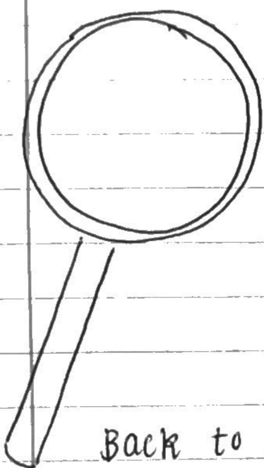
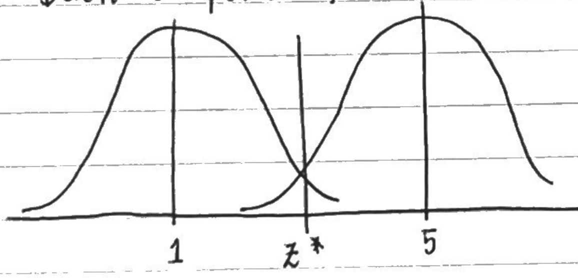
Testing Means



Previously, Exam 2 covered:

- Making H_0, H_1
- using normal-based z-test
- using normal-based CI
- p-value
- power

Back to power for one second, look at these sampling dists.



$H_0: \mu = 1$

$H_A: \mu = 5$

Power = $1 - \beta$

= 1 - Type II Error Rate

You can increase power by

1. Increasing n
   (creates tighter sampling distributions)
2. Increase effect size
   (change H_A)
3. Increase α (Prob. of Type I Error), increase overall rejection region (z^* will be deeper in H_0 dist than H_A)

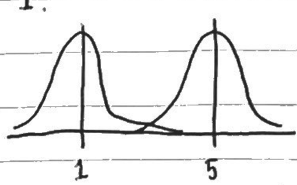
	H_0 True	H_A True
FTR H_0	✓	β (Type II) P of
Reject H_0	α (Type I) Prob of making	✓

$$\boxed{\text{Power} = P(\text{Reject } H_0 \mid H_A \text{ True})}$$

This is overlap between the 2 distributions.

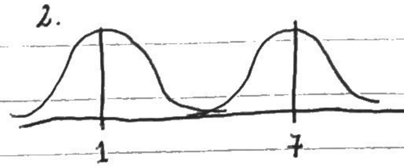
Visually,

1.



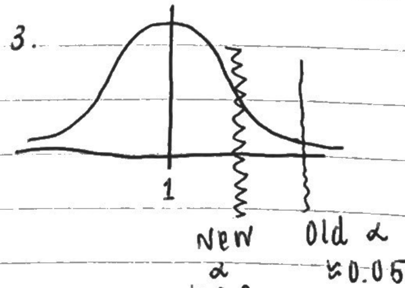
sampling distributions stay wound up near mean

2.



still needs to be reasonable

3.



$\alpha = P(\text{Reject } H_0 \mid H_0 \text{ True})$

What is the first thing you must do when you want to test some H_0, H_A (H_1)?

- Check assumptions!
- These vary...

<u>Example</u> Bob the builder says that on average he gets "the job done" twice per day. Since Bob's team made a big switch and hired lots of new grad engineers, he wants to know if that average has changed. We have an SRS to check.

	4.75	4.4	3.8	5.2	4.2	
4.7	5.12	4.9	6	2	2.3	
	1.5	2.2	3.8	3.7	6.5	6.2

<u>Can we z-test (CI)?</u>

- Random samples that are independent
- Samples are approx. normal
- Population std. dev. is known

OKAY. PH142 IS OUT FOR THE SEMESTER. STATISTICS IS BROKE!!! SORRY, BOB!!!

<u>"Can we fix it?"</u>

$$z = \frac{\bar{x} - \mu_0}{\sqrt{\frac{\sigma^2}{n}}} = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$$

What's the problem?

We don't have σ^2!

(The population parameter)

We can estimate σ (the parameter) with s (the sample sd).

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

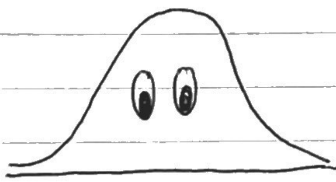
But did we fix it?
Can we use this under the same instances/assumptions as before?

SHORT ANSWER : NO.

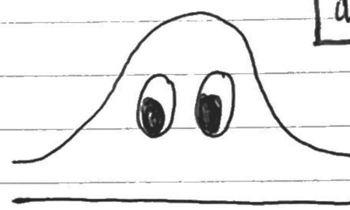
Meet Zee's cousin Tee

$\boxed{df = n-1}$

Normal Distribution "Z"



T



- The older cousin
- 2 parameters (μ, σ)
- Long, skinny tails
- Tall in the center

- The younger one
- Chubbier in the tails
- 1 parameter (degrees of freedom)
- When more free, looks more like Z

Assumptions for t-test

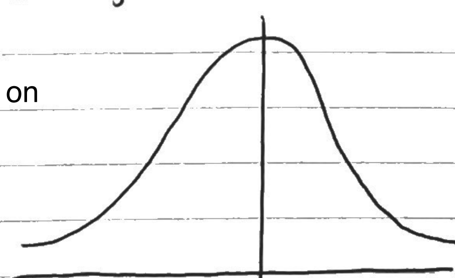
- Random sample, independent, continuous
- underlying population ≈ Normal-looking
- No outliers when n is "small"
- See lecture notes for exact assumptions based on how big sample size (n) is

Back to Bob

- Plot histogram (R)
- calculate $t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$
- Calculate p-value (R) → pt()

?pt

- Interpret
- Conclude



← all your possible t-values here →

Centered at 0

Approximate, not to scale

T  is versatile!

Soon, we will also be testing hypotheses

$\ast$ $H_0: \mu_1 = \mu_2$

$H_1: \mu_1 \neq \mu_2$

"Do the means of 2 groups equate to each other?"

TWO-SAMPLE T-TEST

$\ast$ H_0: All $D_i = 0$

H_1: Otherwise

"pairwise comparisons"

PAIRED T-TEST

What is the first step to testing H_0, H_1?

AGAIN: CHECK ASSUMPTIONS!!!

Two Sample

- Two SRS's, independent
- Both populations ~ Normal distribution

↓

$\left.\begin{array}{l}\text{- Similar shapes} \\ \text{- No outliers}\end{array}\right\}$ Also suffices

Paired

- subjects / individuals that can be matched
- 1 sample t-test

Bob the builder (One sample t-test)

We wish to test the hypotheses of getting the job done. Does Bob the builder get the job done 2 times per day or not?

$H_0 : \mu = 2$
$H_1 : \mu \neq 2$

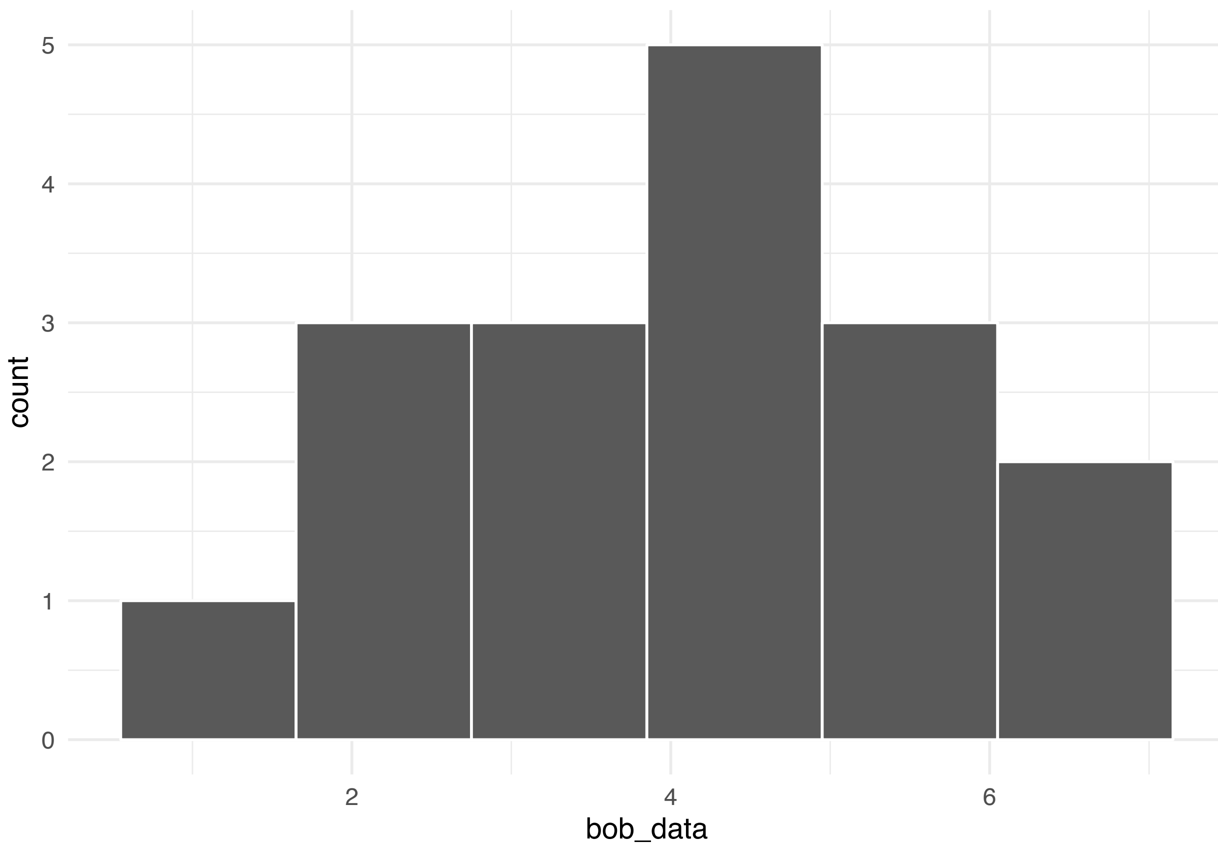
```
mu_0 <- 2
```

We have a random sample of independent observations.

```
bob_data <- c(4.75, 4.4, 3.8, 5.2, 4.2, 4.7, 5.12, 4.9, 6, 2, 2.3, 1.5, 2.2, 3.8, 3.7, 6.5, 6.2)
```

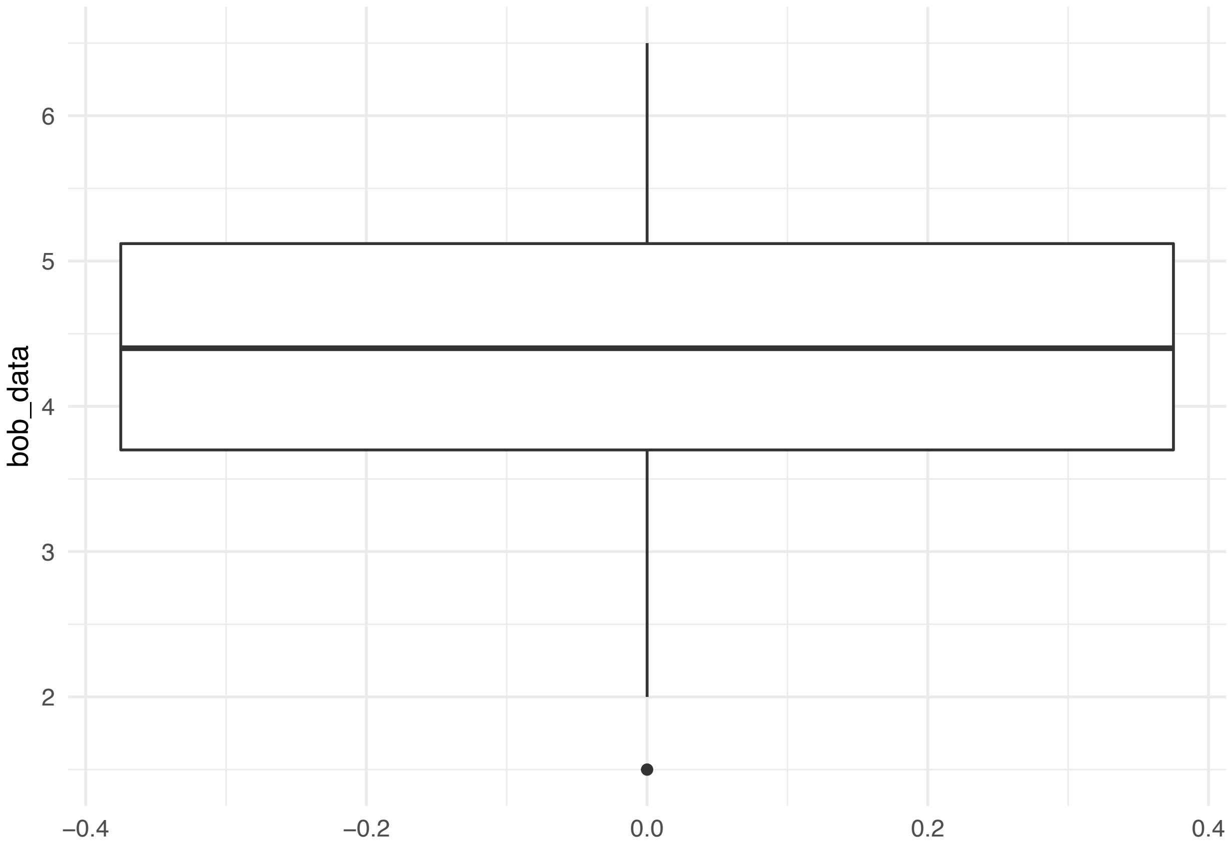
The histogram doesn't look too bad? We have enough data (n=17).

```
library(ggplot2)
ggplot(data.frame(bob_data=bob_data), aes(x=bob_data)) +
  geom_histogram(binwidth=1.1, col="white", lwd=0.5) +
  theme_minimal()
```



There is one outlier.

```
ggplot(data.frame(bob_data=bob_data), aes(y=bob_data)) +
  geom_boxplot() +
  theme_minimal()
```



A t-test is robust, so with caution from above, we'll proceed.

"By Hand" Calculation

Meaning: Use `R` like it is a simple calculator.

```
# * THIS IS BY HAND
n         <- length(bob_data)
x_bar     <- mean(bob_data)
sample_sd <- sd(bob_data)
c(n=n, x_bar=x_bar, sample_sd=sample_sd)
```

```
##         n     x_bar sample_sd
## 17.000000  4.192353  1.495157
```

We are using $t = \frac{\bar{x}-\mu_0}{s/\sqrt{n}}$. Check out what t equals.

```
# * THIS IS CONTINUING THE BY HAND CALCULATION
t     <- (x_bar-mu_0) / (sample_sd/sqrt(n))
t
```

```
## [1] 6.045722
```

By definition, we have degrees of freedom as 1 minus the number of observations.

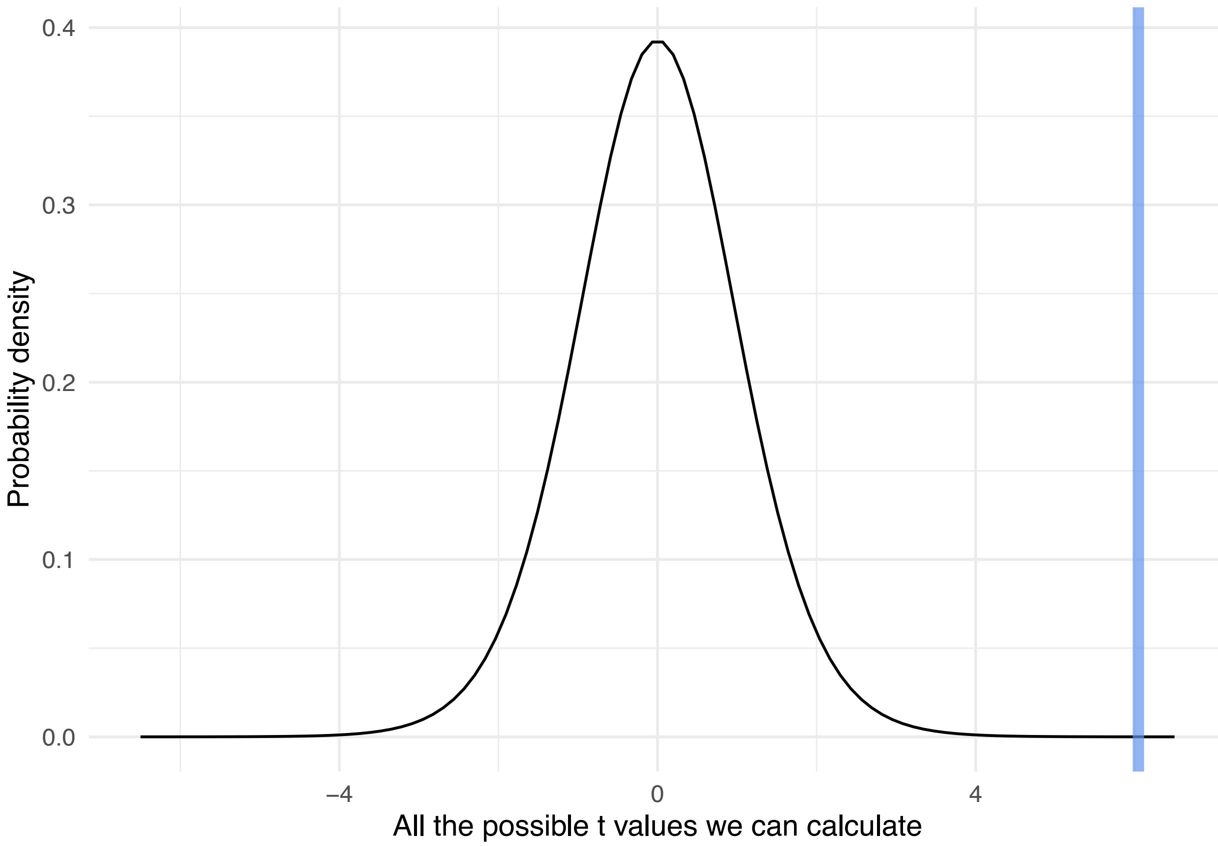
```
# * THIS IS CONTINUING THE BY HAND CALCULATION
df    <- n-1
df
```

```
## [1] 16
```

Let's see where t lands on our distribution. I am plotting a t-distribution with $df = n - 1 = 16$.

```
# * THIS IS THE T-DISTRIBUTION WE ARE COMPARING AGAINST
x  <- seq(-6.5, 6.5, length=100)
hx <- dt(x, df=n-1)
t_dist <- data.frame(cbind(x,hx))

ggplot(t_dist, aes(x=x, y=hx)) +
  geom_line() +
  geom_vline(xintercept=t, col="cornflowerblue", lwd=2, alpha=0.7) +
  xlab("All the possible t values we can calculate") +
  ylab("Probability density") +
  theme_minimal()
```



Can you guess what our p-value will be? (Big? Small?) We're going to take the area of being above the blue line on the above distribution as our p-value. (The probability of rejecting H_0 given thatn H_0 is actually the truth.)

```
# * THIS IS CONTINUING THE BY HAND CALCULATION
p_val <- 2*(1 - pt(q=t, df=df))
p_val
```

```
## [1] 1.699117e-05
```

Look at slides to see interpretation of p-value!

Also, question: Would the corresponding confidence interval include or not include $\mu_0 = 2$?

“Using R” Calcuation

Meaning: Use more than just simple `R` functions.

```r
# * THIS IS "USING R"
test <- t.test(x=bob_data, alternative="two.sided", mu=2)
test$p.value
```

```
## [1] 1.699117e-05
```