hypothesis testing (for a mean)

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

Dr. Mine Çetinkaya-Rundel
Duke University

Duke
UNIVERSITY

hypotheses

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.	$<, >, \neq$

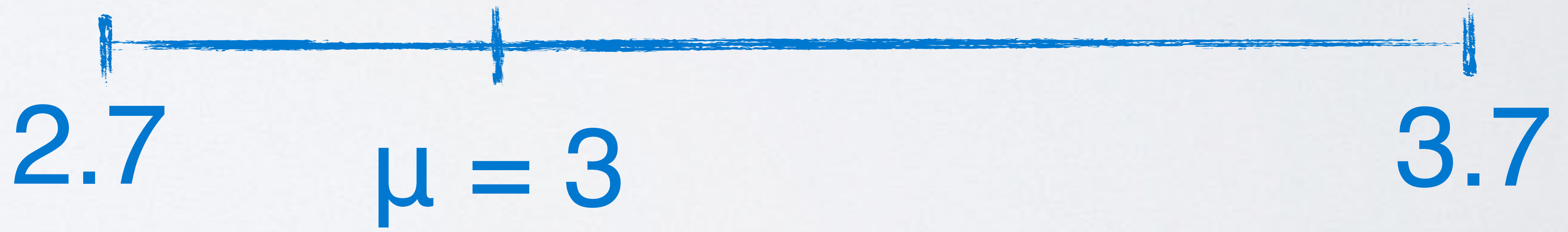
The skeptic will not abandon the H_0 unless the evidence in favor of the H_A is so strong that she rejects H_0 in favor of H_A.

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 exclusive relationships, on average.

always about pop. parameters,
never about sample statistics

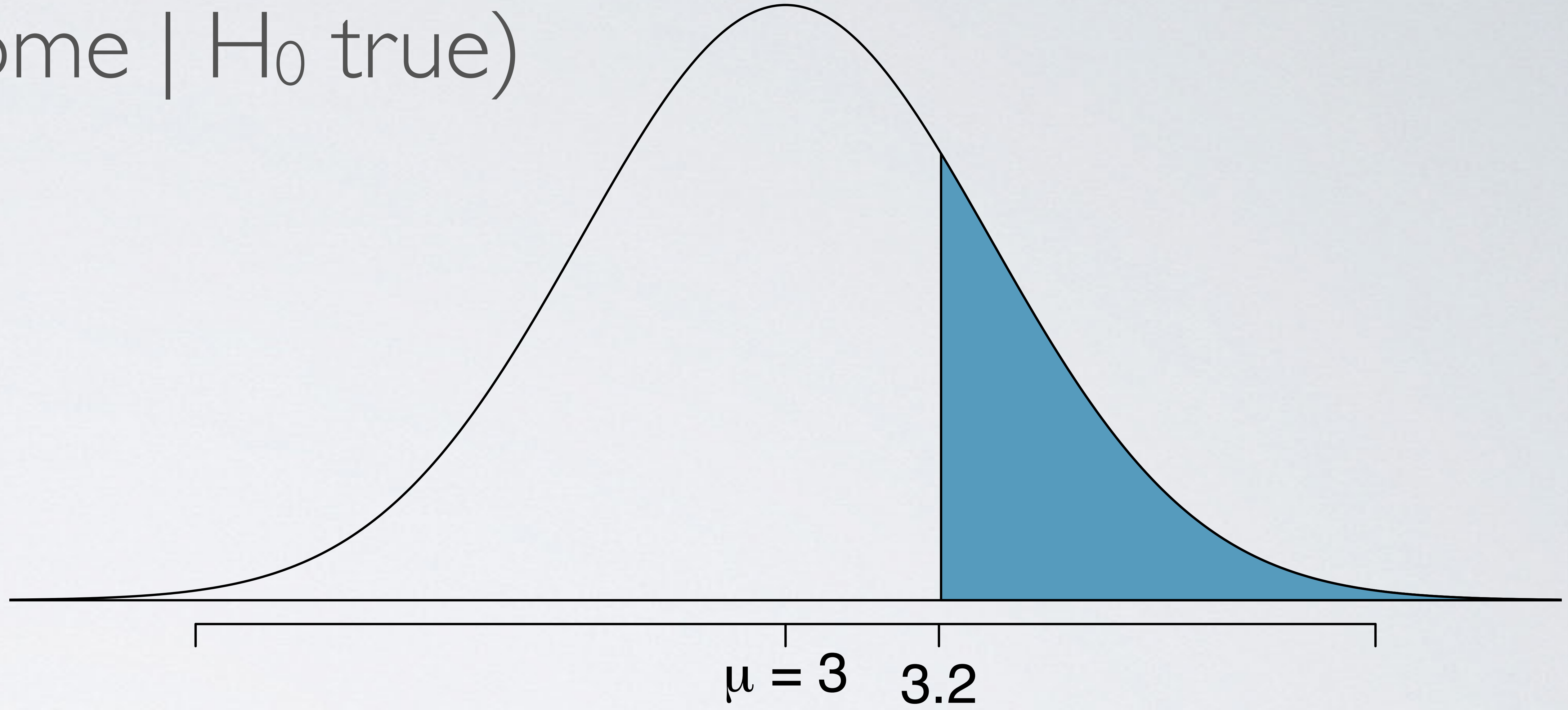


p-value

P(observed or more extreme outcome | H_0 true)

$P(\bar{X} > 3.2 \mid H_0: \mu = 3)$

$\bar{X} \sim N(\mu = 3, SE = 0.246)$



$n = 50$

$\bar{x} = 3.2$

$s = 1.74$

$SE = 0.246$

test statistic

$$Z = \frac{3.2 - 3}{0.246} = 0.81$$

$$\text{p-value} = P(Z > 0.81) = 0.209$$

decision based on the p-value

- We used the test statistic to calculate the p-value, the probability of observing data at least as favorable to the alternative hypothesis as our current data set, if the null hypothesis was 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.

$P(\bar{X} > 3.2 \mid H_0: \mu = 3)$

$\bar{X} \sim N(\mu = 3, SE = 0.246)$

$$Z = \frac{3.2 - 3}{0.246} = 0.81$$

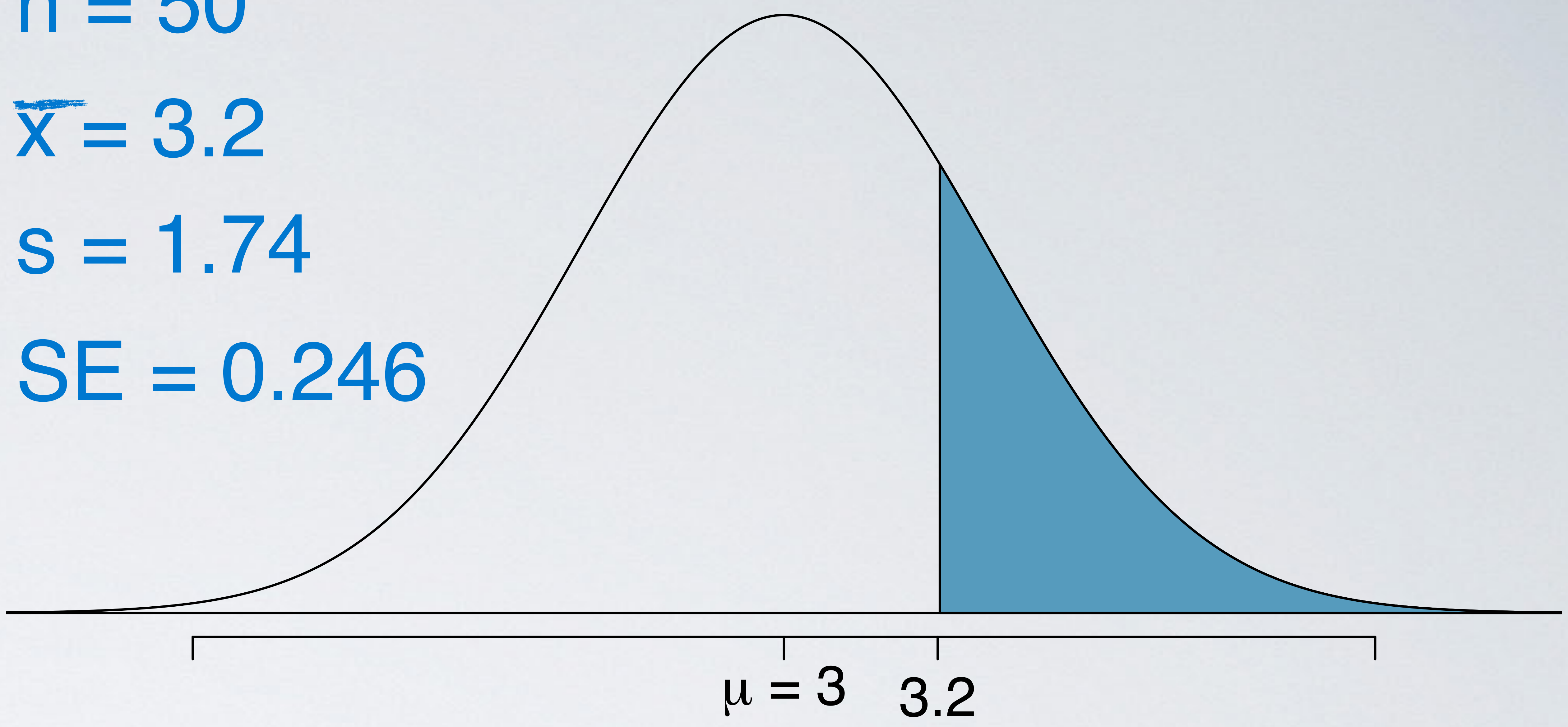
p-value = $P(Z > 0.81) = 0.209$

$n = 50$

$\bar{x} = 3.2$

$s = 1.74$

$SE = 0.246$

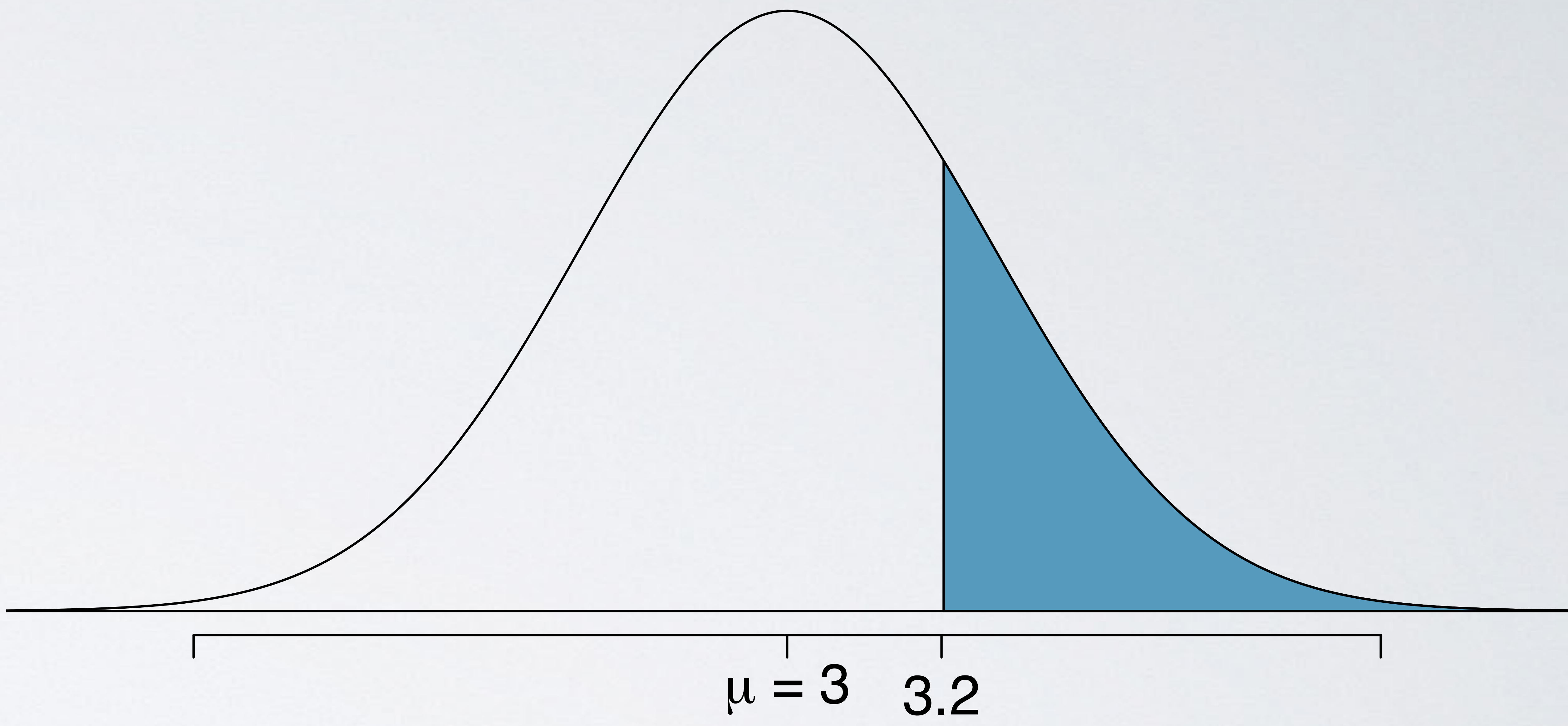


Since p-value is high, we do not reject H_0.

interpreting the p-value

- If in fact 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



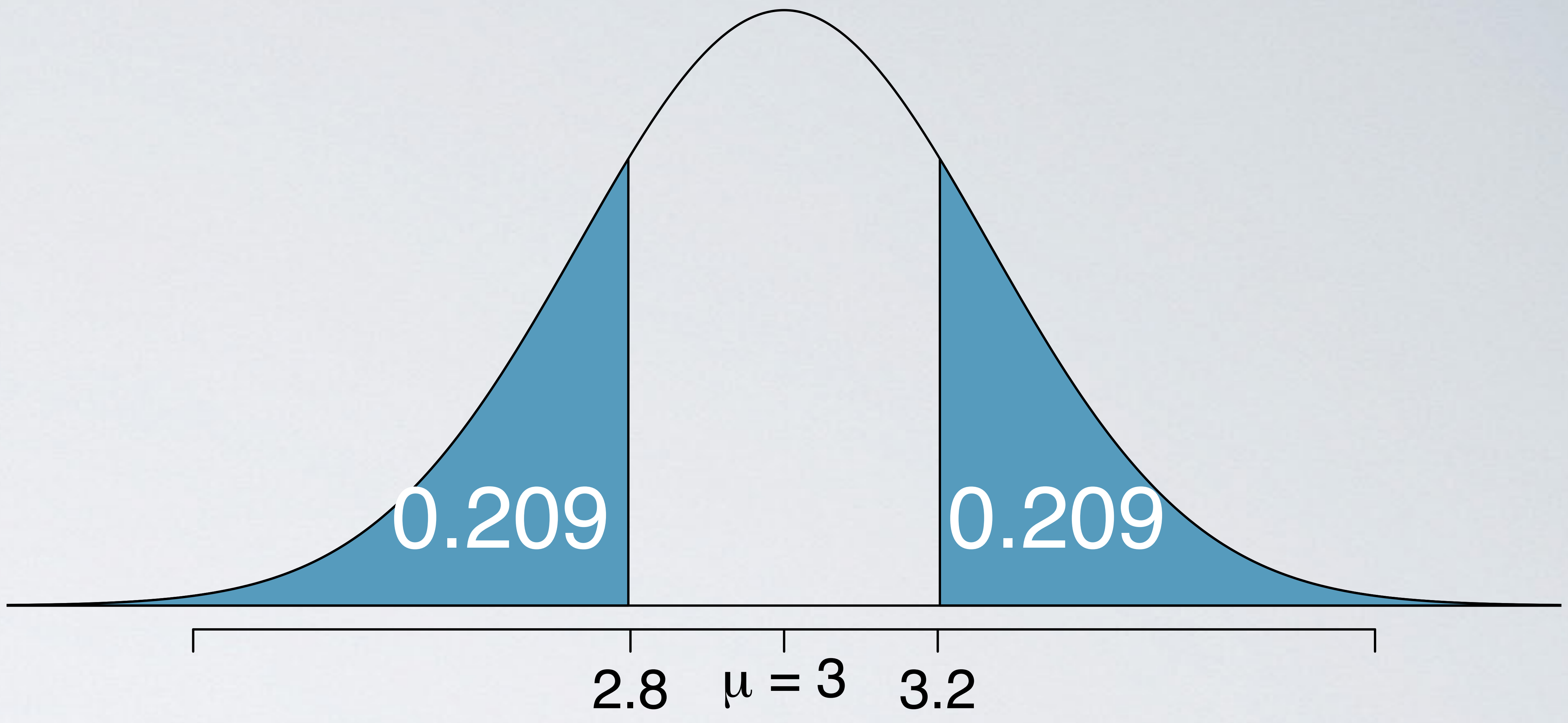
making a decision

- Since p-value is high (higher than 5%) we fail to reject H_0.
- The 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 direction, 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 different since we need to consider "at least as extreme as the observed outcome" in both directions.

$P(\bar{X} > 3.2 \text{ OR } \bar{X} < 2.8 \mid H_0: \mu = 3)$



p-value =

$= P(Z > 0.81) + P(Z < -0.81)$

$= 2 \times 0.209$

$= 0.418$

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:
   1. ***Independence:*** Sampled observations must be independent (random sample/assignment & if sampling without replacement, n < 10% of population)
   2. ***Sample size/skew:*** n ≥ 30, larger if the population distribution is very skewed.

4. Draw sampling distribution, shade p-value, calculate test statistic $Z = \frac{\bar{x} - \mu}{SE}, \quad SE = \frac{s}{\sqrt{n}}$

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.
   - If p-value > α, fail to reject H_0 the data *do not* provide convincing evidence for H_A.