z-scores

Topic 4 · Thursday, January 14

EDUC 15: Statistics for Education Research
UCI School of Education

Where are we?

Description

Summarize data

Visualize data

We are here

Prediction

Explanation



Peter the Anteater at a UC Irvine basketball game.

Plan for today

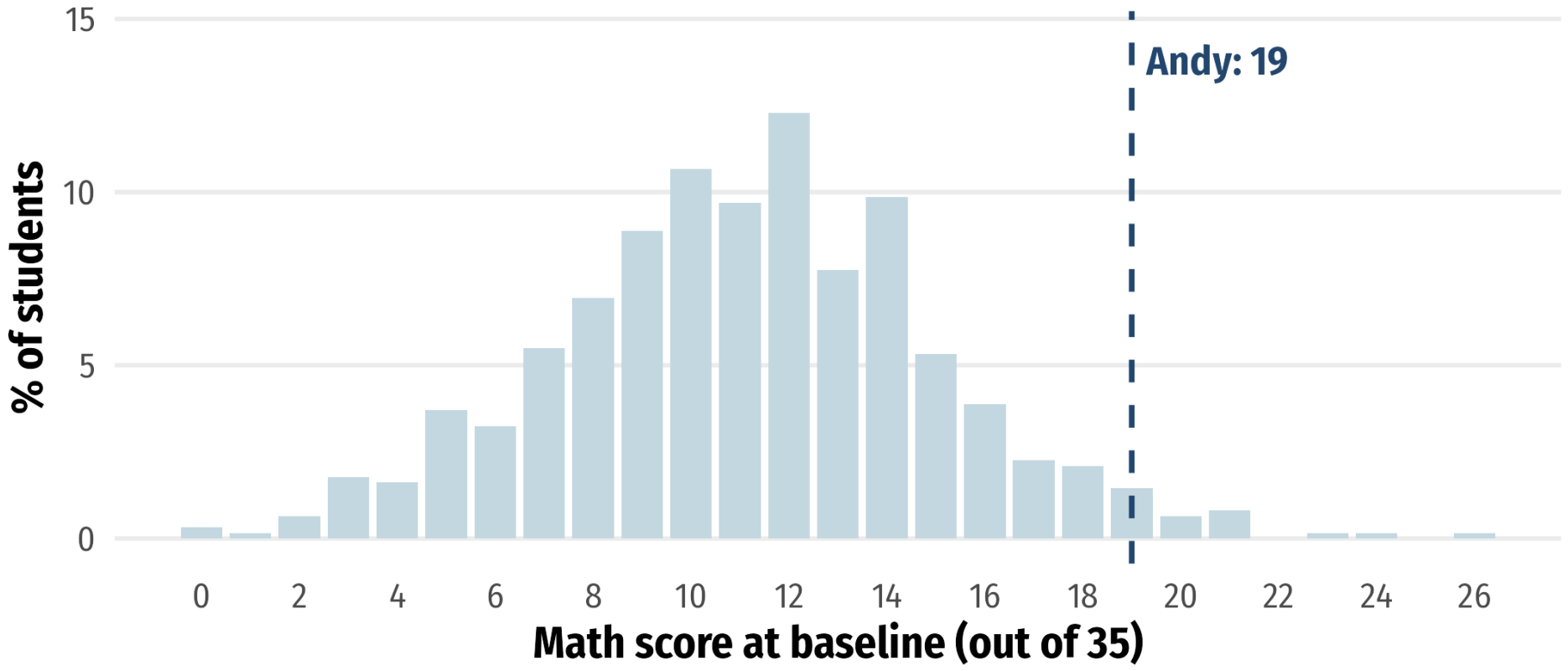
z-scores and locations

Transforming distributions

Standardized distributions

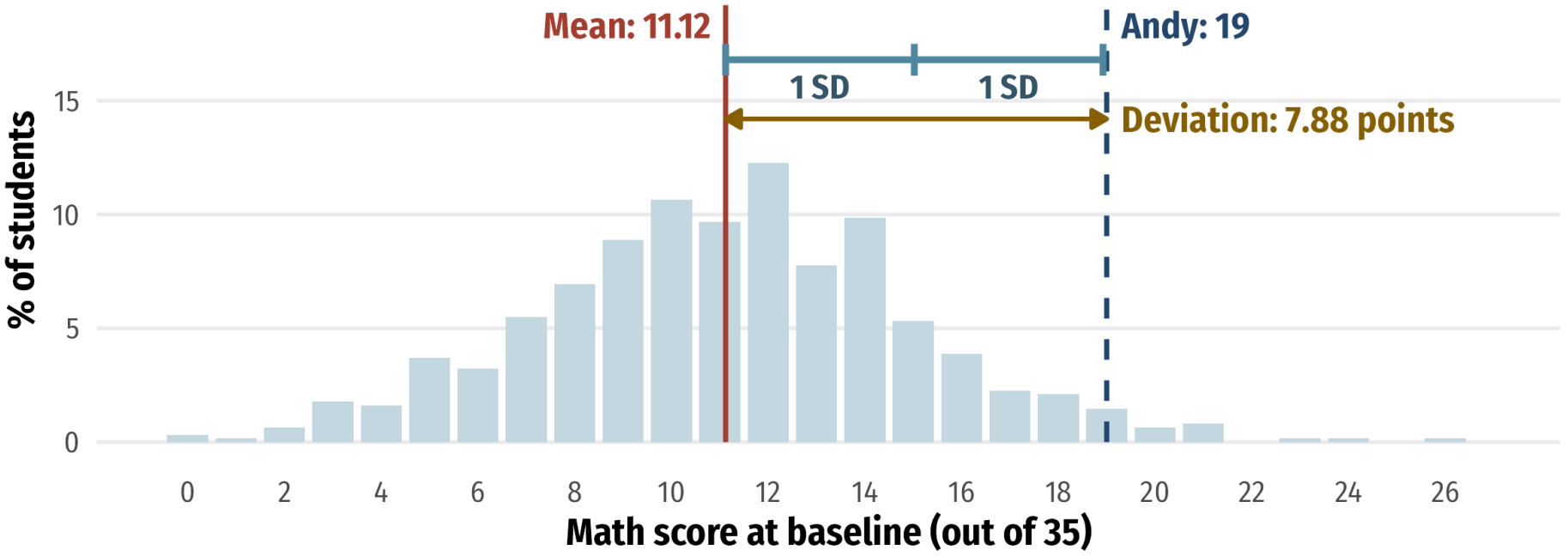
z-scores and locations

Andy scored 19. Is that good?



Data: Muralidharan, Singh, and Ganimian (2019). 619 students in Delhi.

Measure the distance in SDs



Each SD is 3.90 points. Andy's 7.88 points are 7.88 / 3.90 = 2.02 SDs

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

A z-score counts SDs from the mean

$$z = \frac{X - \mu}{\sigma} = \frac{19 - 11.12}{3.90} = 2.02$$

X − μ

How far the score is from the mean, in points. This is the deviation from Tuesday

σ

The ruler. It says how big a step is in this distribution

Treat the 619 students as a population: μ = 11.12 and σ = 3.90

Sign: which side. Size: how far

Score X	Deviation X − μ	z	In words
19	+7.88	+2.02	About two SDs above the mean
15	+3.88	+0.99	About one SD above the mean
11	−0.12	−0.03	Almost exactly at the mean
5	−6.12	−1.57	About one and a half SDs below the mean

z = 0 is the mean. Most of these students sit between z = −2 and z = +2: 94% of them

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

From a z-score back to a score

$$X = \mu + z\sigma = 11.12 + 2.02 \times 3.90 = 19.00$$

First, read the z

z = 2.02, so Andy is above the mean

Then go back to points

2.02 × 3.90 = 7.88 points away

Last, add the mean

11.12 + 7.88 = 19.00

A sample works the same way

	Score to z	z to score
Population	$z = \frac{X - \mu}{\sigma}$	$X = \mu + z\sigma$
Sample	$z = \frac{X - M}{\hat{\sigma}}$	$X = M + z\hat{\sigma}$

Tuesday's five: M = 12.2 and $\hat{\sigma}$ = 2.86. In this sample, the 16 has z = 1.33

In the book: $\hat{\sigma}$ is written s

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

Try it: Go there and back

For the math test: $\mu = 11.12$ and $\sigma = 3.90$

1. A student scored 9. What's the z-score?

2. A student has z = +1.50. What's the score?

3. In Tuesday's sample, M = 12.2 and $\hat{\sigma} = 2.86$. Find the z of the 8.

Work alone for two minutes. Compare with a neighbor. Then we vote

Try it: Answer

Question 1

(9 − 11.12) / 3.90 =

−0.54

Question 2

11.12 + 1.50 × 3.90 =

16.97

Question 3

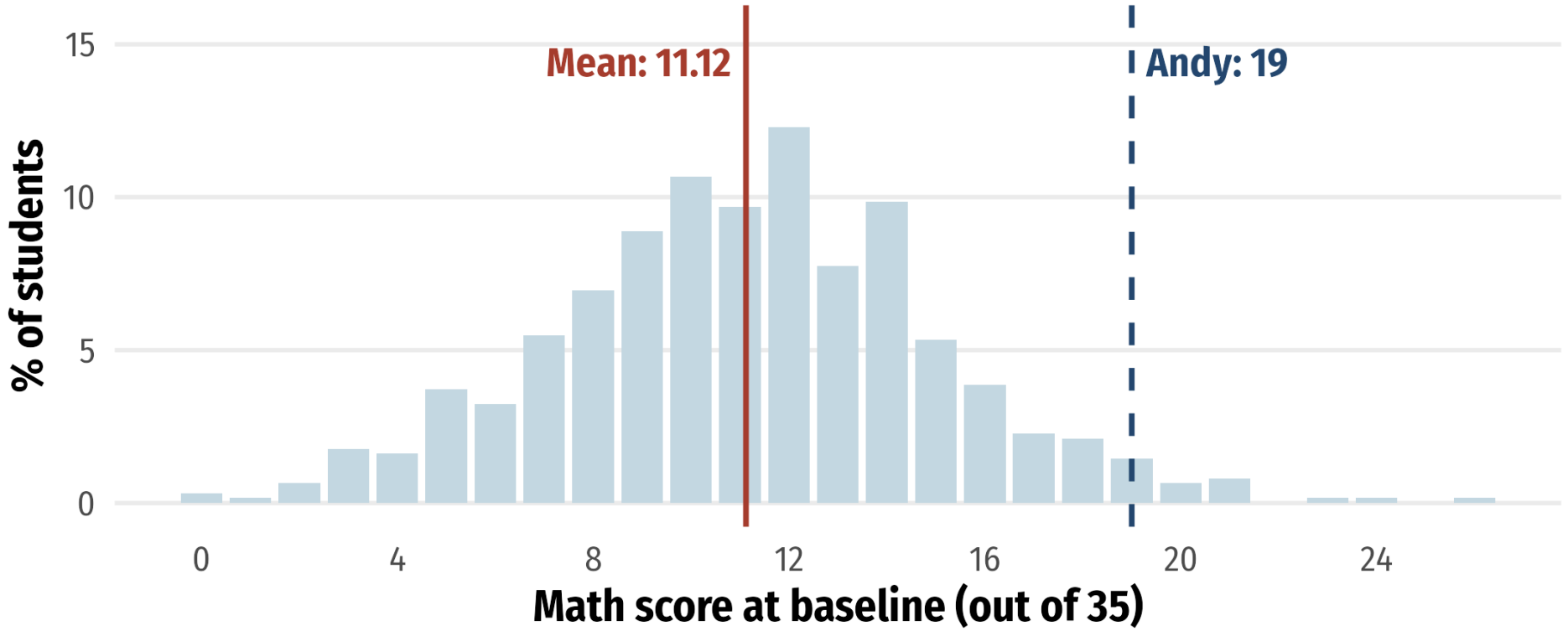
(8 − 12.2) / 2.86 =

−1.47

A negative z-score means below the mean. It doesn't mean a bad score

Transforming distributions

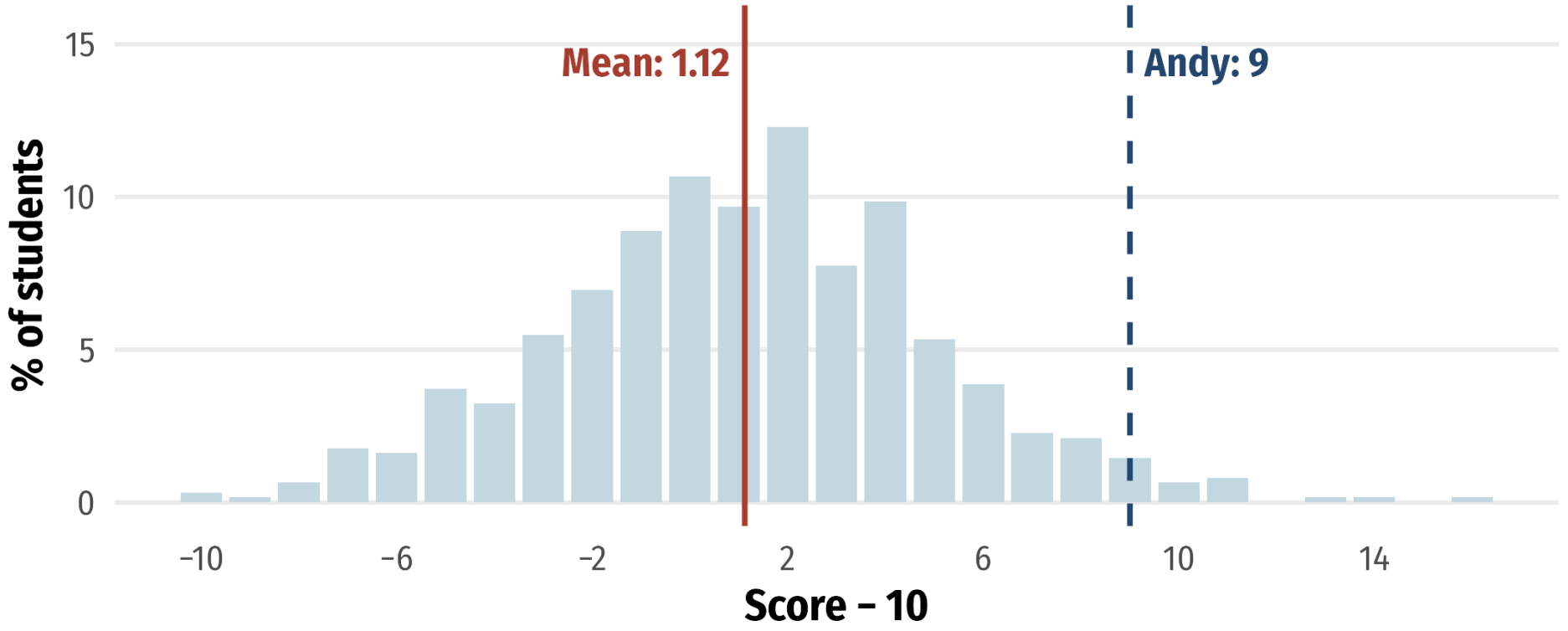
A 19 means nothing by itself



It's only a label. We could relabel every score, and everyone would keep their place

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

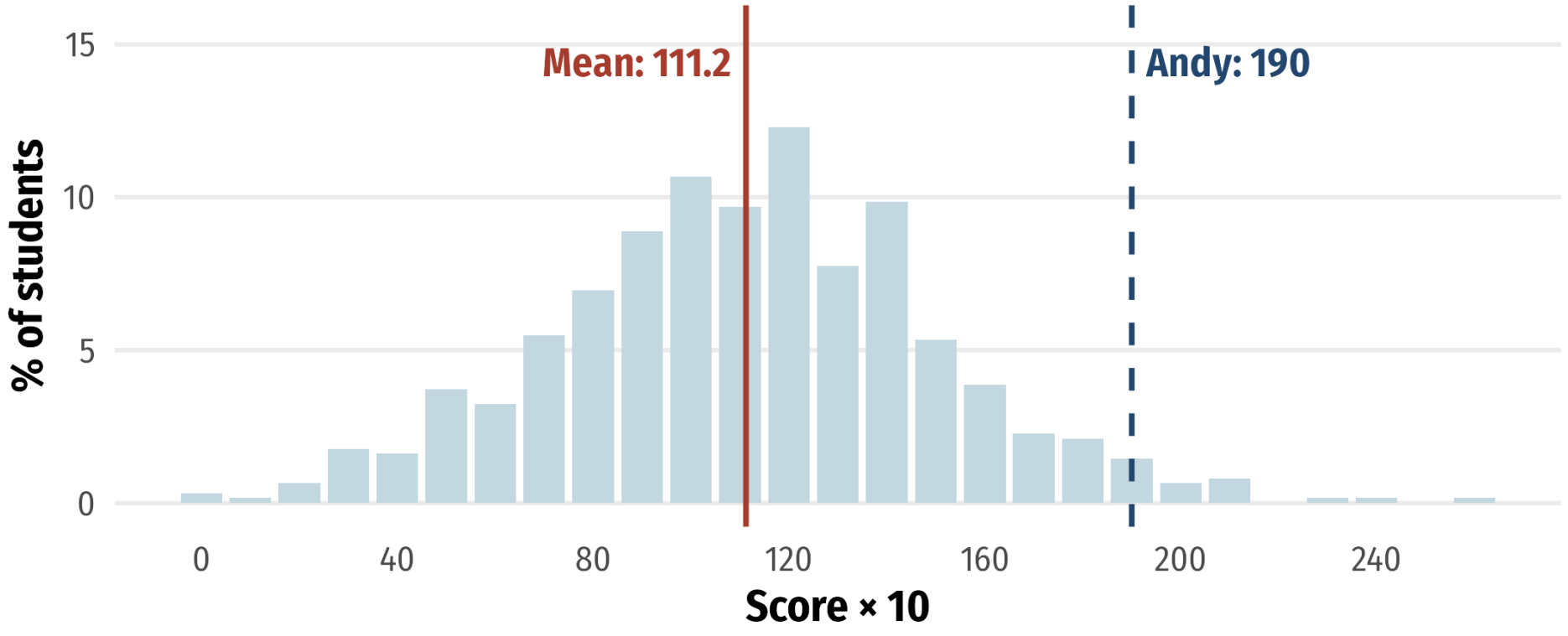
Subtract 10: the labels slide



Mean: 11.12 ⟶ 1.12. The SD stays at 3.90. Andy is still 2.02 SDs above the mean

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

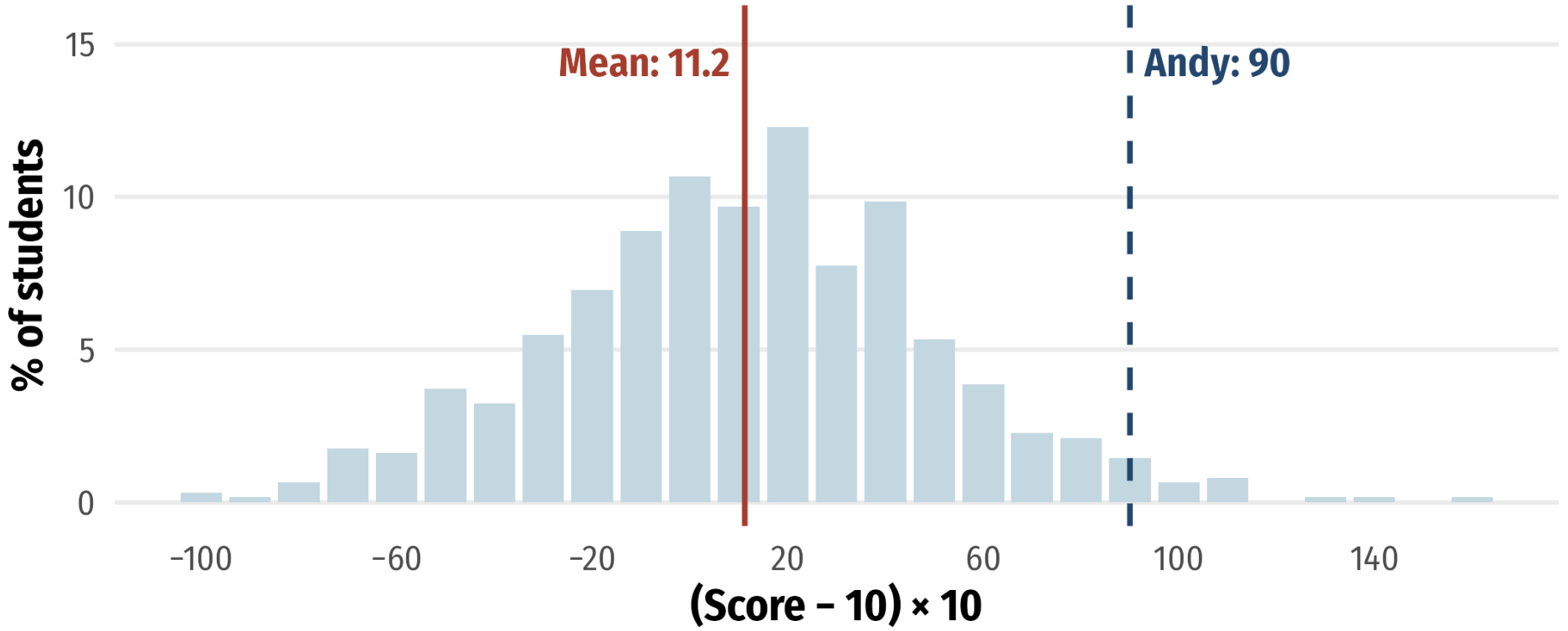
Multiply by 10: the labels stretch



Mean: 11.12 → 111.2. SD: 3.90 → 39.0. Andy is still 2.02 SDs above the mean

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

Do both: subtract, then multiply



Mean: 11.12 ⟶ 11.2. SD: 3.90 ⟶ 39.0. Andy is still 2.02 SDs above the mean

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

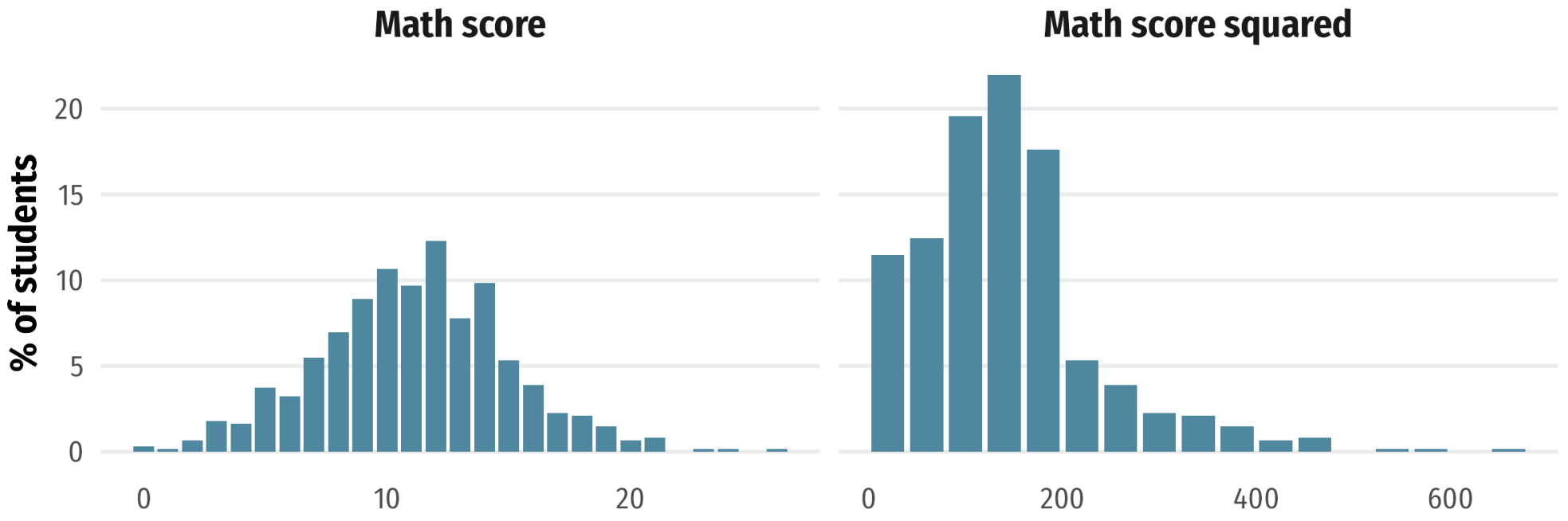
These only change the labels

Do this to every score	Mean	SD	Shape
Add or subtract a constant	Moves by the constant	Stays the same	Stays the same
Multiply or divide by a positive constant	Multiplied or divided	Multiplied or divided	Stays the same

We call these "scale-invariant" transformations

They keep the shape and the relative distances between students

Squaring changes the shape



Now 4 and 6 are 20 apart, but 18 and 20 are 76 apart

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

Predict it: Two steps at once

Respond at **PollEv.com/adb**

Take every math score. Subtract 11.12, the mean. Then divide by 3.90, the SD. What are the new mean and SD?

A Mean 11.12, SD 3.90

B Mean 0, SD 3.90

C Mean 0, SD 1

D Mean 1, SD 0

Standardized distributions

A z-score is the same kind of relabeling

$$z = \frac{X - \mu}{\sigma}$$

Subtract one constant

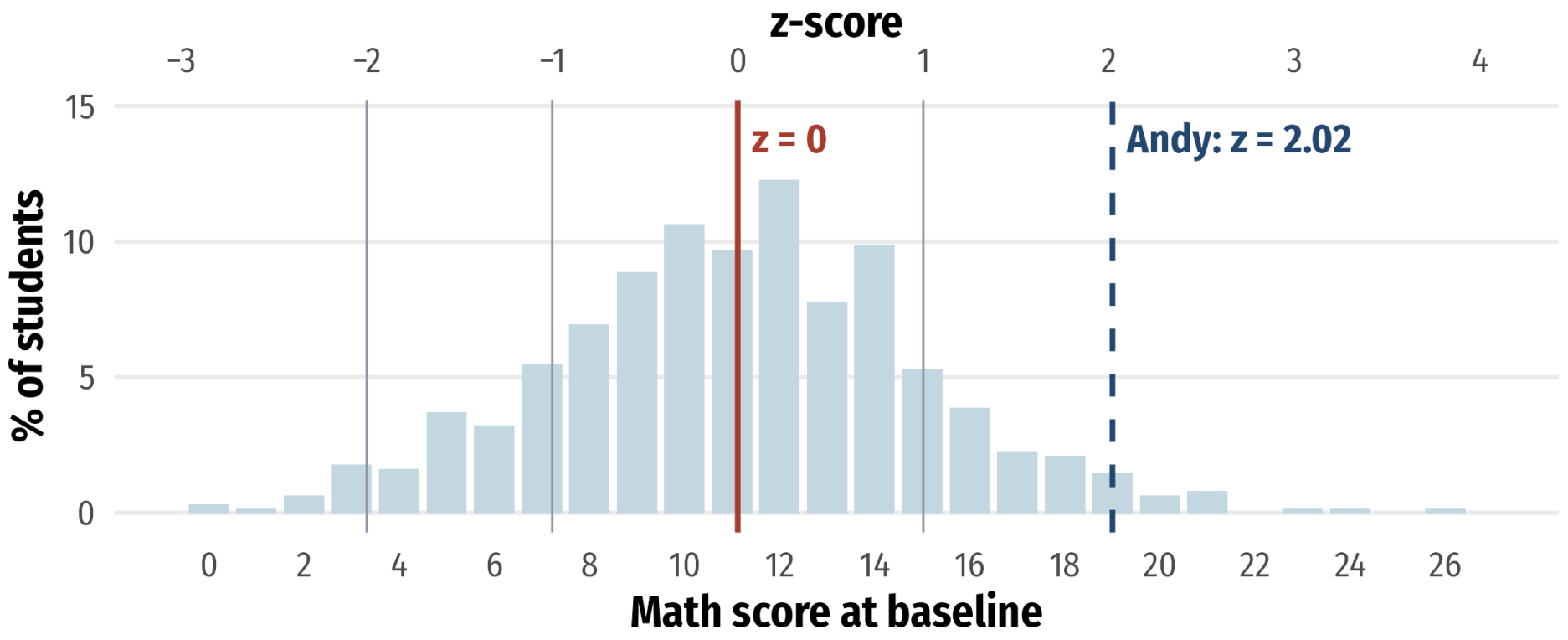
That constant is the mean

Divide by another constant

That constant is the standard deviation

Do it to every score and you get a *standardized* distribution

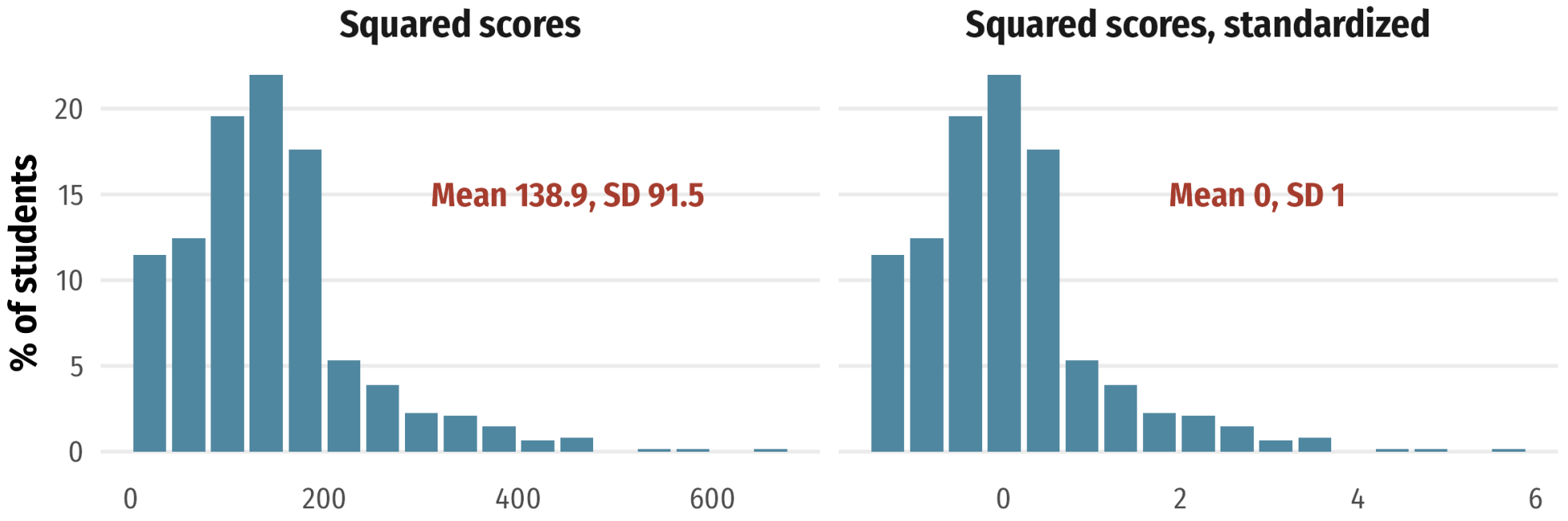
Same picture, new axis: mean 0, SD 1



The z-scores have mean 0.00 and SD 1.00. Every student keeps their place

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

z-scores don't make a shape normal



A skewed distribution stays skewed. Only the axis changes

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

We can add labels we like better

$$\text{New score} = \text{new mean} + z \times \text{new SD}$$

Mean 500, SD 100

Andy: 500 + 2.02 × 100 = **702**

Mean 100, SD 15

Andy: 100 + 2.02 × 15 = **130**

Same student, same place. Nobody needs to see a negative score or a decimal

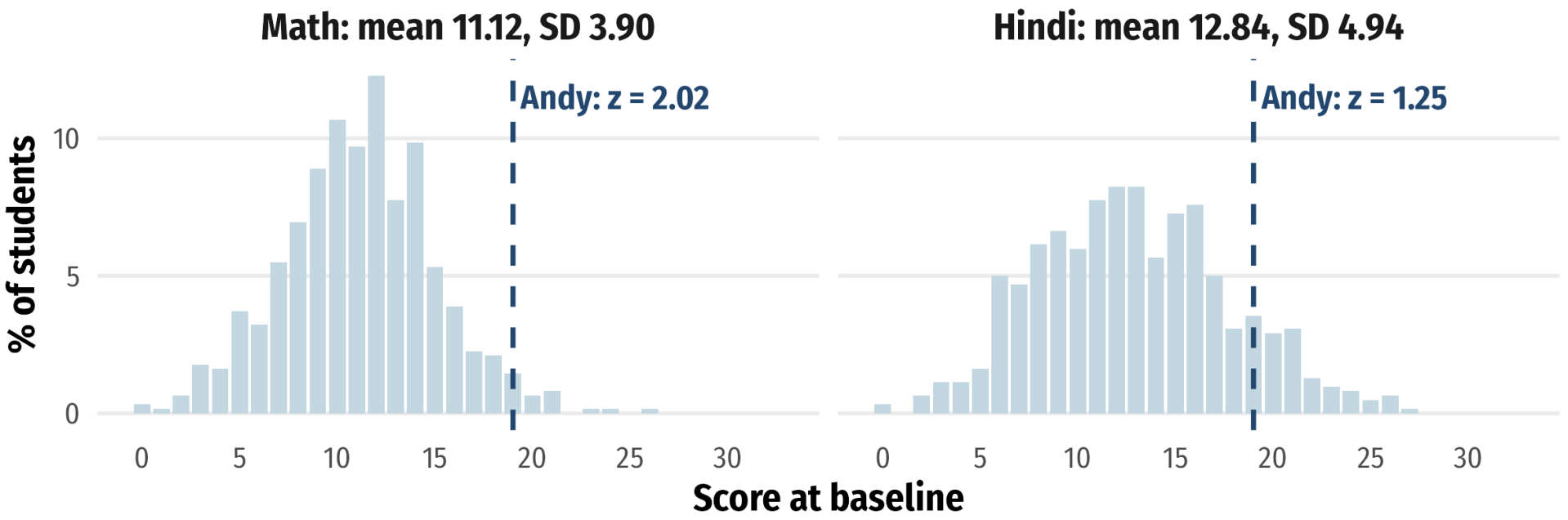
Score reports are tables like this one

Math score	z	Mean 500, SD 100	Mean 100, SD 15	% who scored lower
3	−2.08	292	69	1
7	−1.06	394	84	11
11	−0.03	497	100	43
15	+0.99	599	115	83
19	+2.02	702	130	97

The study's scaled math scores do this too: mean 0.00, SD 1.00 at baseline

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

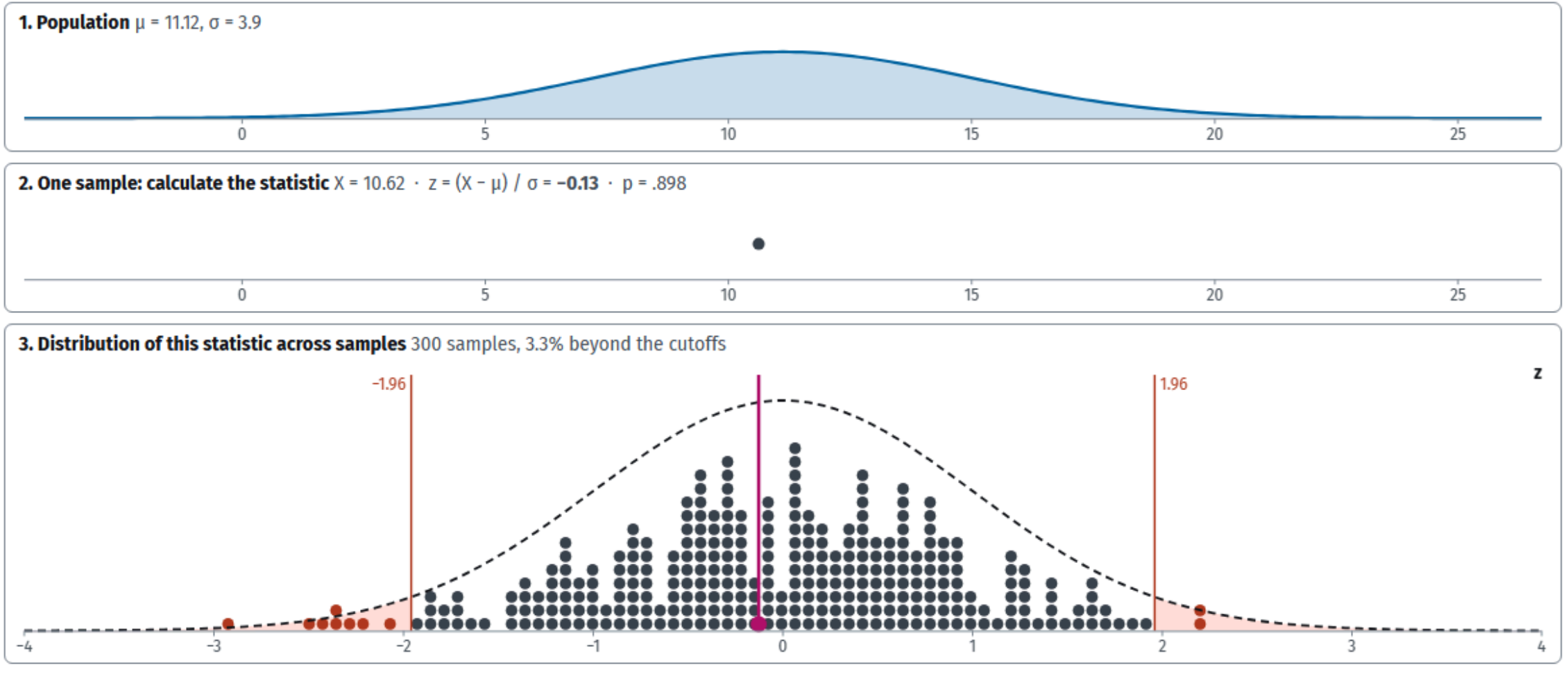
A 19 in math beats a 19 in Hindi



Percent correct says Hindi: 63% against 54%. z-scores say math: 2.02 against 1.25

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

The blueprint starts here



Simulator after Kristoffer Magnusson, rpsychologist.com.

Play it: z-scores

100

The mean of any full set of z-scores

200

On a test with $\mu = 40$ and $\sigma = 8$, the score with $z = -1.5$

300

What 5 bonus points for everyone do to a z-score

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

Play it: Answers

100

What is 0?

The deviations add up to zero

200

What is 28?

$40 - 1.5 \times 8 = 28$

300

What is nothing?

X and μ both rise by 5, and σ stays

What you can do now

Turn a score into a z-score, and a z-score back into a score

Say what the sign and the size of a z-score mean

Say what adding, multiplying and standardizing do to a distribution

Compare scores from two different tests

Next steps

Before Tuesday

Read Chapter 6

Tomorrow: Lab 2

Bring a calculator and a laptop. It includes a 15-minute individual check

Mastery and Problem Set for Chapter 5 are due Sunday at 6 PM