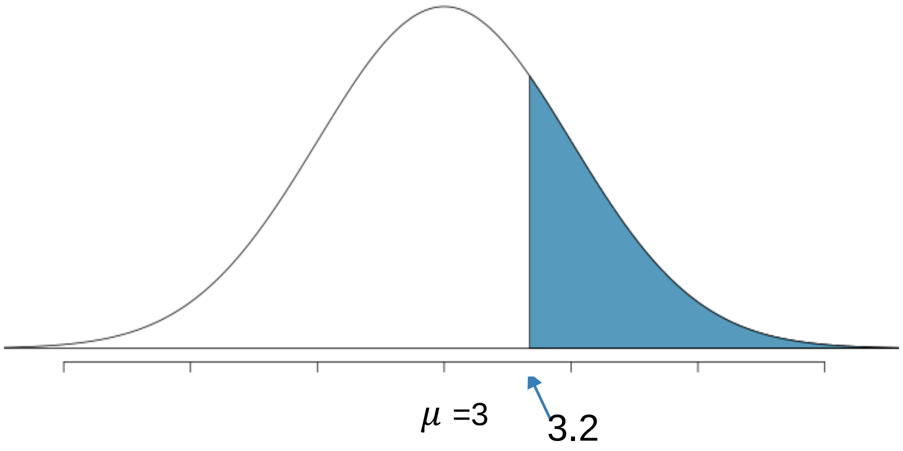
- use the test statistic to calculate the p-value, the probability of observing data at least as favorable of the alternative hypothesis as our current data set, if the null hypothsis is true.

- If the p-value is low (lower than the significance level α, which is usually 5%) we say that it would be very unlikely to observe the data if the null hypothesis were true, and hence reject H_0

- If the p-value is high (higher than α) we say that it is likely to observe the data even if the null hypothesis were true, and hence do not reject H_0

p-value

P(observed or more extreme outcome| H_0 true)



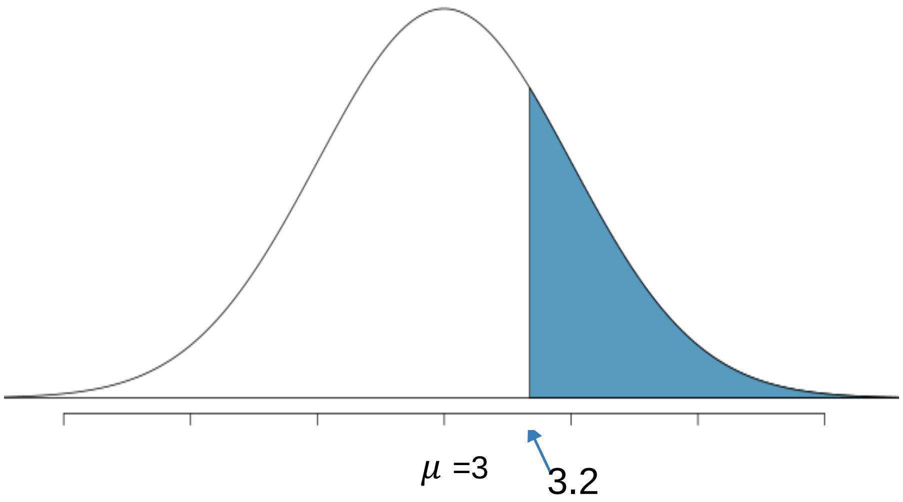
test statistic

interpreting p-value

- If infact college students have been in 3 exclusive relationships on average, there is a 21% chance that a random sample of 50 college students would yield a sample mean of 3.2 or higher.

- This is a pretty high probability, so we think that a sample mean of 3.2 or more exclusive relationships is likely to happen simply by chance.

p-value=0.209≈ 0.21

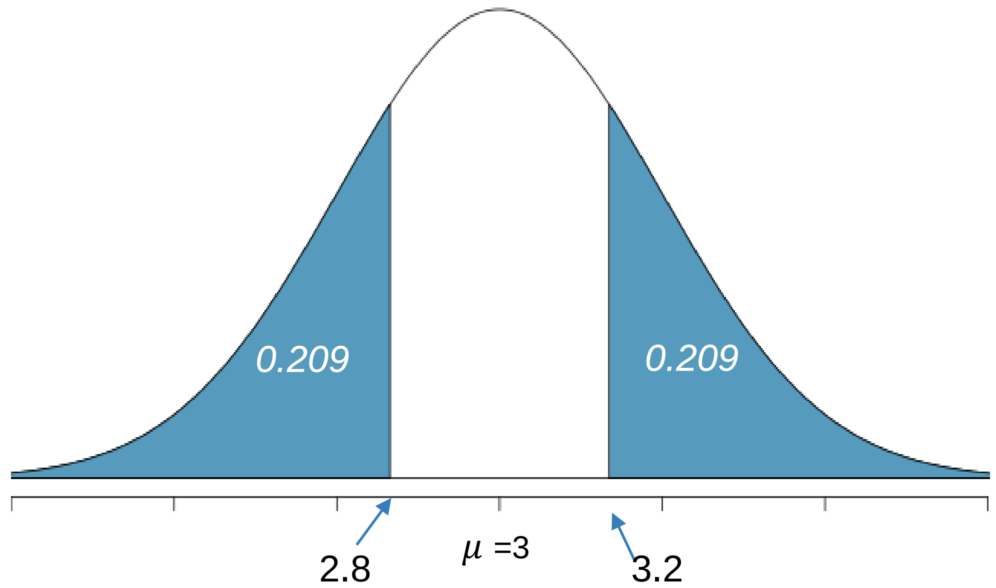


making a decision

- Since p-value is high (higher than 5%) we fail to reject H_0
- These data do not provide convincing evidence that college students have been in more than 3 relationships on average.
- The difference between the null value of 3 relationships and the observed sample mean of 3.2 relationships is due to chance or sampling variability.

two-sided tests

- Often instead of looking for a divergence from the null in a specific dirction, we might be interested in divergence in any direction.
- We call such hypothesis tests two-sided (or two-tailed)
- The definition of a p-value is the same regardless of doing a one or two-sided test, however the calculation is slightly differenct since we need to consider “at least as extremme as the observed outcome” in both directions.


0.209
0.209
2.8
μ =3
3.2

Hypothesis testing for a single mean

1. Set the hypotheses : $H_0: \mu =$ *null value*

   $H_A: \mu < or > or \neq$ *null value*

2. Calculate the point estimate: $\bar{x}$

3. Check conditions:

Independence: sampled observations must be independent(random sample/assignment & if sampling without replacement, n<10% of population)

Sample size/skew: n≥30, larger if the population distribution is skewed

4. Draw sampling distribution, shade p-value, calculate test statistic

$$Z = \frac{\bar{x} - \mu}{SE}, \qquad SE = \frac{s}{\sqrt{n}}$$

5. Make a decision, and interpret it in context of the research question:

   if p-value < α, reject H_0; the data provide convincing evidence for H_A

Assignment

Randomly selected 36 cans of Coke are measured for the amount of cola, in ounces. The sample has a mean of 12.19 ounces and a standard deviation of 0.11 ounce. Assume that we want to use a 0.05 significance level to test the claim that cans of Coke have a mean amount of cola greater than 12 ounces.

(a) What is the null hypothesis?

(b) What is the alternative hypothesis?

(c) What is the value of the standard score for the sample mean of 12.19 ounces?

(d) What is the p-value?

(e) What do you conclude?

(f) Find the p-value if the test is modified to test the claim that the mean is different from 12 ounces.