Hypothesis Testing cont.

Grinnell College

March 27, 2026

Warm-up

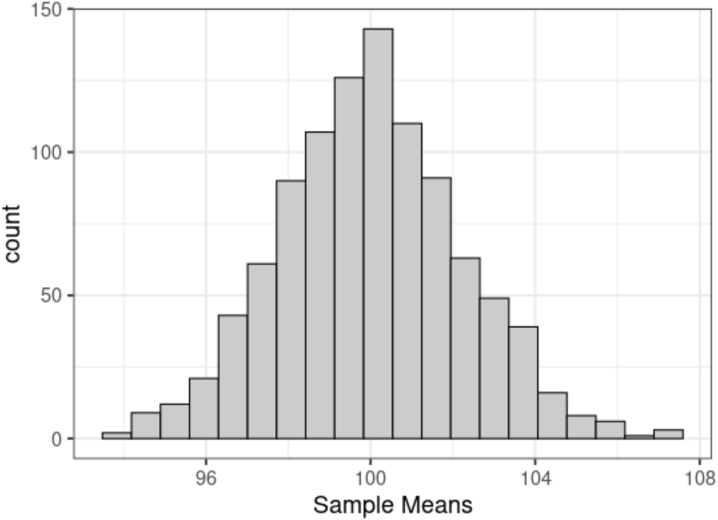
How t dist similar to normal? how diff?
Why does it make sense that wiggle of t function of n
How does changing n change our 90% critical value
What is the point of doing things here when we know parameters. Why is knowing distribution important even though we never actually see that

Our sample

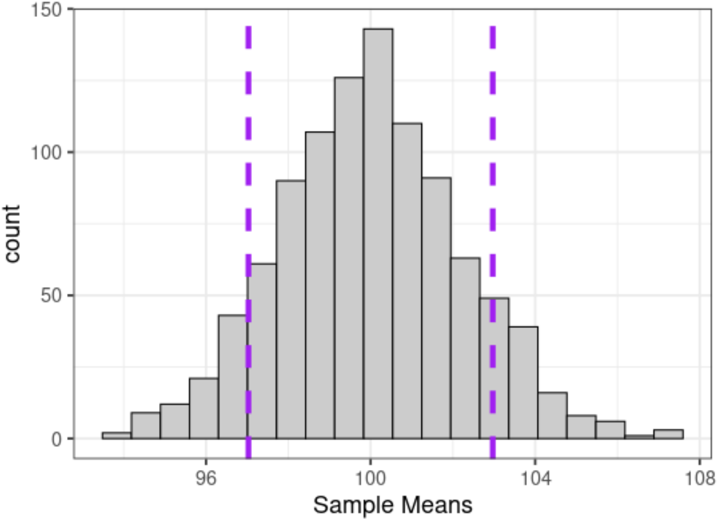
ok here, world where we know the truth. $\mu = 100$, $\sigma = 10$ and $n = 20$. I personally go out, I get $\overline{x} = 97.99$ and find $\hat{\sigma}/\sqrt{n} = 1.95$
Is this a crazy value to get? Is it unlikely? Or is it what I would expect?
How do I answer this? I can check my standardized value (using t, why?) to see if this is crazy or not

Ok everybody sample

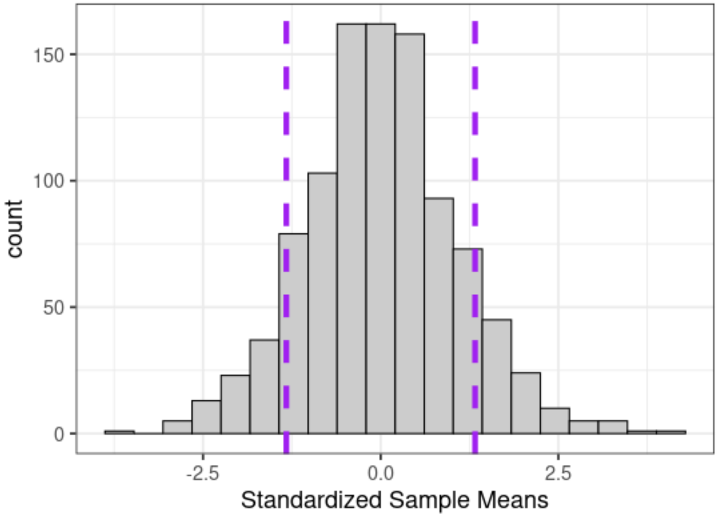
If every single one of us were to sample things from this distribution (where we know the truth), we would find this distribution of values

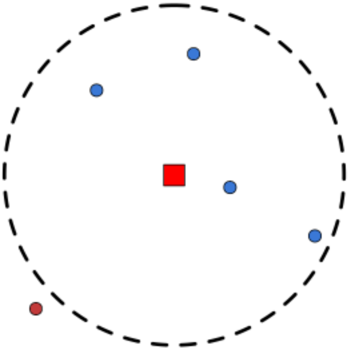


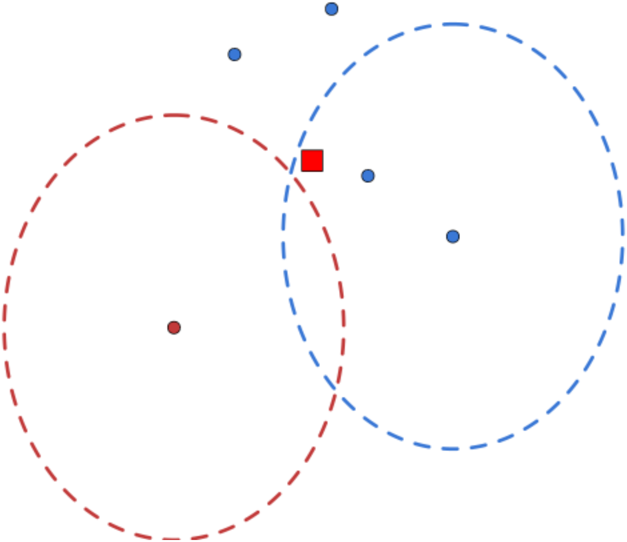
We can find critical values C to give us middle 80% (97.031, 102.969)

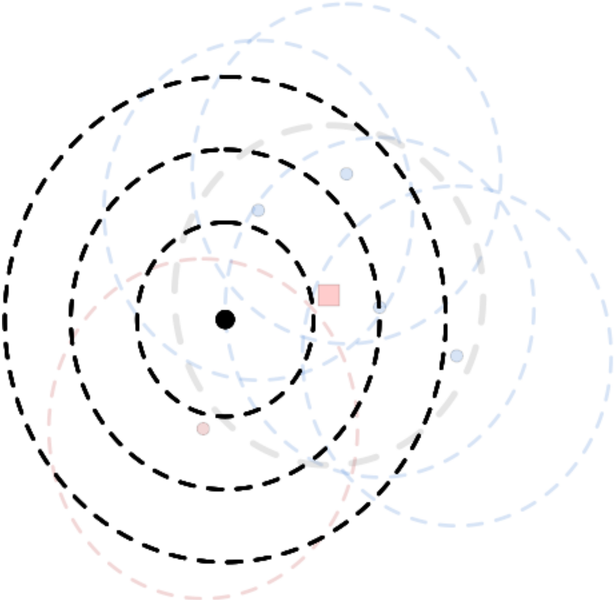


We standardize each mean like this: subtract real μ and divide by $\hat{\sigma}/\sqrt{n}$.
What if we used σ instead? we get this distribution









Repeated Samples

First sample one we got, but see how each other one behaves

Sample	$\overline{x}$	$\hat{\sigma}$	80% CI	Contains μ
1	97.99	1.95	(95.4, 100.59)	TRUE
2	98.91	2.44	(95.67, 102.15)	TRUE
3	97.42	1.83	(94.99, 99.85)	FALSE
4	100.37	2.4	(97.18, 103.55)	TRUE
5	100.53	2.01	(97.86, 103.19)	TRUE
6	104.99	2.89	(101.16, 108.83)	FALSE
7	101.92	1.73	(99.62, 104.21)	TRUE
8	97.5	2.27	(94.49, 100.51)	TRUE
9	100.96	2.73	(97.34, 104.58)	TRUE
10	99.93	2.4	(96.75, 103.11)	TRUE

80% critical values $= 1.3277$

Sample	$\overline{X}$	$\hat{\sigma}/\sqrt{n}$	$t = \frac{\overline{X}-\mu}{\hat{\sigma}/\sqrt{n}}$	$\lvert t \rvert < C$
1	97.99	1.95	-1.03	TRUE
2	98.91	2.44	-0.45	TRUE
3	97.42	1.83	-1.41	FALSE
4	100.37	2.40	0.15	TRUE
5	100.53	2.01	0.26	TRUE
6	104.99	2.89	1.73	FALSE
7	101.92	1.73	1.11	TRUE
8	97.50	2.27	-1.10	TRUE
9	100.96	2.73	0.35	TRUE
10	99.93	2.40	-0.03	TRUE

Why this works

1. We know from dashed circles, if $\overline{X}$ in 80% interval of μ then μ will be in 80% interval around $\overline{X}$
2. If μ is in the interval around $\overline{X}$ then

$$\overline{X} - C\frac{\hat{\sigma}}{\sqrt{n}} < \mu < \overline{X} + C\frac{\hat{\sigma}}{\sqrt{n}}$$

3. Doing some arithmetic we find that this is equivalent to

$$-C < \frac{\overline{X} - \mu}{\hat{\sigma}/\sqrt{n}} < C \qquad \Leftrightarrow \qquad -C < t < C$$

4. We can check this quickly by checking if $|t| < C$

This is a big deal

Simulations are great except we don't actually ever know μ

Things we know from sample

1. Sample mean $\overline{X}$
2. Sample standard deviation $\hat{\sigma}$
3. Sample size n

Things we have to guess

- The true mean μ. But we can make a guess with μ_0

So important: the t statistic combines these two pieces of information. The t statistic is a statement relating what is known to what we guess

t Statistic

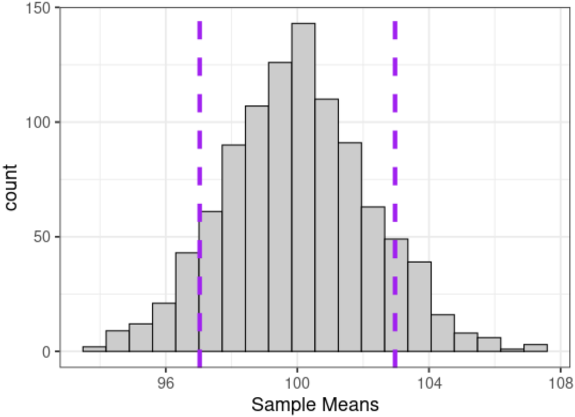
The t statistic is a statement relating our null hypothesis μ_0 to our observed data

$$t = \frac{\overline{X} - \mu_0}{\hat{\sigma}/\sqrt{n}}$$

- If t is large, our observed data is far from our hypothesis
- If t small, observed data more consistent with hypothesis
- What is large? Depends on our critical value

Errors

Remember, ultimately, we have to make a decision: either data is consistent with hypothesis or it is not. By definition, if we have 80% confidence, 20% of our data will fall beyond the true critical values. With 80% confidence, *if the null hypothesis is correct, we will incorrectly reject 20% of the time*



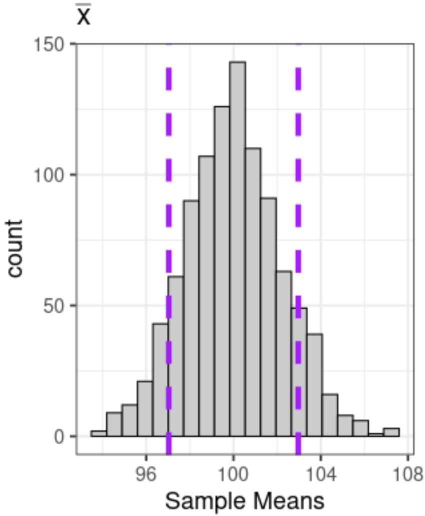
Decisions decisions

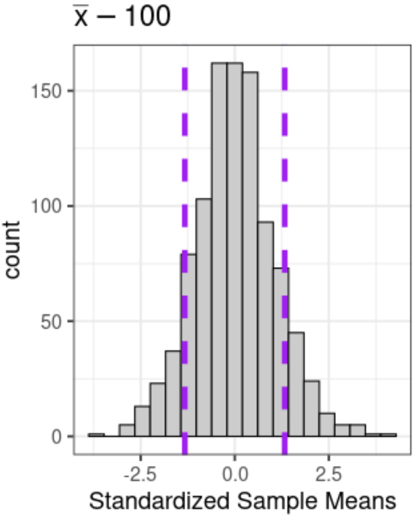
All of this is predicated on our null hypothesis being correct. If our null hypothesis is false, then our sample mean is likely to be even further away. This will be reflected in our t-statistics. Why?

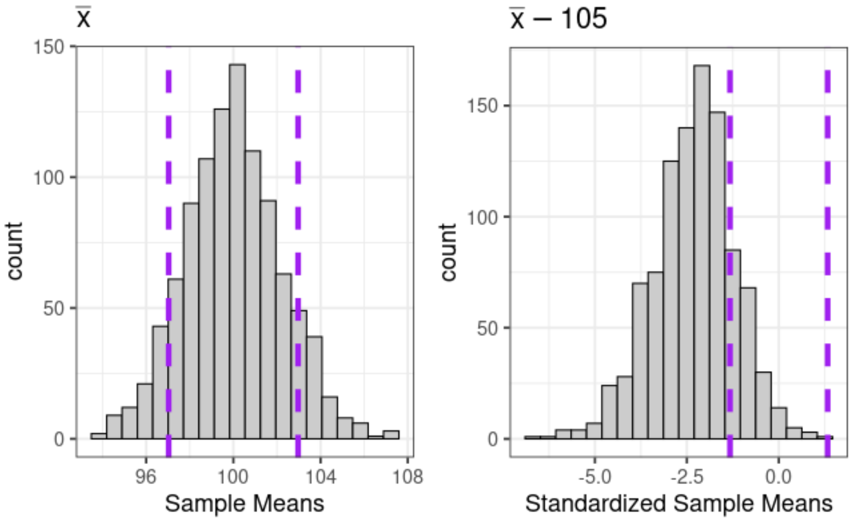
Relating data to hypothesis

For 80% CI and df = 19, we still have C = 1.33

Sample	$\overline{X}$	$\hat{\sigma}/\sqrt{n}$	$t_1 = \frac{\overline{X}-100}{\hat{\sigma}/\sqrt{n}}$	$\lvert t_1 \rvert < C$	$t_2 = \frac{\overline{X}-105}{\hat{\sigma}/\sqrt{n}}$	$\lvert t_2 \rvert < C$
1	97.99	1.95	-1.03	TRUE	-3.59	FALSE
2	98.91	2.44	-0.45	TRUE	-2.49	FALSE
3	97.42	1.83	-1.41	FALSE	-4.14	FALSE
4	100.37	2.40	0.15	TRUE	-1.93	FALSE
5	100.53	2.01	0.26	TRUE	-2.23	FALSE
6	104.99	2.89	1.73	FALSE	-0.00	TRUE
7	101.92	1.73	1.11	TRUE	-1.78	FALSE
8	97.50	2.27	-1.10	TRUE	-3.31	FALSE
9	100.96	2.73	0.35	TRUE	-1.48	FALSE
10	102.43	2.46	0.99	TRUE	-1.05	TRUE


$\bar{x}$
150
100
50
0
count
96
100
104
108
Sample Means


$\bar{x} - 100$
150
100
50
0
count
-2.5
0.0
2.5
Standardized Sample Means


x̄
count
150
100
50
0
96
100
104
108
Sample Means
x̄ − 105
count
150
100
50
0
-5.0
-2.5
0.0
Standardized Sample Means

Harping again on t

t-distribution relate observed data to hypothesis

Different μ_0 means different t

Think about what it's saying

What do the pieces mean? They all fit together

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 $|t| < C$