Hypothesis Testing cont.

Grinnell College

October 26, 2026

Warm-up

I want definitions for these things:

1. Sampling distribution
2. Critical values
3. Null hypothesis
4. Test statistic

Bonus Question: In 2-3 sentences, what do these things have to do with each other?

What are we doing?
What are our goals?
Why are we even here?

Hypothesis Testing

Recall the framework we introduced for hypothesis testing

Hypothesis testing involves:

1. Formulating an *unambiguous* statement about a population parameter, called our **null hypothesis**
2. Collecting observational or experimental data
3. Determining if the data collected is consistent with our hypothesis
4. Either *rejecting* or *failing to reject* our hypothesis based on the *strength* of the evidence

How did we decide we were going to do this?

High Level Business

1. CLT: $\overline{X} \sim N(\mu, \sigma/\sqrt{n})$
   - Standard normal (Z-score)
   - t statistics (based on n)
   - Critical values
2. In theory, we can use this to determine if $\overline{X}$ is far from where it should be
3. In practice, this is equivalent to determining if a hypothesized value μ_0 is inconsistent with $\overline{X}$
4. We are using these ideas to *test a hypothesis*, either rejecting or failing to reject, at a specified level of confidence

What happens when we sample repeatedly?

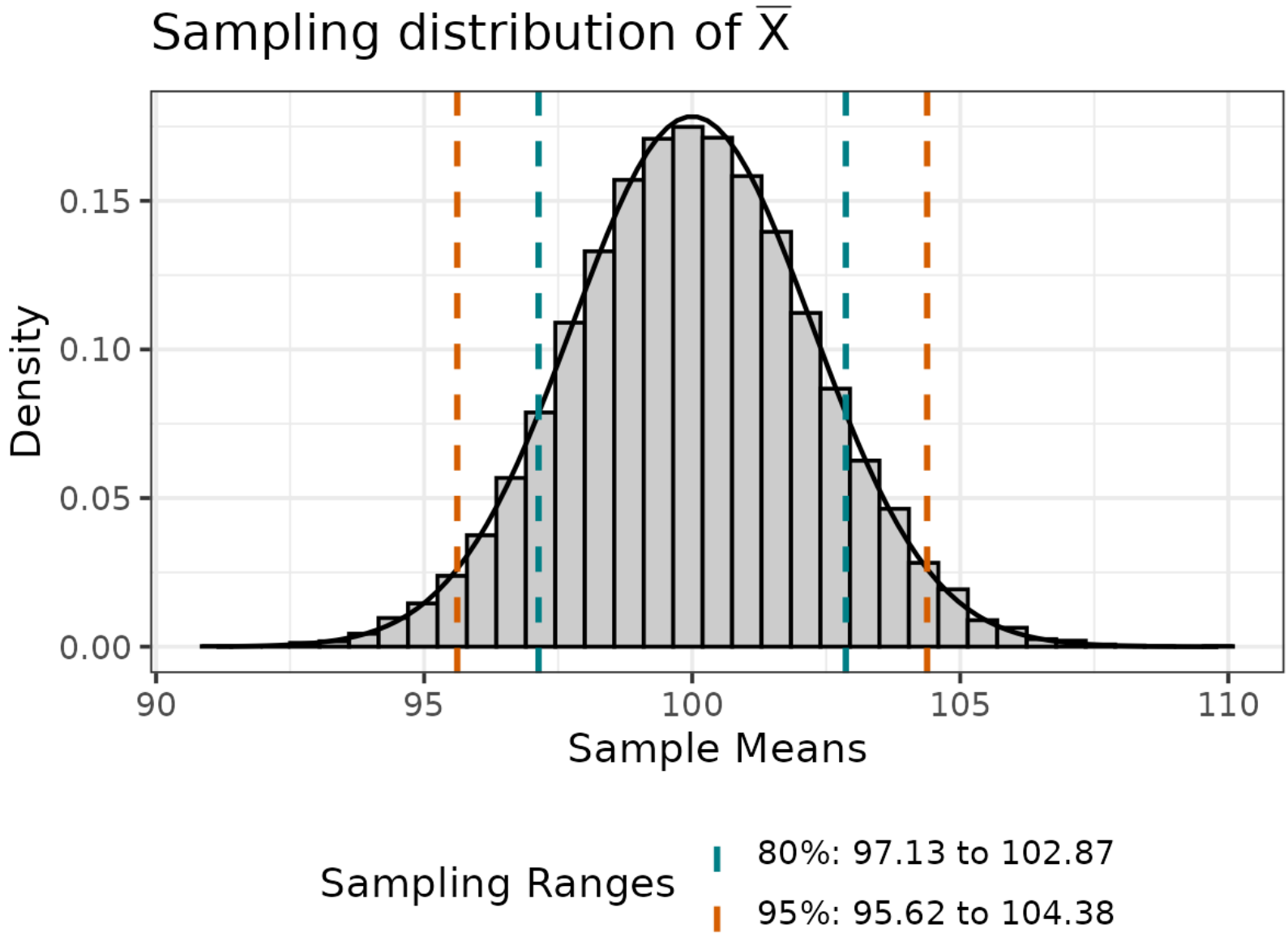
Suppose we know the population is normal, with

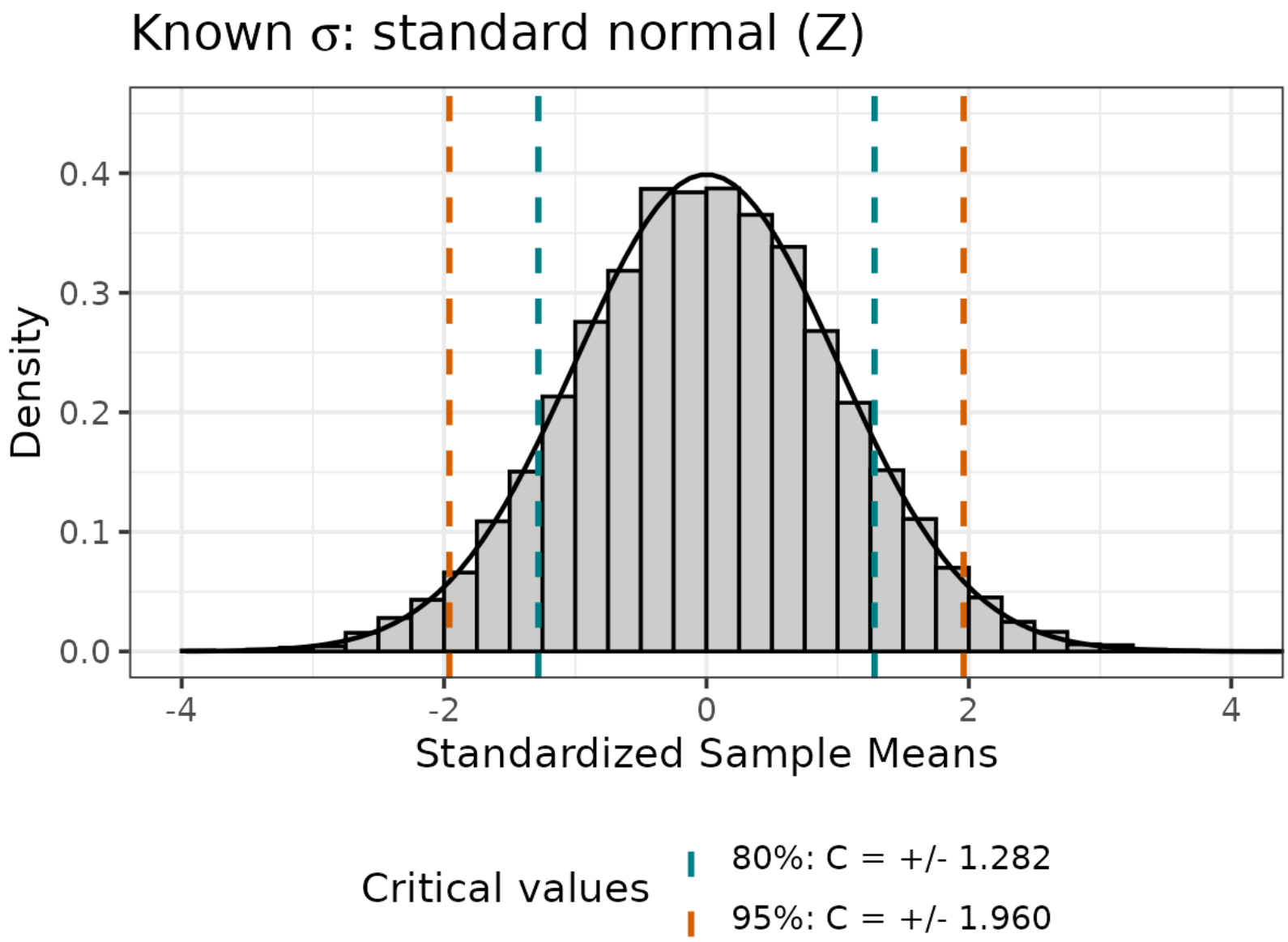
$$\mu = 100, \qquad \sigma = 10.$$

Repeatedly take samples of size $n = 20$.

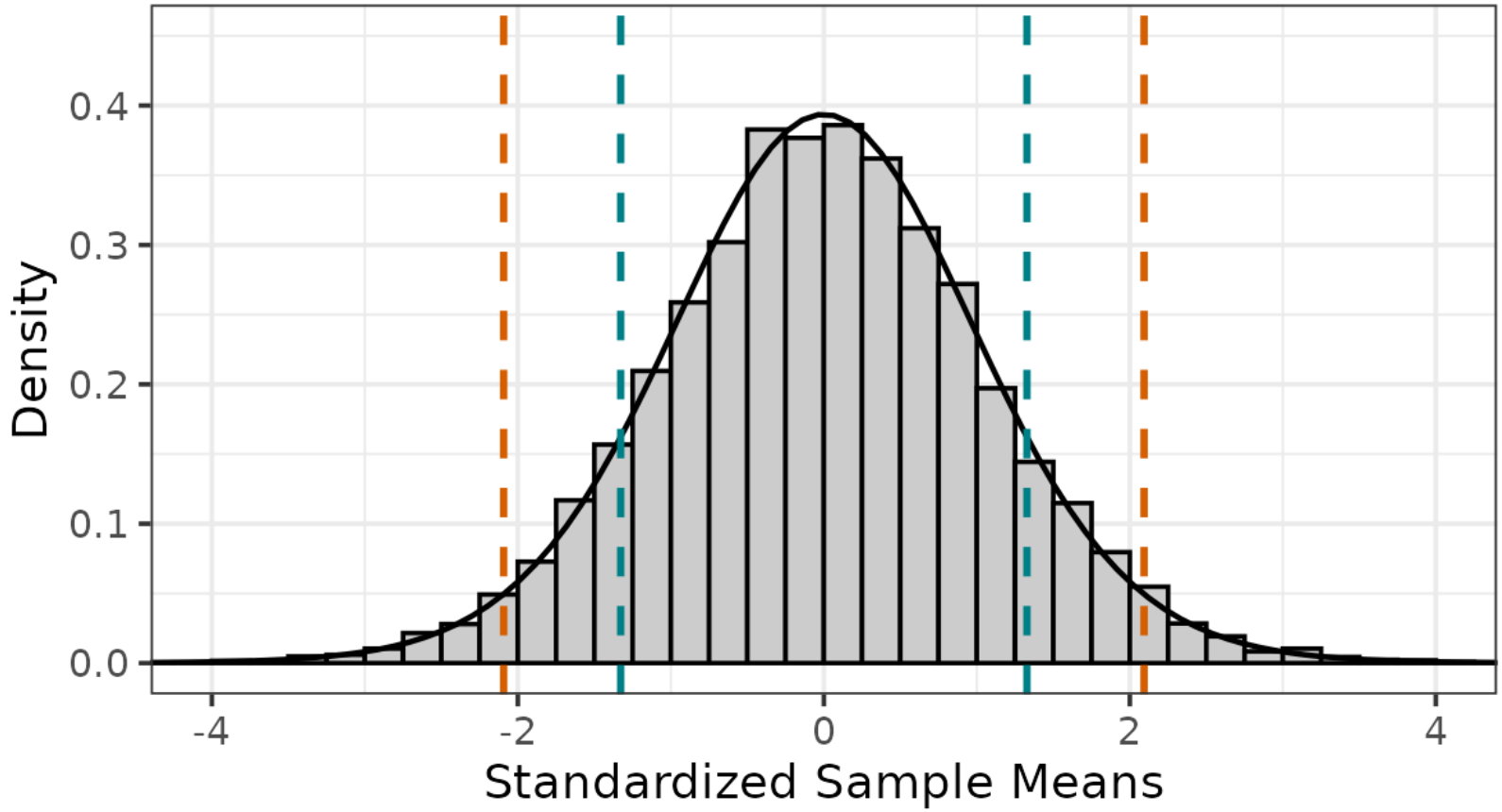
1. Look at the sampling distribution of $\overline{X}$.
2. Standardize using the known σ: we get Z.
3. Replace σ with each sample's s: we get t.

In each case, we can use our distributions to find the cutoffs enclosing the middle 80% and 95%.


Sampling distribution of $\overline{X}$
Density
0.15
0.10
0.05
0.00
90
95
100
105
110
Sample Means
Sampling Ranges
80%: 97.13 to 102.87
95%: 95.62 to 104.38


Known σ: standard normal (Z)
Density
0.0
0.1
0.2
0.3
0.4
-4
-2
0
2
4
Standardized Sample Means
Critical values
80%: C = +/- 1.282
95%: C = +/- 1.960

Unknown σ: t distribution (19 df)



Critical values

80%: C = +/- 1.328

95%: C = +/- 2.093

Flipping the script

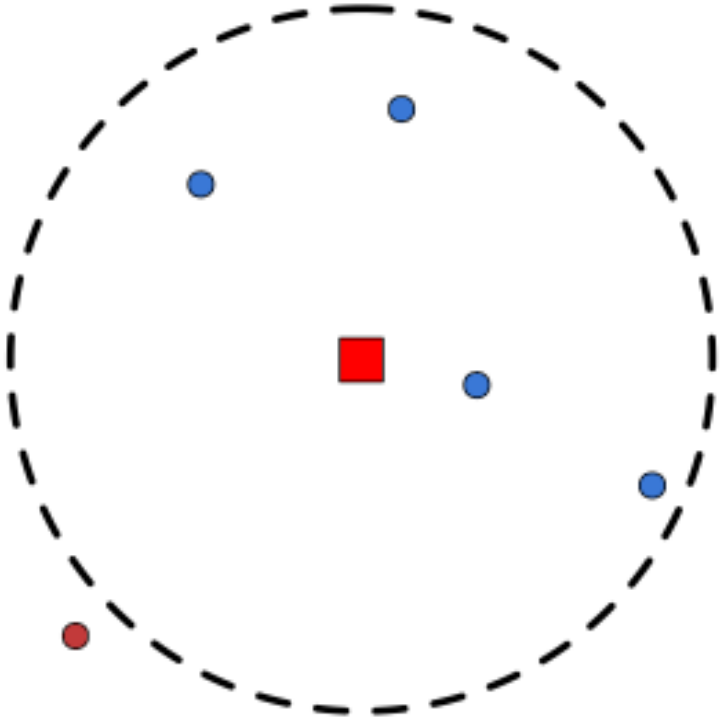
In practice, we don't know the true value μ, but we can hypothesis a value with

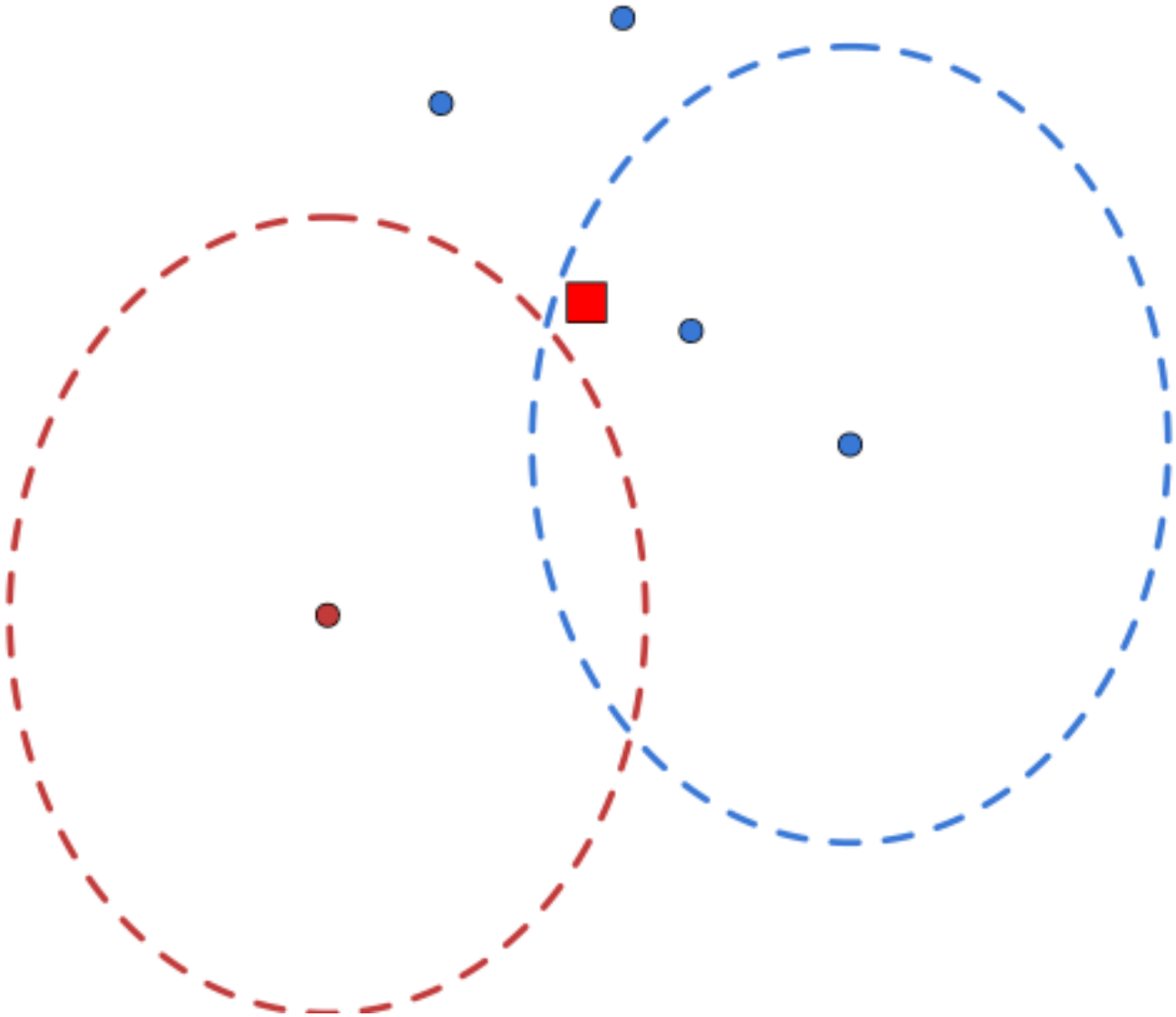
$$H_0 : \mu = \mu_0$$

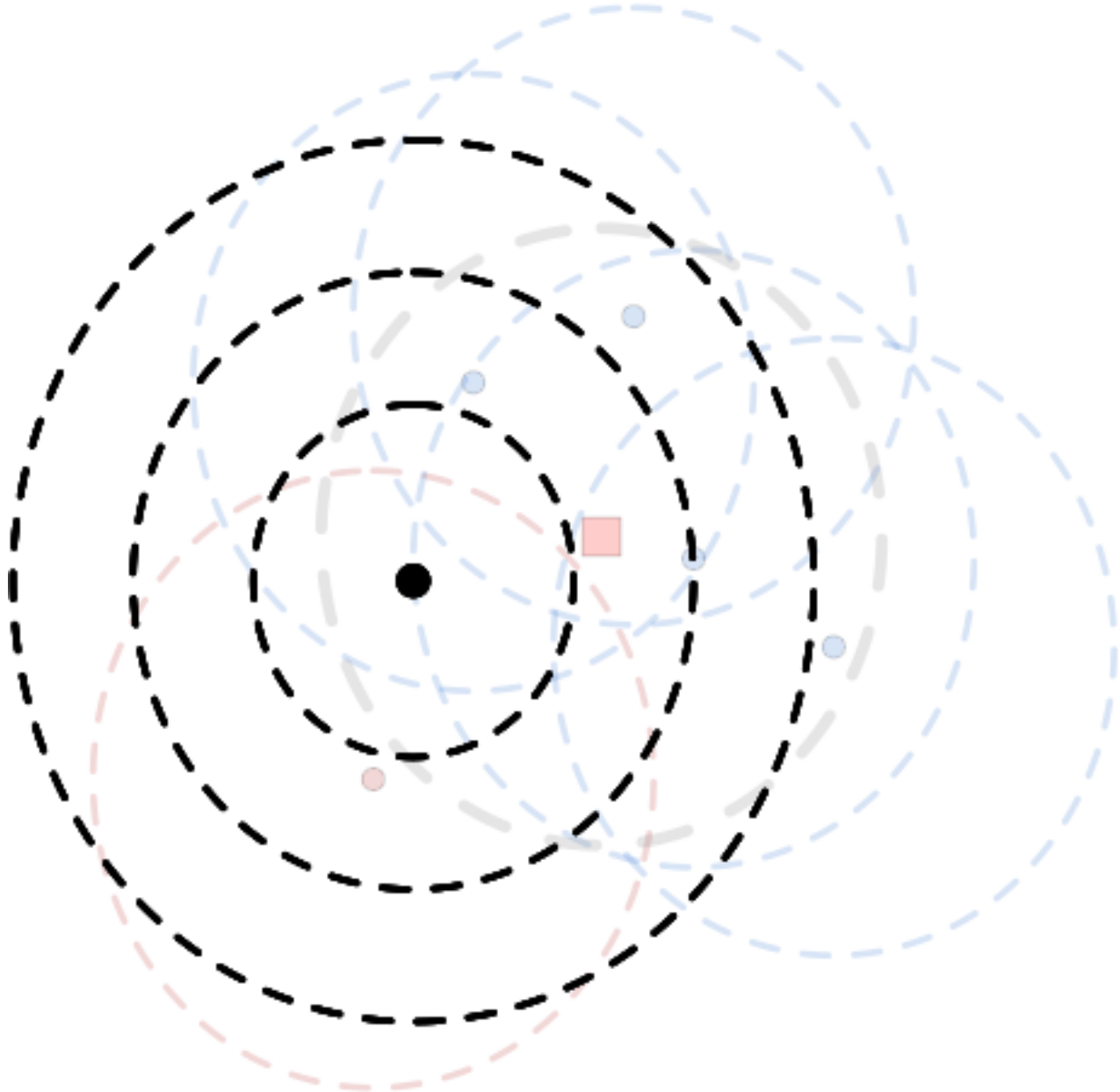
This constitutes our **null hypothesis**

We then go about assuming that H_0 is true, using standardized values to determine if our observed $\overline{X}$ is either within or outside the bounds of what we expect.

This is equivalent to asking: is our hypothesized value μ_0 consistent with our observed $\overline{X}$, at a specified level of confidence?







Why this works

1. The dashed circles illustrate the idea: being close to μ_0 means μ_0 is also close to us.
2. If μ_0 is inside the confidence interval around $\overline{x}$, then

$$\overline{x} - C\frac{s}{\sqrt{n}} < \mu_0 < \overline{x} + C\frac{s}{\sqrt{n}}$$

3. Doing some arithmetic, this is equivalent to

$$-C < \frac{\overline{x} - \mu_0}{s/\sqrt{n}} < C \qquad \Longleftrightarrow \qquad -C < t < C$$

4. So checking whether μ_0 is inside the interval is the same as checking whether $|t| < C$.

Testing our Hypothesis

In terms of assessing our hypothesis, we have two equivalent methods for a given level of confidence:

1. Create a confidence interval around $\overline{X}$
2. Construct a test statistic to compare to a critical value

To illustrate, suppose we collect a sample of $n = 20$, and we find $\overline{x} = 103.5$ with $s = 10$. Our null hypothesis is $H_0 : \mu_0 = 100$, and we'll test at both 80% and 90% confidence

Confidence Interval Method

$$\overline{x} = 103.5, \qquad s = 10, \qquad n = 20, \qquad \mu_0 = 100$$

1. Recall that our confidence interval is of the form
$$\overline{X} \pm C \times \frac{s}{\sqrt{n}}$$
2. For $df = 19$, we have 80% and 90% critical values of $C_{80} = 1.328$ and $C_{90} = 1.729$. Our standard error is equal to $s/\sqrt{n} = 2.236$
3. This gives us confidence intervals of:
$$\begin{aligned} 80\%: &\quad (100.53, 106.47) \\ 90\%: &\quad (99.63, 107.37) \end{aligned}$$
4. Because the 80% interval *does not* contain $\mu_0 = 100$, we *reject the null hypothesis* with 80% confidence
5. Because the 90% confidence interval *does* contain $\mu_0 = 100$, we *fail to reject the null hypothesis* with 90% confidence

Test Statistic Method

$$\overline{x} = 103.5, \qquad s = 10, \qquad n = 20, \qquad \mu_0 = 100$$

1. Recall that a **test statistic** is a statistic that relates our observed data to our null hypothesis. In this case, we have
$$t = \frac{103.5 - 100}{2.2361} = 1.57$$
2. From the last slide, we found 80% and 90% critical values of $C_{80} = 1.328$ and $C_{90} = 1.729$
3. To test our hypothesis, we need only check $|t| < C$:
$$\begin{aligned} 80\%: &\qquad |1.57| > 1.327 \\ 90\%: &\qquad |1.57| < 1.729 \end{aligned}$$
4. Based on this, we *reject our hypothesis* with 80% confidence but *fail to reject our hypothesis* with 90% confidence

Questions

- What are the pieces of the t
- Why does changing H_0 change t
- Why is larger $|t|$ more evidence against null?
- Why do we fail to reject instead of accept null hypothesis if $|t| < C$