Regression: is the slope real?

Topic 13 · Tuesday, March 9

EDUC 15: Statistics for Education Research
UCI School of Education

Where are we?

Description

Summarize data

Visualize data

Prediction

Does one score go with another?

Use one score to predict another

Is the slope real?

We are here



Explanation

Peter the Anteater at a UC Irvine basketball game.

Same four steps, one new statistic

Step 1: State a hypothesis about the slope in the population

Step 2: Set the criteria: how far from 0 must our slope be to count as unexpected?

Step 3: Observe a sample and find b. How unexpected is it, if the true line is flat?

Step 4: Reject or maintain H_0

Plan for today

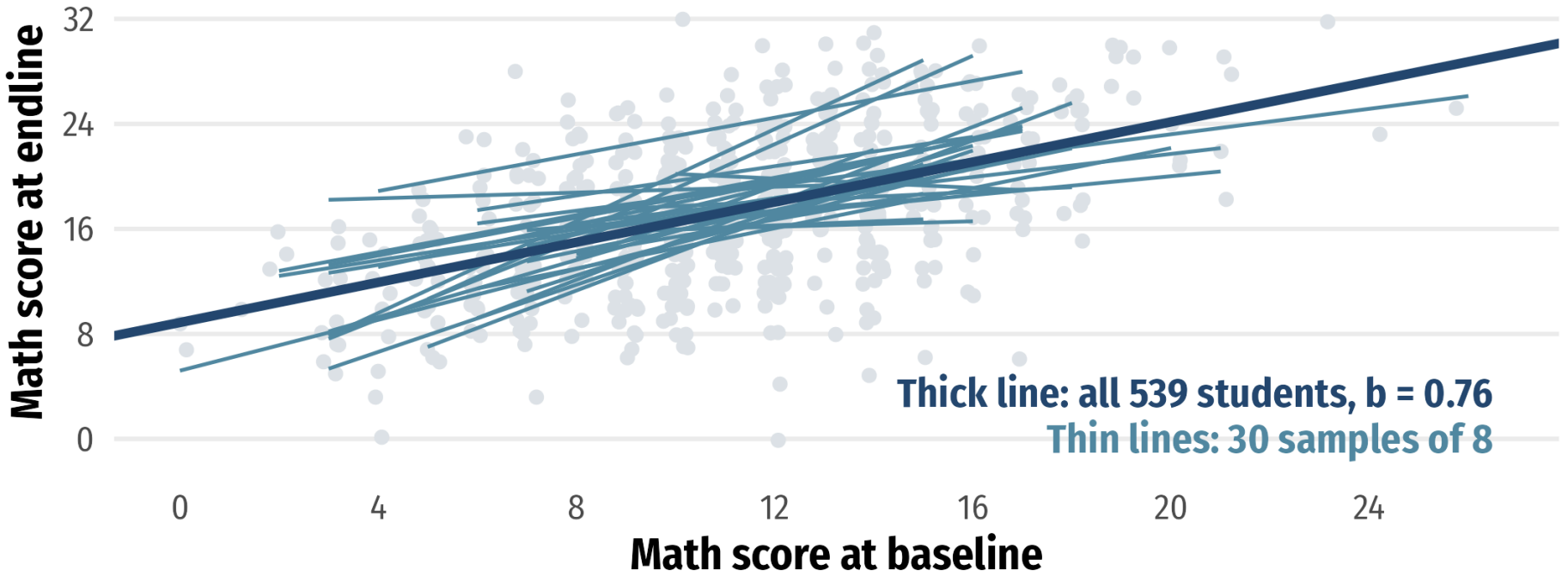
A slope bounces

Testing the slope

F, r^2 and R's output

A slope bounces

Eight students at a time, the line swings



Treat our 539 students as a population. Each sample of 8 gets its own line

Data: Muralidharan, Singh, and Ganimian (2019). Samples drawn by computer, with replacement. Dots are nudged.

β is the slope for everyone

	Our estimate	In the population
Mean	M	μ
Correlation	r	ρ
Slope	**b**	**β**

2.25 and 0.76 are values of b. Never call an estimate beta

A population line uses plain numbers, such as Y = 2.00 + 0.50X, so β = 0.50 here

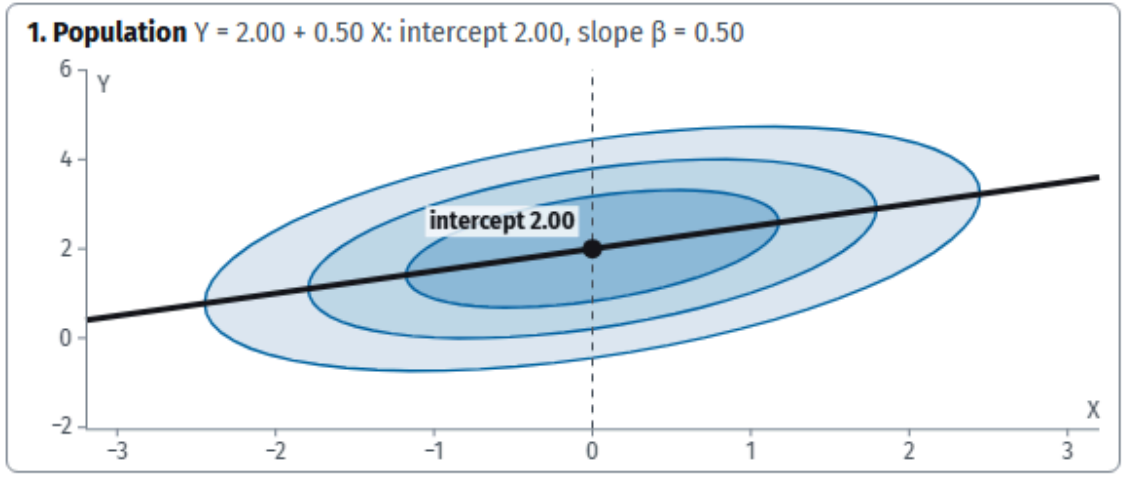
Careful: in Topic 7, β was the Type 2 error rate. Here it's a slope. Same letter, different job

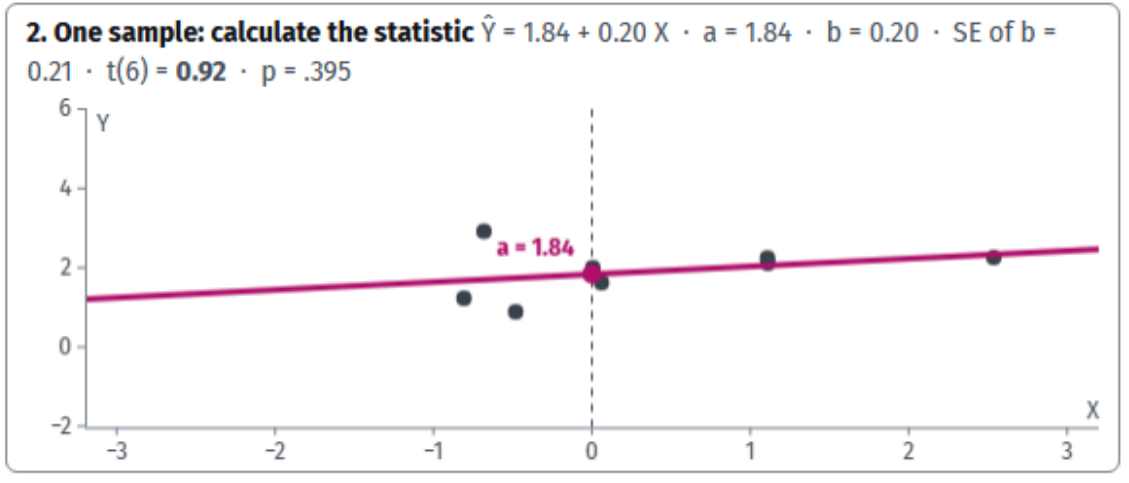
Same frame, new statistic

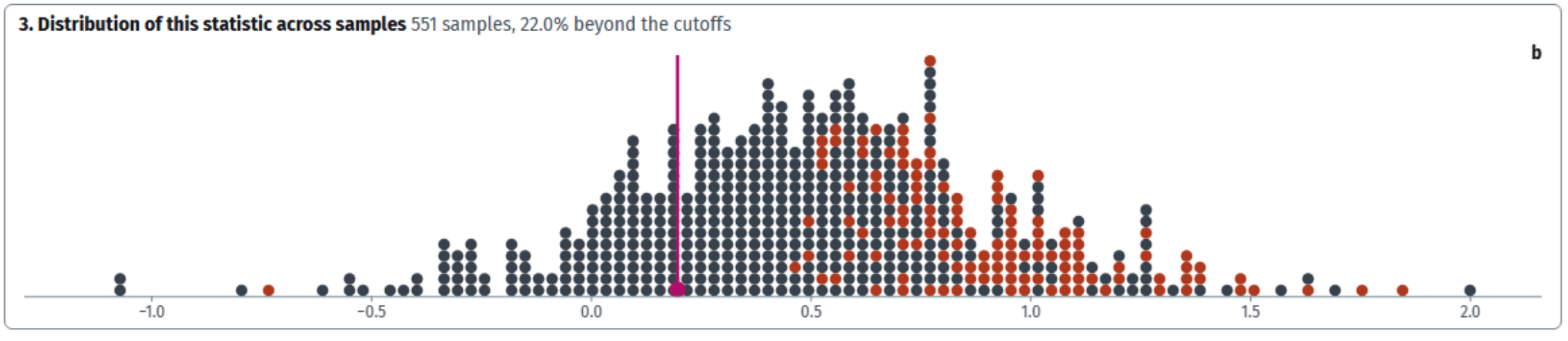
Blueprint step	Panel 1: population	Panel 2: one sample	Panel 3: many samples
2. One mean	A curve with μ	n scores, one M	A pile of M
4. Two means	Two curves	Two groups, $M_1 - M_2$	A pile of $M_1 - M_2$
5. Two or more means	Three curves	Three groups, one F	A pile of F
6. A correlation	A cloud with ρ	n dots, one r	A pile of r
7a. A slope	**A line with β**	**n dots, one b**	**A pile of b**

The three panels never change. Only the statistic does

Step 7a: Every sample draws its own line

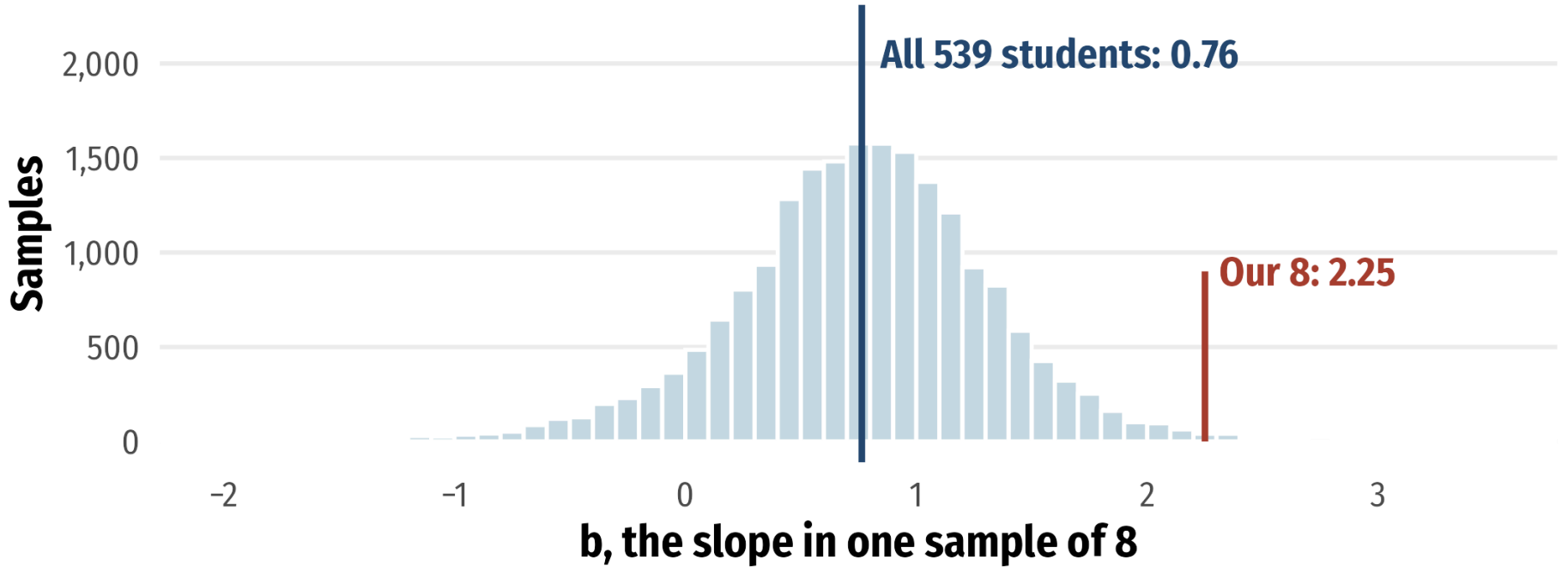






Simulator after Kristoffer Magnusson, rpsychologist.com. On your phone: [the simulator page](#).

Our 8 students' b came from the far tail



The middle 95% runs from −0.48 to 1.87, and only 0.9% of samples reach 2.25. The SD of this pile is the standard error of b

Data: Muralidharan, Singh, and Ganimian (2019). 20,000 samples of 8, drawn with replacement.

Predict it: Which slope bounces more?

Respond at **PollEv.com/adb**

Two studies draw 30 Mindspark students, again and again. Study A takes any baseline score. Study B takes only baseline scores from 8 to 14. Whose b bounces more from sample to sample?

A Study A

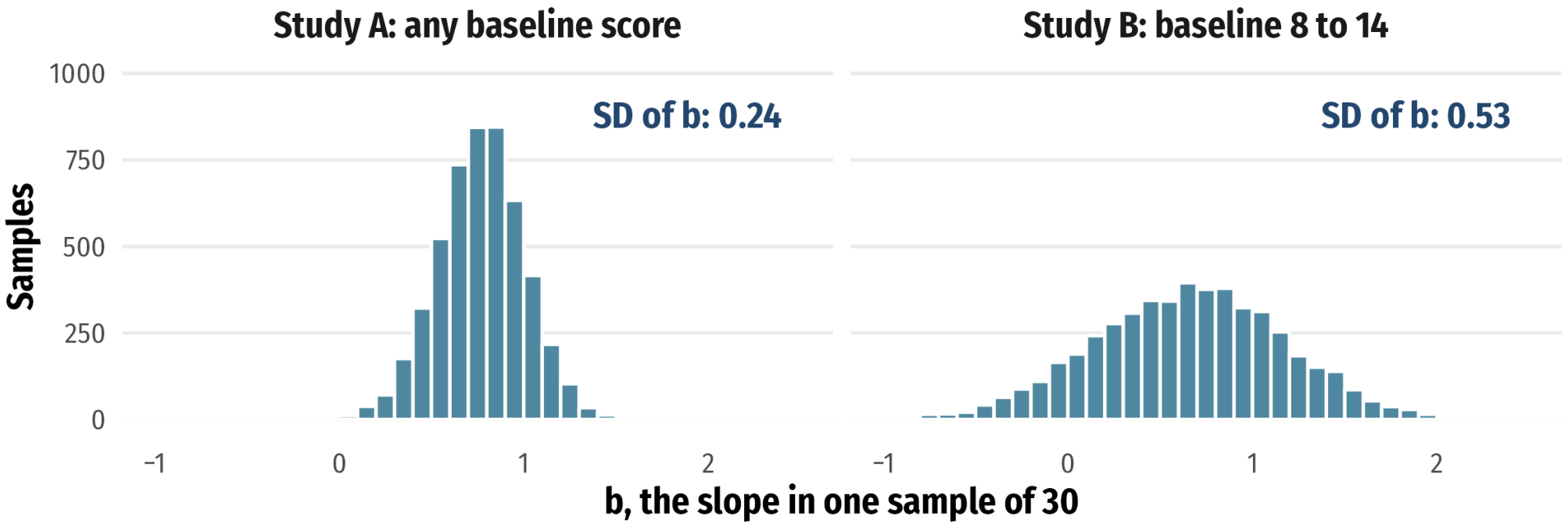
B Study B

C About the same

D We can't tell without β

Data: Muralidharan, Singh, and Ganimian (2019).

Predict it: Narrow X, wobbly slope



Same number of students. The only change is how far apart they are on X

Data: Muralidharan, Singh, and Ganimian (2019). 5,000 samples of each, drawn with replacement.

Testing the slope

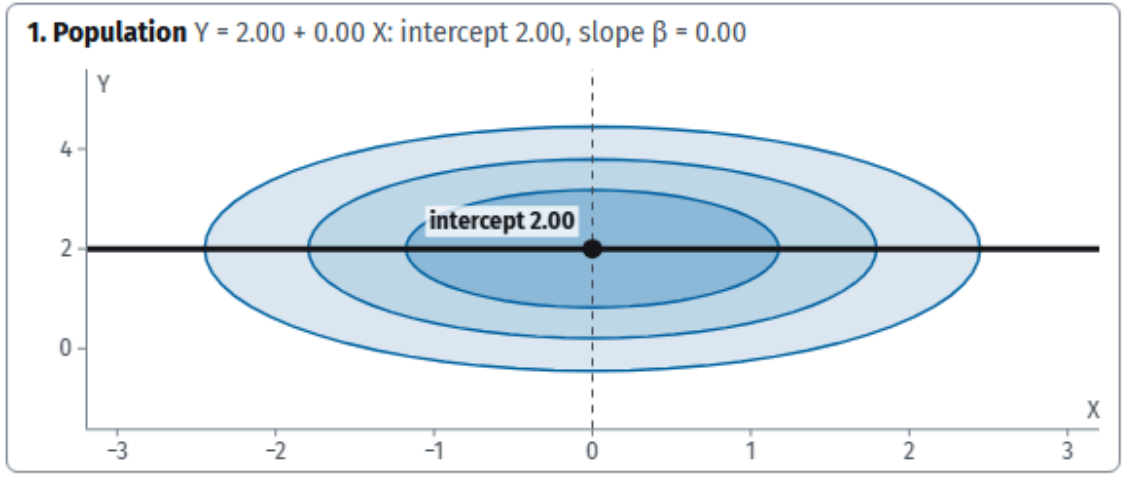
Steps 1 and 2: H_0 says the line is flat

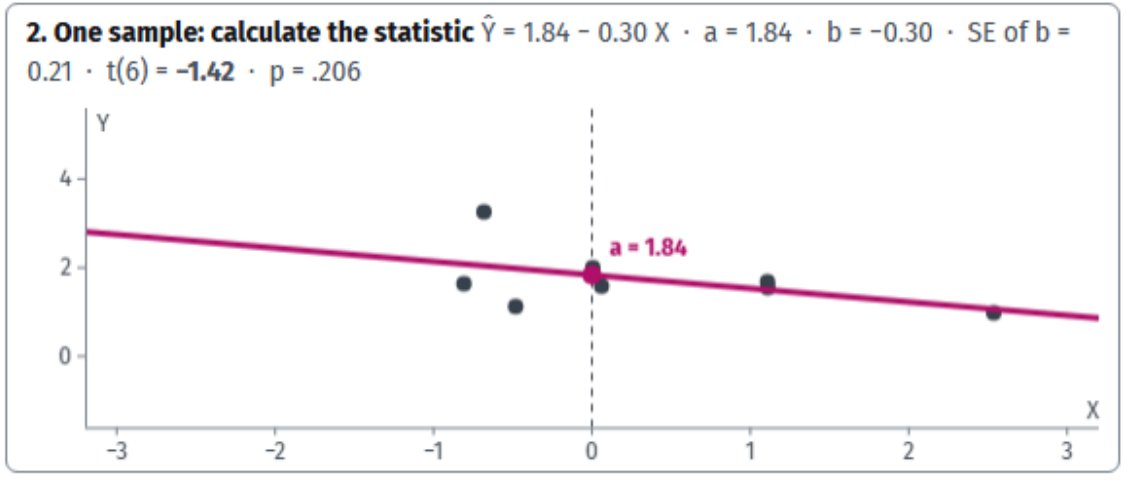
Step	In words	Our 8 students
1. State H_0	Baseline predicts nothing about endline, in the population	H_0: $\beta = 0$ H_1: $\beta \neq 0$
2. Set the criteria	Two tails, $\alpha = 0.05$, and a t curve	$df = n - 2 = 6$ t(6) beyond ±2.447

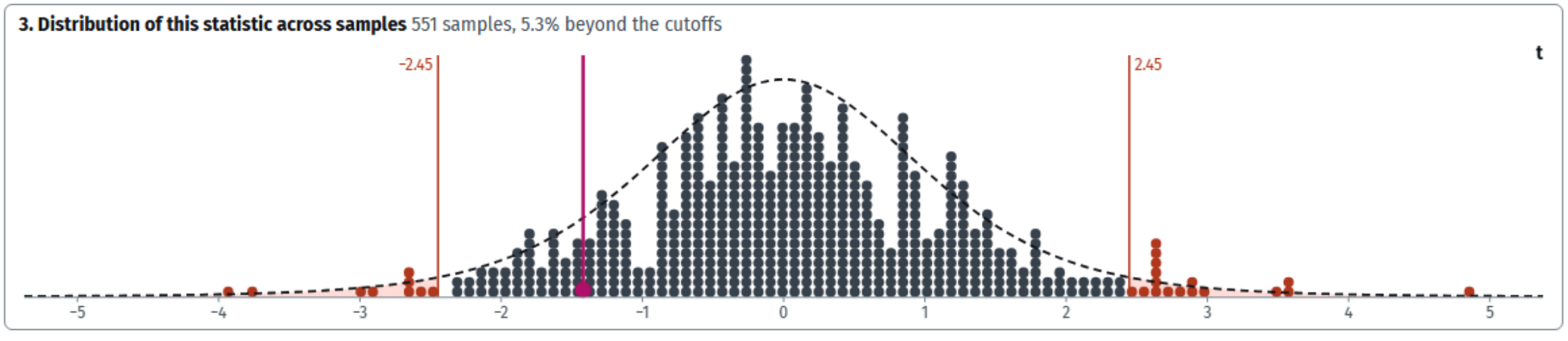
We found b = 2.25. Would a slope that steep be surprising, if the true line were flat?

Data: Muralidharan, Singh, and Ganimian (2019). Students 25 to 32.

Under H_0, sample lines still tilt







Simulator after Kristoffer Magnusson, rpsychologist.com.

Three things narrow the pile

$$\hat{\sigma}_b = \sqrt{\frac{MS_{residual}}{SS_X}} \qquad \text{where} \qquad MS_{residual} = \frac{SS_{residual}}{n-2}$$

Tight dots: small $MS_{residual}$

Wide X: big SS_X

More students: big SS_X

It has the shape of $\hat{\sigma}_M = \hat{\sigma} / \sqrt{n}$: the scatter, over how much information we have

In the book: $\sqrt{MS_{residual}}$ is the standard error of estimate. Software calls $\hat{\sigma}_b$ the "Std. Error"

Step 3: The same fraction again

$$MS_{residual} = \frac{58.5}{6} = 9.75 \qquad \hat{\sigma}_b = \sqrt{\frac{9.75}{24}} = 0.637$$

$$t = \frac{b - \beta}{\hat{\sigma}_b} = \frac{2.25 - 0}{0.637} = 3.53$$

Top: how far b is from H_0's β of 0

Bottom: how far a b typically falls

3.53 is beyond ±2.447, so we reject H_0: t(6) = 3.53, p = 0.012

Data: Muralidharan, Singh, and Ganimian (2019). Students 25 to 32. $SS_{residual}$ is from Topic 12, SS_X from Topic 11.

It's Topic 11's t, to the last decimal

Our 8 students	Bottom of the fraction	t
Topic 11: test r	$\hat{\sigma}_r = \sqrt{((1 - r^2) / (n - 2))} = 0.233$	0.822 / 0.233 = **3.53**
Today: test b	$\hat{\sigma}_b = \sqrt{(MS_{residual} / SS_X)} = 0.637$	2.25 / 0.637 = **3.53**

b is r, rescaled into points. The same rescaling turns $\hat{\sigma}_r$ into $\hat{\sigma}_b$, so the fraction doesn't change

H_0: $\rho = 0$ and H_0: $\beta = 0$ are one question. If r is 0, b is 0

Data: Muralidharan, Singh, and Ganimian (2019). Students 25 to 32.

Eight students or 539: reject H_0

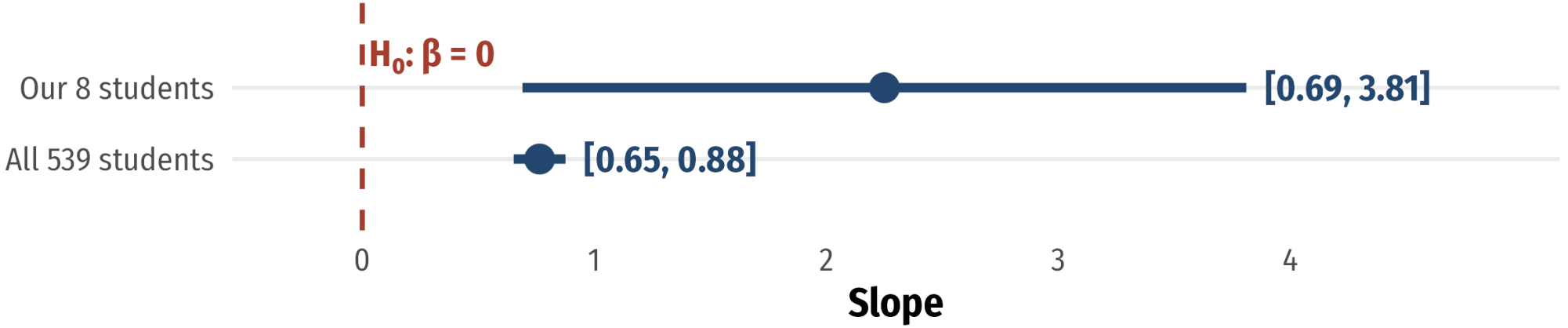
Step	Our 8 students	All 539 students
1. State H_0	$\beta = 0$	$\beta = 0$
2. Set the criteria	$t(6)$ beyond ± 2.447	$t(537)$ beyond ± 1.964
3. Find the spread	$\hat{\sigma}_b = \sqrt{(9.75 \,/\, 24)} = 0.637$	$\hat{\sigma}_b = \sqrt{(26.66 \,/\, 8{,}269)} = 0.057$
3. Compute t	$b = 2.25$, $t = 3.53$	$b = 0.76$, $t = 13.46$
4. Decide	Reject H_0: $3.53 > 2.447$	Reject H_0: $13.46 > 1.964$

Eight students: $b = 2.25$, $t(6) = 3.53$, $p = 0.012$, $r^2 = 0.675$

All 539: $\hat{Y} = 8.87 + 0.76X$, $t(537) = 13.46$, $p < 0.001$, $r^2 = 0.252$

Data: Muralidharan, Singh, and Ganimian (2019).

An interval for β, the Topic 8 way



8 students: 2.25 ± 2.447 × 0.637

All 539: 0.76 ± 1.964 × 0.057

Neither interval holds 0, so both tests reject. It's the rule from Topic 8

Data: Muralidharan, Singh, and Ganimian (2019).

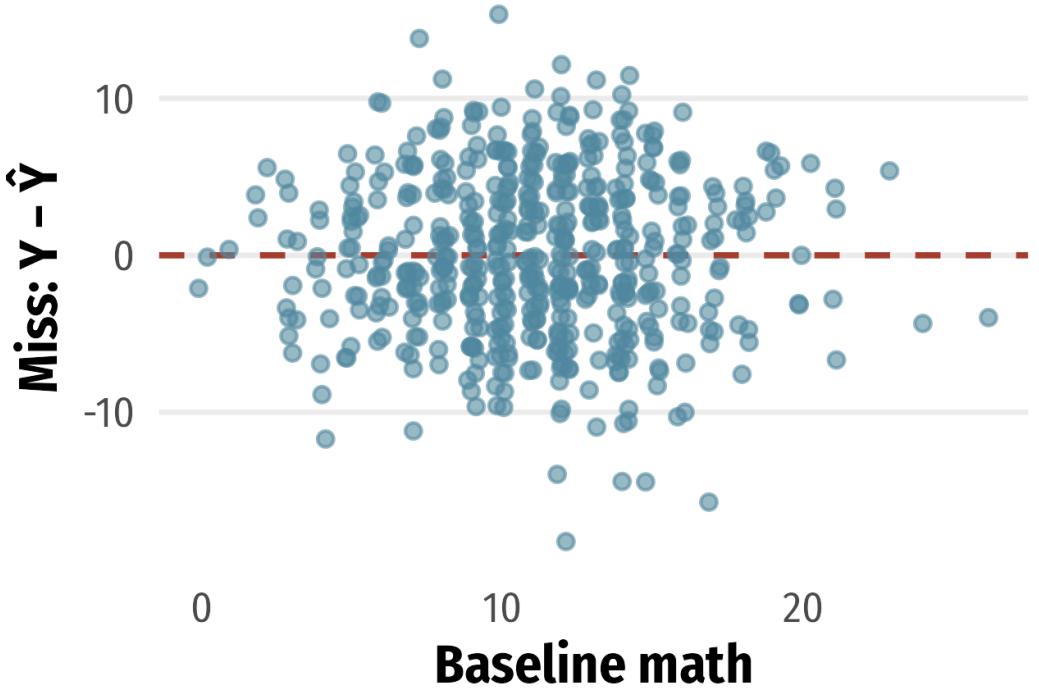
The test trusts four things

Straight: a line, not a curve

Even spread: as wide at every X

Independent: one miss says nothing about another

Normal misses: a bell around 0. Matters less with many students



Plot the misses first. Ours form an even band around 0, so the test is fair

Data: Muralidharan, Singh, and Ganimian (2019). All 539 students, jittered so tied scores show.

Real, big and causal are three questions

Question	What answers it	All 539 students
Is the slope real?	The test: t and p	Yes: $t(537) = 13.46$, $p < 0.001$
How big is it?	b in points, and r^2	0.76 points per point, $r^2 = 0.252$
Did X cause Y?	The design, not the test	We can't say: no lottery set baseline scores

Rejecting H_0 says the line isn't flat, not that it's steep. With enough students, a tiny slope passes too

Data: Muralidharan, Singh, and Ganimian (2019).

Try it: Nine fifth graders' slope

$n = 9$

$b = 0.489$

$SS_X = 412$

$SS_{residual} = 191.2$

Test H_0: $\beta = 0$ with $\alpha = 0.05$ and two tails. The cutoff is $t(7) = \pm 2.365$

Find $MS_{residual}$, then $\hat{\sigma}_b$, then t. Then decide

First, predict: Topic 11 tested these students' r. What will t be today? Then work alone, compare with a neighbor, and vote with hands: reject or maintain?

Data: Muralidharan, Singh, and Ganimian (2019).

Try it: Answer

$MS_{residual}$ = 191.2 / 7 = **27.31**

$\hat{\sigma}_b = \sqrt{(27.31 / 412)} =$ **0.257**

t = 0.489 / 0.257 = **1.90**

1.90 is inside ±2.365, so we maintain H_0: t(7) = 1.90, p = 0.099

It's the same t as Topic 11, so it's the same decision

Maintaining H_0 doesn't mean the true line is flat. It means 9 students can't tell

Data: Muralidharan, Singh, and Ganimian (2019).

F, r^2 and R's output

ANOVA's split works for a line

$SS_Y = 180.0$

$df_Y = n - 1 = 7$

↙ ↘ ↙ ↘

$SS_{regression}$	$SS_{residual}$	$df_{regression}$	$df_{residual}$
$r^2SS_Y = 121.5$	$(1 - r^2)SS_Y = 58.5$	1	$n - 2 = 6$

Regression plays the part of between. Residual plays the part of within

After Gravetter et al., *Essentials of Statistics for the Behavioral Sciences*, 10e. Data: Muralidharan et al. (2019).

F agrees: 12.46 is beyond 5.99

Source	SS	df	MS	F
Regression	121.50	1	121.50	**12.46**
Residual	58.50	6	9.75	
Total	180.00	7		

H_0: $\beta = 0$. One cutoff: $F(1, 6) = 5.99$

$F = 121.50 / 9.75 = 12.46$

12.46 is beyond 5.99, so we reject H_0: $F(1, 6) = 12.46$, $p = 0.012$

Data: Muralidharan, Singh, and Ganimian (2019). Students 25 to 32.

One test, three ways

	Formula	Our 8	All 539
Test r	$t = r / \hat{\sigma}_r$	$t(6) = 3.53$	$t(537) = 13.46$
Test b	$t = b / \hat{\sigma}_b$	$t(6) = 3.53$	$t(537) = 13.46$
Test with F	$F = MS_{regression} / MS_{residual}$	$F(1, 6) = 12.46$	$F(1, 537) = 181.2$
Effect size	$r^2 = SS_{regression} / SS_Y$	0.675	0.252

$F = t^2$: $3.53^2 = 12.46$, and the cutoffs match too: $2.447^2 = 5.99$

With one predictor, these are the same test, so they always agree

Data: Muralidharan, Singh, and Ganimian (2019).

R prints all of it at once

```
Call:
lm(formula = endline ~ baseline, data = mindspark)

Residuals:
     Min       1Q   Median       3Q      Max
-18.0384  -3.5099  -0.1523   3.8722  15.4901

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  8.86764    0.66623   13.31   <2e-16 ***
baseline     0.76423    0.05678   13.46   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 5.163 on 537 degrees of freedom
Multiple R-squared:  0.2523,  Adjusted R-squared:  0.2509
F-statistic: 181.2 on 1 and 537 DF,  p-value: < 2.2e-16
```

← **a**

← **b**, $\hat{\sigma_b}$, **t**, **p**

← $\sqrt{MS_{residual}}$, **df**

← r^2

← $F = t^2$

Data: Muralidharan, Singh, and Ganimian (2019). Output from R, `summary(lm())`. You read this. You won't need to write it.

Every number, in our words

R prints	We say	Here
Estimate, both rows	a and b	8.87 and 0.76
Std. Error	$\hat{\sigma}_b$	0.057
t value	$t = b / \hat{\sigma}_b$	13.46
Pr(>|t|)	p, with two tails	Below 0.001
Residual standard error	$\sqrt{MS_{residual}}$, the typical miss	5.16 points
degrees of freedom	$n - 2$	537
Multiple R-squared	r^2	0.252
F-statistic	$F = t^2$	181.2

<2e-16 is a decimal point, 15 zeros, then a 2. We report it as p < 0.001

Data: Muralidharan, Singh, and Ganimian (2019).

A paper shows the same numbers

	Endline math
Baseline math	0.76*** (0.06)
Constant	8.87*** (0.67)
Observations	539
R^2	0.25

Standard errors in parentheses. Three stars: p < 0.001

Each column is one regression, and it names the outcome

The number in parentheses is usually $\hat{\sigma}_b$. Find t yourself: 0.76 / 0.06

Stars stand in for p. The constant is a

Table made for this slide. Data: Muralidharan, Singh, and Ganimian (2019).

Play it: Read the output

```
lm(formula = endline ~ ses_index, data = mindspark)
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  17.3651     0.2531  68.605  < 2e-16 ***
ses_index     0.6557     0.1534   4.275 2.26e-05 ***
Residual standard error: 5.872 on 537 degrees of freedom
Multiple R-squared:  0.03292,  Adjusted R-squared:  0.03112
F-statistic: 18.28 on 1 and 537 DF,  p-value: 2.259e-05
```

100 The predicted endline score at an SES index of 0

200 This, divided by 0.153, gives 4.275

300 The share of endline variability that goes with SES

400 Real? Big? Answer both with numbers from here

Answer with a neighbor, as a question. Nothing is scored

Data: Muralidharan, Singh, and Ganimian (2019). The SES index is a score for each family.

Play it: Answers

100: What is 17.37?
It's a, the intercept

200: What is b = 0.656?
Estimate / Std. Error = t

300: What is 3.3%?
Multiple R-squared is 0.033

400: What is real, but small?
$t(537) = 4.275$, $p < 0.001$, $r^2 = 0.033$

$4.275^2 = 18.28$, the F-statistic. One predictor, one test

Baseline math has $r^2 = 0.252$. SES has 0.033. Both slopes are real, and they aren't equally useful

Data: Muralidharan, Singh, and Ganimian (2019).

The blueprint so far

Blueprint step	Statistic	Bottom of the fraction	df
1. One score	z	σ	None
2. One mean, σ known	z	σ_M	None
3a. One mean, σ unknown	t	$\hat{\sigma}_M$	n − 1
3b. Mean gain	t	$\hat{\sigma}_M$ of the gains	n − 1
4. Two means	t	$\hat{\sigma}_{M_1-M_2}$	$n_1 + n_2 - 2$
5. Two or more means	F	MS_{within}	k − 1 and N − k
6. A correlation	t	$\hat{\sigma}_r$	n − 2
7a. A slope	**t**	$\hat{\sigma}_b$	**n − 2**
7b. A slope for a yes/no X: next	t	Topic 14	

Same idea every time: what we saw, over what chance alone would give

What you can do now

Explain the sampling distribution of b, and what narrows it

Test H_0: $\beta = 0$ with $\hat{\sigma}_b$, t and $n - 2$ degrees of freedom

Split SS and df, and run the same test with F

Say why the tests of r, b and F agree, and why $F = t^2$

Name the four things the test assumes, and check them in a plot of the misses

Read a, b, $\hat{\sigma}_b$, t, p, r^2 and F from R's output or a table in a paper

Next steps

Thursday, March 11: Topic 14

Regression: groups and interactions. Read section 13-1

Friday, March 12: Lab 10

Quiz: final review. Bring a calculator

Finals week, March 13 to 19: the final exam. Date to come