

Modeling data



Chat

What's your favorite season and in which place?

To: Everyone ▾ More ▾

Type message here...

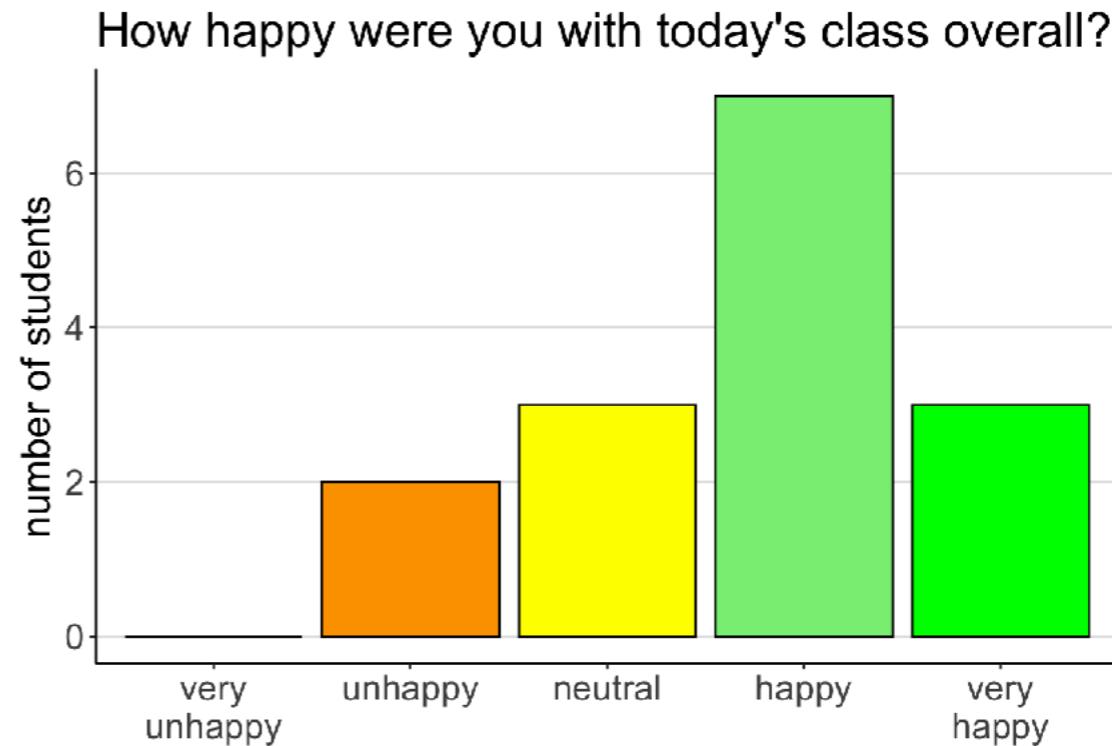
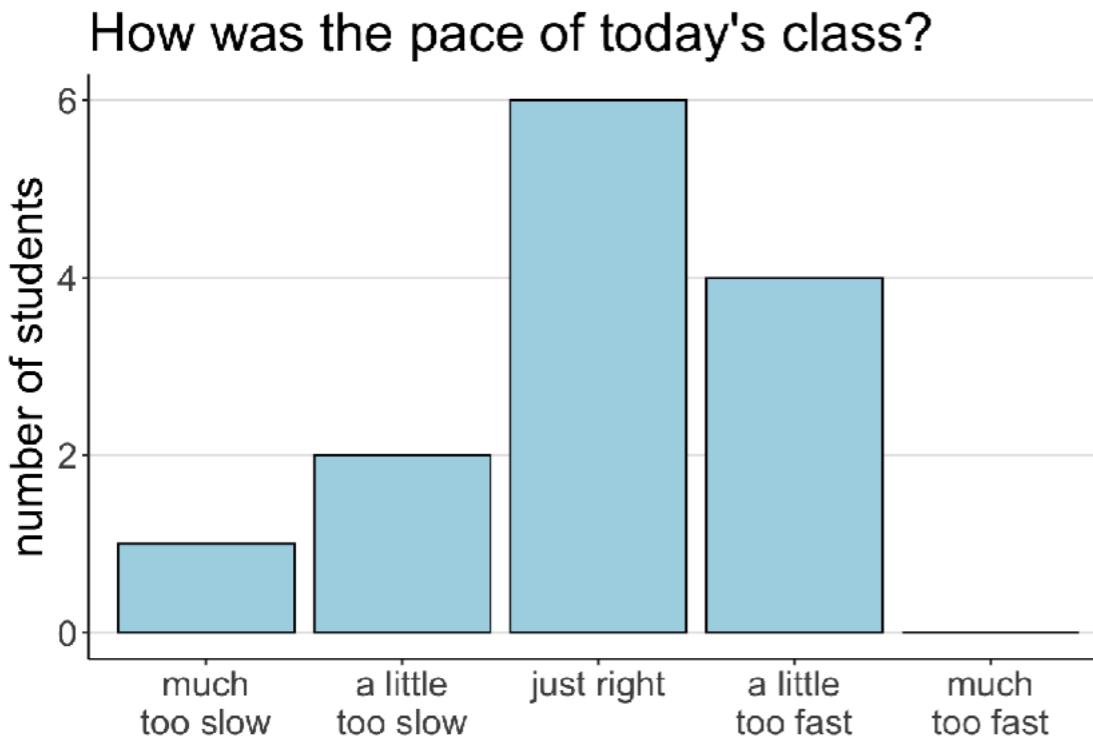
A screenshot of a Spotify interface showing a collaborative playlist titled "psych252". The interface includes a play button and a link to the playlist: <https://tinyurl.com/psych252spotify22>. The background features a stylized bar chart with a smiling face.

We're listening to
"I heard" by "Young Fathers", submitted by Tobi

01/24/2022

Feedback

Your feedback

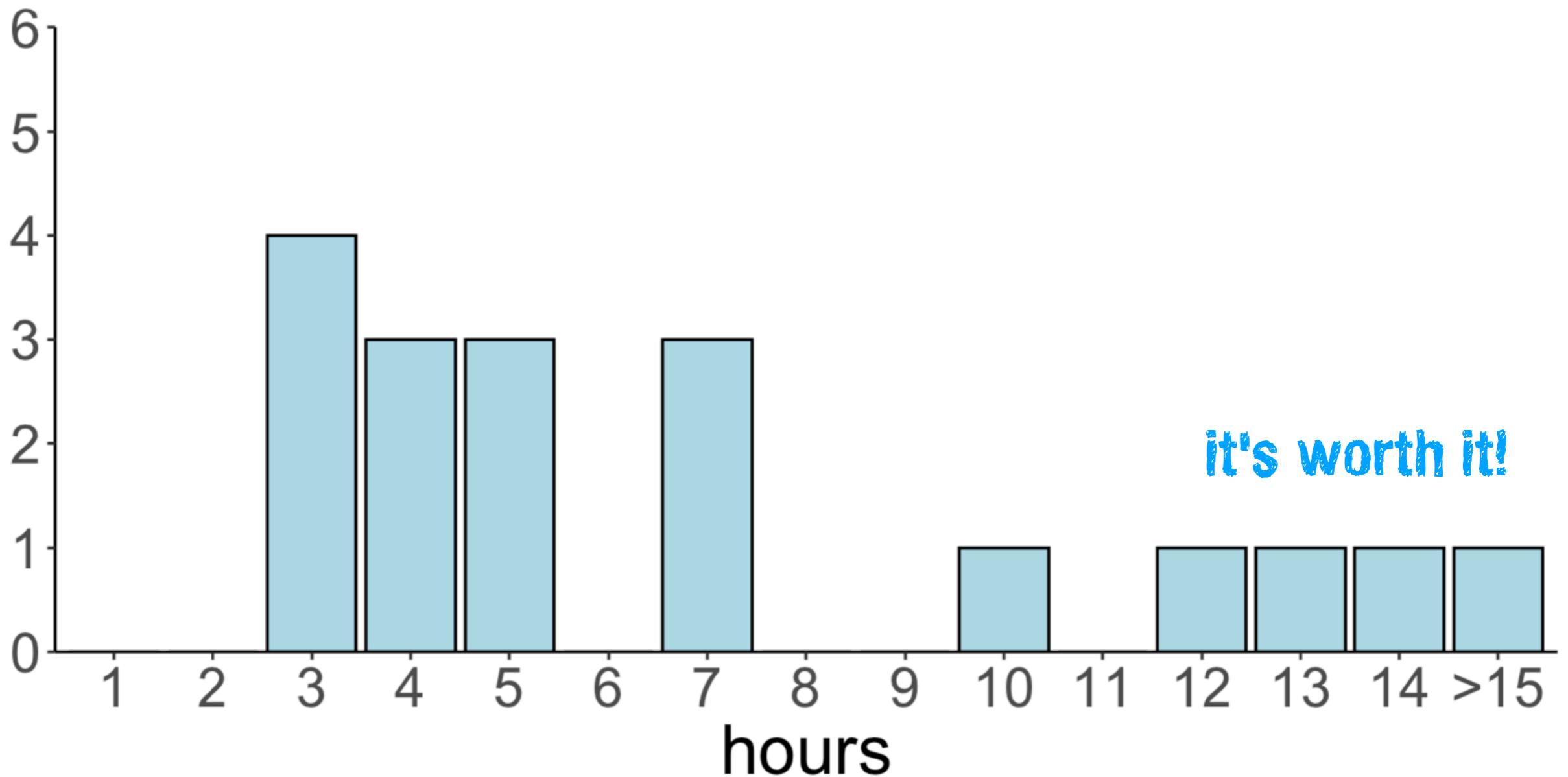


please continue to review key concepts from prior lectures at the beginning of the lecture

This is the clearest time I learned p-value. Thank you!

This may just be a reminder, if we get lost we ask less or no questions, so more questions are indicative of a good pace (and waiting in silence during q breaks is good so we can catch up).

Homework 2



Plan for today

- Quick recap
- Statistical concepts
 - Confidence intervals
 - Bootstrapping
- Cookbook vs. Model Comparison
- Modeling data
- Statistical inferences about parameter values

Quick recap

Quick recap: Bayesian inference



twice as many kids go to the basketball camp

$$X \sim \text{Normal}(\mu = 170, \sigma = 8)$$



$$X \sim \text{Normal}(\mu = 180, \sigma = 10)$$



Can you feel the Bayes?

$$H = \{\text{basketball, chess}\}$$

$$D = 175 \text{ cm}$$

$$p(H|D) = \frac{\text{likelihood} \cdot \text{prior}}{p(D)} \quad H = \text{Hypothesis}$$

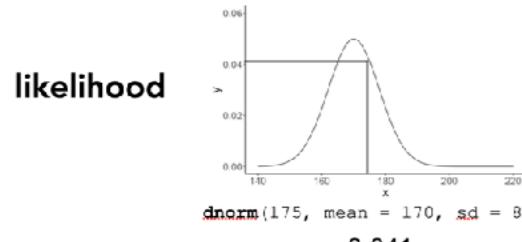
$$D = \text{Data}$$

probability of the data?!

Summer camp

$$\text{prior} \quad p(\text{chess}) = \frac{1}{3}$$

$$p(\text{basketball}) = \frac{2}{3}$$



Analytic solution

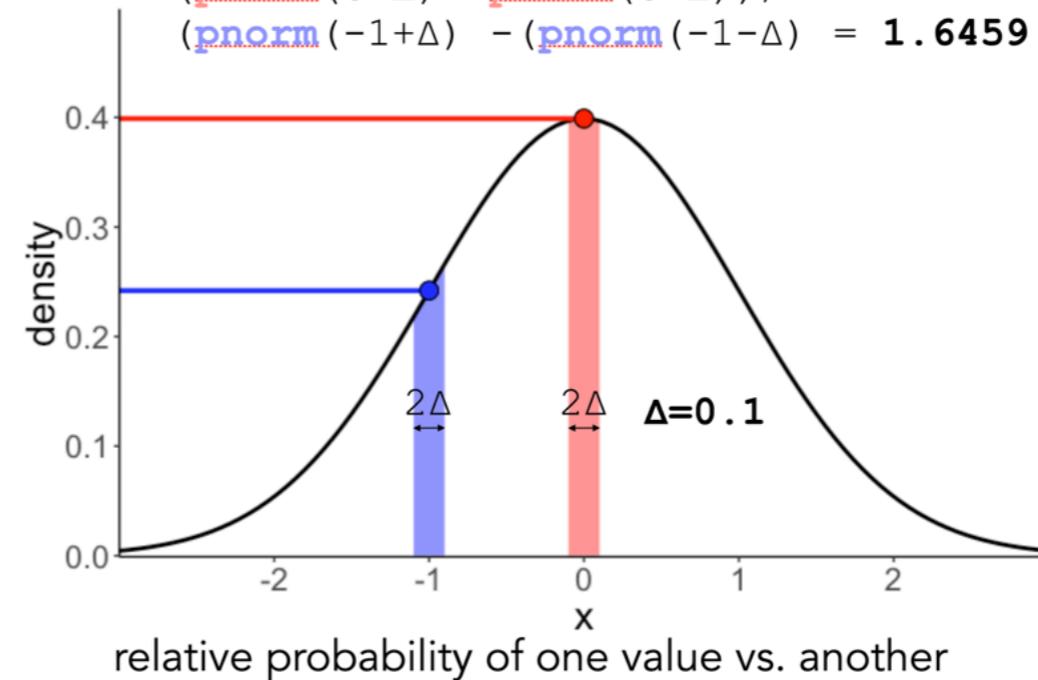
$$\text{posterior} \quad p(\text{sport} = \text{basketball} | \text{height} = 175) = \frac{p(175 | \text{basketball}) \cdot p(\text{basketball})}{p(175)}$$

$$p(\text{basketball} | 175) = \frac{p(175 | \text{basketball}) \cdot p(\text{basketball})}{p(175 | \text{basketball}) \cdot p(\text{basketball}) + p(175 | \text{chess}) \cdot p(\text{chess})}$$

Probability vs. likelihood

$$\text{dnorm}(0) / \text{dnorm}(-1) = 1.6487$$

$$\frac{(\text{pnorm}(0+\Delta) - \text{pnorm}(0-\Delta))}{(\text{pnorm}(-1+\Delta) - \text{pnorm}(-1-\Delta))} = 1.6459$$



Summer camp: Via sampling

Sampling solution

```
1 df.camp = tibble(
2   kid = 1:100000,
3   sport = sample(c("chess", "basketball"),
4                 size = 100000,
5                 replace = T,
6                 prob = c(1/3, 2/3))) %>%
7   rowwise() %>%
8   mutate(height = ifelse(test == "chess",
9                         yes = rnorm(., mean = 170, sd = 8),
10                        no = rnorm(., mean = 180, sd = 10))) %>%
11  ungroup()
```

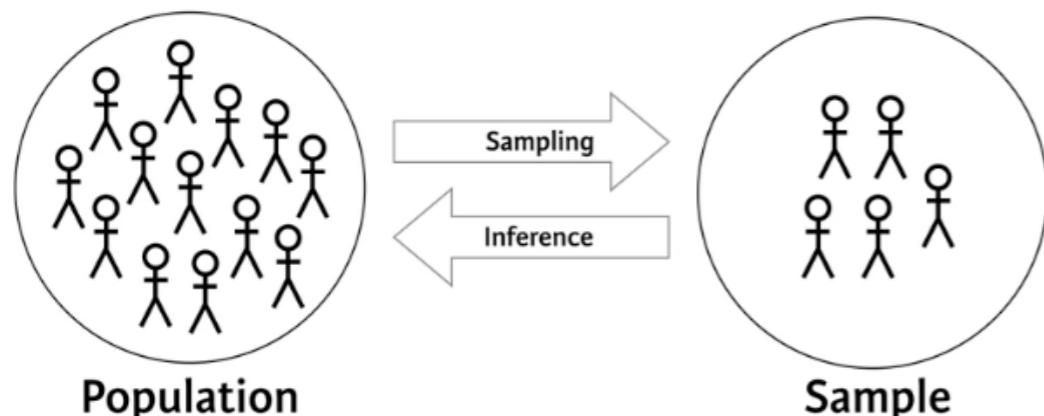
kid	sport	height
1	basketball	164.84
2	basketball	163.22
3	basketball	191.18
4	chess	160.16
5	basketball	182.99
6	chess	163.54
7	chess	168.56
8	basketball	192.99
9	basketball	171.91
10	basketball	177.12

```
1 df.camp %>%
2   filter(between(height,
3                   left = 174,
4                   right = 176)) %>%
5   count(sport)
```

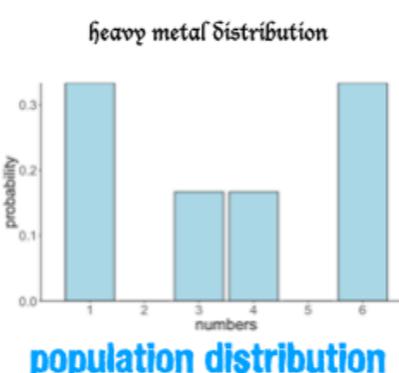
sport	n
basketball	469
chess	273

$$\frac{\text{basketball}}{\text{basketball} + \text{chess}} \approx 0.63$$

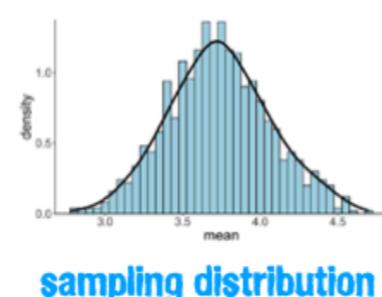
Quick recap: Inference in frequentist statistics



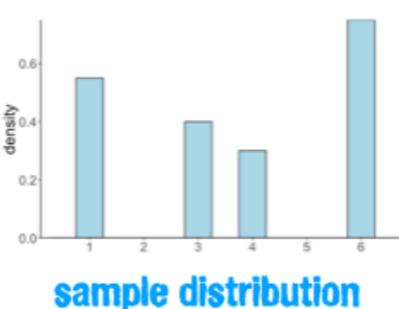
3 distributions in statistical inference



- unknown
- our target for inference
- e.g. we might be interested in the mean of the population distribution



- bridge between sample and population
- derived mathematically / computationally
- asymptotic distribution theory or resampling approaches
- shows how test statistic varies between samples



- our observed sample
- we compute statistics of interest (mean, variance, correlation, ...)
- make an inference about the population via the sampling distribution

Quick recap: What is a p-value?

What is a p-value?

The **p-value** is the probability of finding the observed (or more extreme) results when the null hypothesis (H_0) is true.

$$p(\text{test statistic} \geq \text{observed value} | H_0 = \text{true})$$

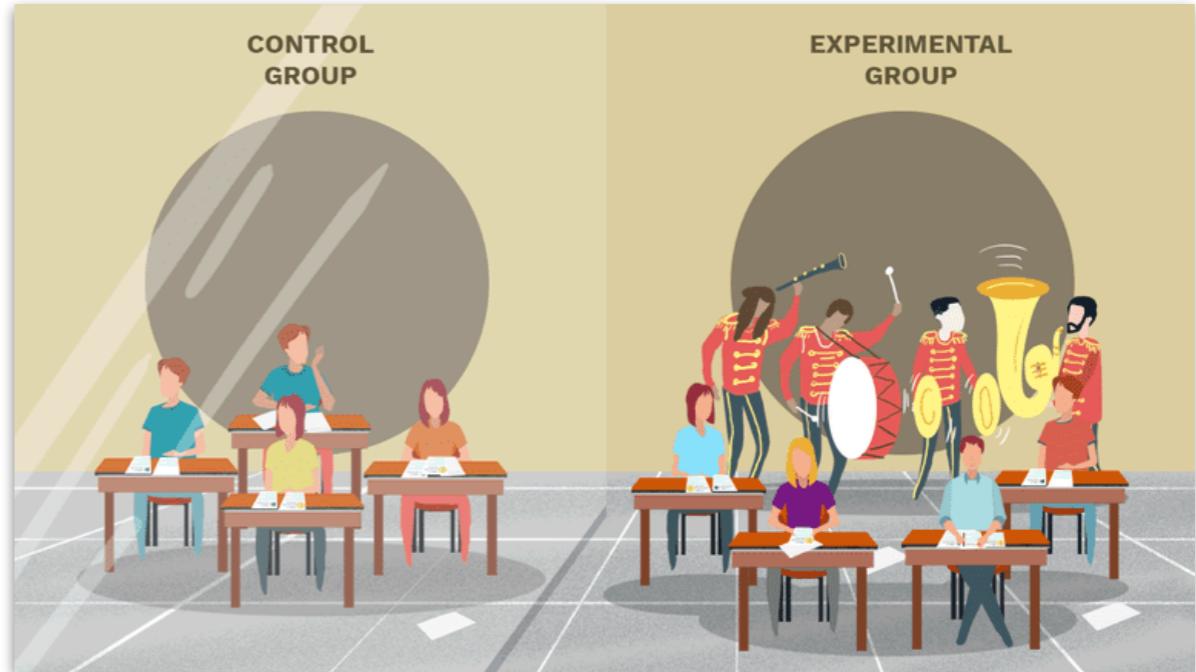
~~what we're actually interested in!~~

~~$p(H_1 = \text{true} | \text{test statistic} \geq \text{observed value})$~~

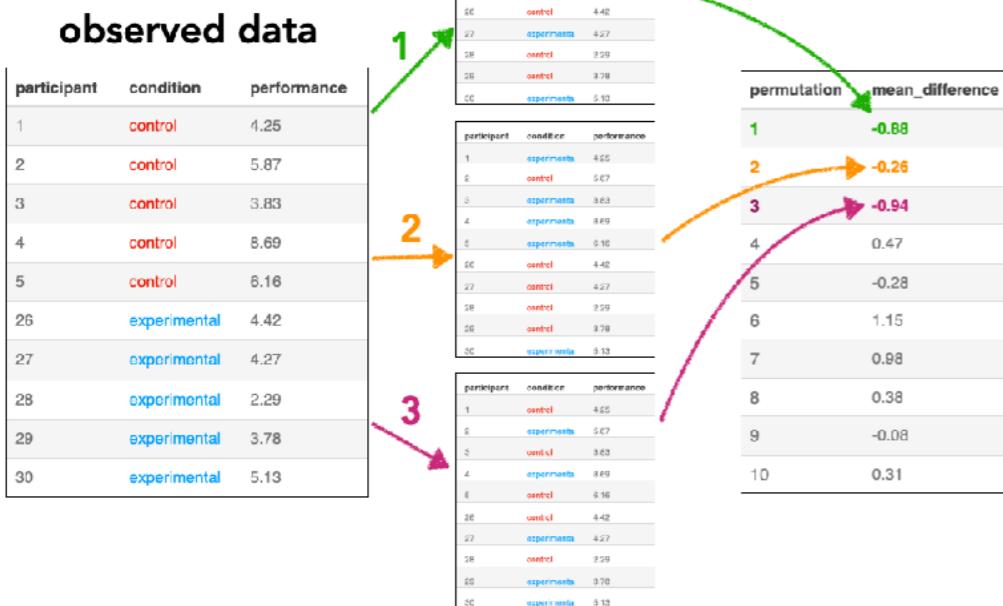
... we'll have to wait for Reverend Bayes

$$p(H|D) = \frac{p(D|H) \cdot p(H)}{p(D)}$$

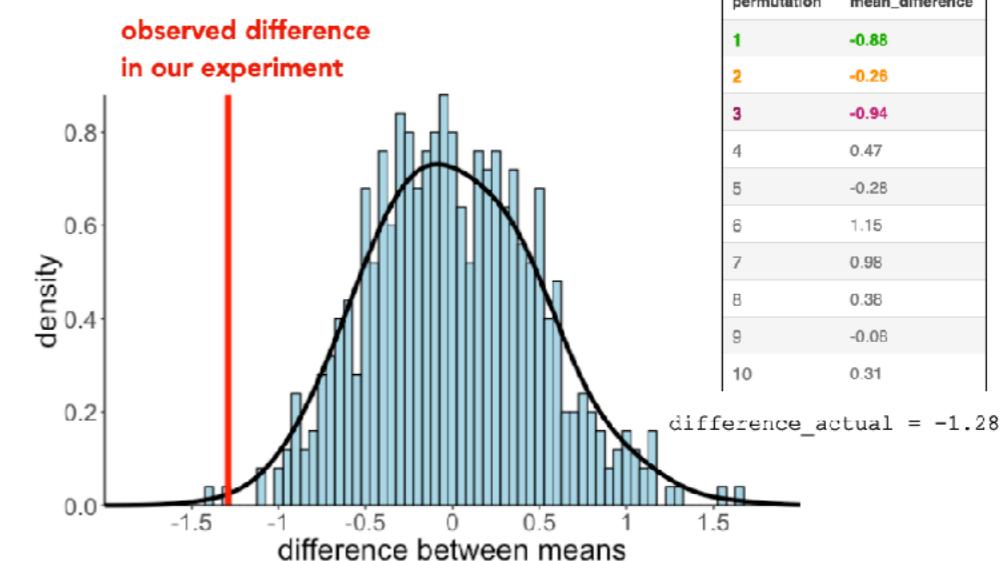
H = Hypothesis
 D = Data



Permutation test



Permutation test



```
1 #calculate p-value of our observed result
2 df.permutations %>%
3   summarize(p_value = sum(mean_difference <= difference_actual)/n())
```

p-value = .002

Statistical concepts

Confidence intervals

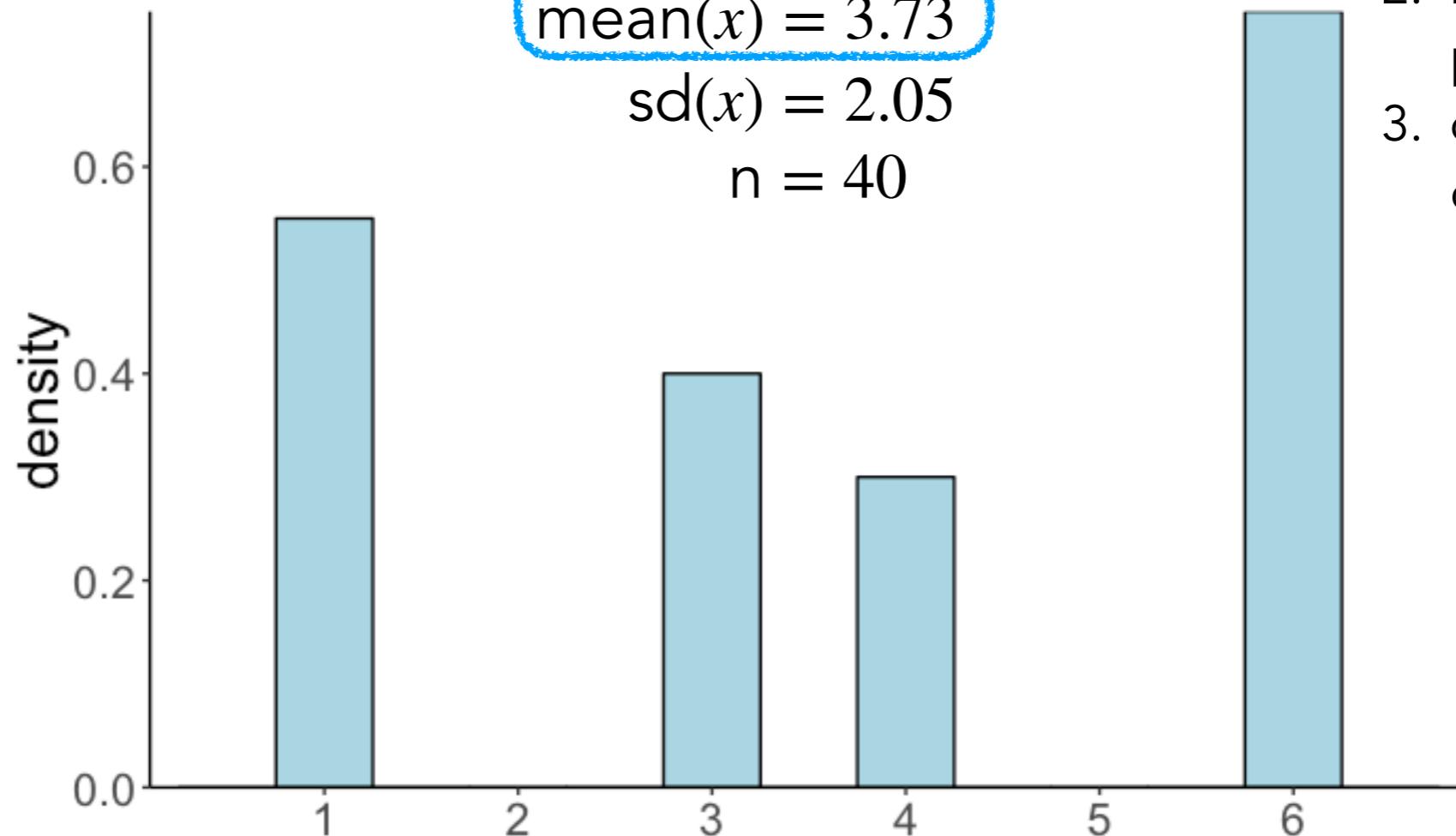


Confidence interval

Goal: Estimate $\mu = \text{the mean of the population distribution}$

Confidence interval of the mean = point estimate \pm critical value

the sample mean is our best guess of the population mean



depends on

1. variance in the data
2. number of data points
3. desired level of confidence

Confidence interval

Goal: Estimate $\mu =$ the mean of the population distribution

what we need:

- sample mean
- sample standard deviation
- sample size
- desired level of confidence

Confidence interval

Confidence interval of the mean = point estimate \pm critical value

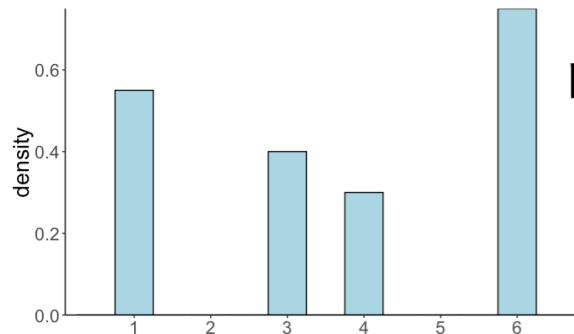
I would like to be
95% confident.

How confident
would you like to be that
you're correct?



Confidence interval

Parametric assumption: The sampling distribution of the mean is a normal distribution.

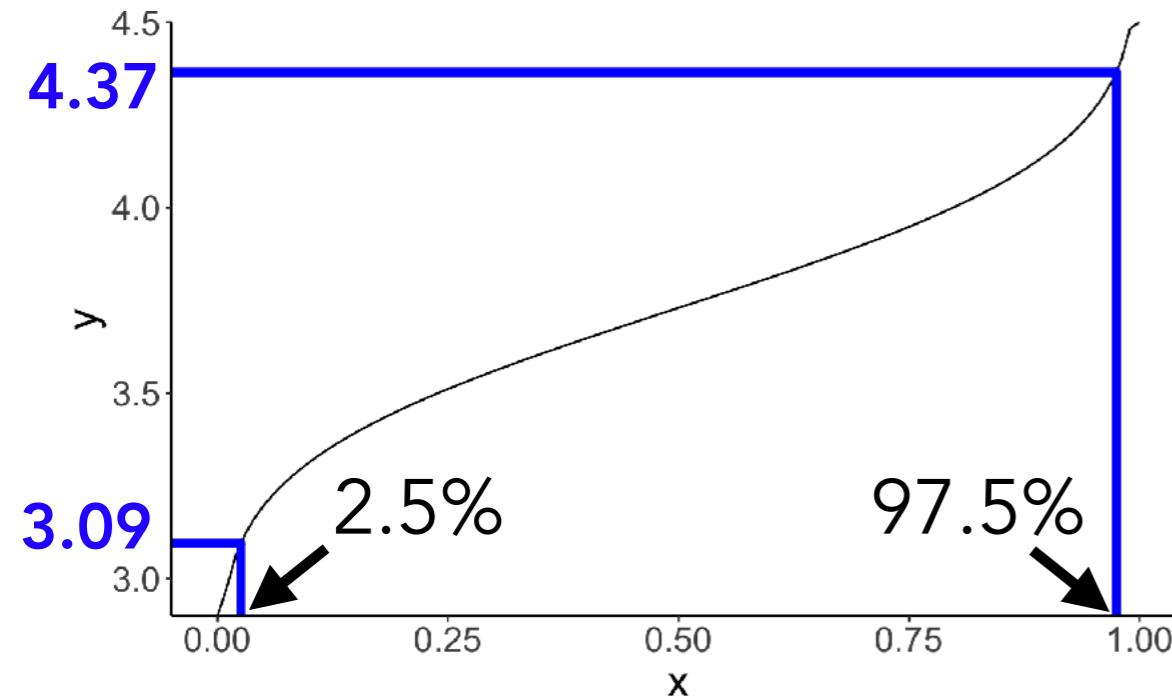


our sample

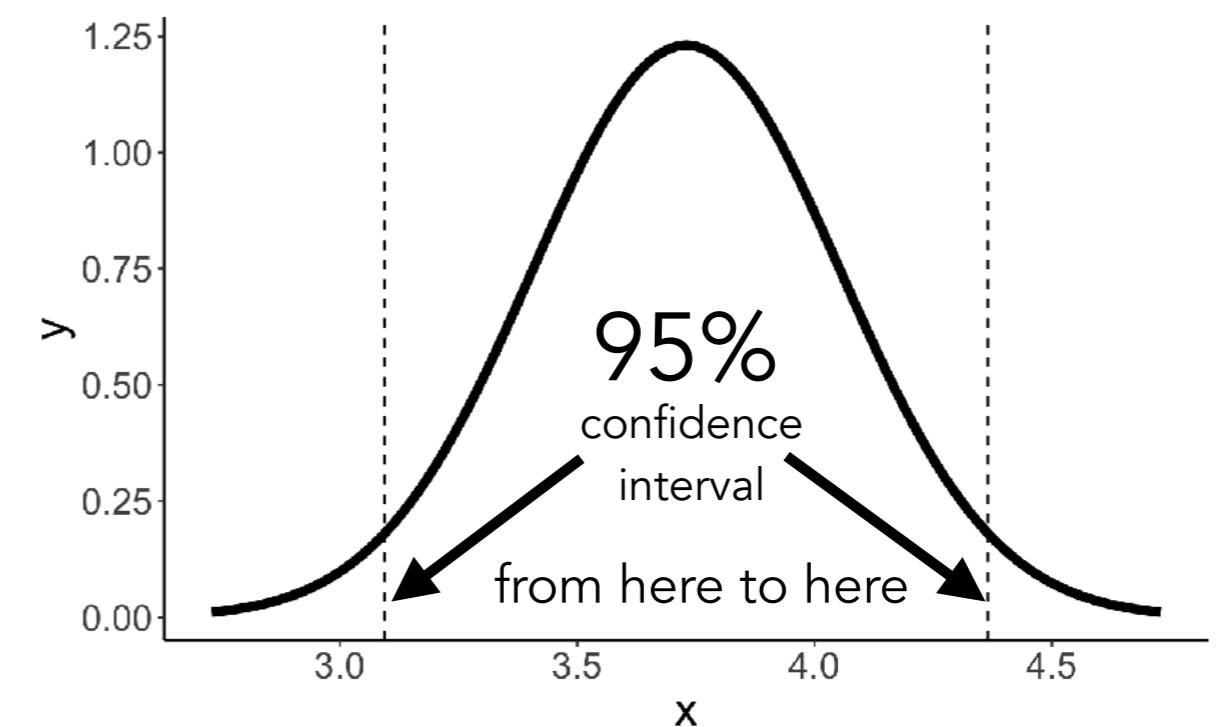
$$\begin{aligned}\text{mean}(x) &= 3.73 \\ \text{sd}(x) &= 2.05 \\ n &= 40\end{aligned}$$

$$\begin{aligned}\text{mean}(\bar{x}) &= 3.73 \\ \text{sd}(\bar{x}) &= \frac{\text{sd}(x)}{\sqrt{n}} = \frac{2.05}{\sqrt{40}} \approx 0.324\end{aligned}$$

sampling distribution of the mean

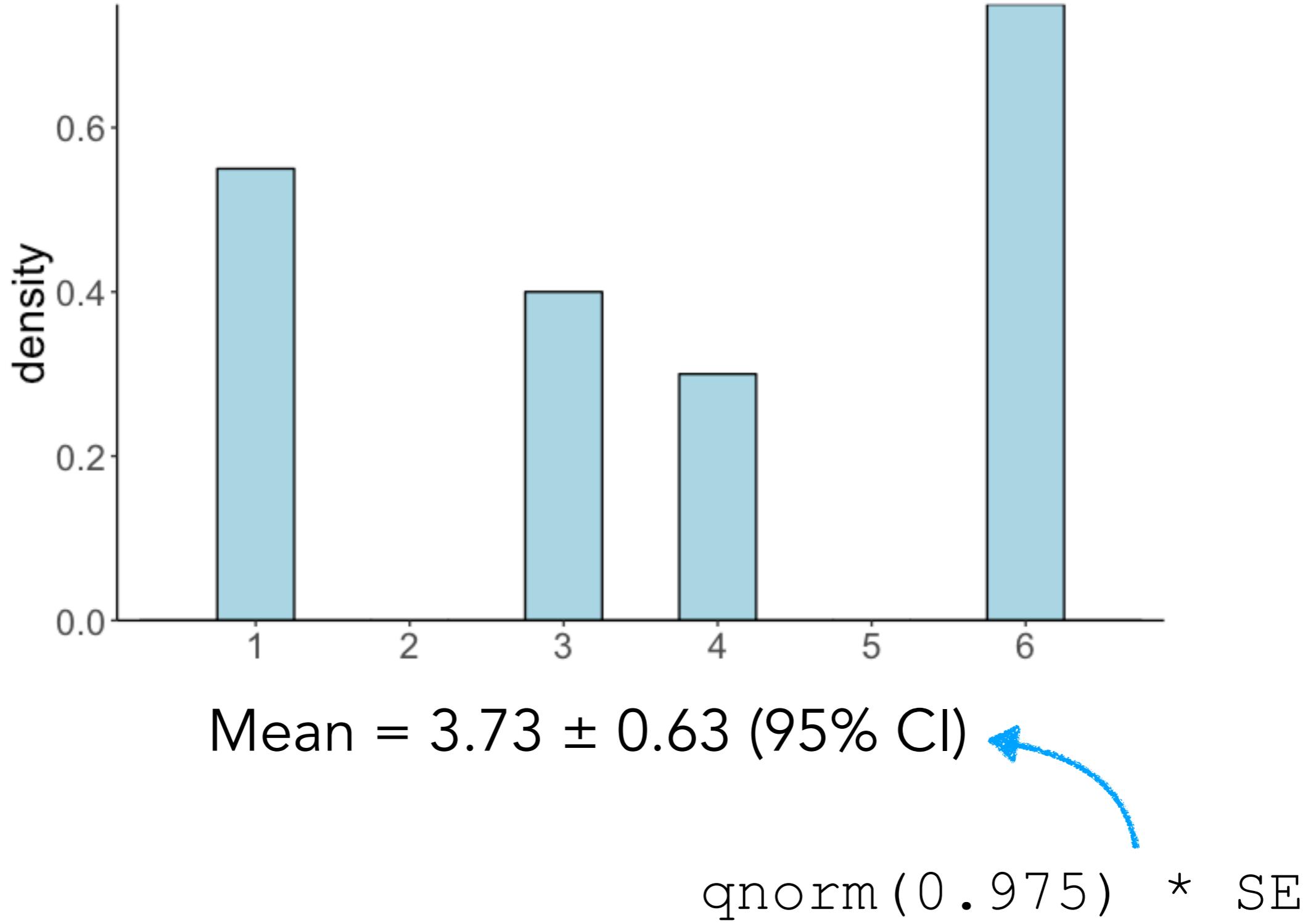


`~ qnorm(., mean = 3.73, sd = 0.324)`



`~ dnorm(., mean = 3.73, sd = 0.324)`

What does the confidence interval mean?



Confidence interval

Confidence interval of the mean = point estimate \pm critical value

I would like to be 95% confident.

How confident would you like to be that you're correct?

**But what does
95% confident even
mean?**



What can we say based on the result of our sample ($N = 40$): Mean = 3.73 ± 0.63 (95% CI)?

95% of the time, the true population mean will be in this interval.

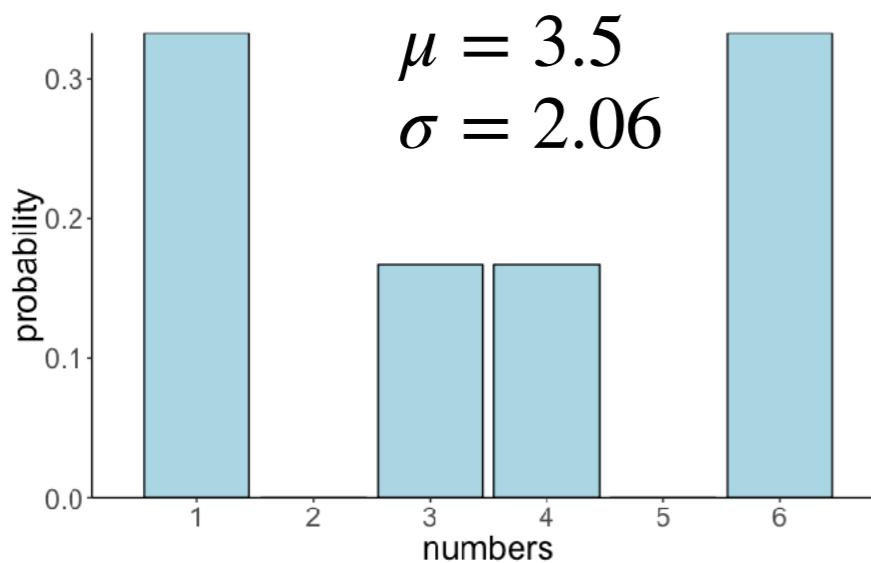
95% of random samples of size 40 will yield confidence intervals that contain the population mean.

The sample means of 95% of the random samples of size 40 will be in this interval.

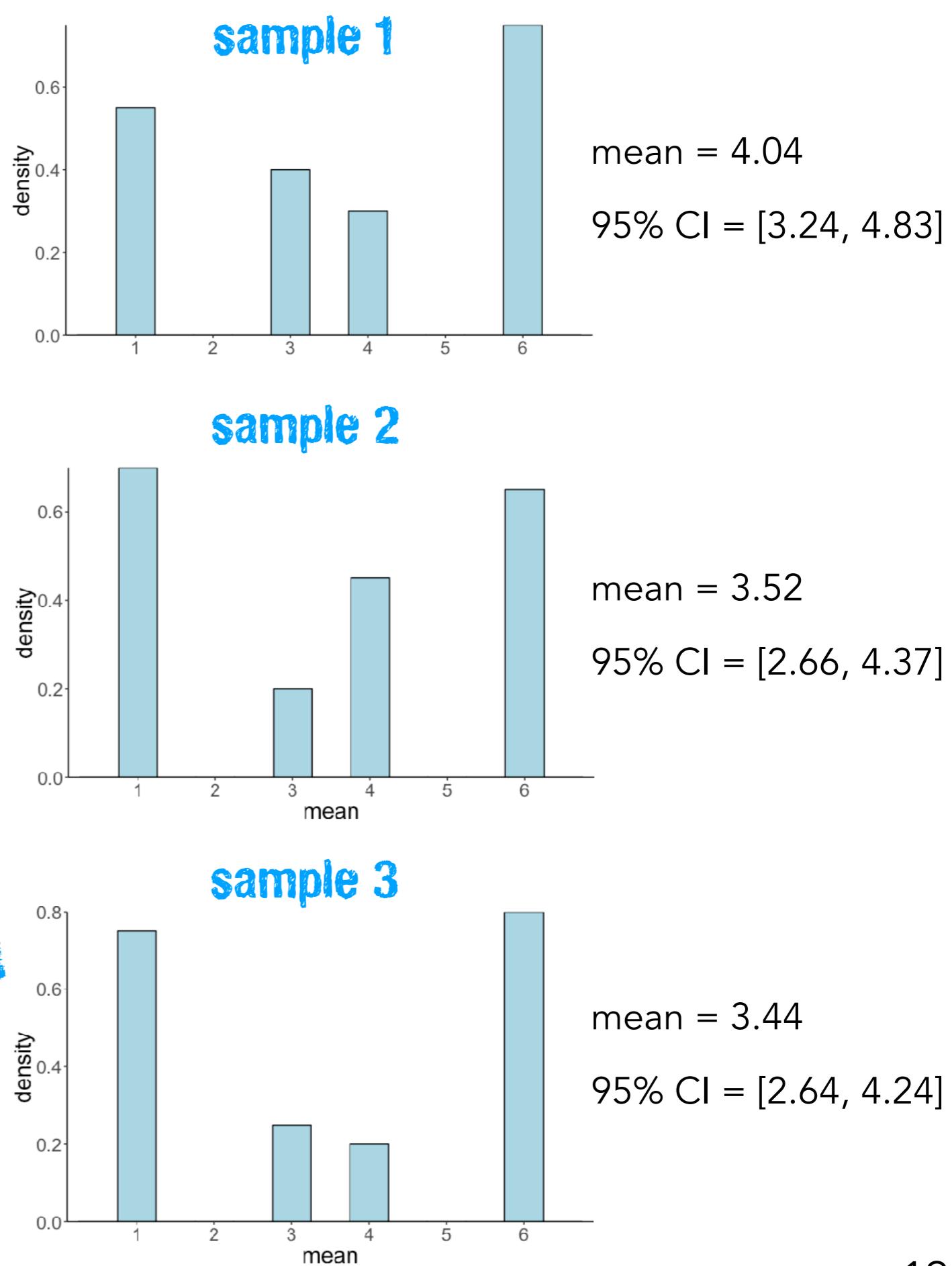
We can be 95% confident that the sample mean is in this interval.

Confidence interval

heavy metal distribution



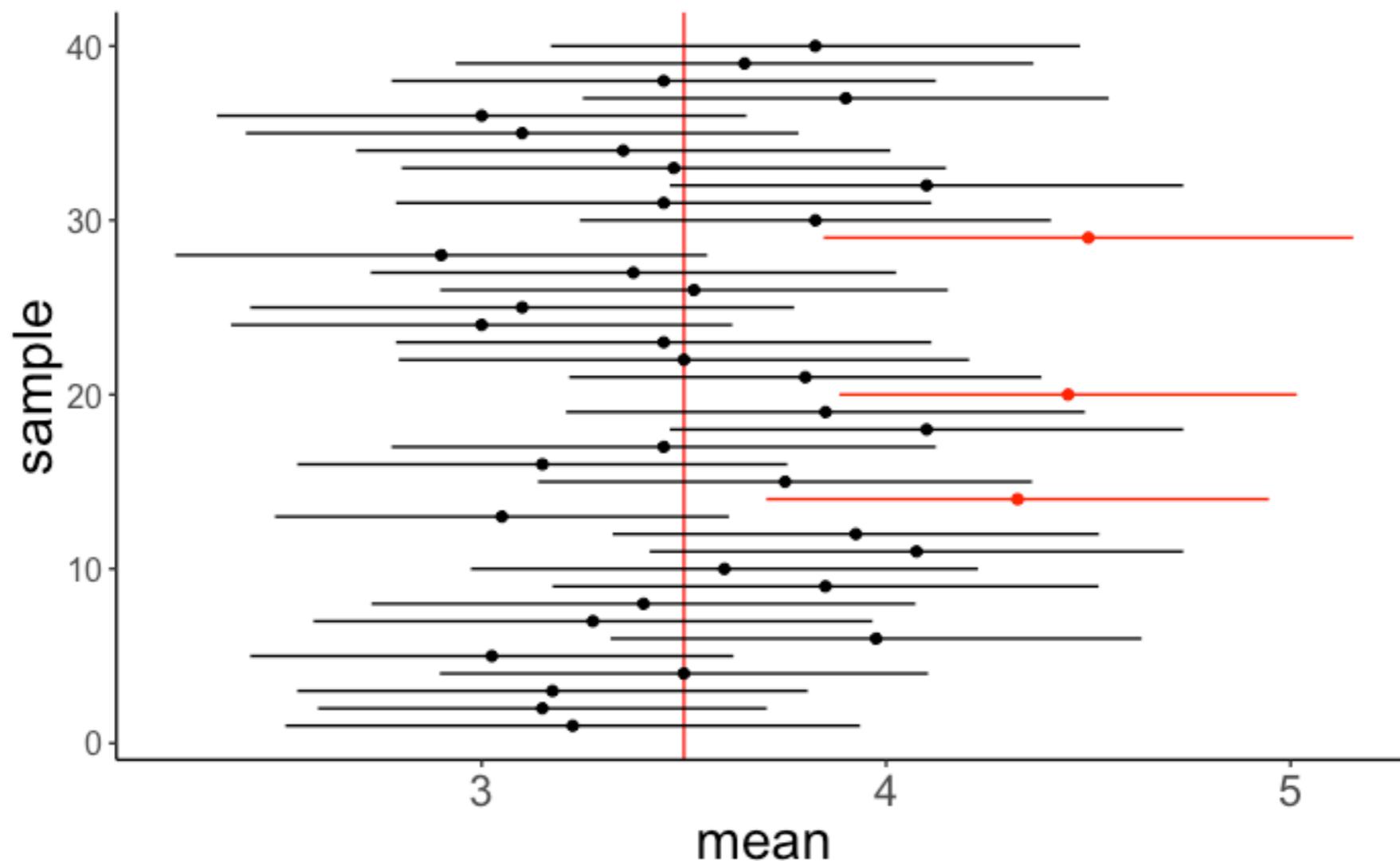
population distribution



95% confidence interval

Definition

"If we were to repeat the experiment over and over, then 95% of the time the confidence intervals contain the estimate of interest."



Hoekstra, R., Morey, R. D., Rouder, J. N., & Wagenmakers, E.-J. (2014). Robust Misinterpretation of Confidence Intervals. *Psychonomic Bulletin & Review*, 21(5), 1157-1164.

What can we say based on the result of our sample ($N = 40$):

Mean = 3.73 ± 0.63 (95% CI)?

95% of the time, the true population mean will be in this interval.

95% of random samples of size 40 will yield confidence intervals that contain the population mean.

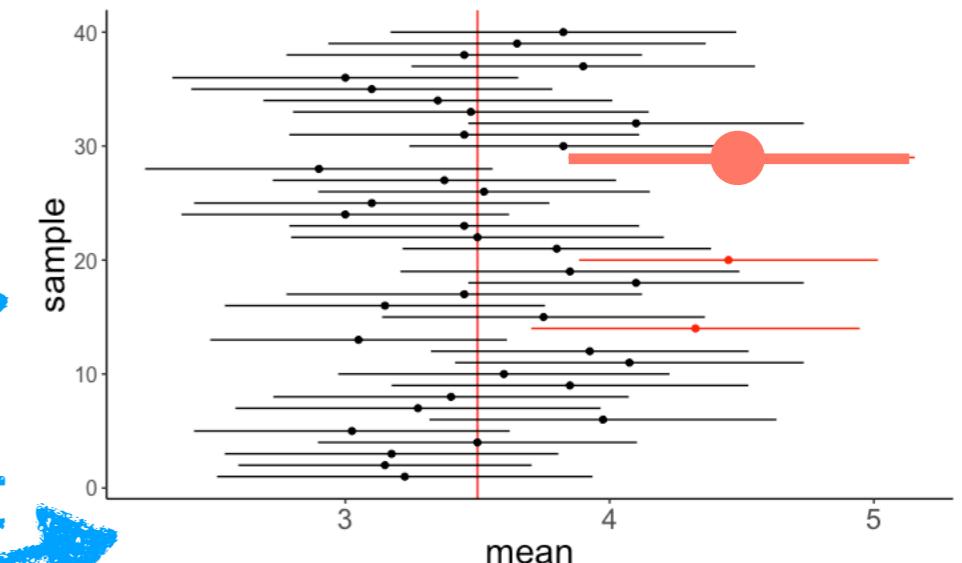
The sample means of 95% of the random samples of size 40 will be in this interval.

We can be 95% confident that the sample mean is in this interval.

It either is in this interval or isn't.

correct

incorrect

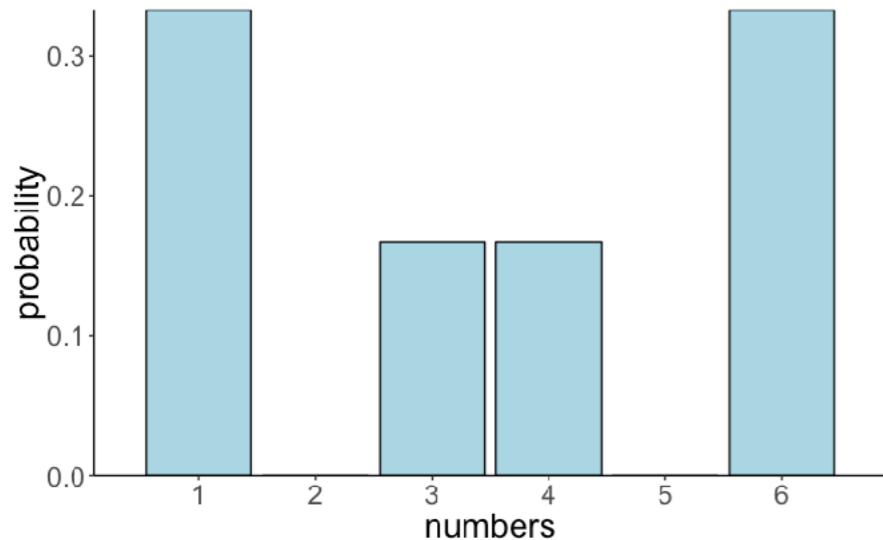


We know what the sample mean is.

Bootstrapping

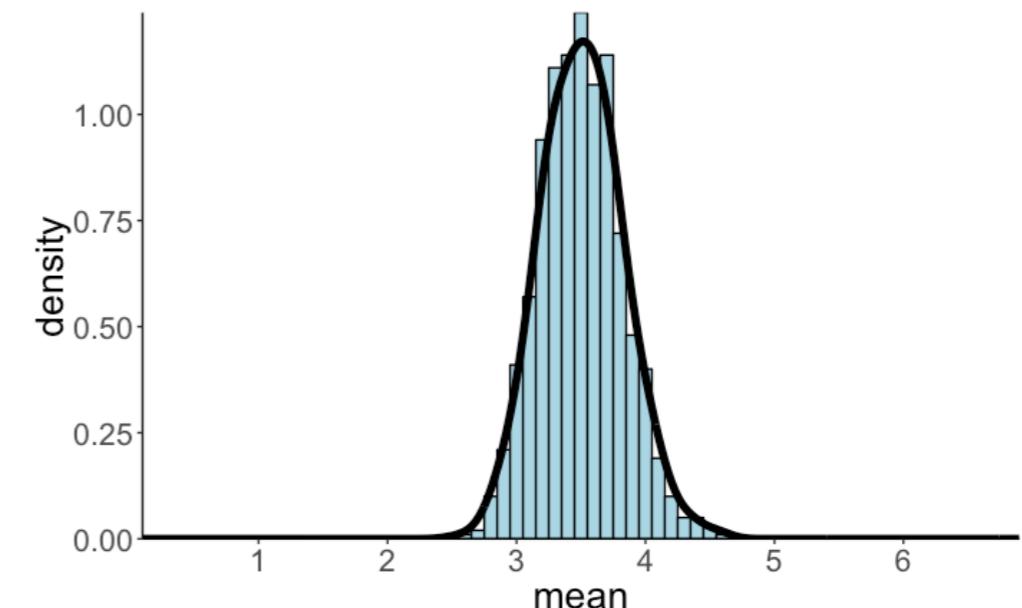


Bootstrap



population distribution

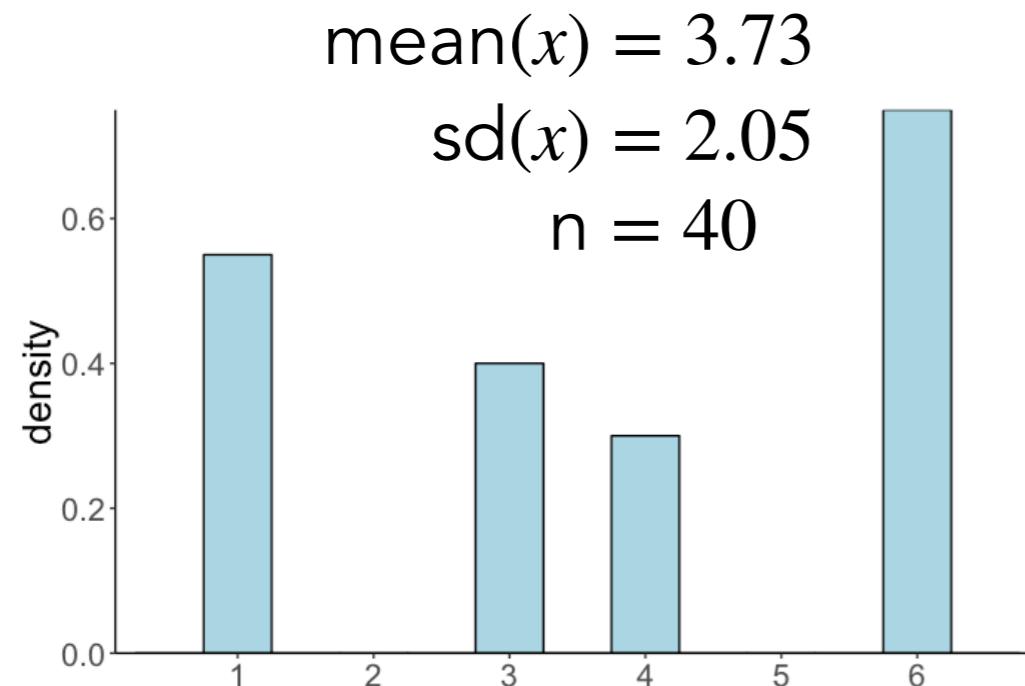
repeated
sampling



sampling distribution

but we don't know the population distribution!

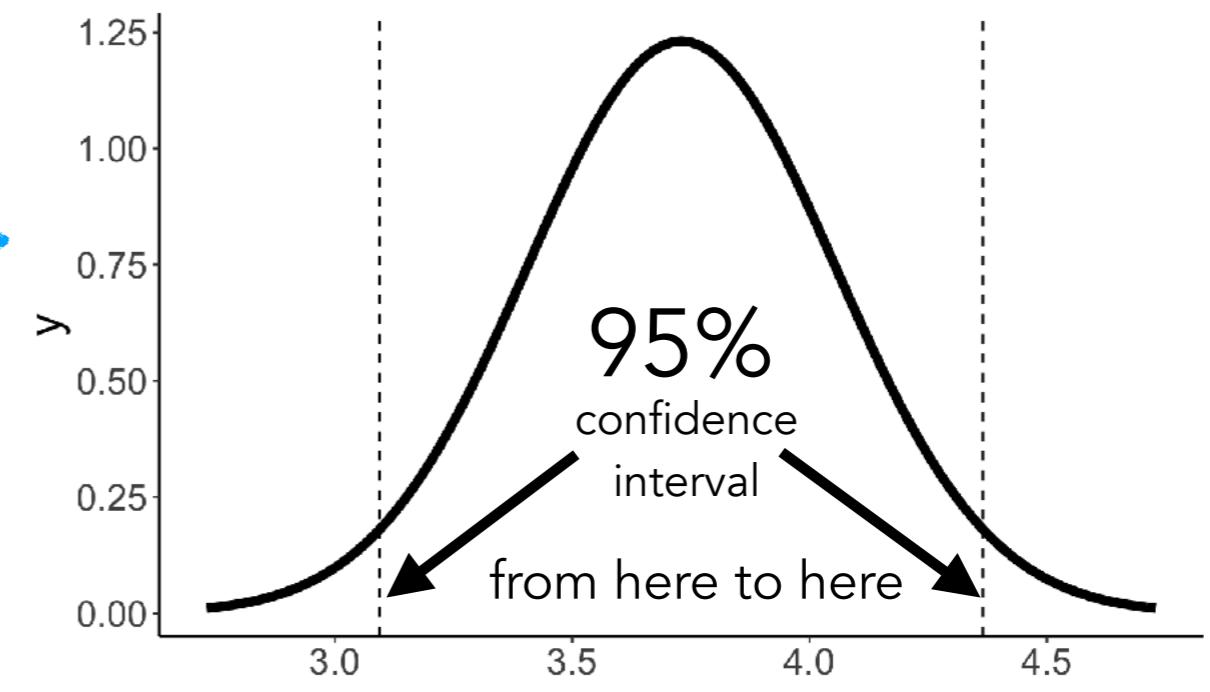
Bootstrap



assuming a
normal distribution

mean(\bar{x}) = 3.73
 $sd(\bar{x}) = \frac{sd(x)}{\sqrt{n}} = \frac{2.05}{\sqrt{40}} \approx 0.324$

sampling distribution of the mean

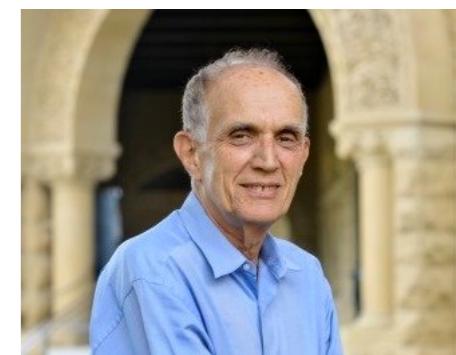


$\sim dnorm(., \text{mean} = 3.73, \text{sd} = 0.324)$ 24

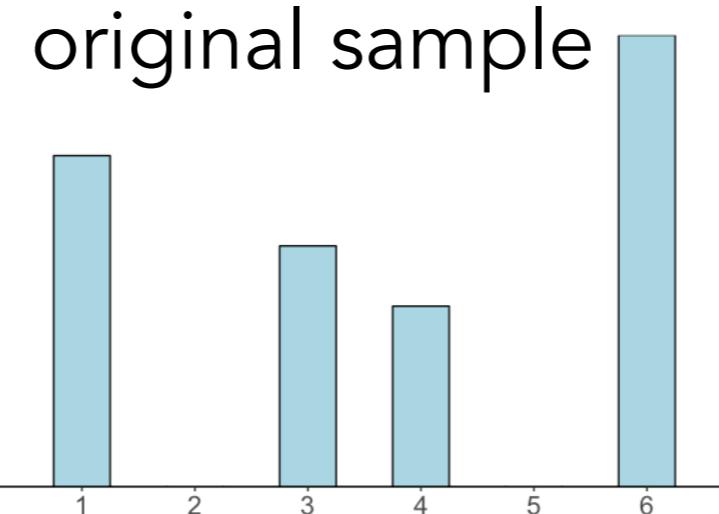
Bootstrap



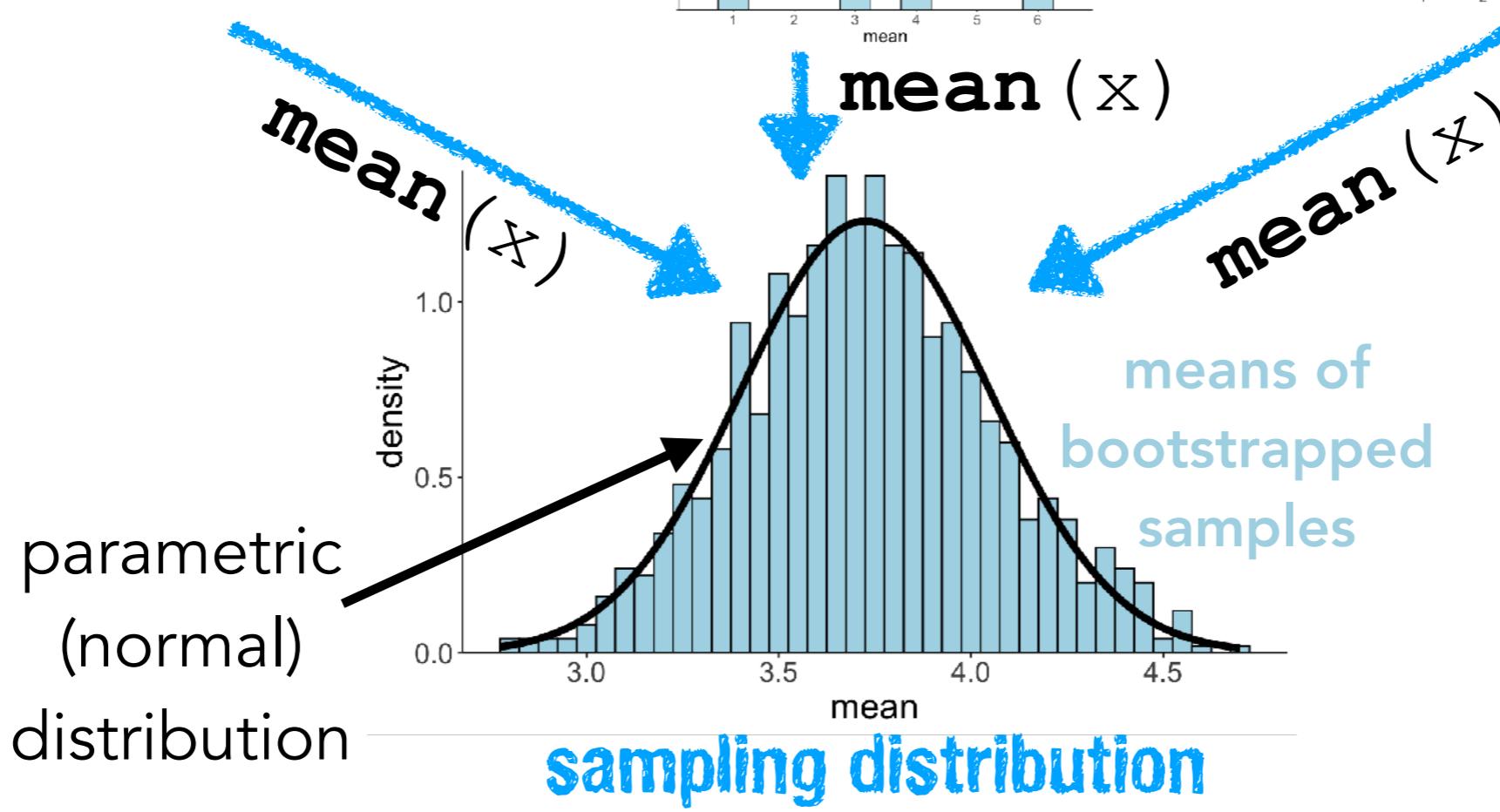
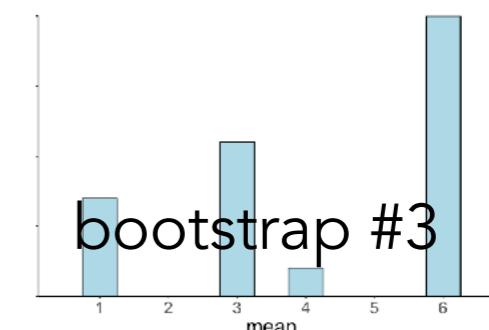
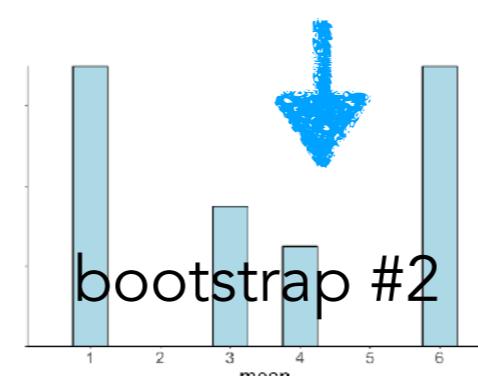
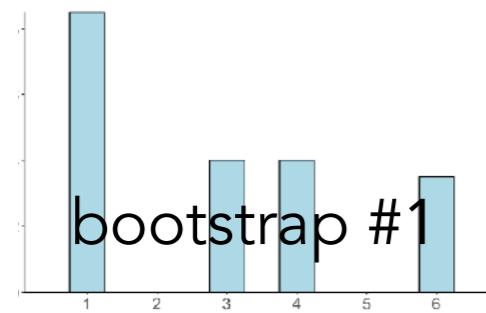
invented at Stanford



repeated sampling **with replacement**



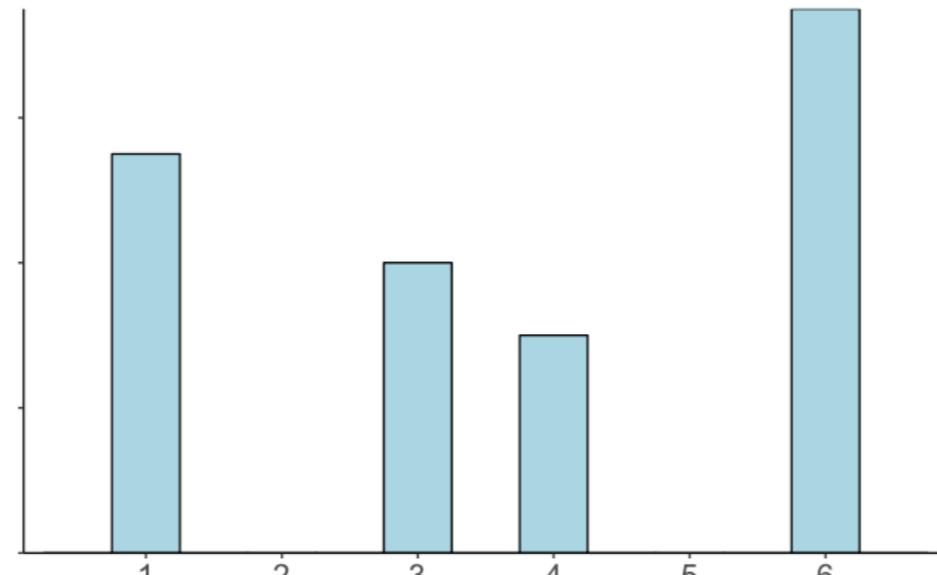
Bradley Efron



Bootstrap

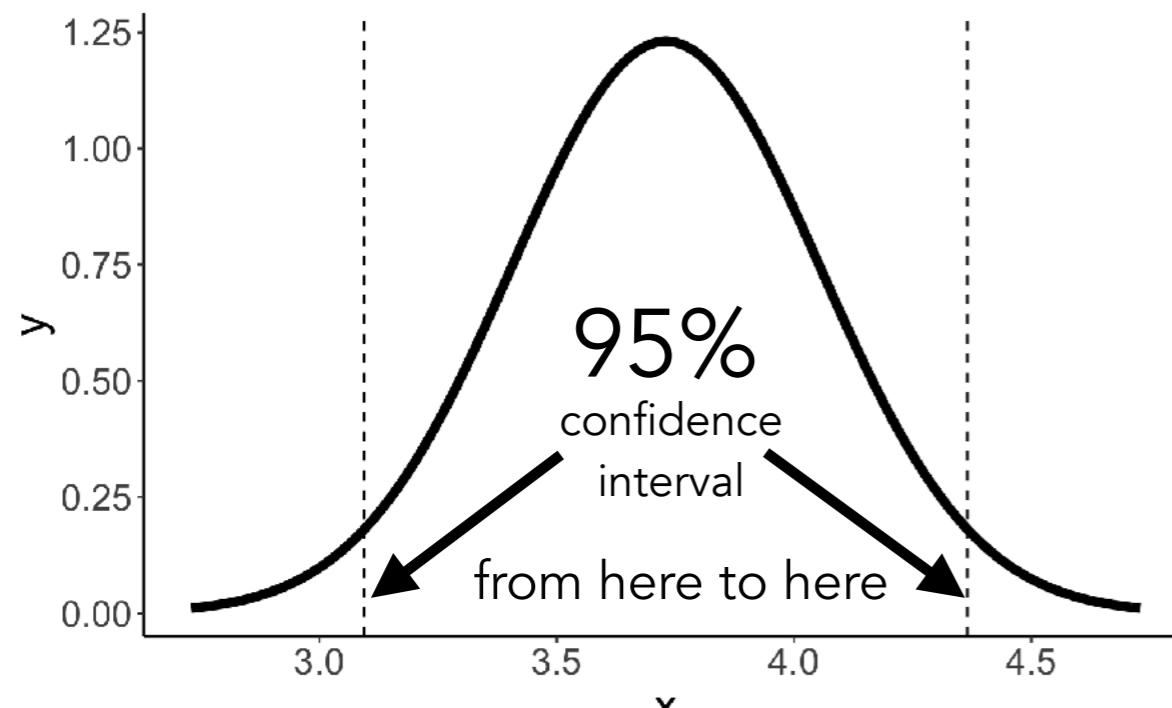
How can I get the confidence interval of a statistical estimate (such as the mean)?

make assumptions

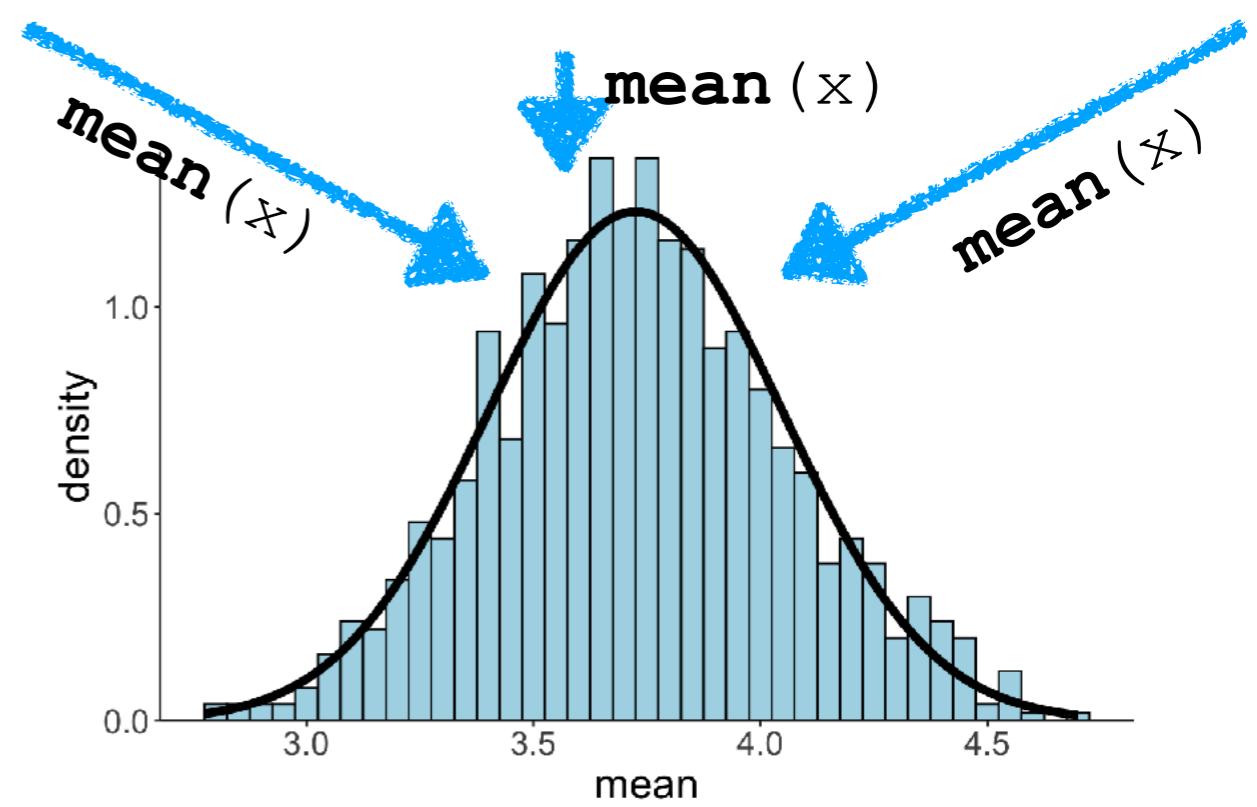


bootstrap

sampling distribution of the mean

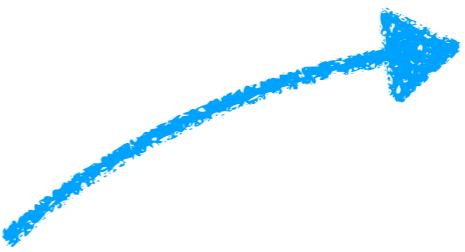


$\sim \text{dnorm}(\cdot, \text{mean} = 3.73, \text{sd} = 0.324)$



mean_cl_boot() explained

```
1 set.seed(1)
2
3 n = 10 # sample size per group
4 k = 3 # number of groups
5
6 df.data = tibble(participant = 1:(n*k),
7                   condition = as.factor(rep(1:k, each = n)),
8                   rating = rnorm(n*k, mean = 7, sd = 1))
```

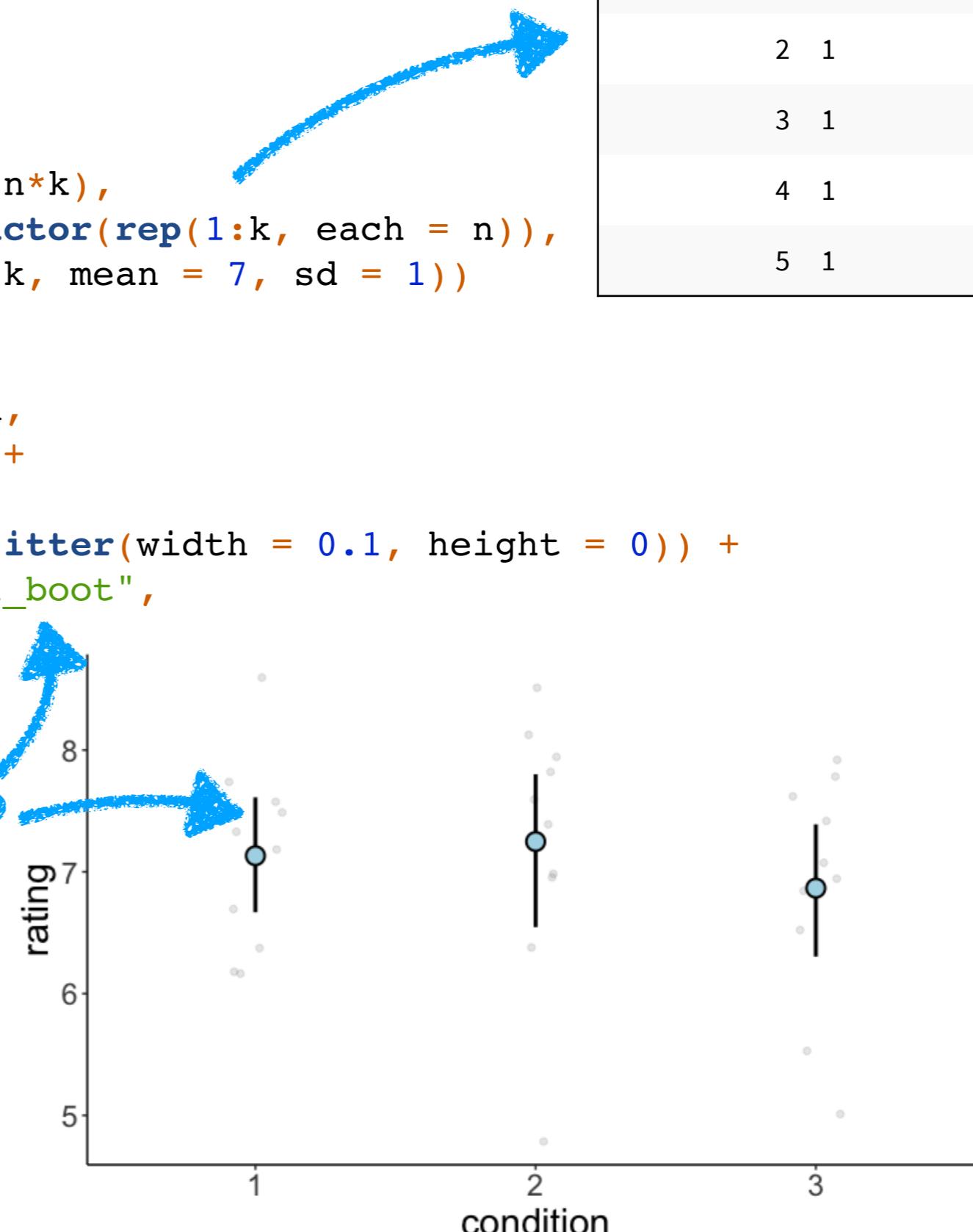


participant	condition	rating
1	1	6.37
2	1	7.18
3	1	6.16
4	1	8.60
5	1	7.33

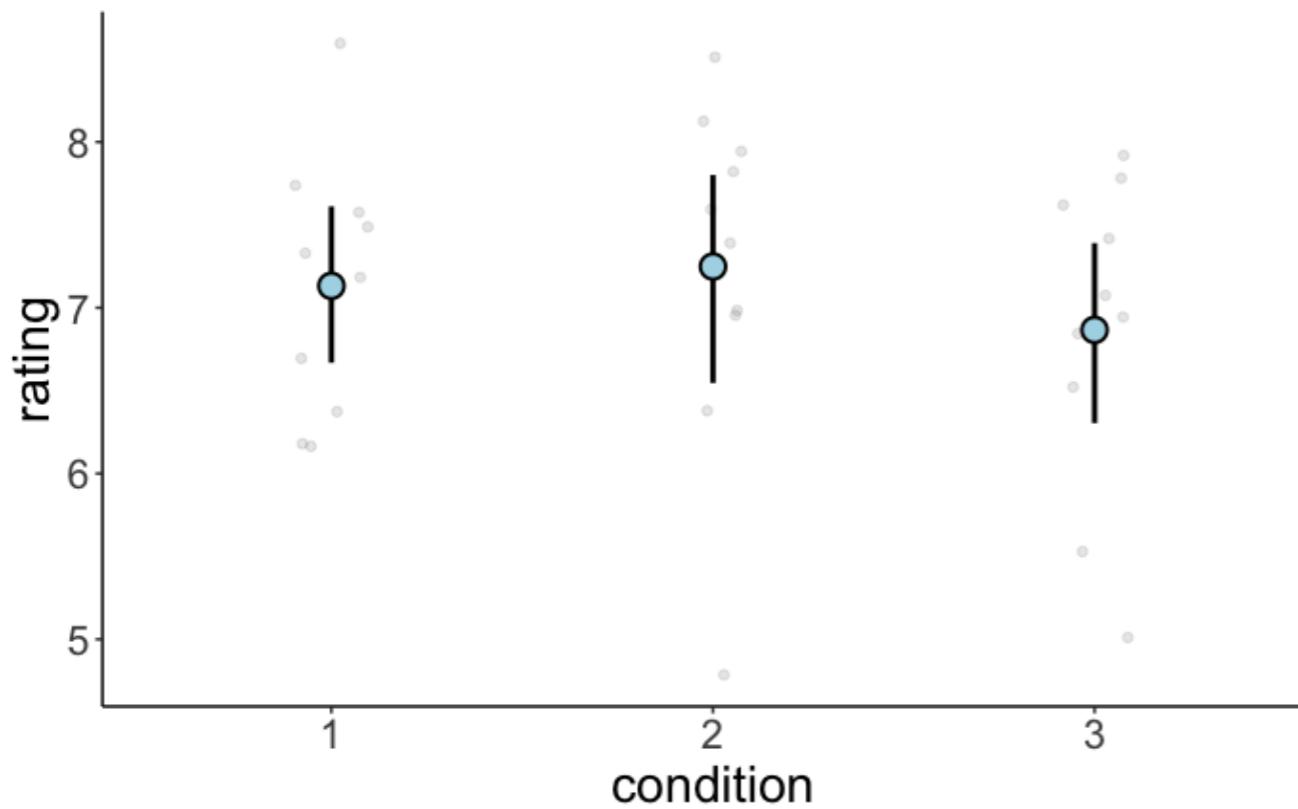
mean_cl_boot() explained

```
1 set.seed(1)
2
3 n = 10 # sample size per group
4 k = 3 # number of groups
5
6 df.data = tibble(participant = 1:(n*k),
7                   condition = as.factor(rep(1:k, each = n)),
8                   rating = rnorm(n*k, mean = 7, sd = 1))
9
10 ggplot(data = df.data,
11           mapping = aes(x = condition,
12                           y = rating)) +
13     geom_point(alpha = 0.1,
14                 position = position_jitter(width = 0.1, height = 0)) +
15     stat_summary(fun.data = "mean_cl_boot",
16                  shape = 21,
17                  size = 1,
18                  fill = "lightblue")
```

what is this magic?



mean_cl_boot() explained

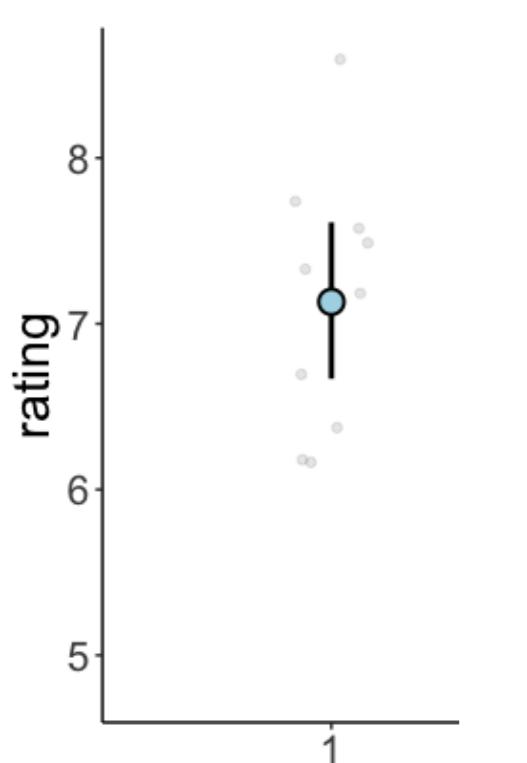


`ggplot_build(p)`

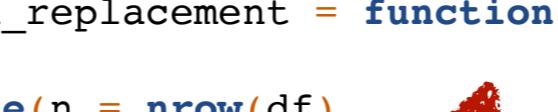
nice function for peeking
behind the scenes of ggplots

x	group	y	ymin	ymax	PANEL	flipped_aes	colour	size	linetype	shape	fill	alpha	stroke
1	1	7.13	6.70	7.59	1	FALSE	black	1	1	21	lightblue	NA	1
2	2	7.25	6.54	7.83	1	FALSE	black	1	1	21	lightblue	NA	1
3	3	6.87	6.26	7.39	1	FALSE	black	1	1	21	lightblue	NA	1

mean_cl_boot() explained



```
1 set.seed(1)
2
3 df.condition1 = df.data %>%
4   filter(condition == 1)
5
6 fun.sample_with_replacement = function(df) {
7   df %>%
8     slice_sample(n = nrow(df),
9                   replace = T) %>%
10    summarize(mean = mean(rating)) %>%
11    pull(mean)
12 }
```



this is what makes
this a bootstrap

original

participant	condition	rating
1	1	6.37
2	1	7.18
3	1	6.16
4	1	8.60
5	1	7.33
6	1	6.18
7	1	7.49
8	1	7.74
9	1	7.58
10	1	6.69

bootstrap #1

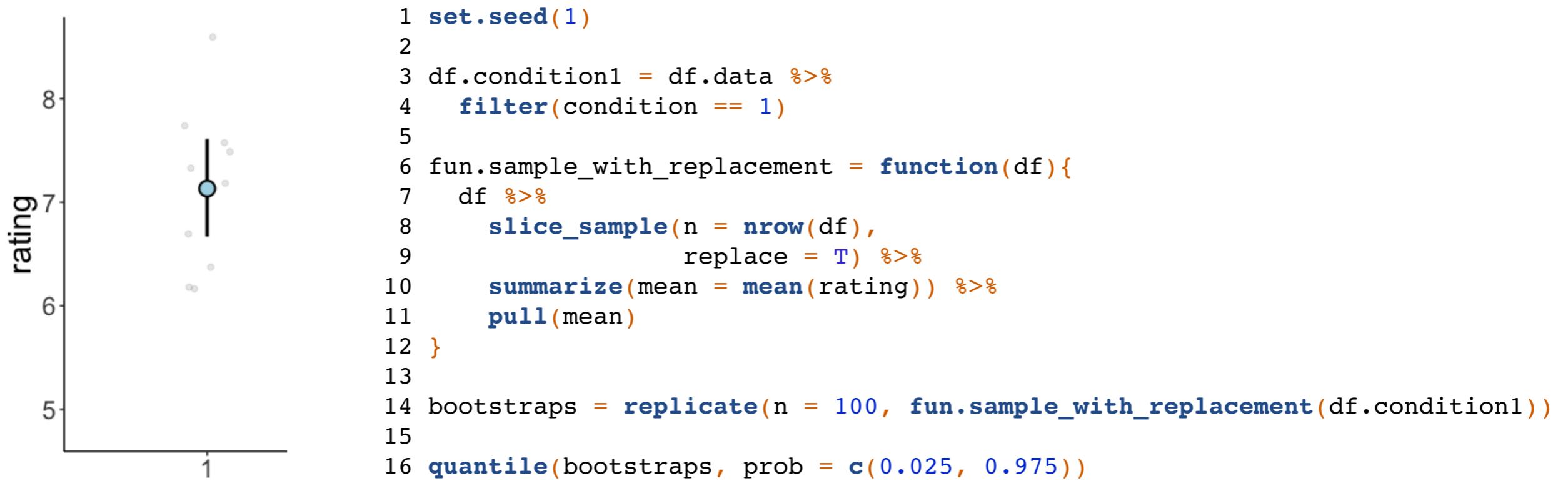
participant	condition	rating
9	1	7.58
3	1	6.16
3	1	6.16
2	1	7.18
3	1	6.16
note the repeat		
7	1	7.49
4	1	8.60
8	1	7.74
5	1	7.33

bootstrap #2

participant	condition	rating
1	1	6.37
1	1	6.37
6	1	6.18
1	1	6.37
4	1	8.60
Excluded observations		18
1	1	6.37
9	1	7.58
4	1	8.60
3	1	6.16

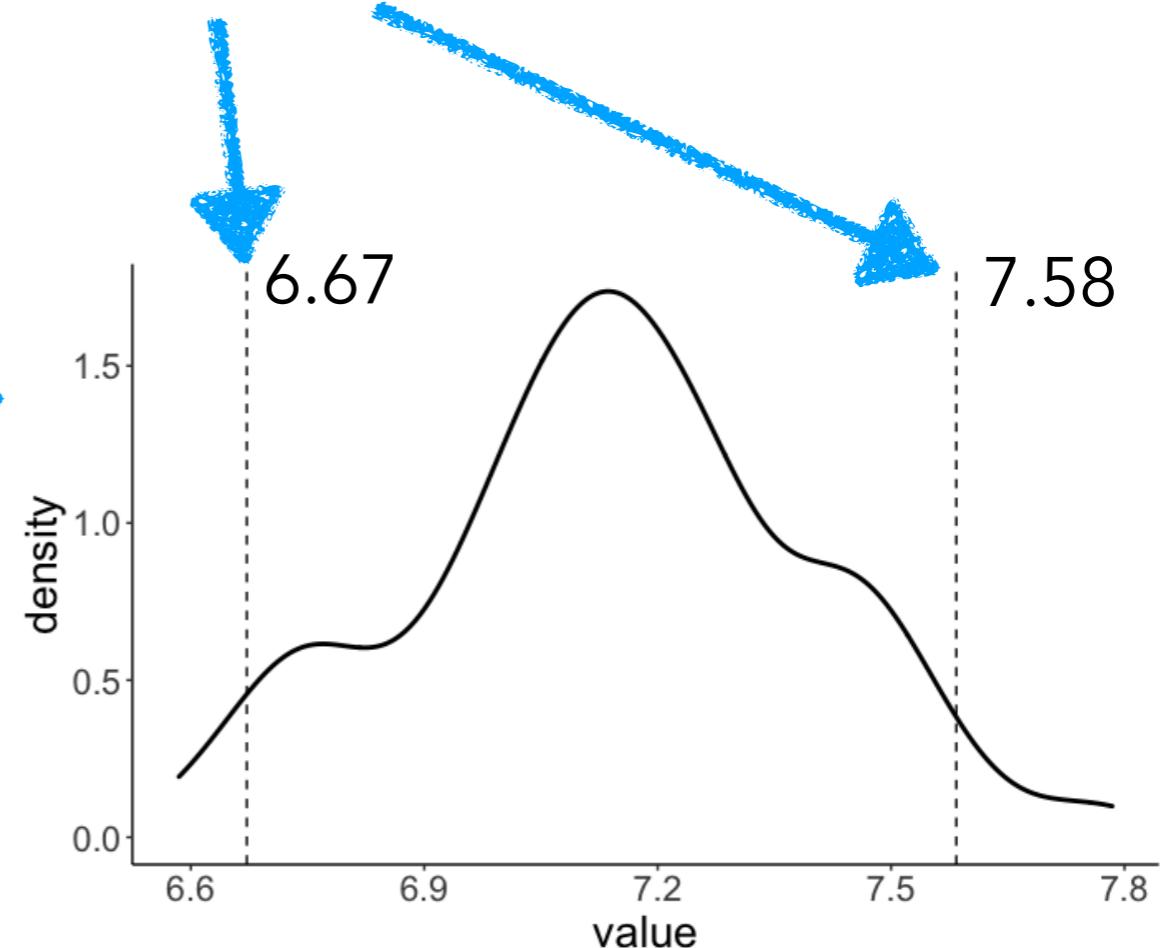
note the repeated observations ¹⁸

mean_cl_boot() explained



x	group	y	ymin	ymax
1	1	7.13	6.70	7.59

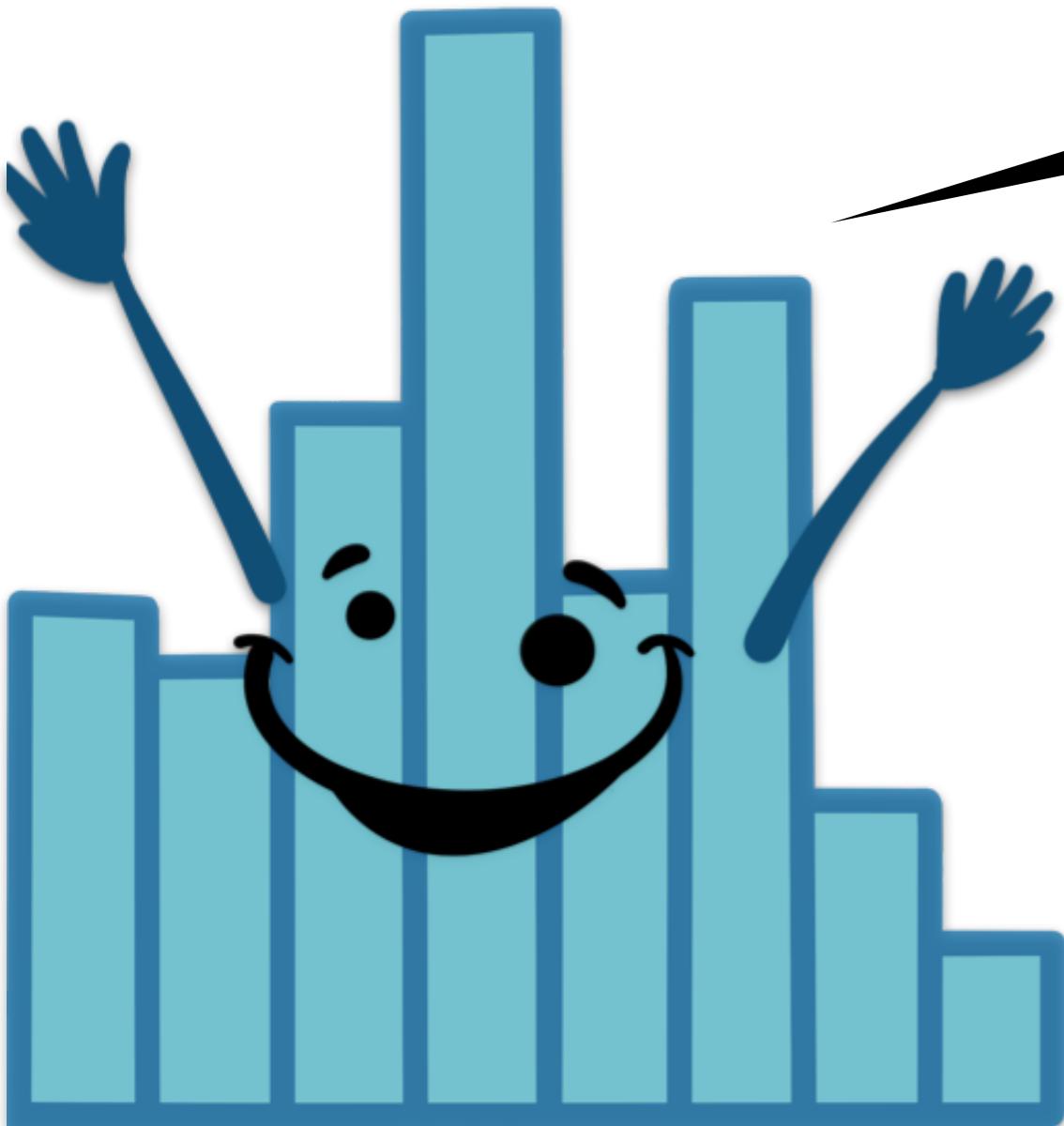
```
1 ggplot(data = as_tibble(bootstraps),
2   mapping = aes(x = value)) +
3   geom_density(size = 1) +
4   geom_vline(xintercept = quantile(bootstraps,
5                             probs = c(0.025, 0.975)),
6   linetype = 2)
```



We're listening to
"Weekend" by
"Clubhouse",
submitted by Sarah Wu

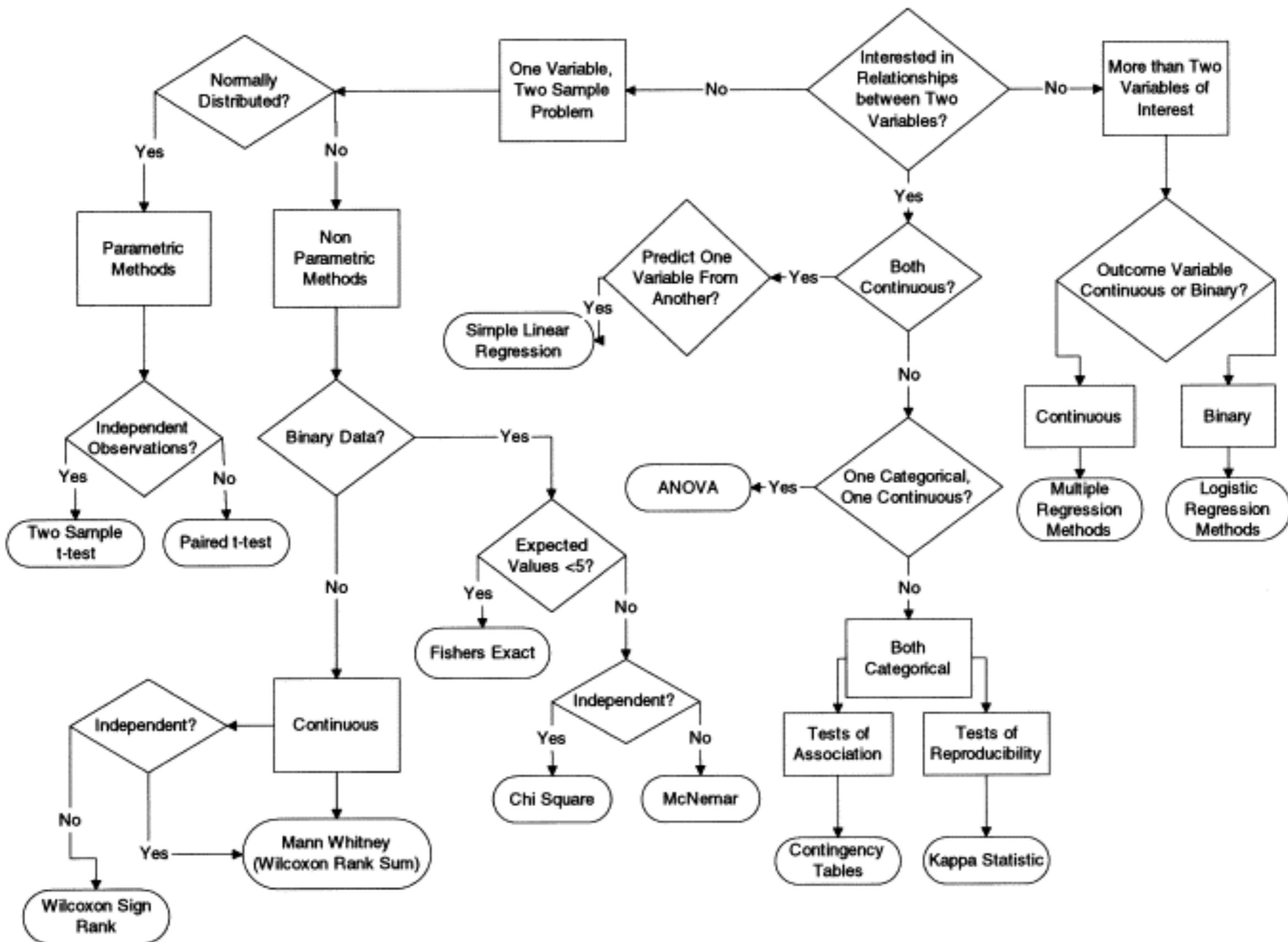
02:00

stretch break!



Cookbook vs. Model Comparison

The cookbook approach

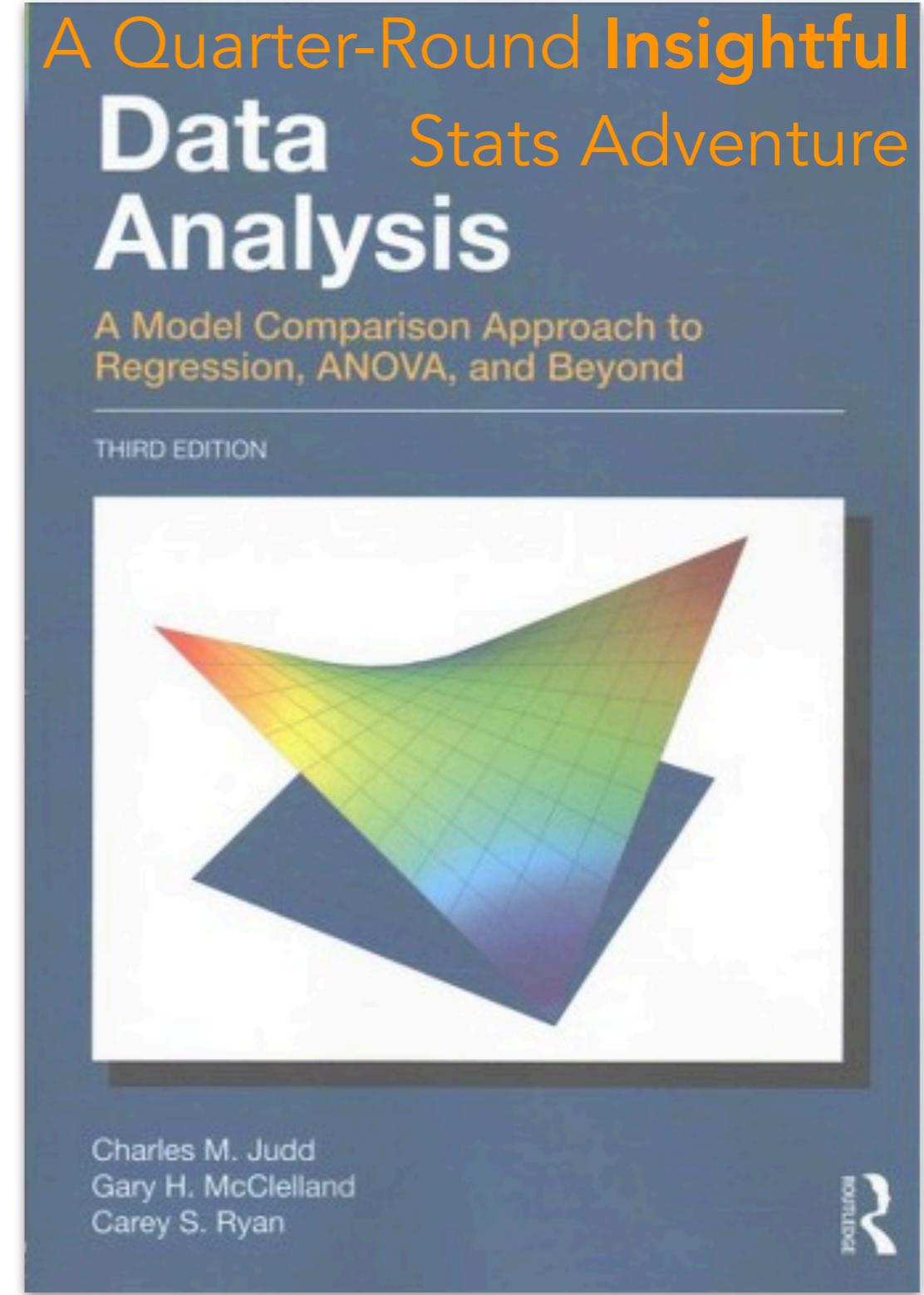


The cookbook approach



- many statistics textbooks are organized in this way
- works reasonably well if what we want to cook is in the book
- leaves us with no idea what to do if we can't find a recipe

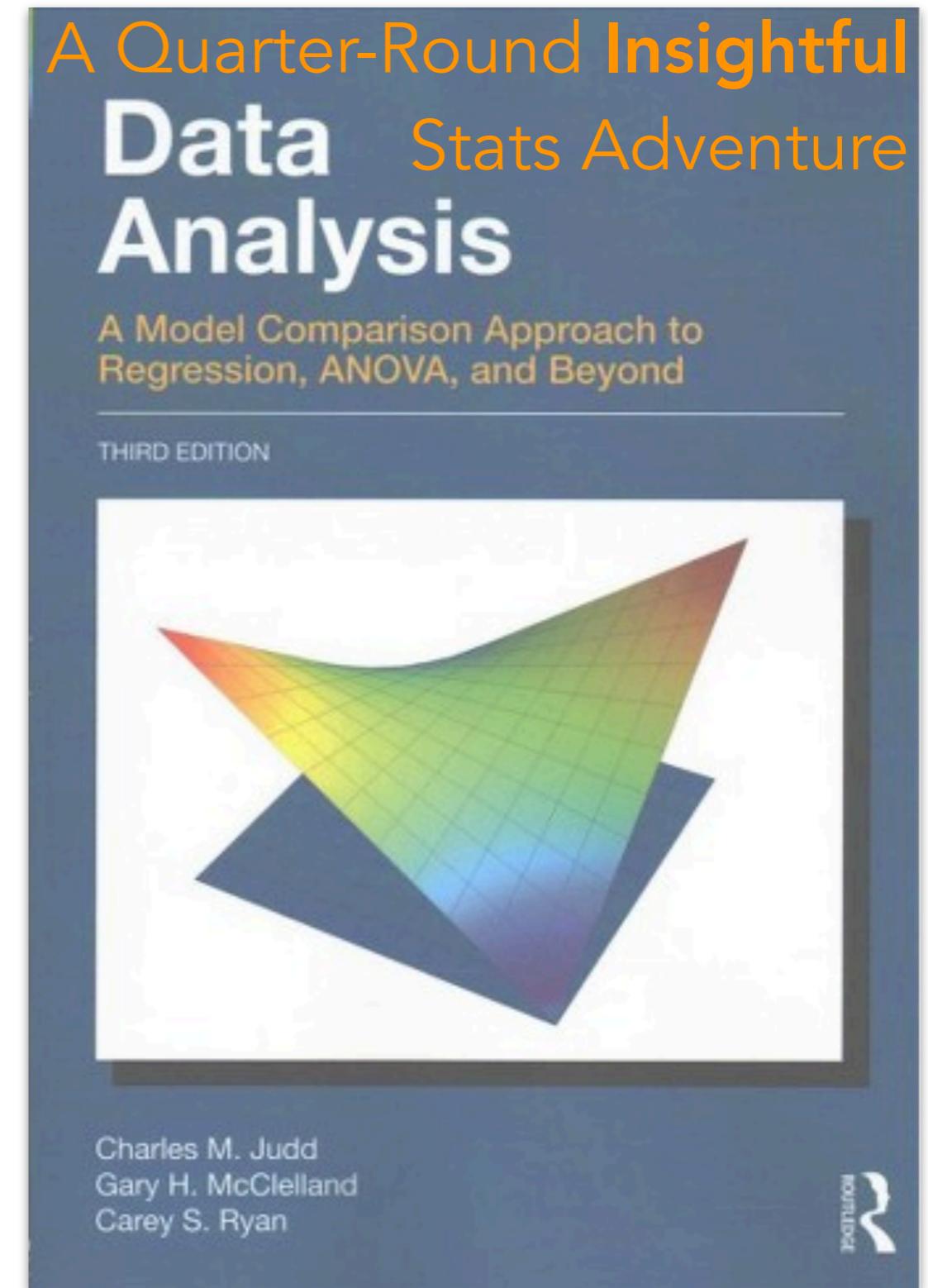
Model comparison approach



Judd, C. M., McClelland, G. H., & Ryan, C. S. (2011). *Data analysis: A model comparison approach*. Routledge.

Model comparison approach

- more flexible approach
- hopefully generates better insight
- thinking of statistical analysis as modeling
- allows for a smoother transition into Bayesian data analysis, and probabilistic modeling more generally



Judd, C. M., McClelland, G. H., & Ryan, C. S. (2011). *Data analysis: A model comparison approach*. Routledge.

Modeling data

Data = Model + Error



what's a good
model?



how shall we
define this?

= residual: the part that's left over after we have used the model to predict/explain the data



$$\text{Error} = \text{Data} - \text{Model}$$

to reduce error we can:

improve the quality of the data

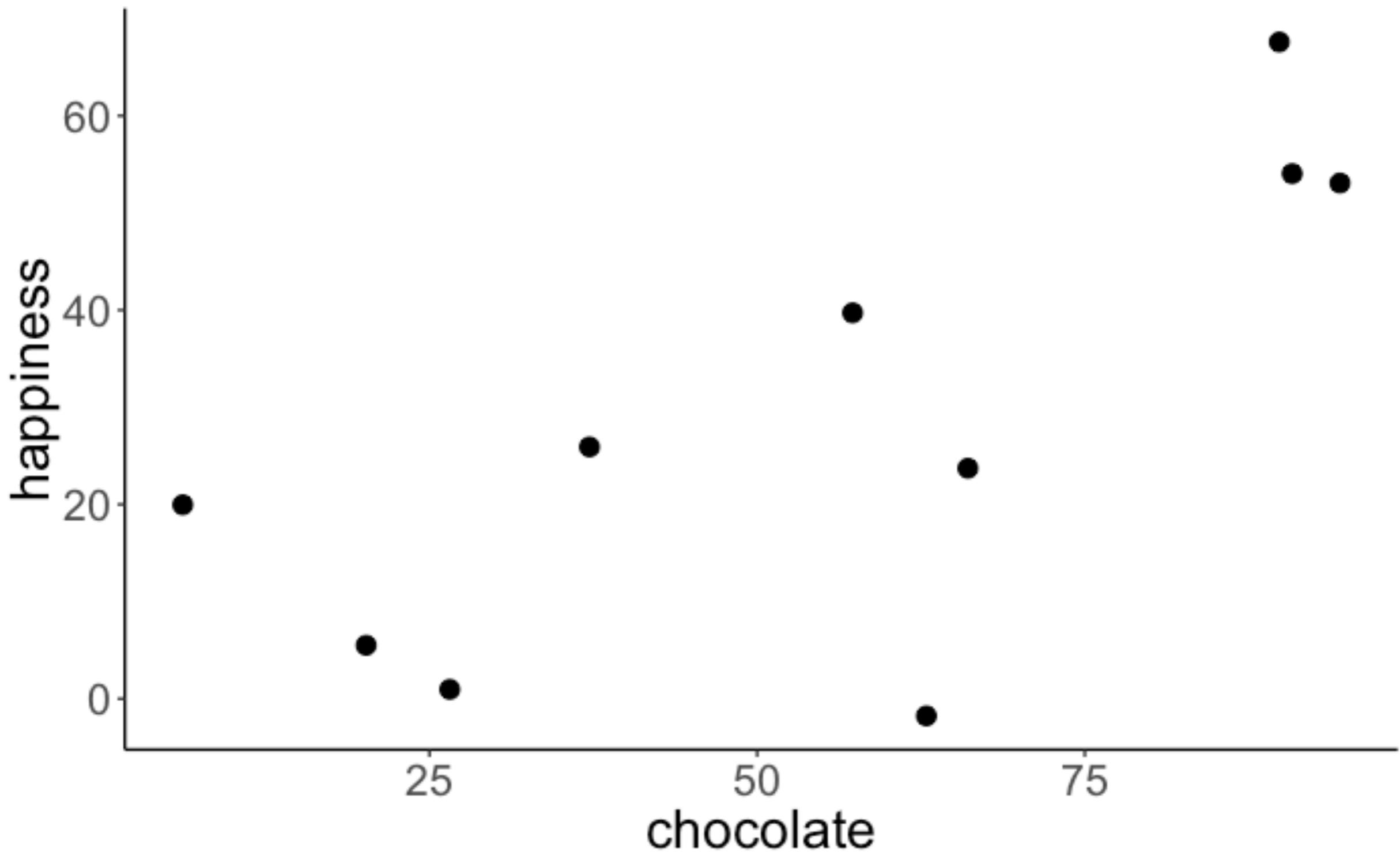
e.g. run good experiments



improve the model

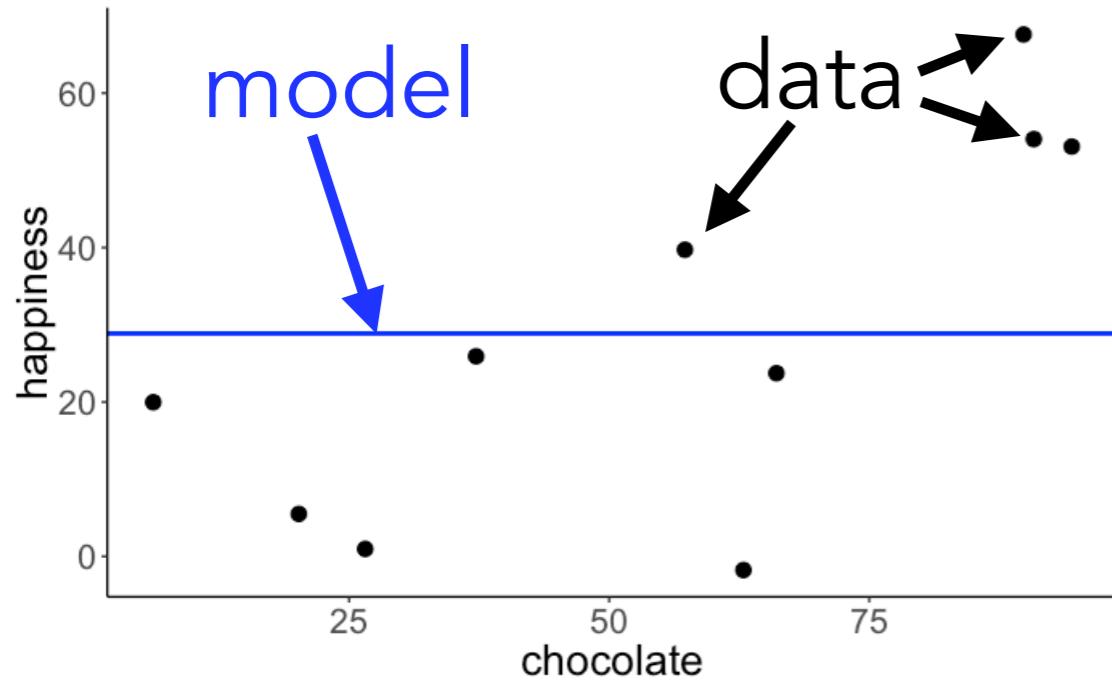
e.g. make predictions conditional on additional information

Is there a relationship between chocolate consumption and happiness?

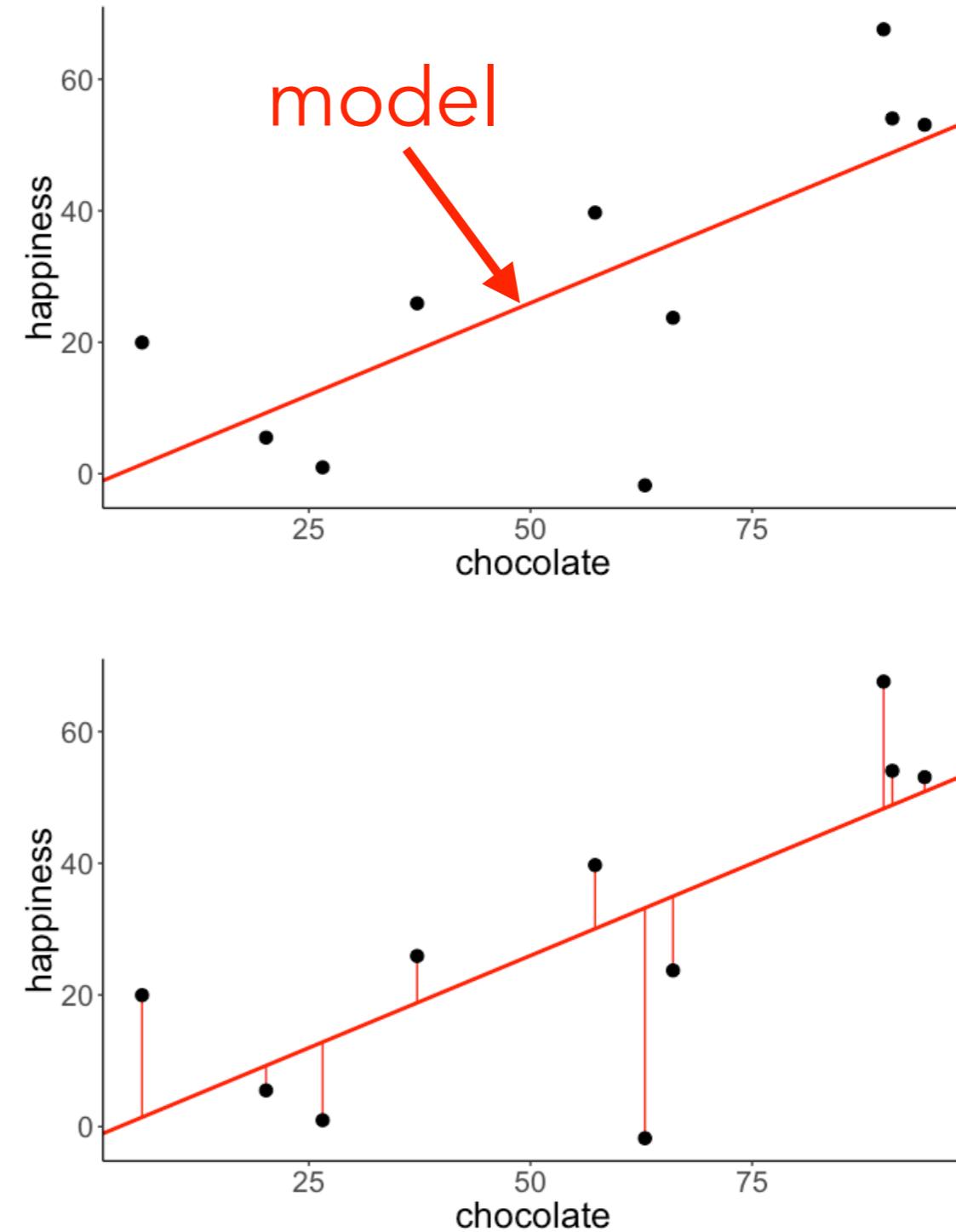
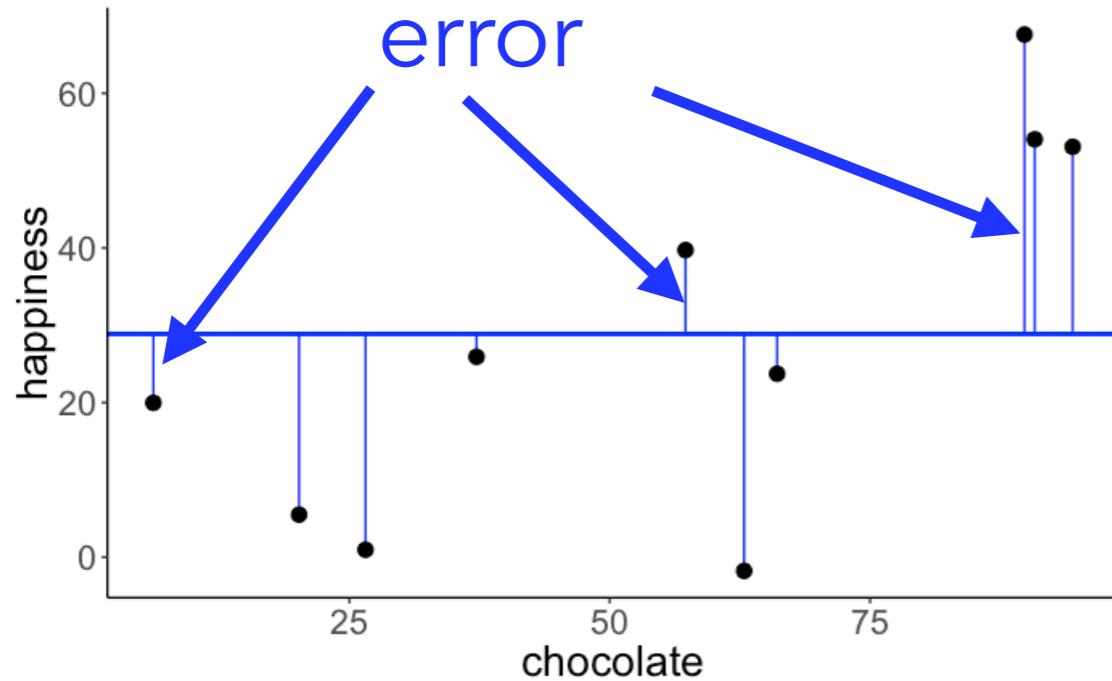


Data = Model + Error

H_0 : Chocolate consumption and happiness are unrelated.



H_1 : Chocolate consumption and happiness are related.



ERROR

Error = Data - Model



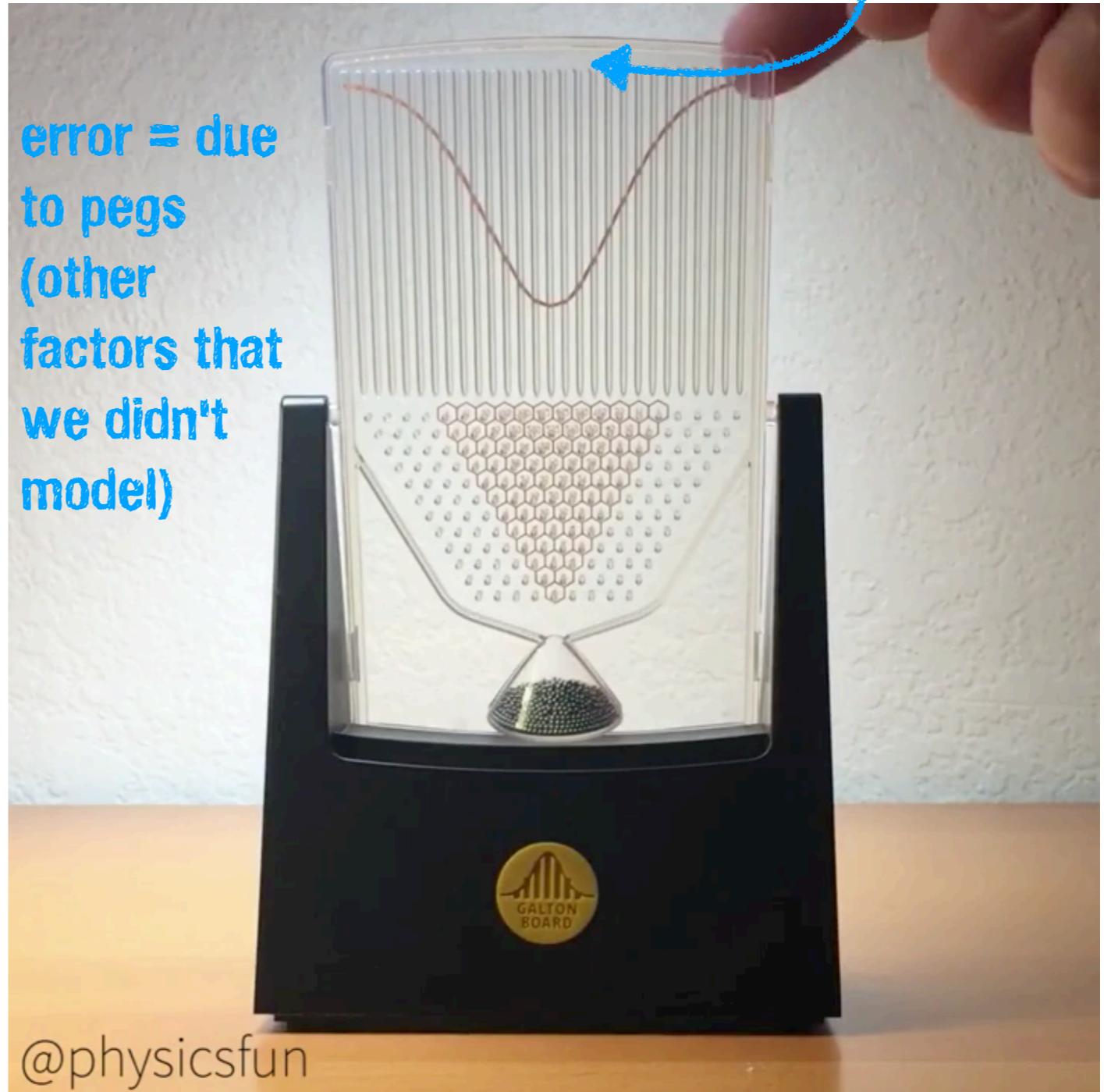
how shall we
define this?

ERROR

1. We assume that the error between model and data is due to (a potentially large number of) factors that we didn't take into account.
2. We assume that each of these factors influences the data in **an additive way** (some pulling in one, others pulling in another direction).

ERROR

1. We assume that the error between model and data is due to (a potentially large number of) factors that we didn't take into account.
2. We assume that each of these factors influences the data in **an additive way** (some pulling in one, others pulling in another direction).



Result: normal distribution

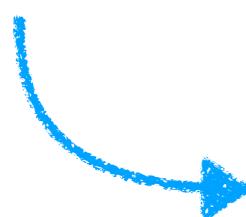
ERROR

$$\text{Error} = \text{Data} - \text{Model}$$



how shall we
define this?

concretely: we will fit our models such that they minimize
the **sum of squared errors**



why squared error?

- we can sum up all the error terms
(positive and negative prediction errors don't cancel out)
- larger errors are weighed more

Assumption of normal distribution

$$\text{Error} = \text{Data} - \text{Model}$$



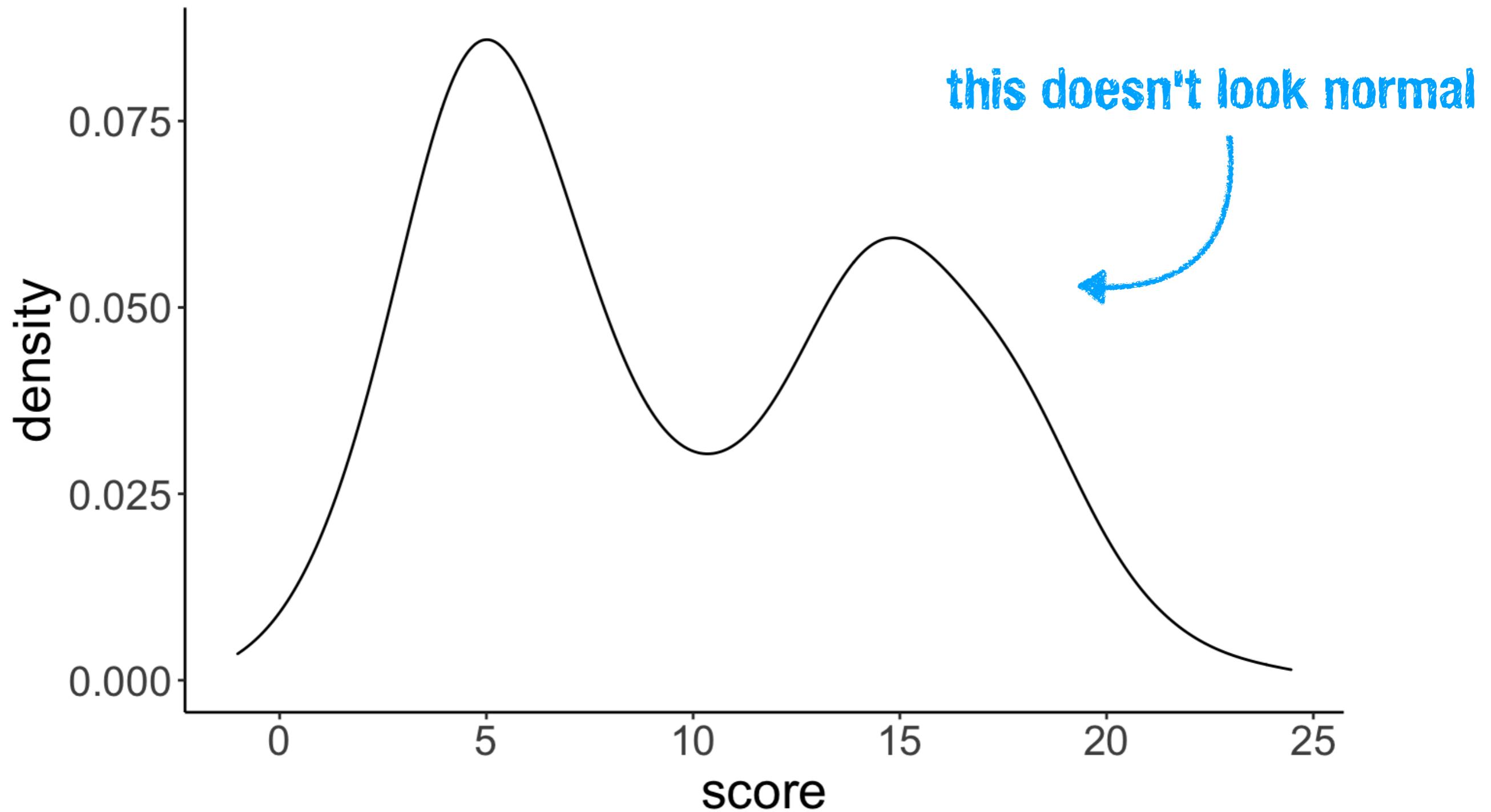
**assumed to be
normally
distributed**



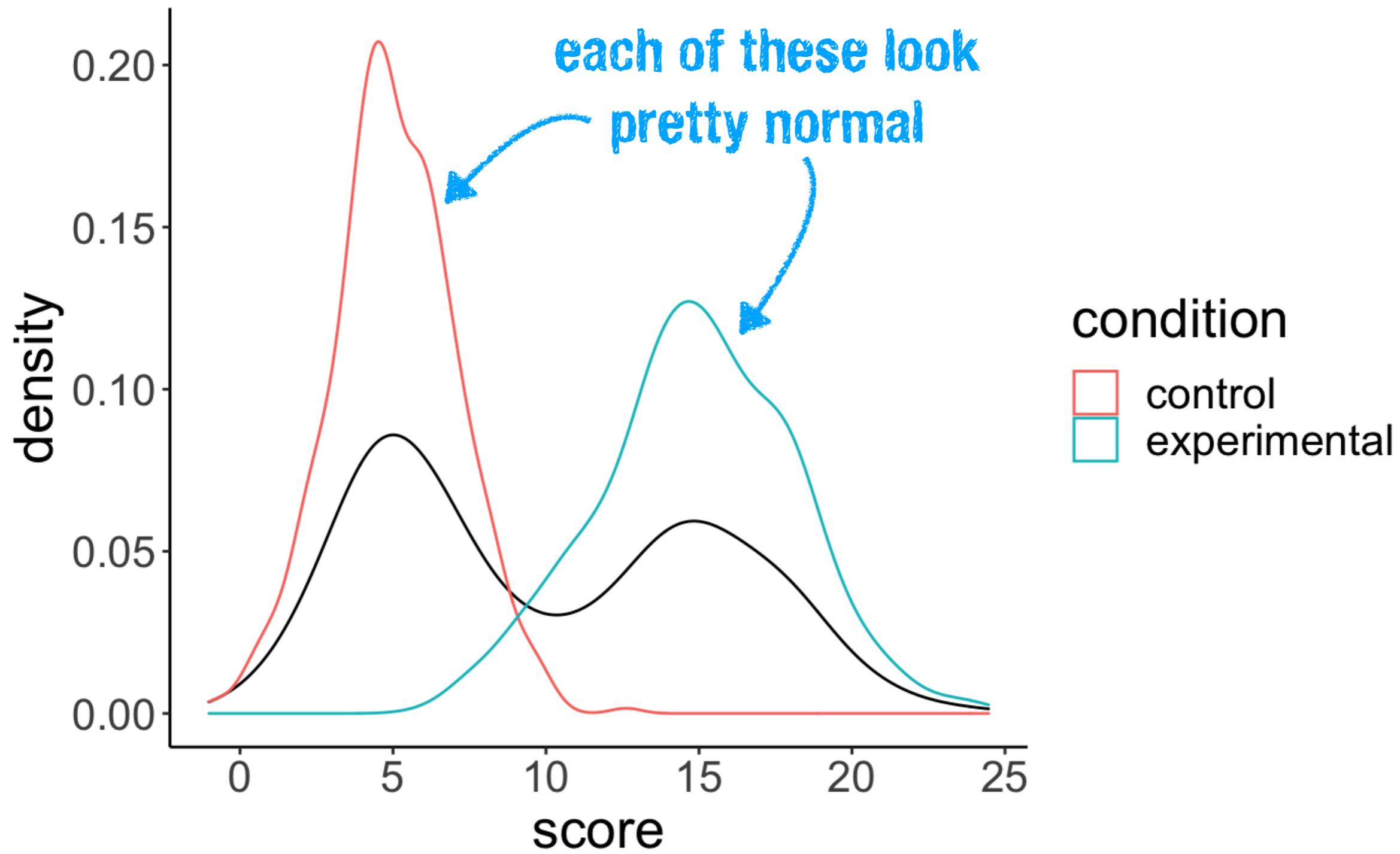
**don't need to
be normally
distributed!!**

very common misconception!!!

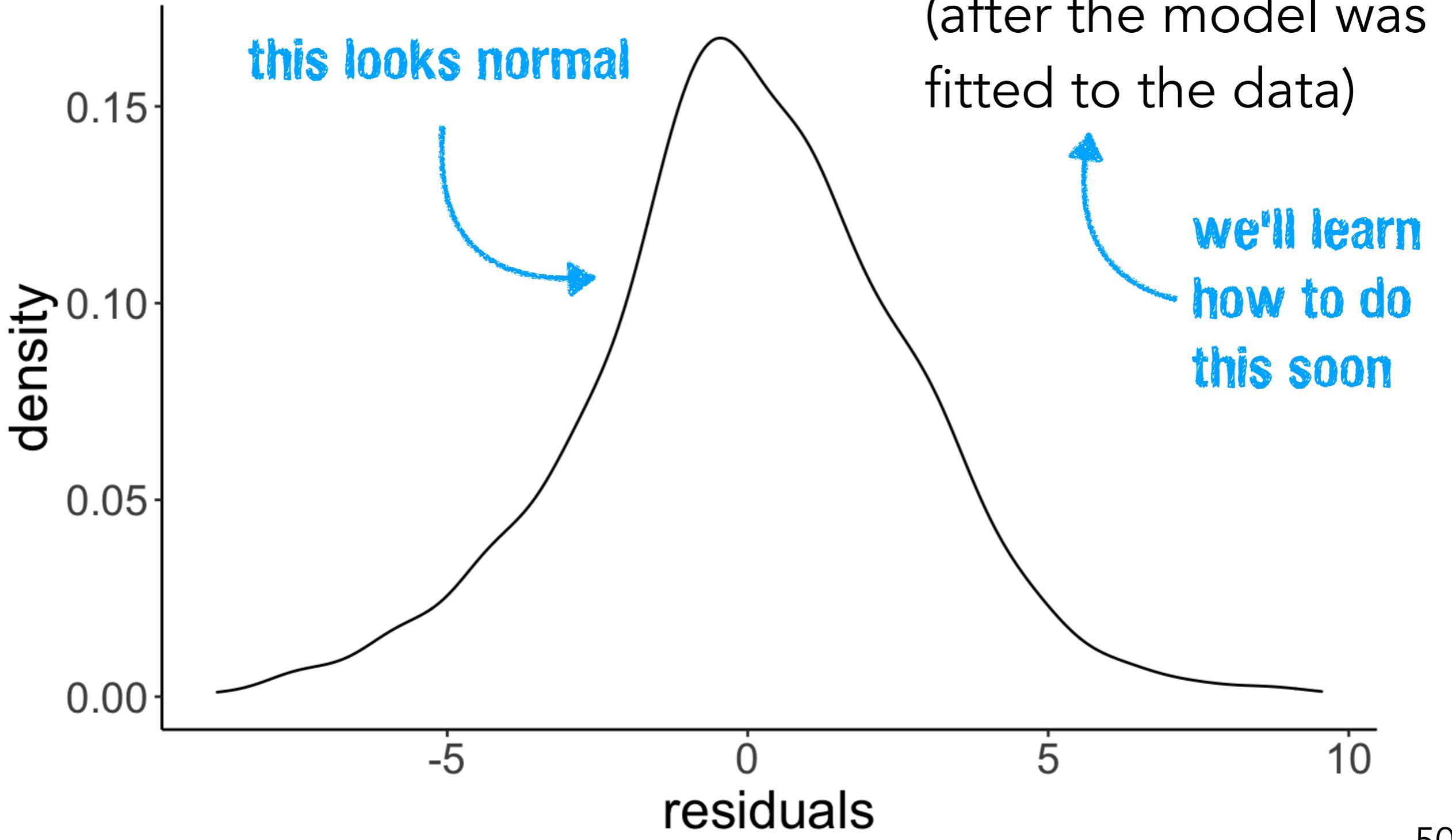
Distribution of **test scores**



Distribution of test scores



Distribution of the residuals



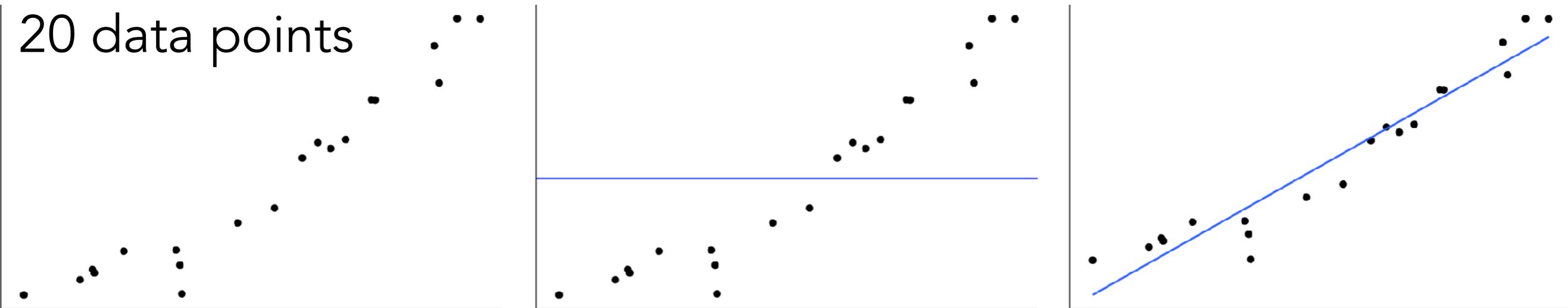
Data = Model + Error



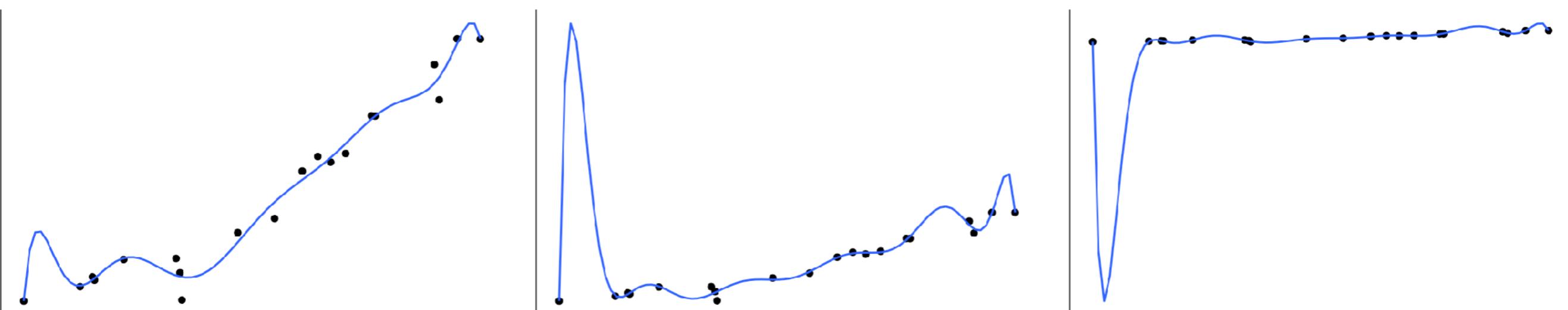
what makes for
a good model?

- we build models with parameters, and fit those parameters to **minimize error**
- adding additional parameters to the model will always improve the model fit and reduce error
- fundamental trade-off between **simplicity** and **accuracy**

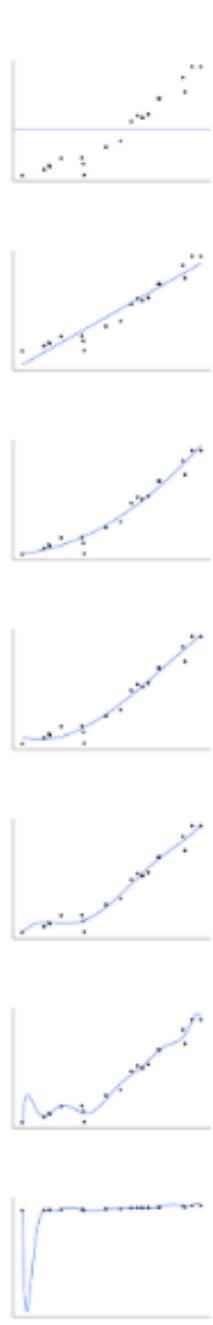
20 data points



Which model describes the data best?



Which model describes the data best





**THE BEST WAY TO
EXPLAIN OVERFITTING**

Example

Percentage of households that had internet access in the year 2013 by US state

i	internet	state	college	auto	density
1	79.0	AK	28.0	1.2	1.2
2	63.5	AL	23.5	1.3	94.4
3	60.9	AR	20.6	1.7	56.0
4	73.9	AZ	27.4	1.3	56.3
5	77.9	CA	31.0	0.8	239.1
6	79.4	CO	37.8	1.0	48.5
7	77.5	CT	37.2	1.0	738.1
8	74.5	DE	29.8	1.1	460.8
9	74.3	FL	27.2	1.2	350.6
10	72.2	GA	28.3	1.1	168.4

set before looking at the data

$$\text{model}_1: Y_i = 75 + \text{ERROR}$$

worth it?

fit to the data

$$\text{model}_2: Y_i = \beta_0 + \text{ERROR}$$

worth it?

additional predictor

$$\text{model}_3: Y_i = \beta_0 + \beta_1 X_i + \text{ERROR}$$

fit to the data

Compact model

model_C: $Y_i = \beta_0 + \text{ERROR}$

Augmented model

model_A: $Y_i = \beta_0 + \beta_1 X_i + \text{ERROR}$

$$\text{ERROR}(C) \geq \text{ERROR}(A)$$

Proportional reduction in error (PRE)

$$\text{PRE} = \frac{\text{ERROR}(C) - \text{ERROR}(A)}{\text{ERROR}(C)}$$

Compact model

$$\text{model}_C: Y_i = \beta_0 + \text{ERROR}$$

$$\text{ERROR}(C) = 50$$

Augmented model

$$\text{model}_A: Y_i = \beta_0 + \beta_1 X_i + \text{ERROR}$$

$$\text{ERROR}(A) = 30$$

Proportional reduction in error (PRE)

$$\begin{aligned} \text{PRE} &= 1 - \frac{\text{ERROR}(A)}{\text{ERROR}(C)} \\ &= 1 - \frac{30}{50} = .40 \end{aligned}$$

Increasing the complexity of the model reduced the error by 40%. **worth it?**

worth it?

Compact model

model_C: $Y_i = \beta_0 + \text{ERROR}$

Augmented model

model_A: $Y_i = \beta_0 + \beta_1 X_i + \text{ERROR}$

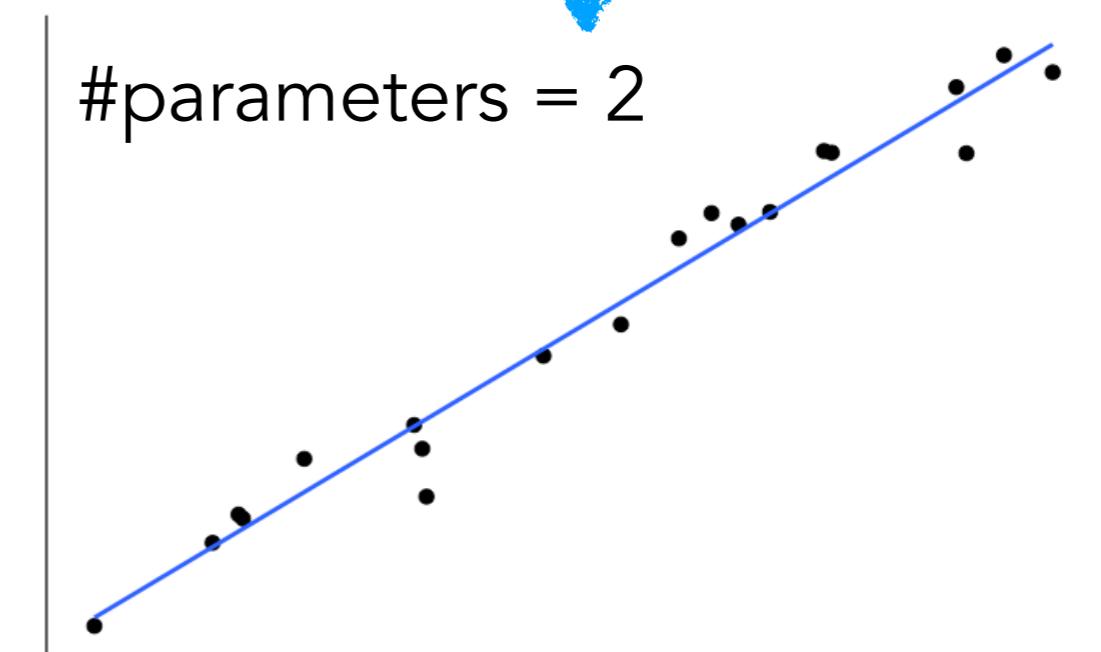
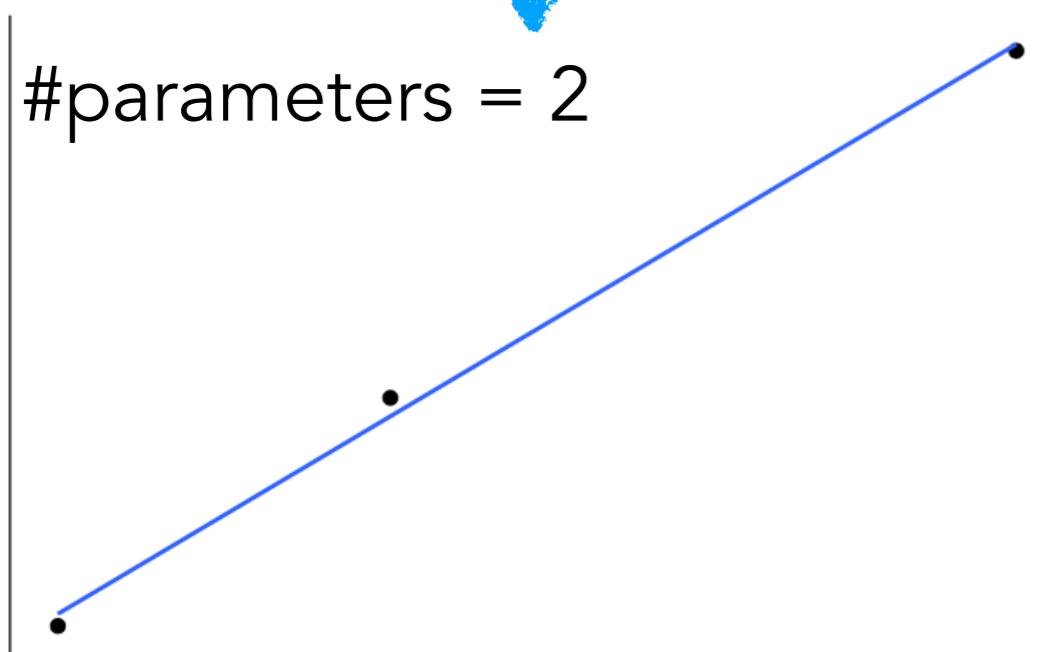
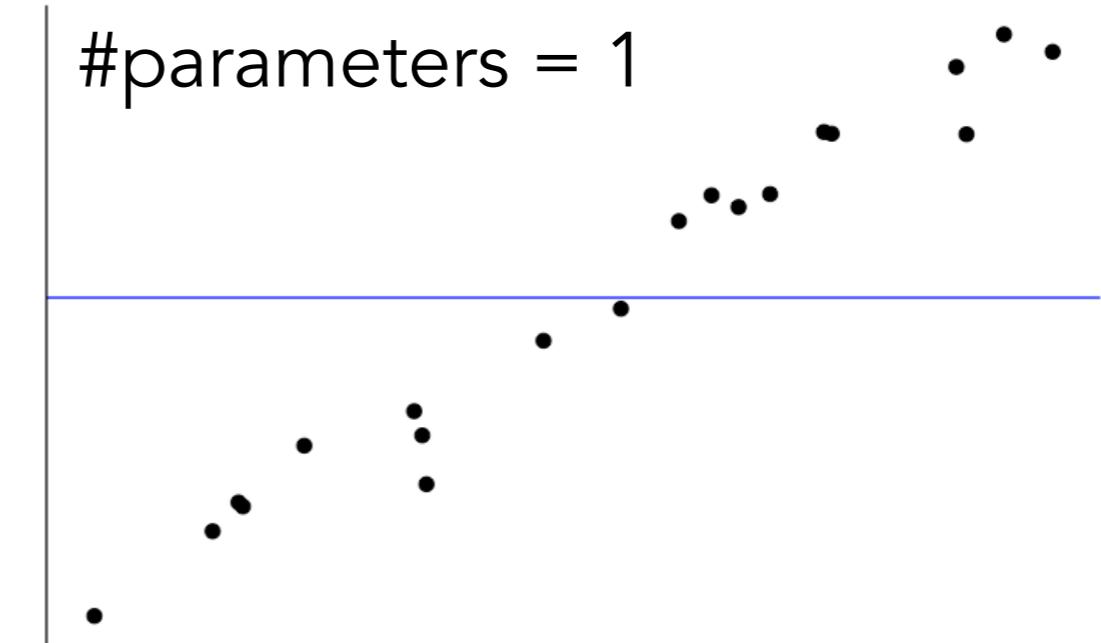
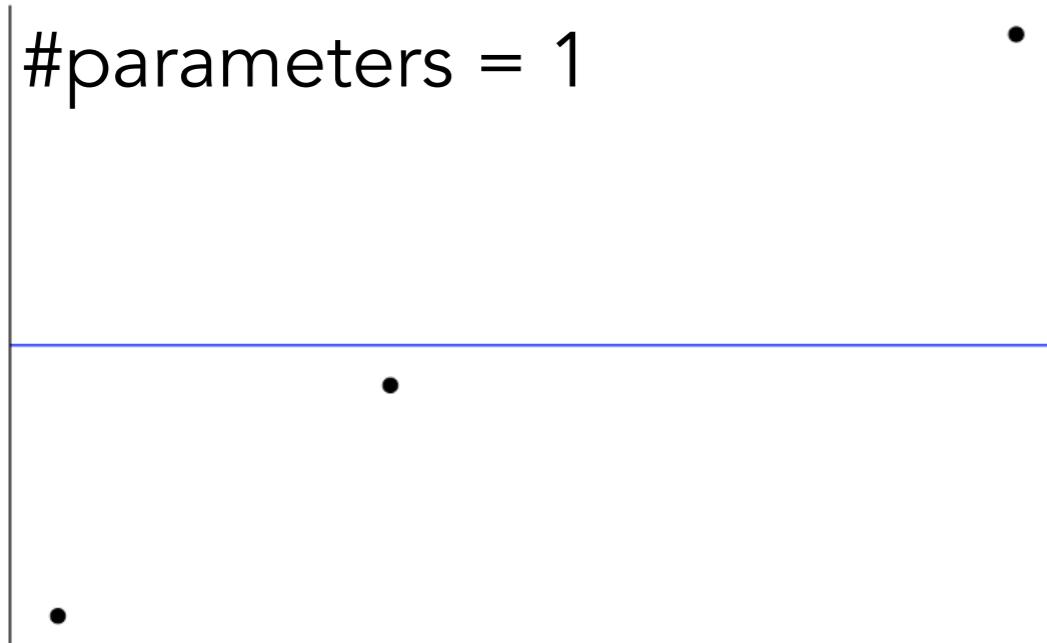
Proportional reduction in error (PRE)

$$\begin{aligned}\text{PRE} &= 1 - \frac{\text{ERROR}(A)}{\text{ERROR}(C)} \\ &= 1 - \frac{30}{50} = .40\end{aligned}$$

- to answer the **worth it?** question we need inferential statistics (define ERROR, sampling distributions, etc.)
- more likely to be **worth it** if:
 1. **PRE** is high
 2. the number of additional parameters in A compared to C is low (**PRE per additional parameter**)
 3. the number of parameters that could have been added to model_C to create model_A but were not is high

more impressed if the number of observations n is much greater than the number of parameters

PRE per parameter for different n



neato!

impressive!

General procedure

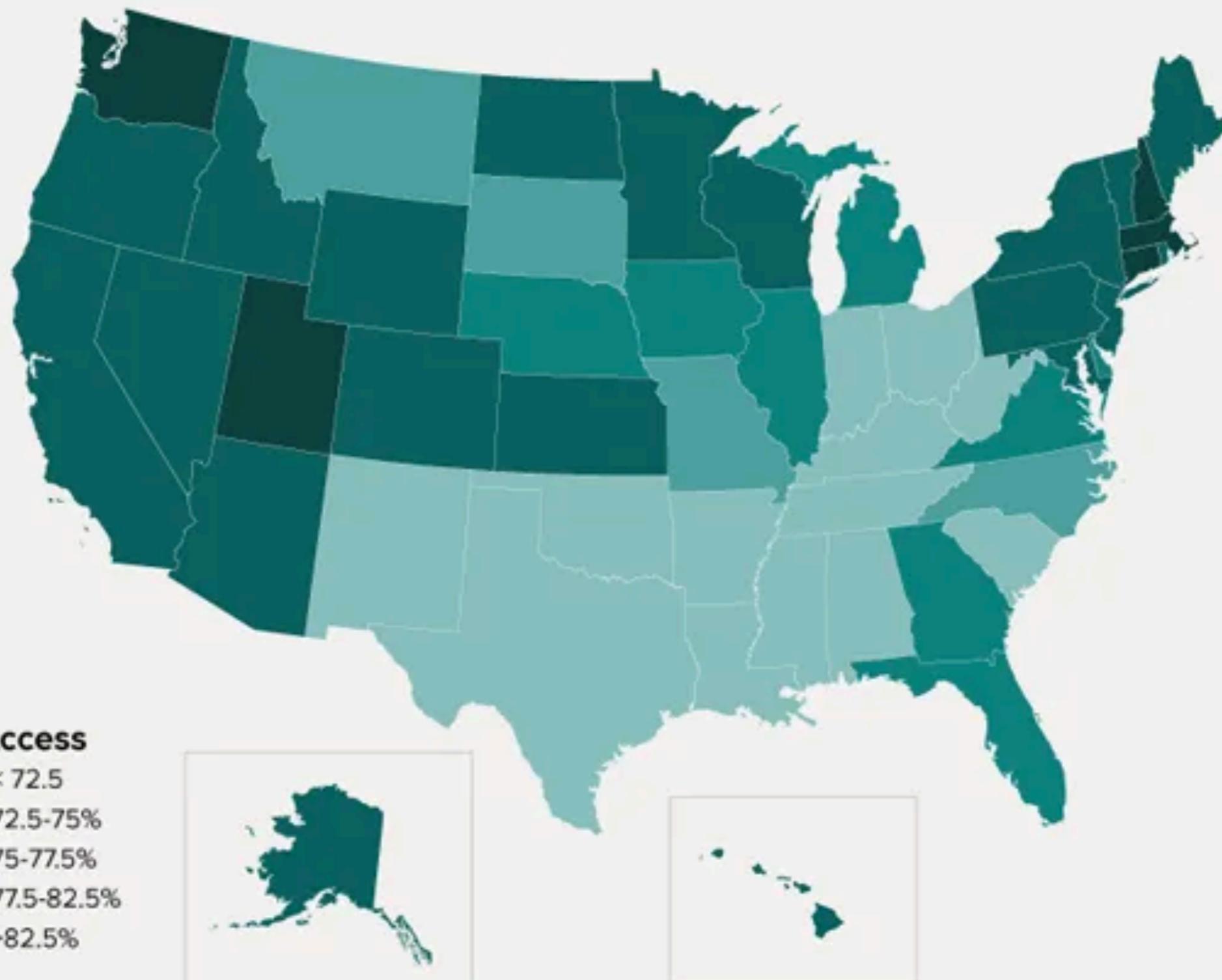
- for any question we want to ask about our DATA
 - we define model_C and model_A
 - compare the models using PRE
 - determine whether PRE is **worth it**
 - in standard frequentist lingo:
 - model_C = H_0 (null hypothesis) 
 - model_A = H_1 (alternative hypothesis) 
 - hypothesis test:
 - H_0 : **all** the parameters that are included in model_A but not in model_C are 0
 - H_1 : **not all** the parameters that are included in model_A but not in model_C are 0
- model comparison**

Statistical inferences about parameter values

Procedure

1. Start with research question
2. Formulate hypothesis as a comparison between compact and augmented model
3. Fit parameters in each model
4. Calculate the proportional reduction of error (PRE)
5. Decide whether PRE is **worth it**

Internet Access At Home



<http://thedataviz.com/2012/07/19/mapping-internet-access-in-u-s-homes/>

Notation

DATA = MODEL + ERROR

$$Y_i = \beta_0 + \epsilon_i \text{ simple model (true parameters)}$$

$$Y_i = b_0 + e_i \text{ simple model (estimated parameters)}$$

$$\hat{Y}_i = b_0$$

density

$$Y_i = b_0 + b_1 X_{i1} + e_i \text{ more complex model}$$

Percentage of households that had internet access in the year 2013 by US state

i	internet	state	college	auto	density
1	79.0	AK	28.0	1.2	1.2
2	63.5	AL	23.5	1.3	94.4
3	60.9	AR	20.6	1.7	56.0
4	73.9	AZ	27.4	1.3	56.3
5	77.9	CA	31.0	0.8	239.1
6	79.4	CO	37.8	1.0	48.5
7	77.5	CT	37.2	1.0	738.1
8	74.5	DE	29.8	1.1	460.8
9	74.3	FL	27.2	1.2	350.6
10	72.2	GA	28.3	1.1	168.4



Greek letters β or ϵ represent the true but unknowable parameters in the population.

Roman letters b or e represent estimates of these parameters using our DATA.

Research question and hypotheses

Is the average percentage of internet users per state significantly different from 75%?

Model_C: $Y_i = B_0 + \epsilon_i$

0 parameters

$$Y_i = 75 + e_i$$

Model_A: $Y_i = \beta_0 + \epsilon_i$

1 parameter

$$Y_i = b_0 + e_i$$

$$= \bar{Y} + e_i$$

i	internet	state	college	auto	density
1	79.0	AK	28.0	1.2	1.2
2	63.5	AL	23.5	1.3	94.4
3	60.9	AR	20.6	1.7	56.0
4	73.9	AZ	27.4	1.3	56.3
5	77.9	CA	31.0	0.8	239.1
6	79.4	CO	37.8	1.0	48.5
7	77.5	CT	37.2	1.0	738.1
8	74.5	DE	29.8	1.1	460.8
9	74.3	FL	27.2	1.2	350.6
10	72.2	GA	28.3	1.1	168.4

Fit parameters and calculate PRE

$$C: Y_i = 75 + e_i \quad A: Y_i = \bar{Y} + e_i$$

i	state	internet	compact_b	compact_se	augmented_b	augmented_se
1	AK	79.0	75	16.00	72.81	38.37
2	AL	63.5	75	132.25	72.81	86.60
3	AR	60.9	75	198.81	72.81	141.75
4	AZ	73.9	75	1.21	72.81	1.20
5	CA	77.9	75	8.41	72.81	25.95
6	CO	79.4	75	19.36	72.81	43.48
7	CT	77.5	75	6.25	72.81	22.03
8	DE	74.5	75	0.25	72.81	2.87
9	FL	74.3	75	0.49	72.81	2.23
10	GA	72.2	75	7.84	72.81	0.37

$$SSE(C) = 1595 \quad SSE(A) = 1355$$

$$\begin{aligned} PRE &= 1 - \frac{\text{ERROR}(A)}{\text{ERROR}(C)} \\ &= 1 - \frac{1355}{1595} \approx .15 \end{aligned}$$

Model A has
15% less error
than Model C.

Decide whether it's **worth it**

- PRE is the estimate of an unknown true reduction of error η^2
- we need a sampling distribution of PRE
 - a distribution of what PRE would look like if Model C (our H_0) were true
 - we could just simulate such a sampling distribution ...
- PRE is closely related to the F statistic!

Decide whether it's **worth it**

- To compute the F statistic, we need:

- PRE
- number of parameters in Model C (PC) and Model A (PA)
- number of observations n

- more likely to be **worth it** if:
 1. PRE is high
 2. the number of additional parameters in A compared to C is low (**PRE per additional parameter**)
 3. the number of parameters that could have been added to model_C to create model_A but were not

**difference in parameters
between models A and C**

$$F = \frac{\text{PRE}/(\text{PA} - \text{PC})}{(1 - \text{PRE})/(n - \text{PA})}$$

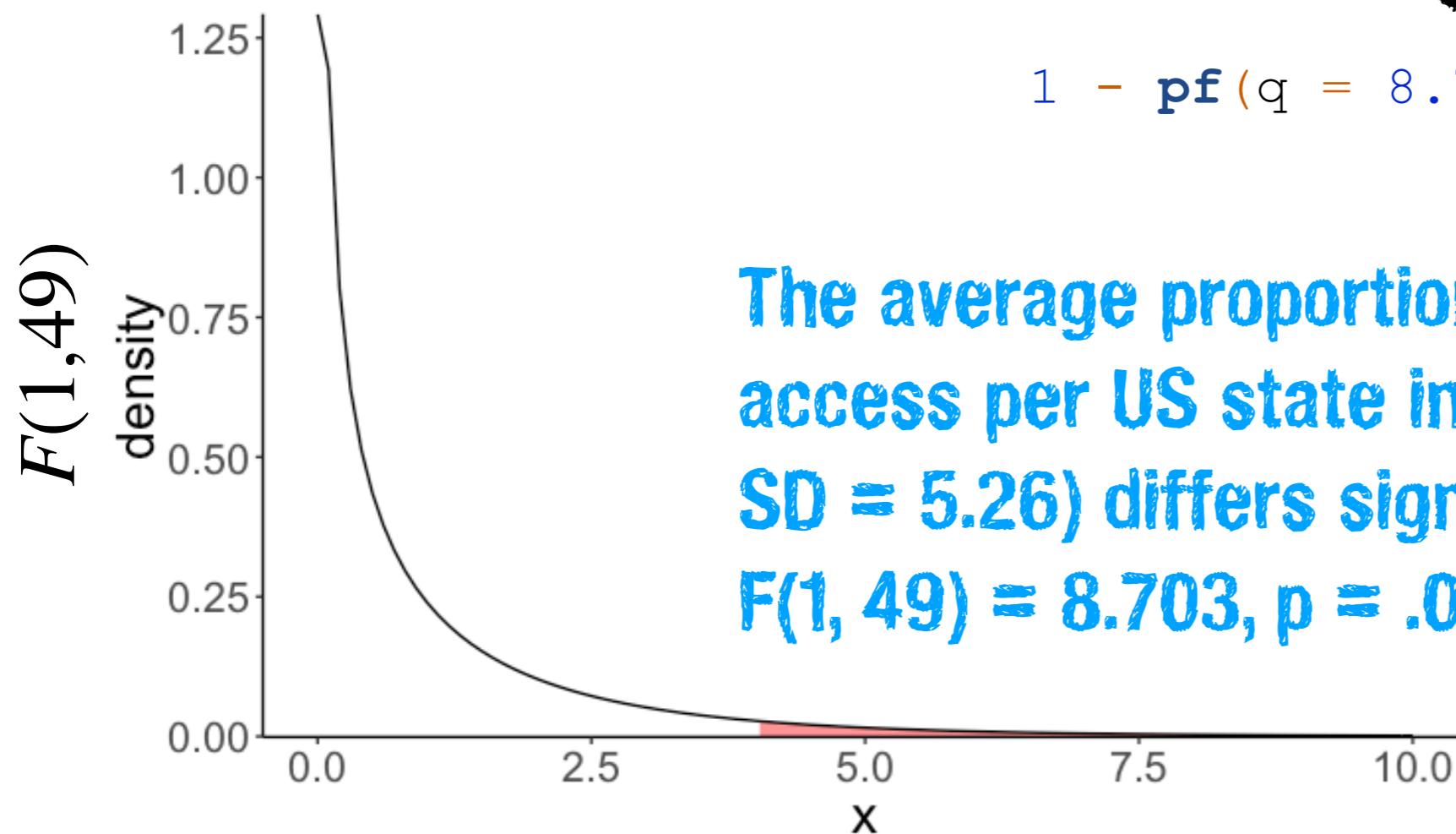


**number of observations
vs. parameters in Model A**

Decide whether it's **worth it**

$$F = \frac{\text{PRE}/(\text{PA} - \text{PC})}{(1 - \text{PRE})/(n - \text{PA})}$$

$$= \frac{.15/(1 - 0)}{(1 - .15)/(50 - 1)} = 8.703 \quad p = .00486$$



Note: I've used the approximated PRE here (PRE = 0.15), but I'm reporting the F value for the non-approximated PRE value.

we just performed a one sample t-test ...

```
t.test(df.internet$internet, mu = 75)
```

One Sample t-test

```
data: df.internet$internet  
t = -2.9502, df = 49, p-value = 0.00486  
alternative hypothesis: true mean is not equal to 75
```

but ...

- we will apply the general procedure we went through to day to a large range of different situations
- thinking of hypothesis testing as model comparison is (hopefully more) intuitive
- this approach still works even if we can't find a recipe in our favorite stats cook book

Summary

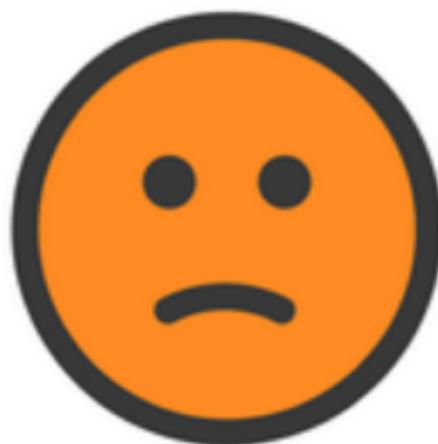
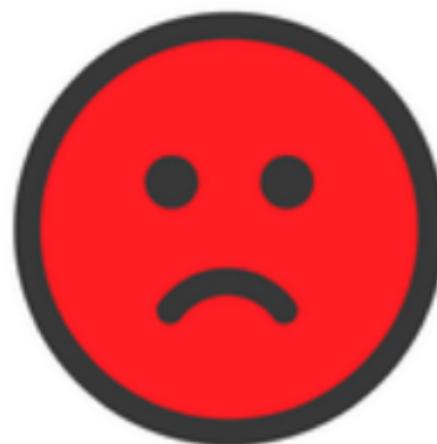
- Quick recap
- Statistical concepts
 - Confidence intervals
 - Bootstrapping
- Cookbook vs. Model Comparison
- Modeling data
- Statistical inferences about parameter values

Feedback

How was the pace of today's class?

much a little just a little much
too too right too too
slow slow fast fast

How happy were you with today's class overall?



What did you like about today's class? What could be improved next time?

Thank you!