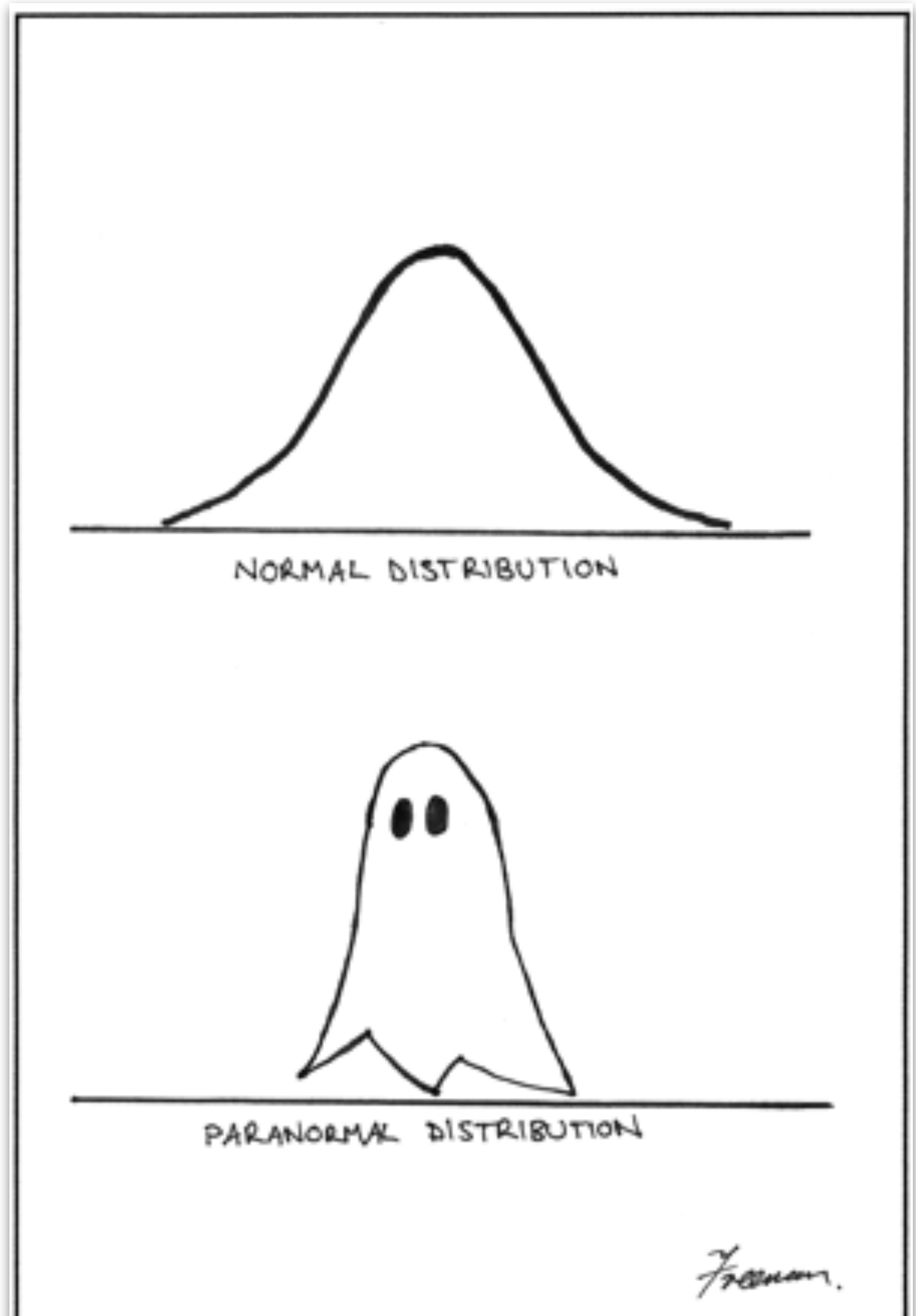
Simulation 1



01/22/2020

Logistics

Final presentation day

Wednesday, March 18th, 3:30pm - 6:30pm

Application section

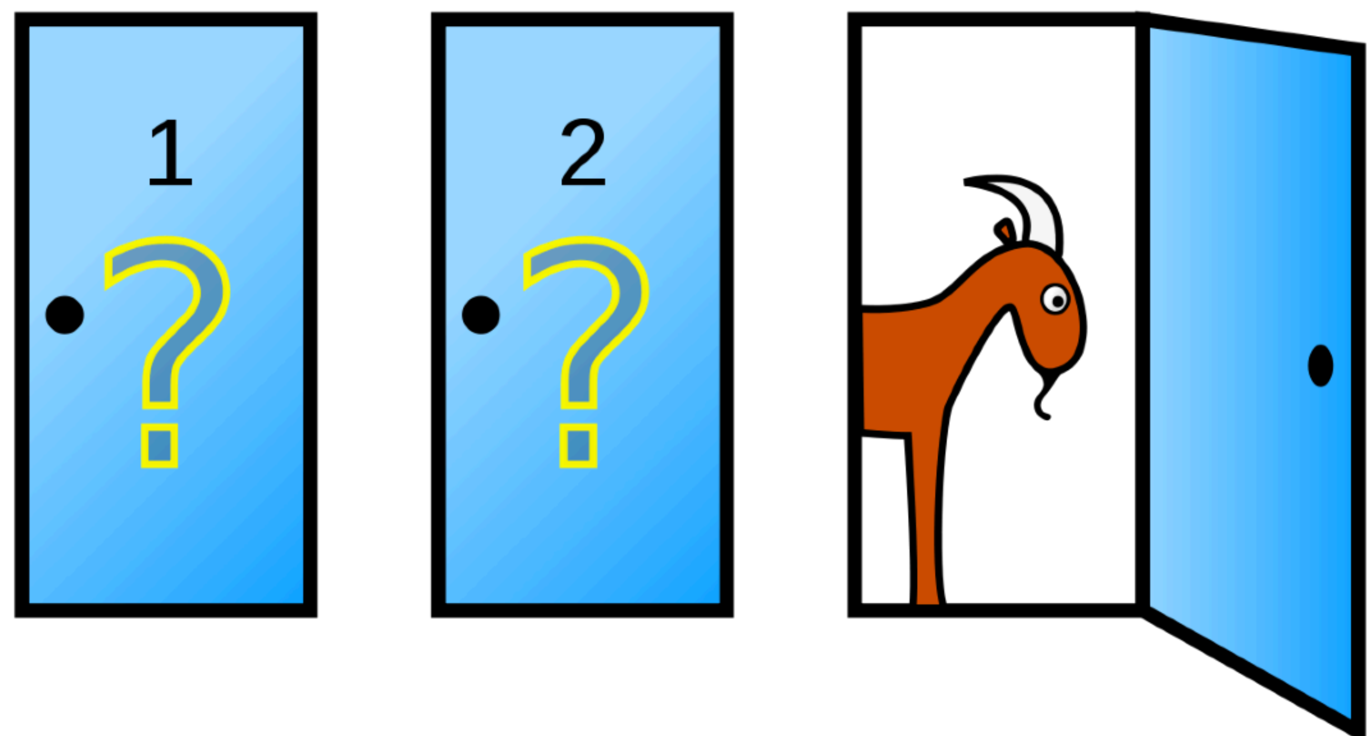
Andrew Nam

01/23/2020

- 1 The Monty Hall Problem
  - 1.1 Estimating the probability
  - 1.2 Analytical solution
- 2 The Birthday Paradox
  - 2.1 Estimating the probability
  - 2.2 Analytic solution
  - 2.3 Hash Functions

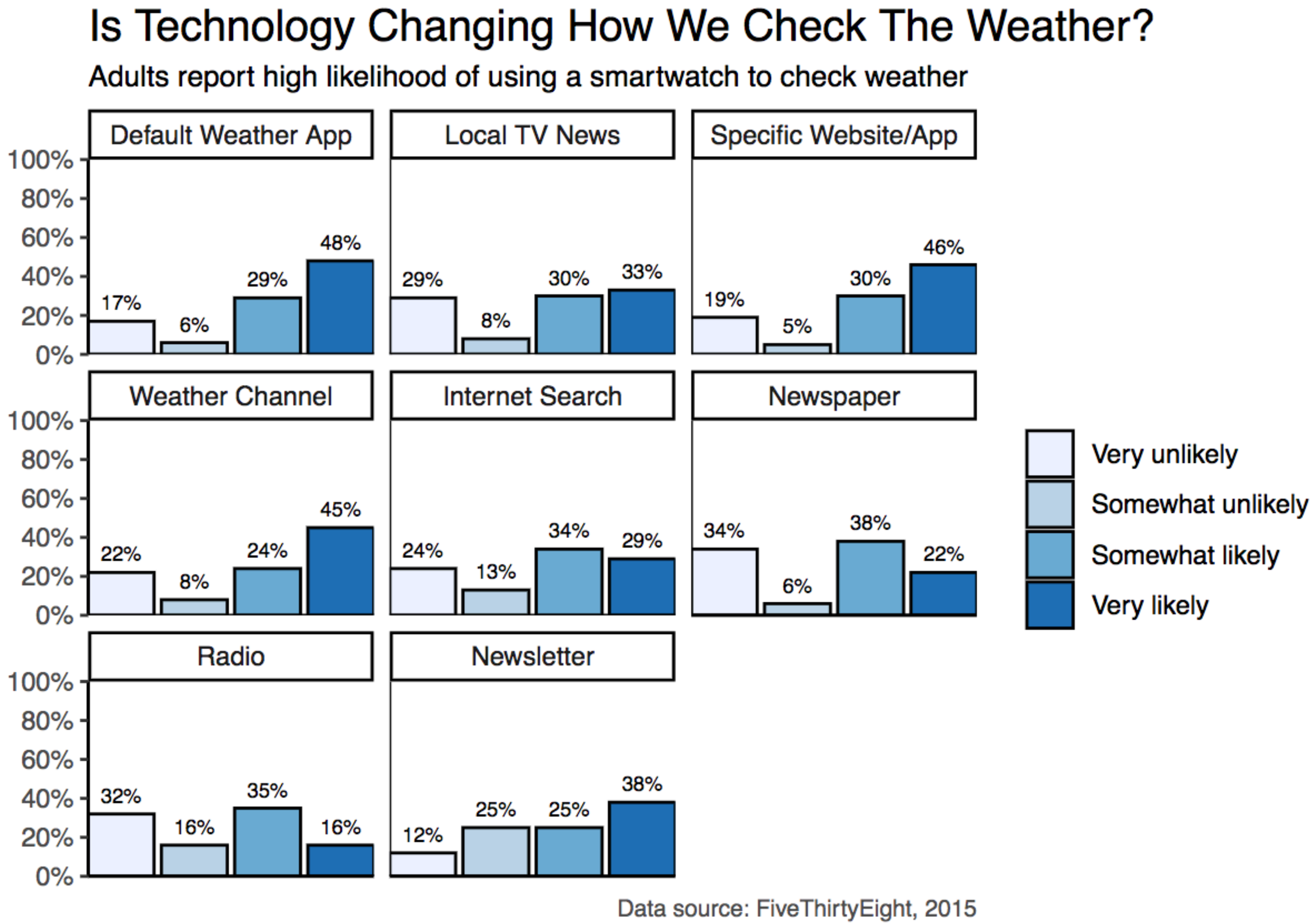
The Monty Hall Problem

Suppose you're on a game show, and you're given the choice of three doors: Behind one door is a car; behind the others, goats. You pick a door, say No. 1, and the host, who knows what's behind the doors, opens another door, say No. 3, which has a goat. He then says to you, "Do you want to pick door No. 2?" Is it to your advantage to switch your choice?

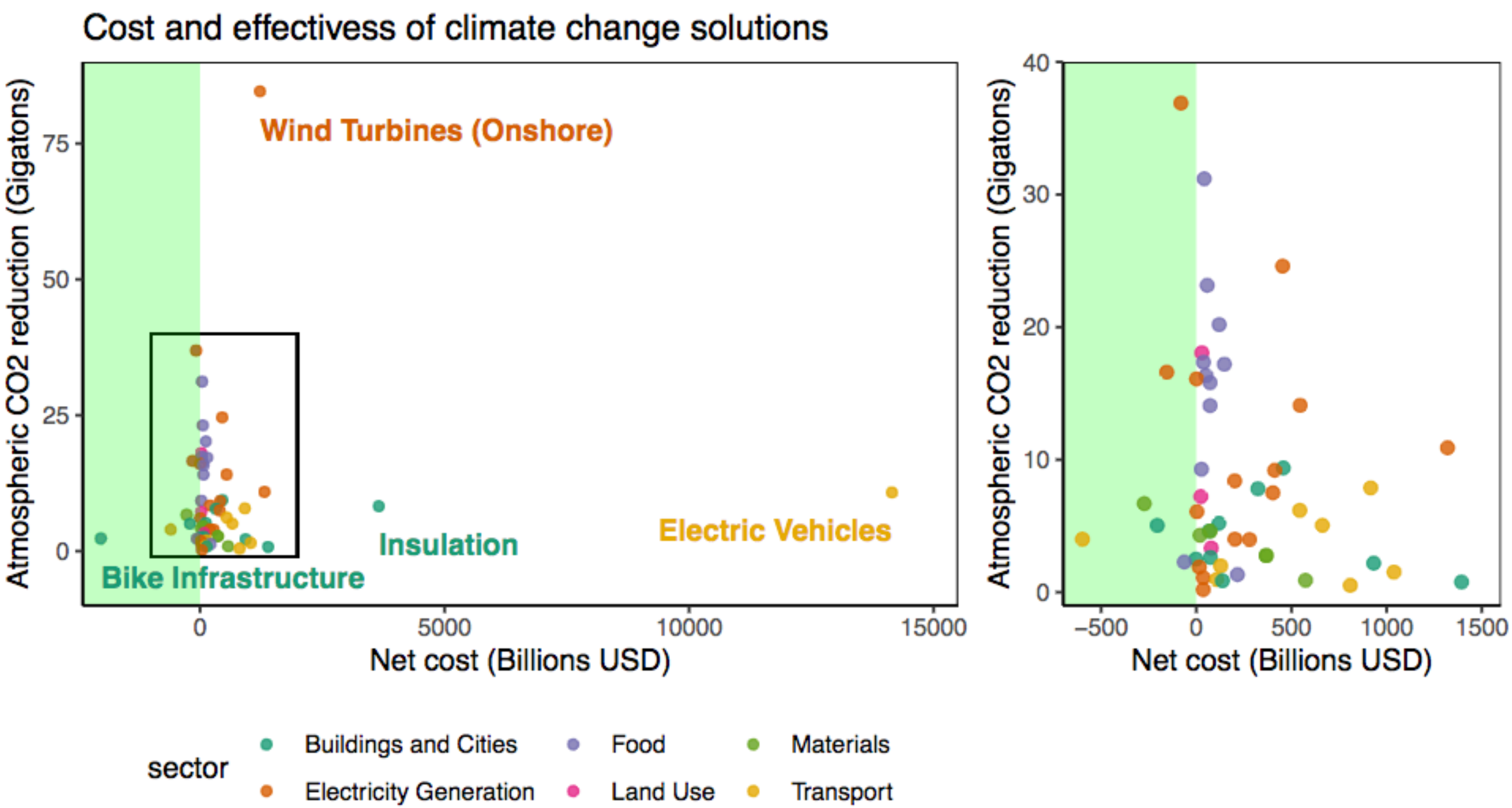


Homework

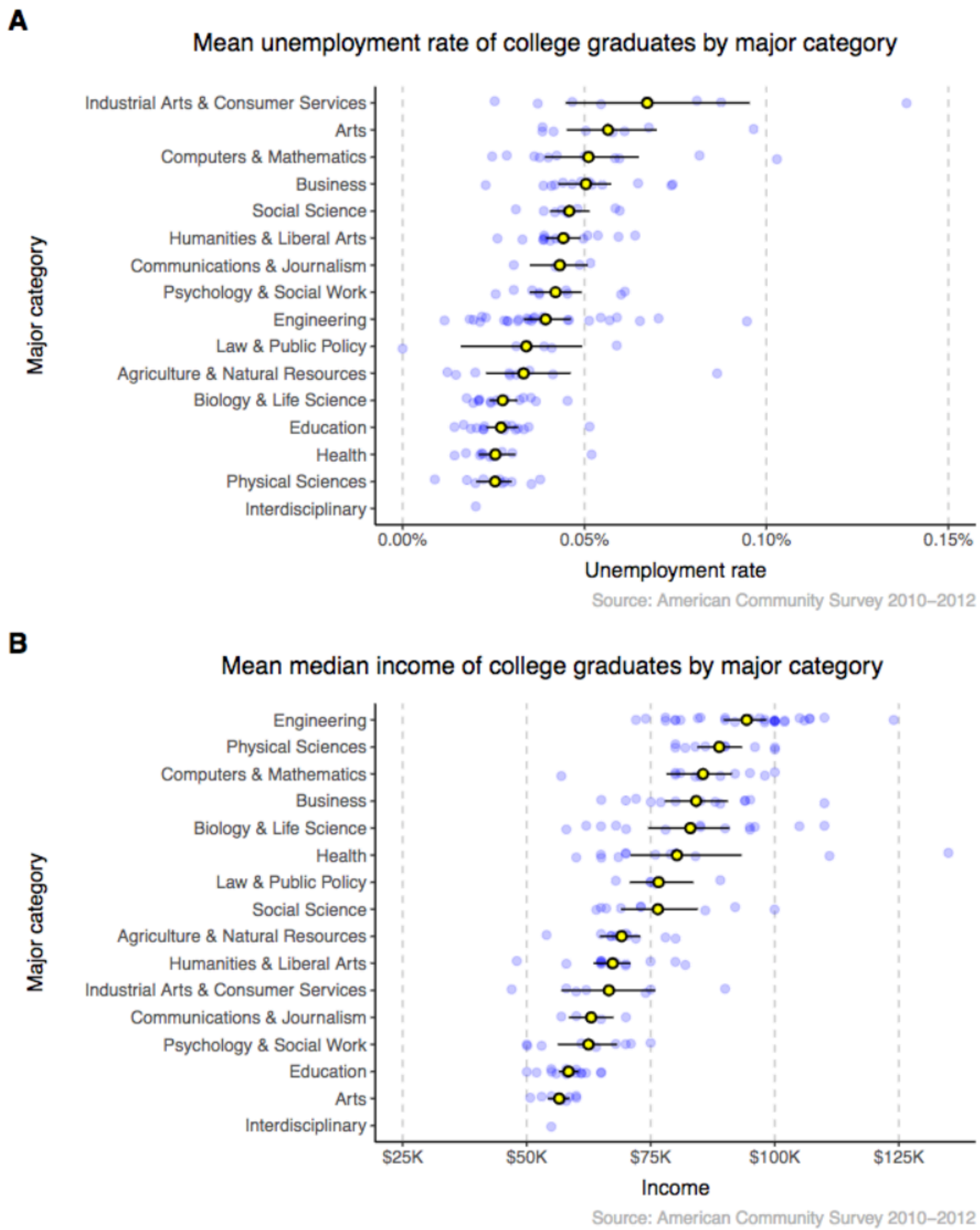

How many hours did you spend on homework 1?
number of students
0
2
4
6
8
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
or more

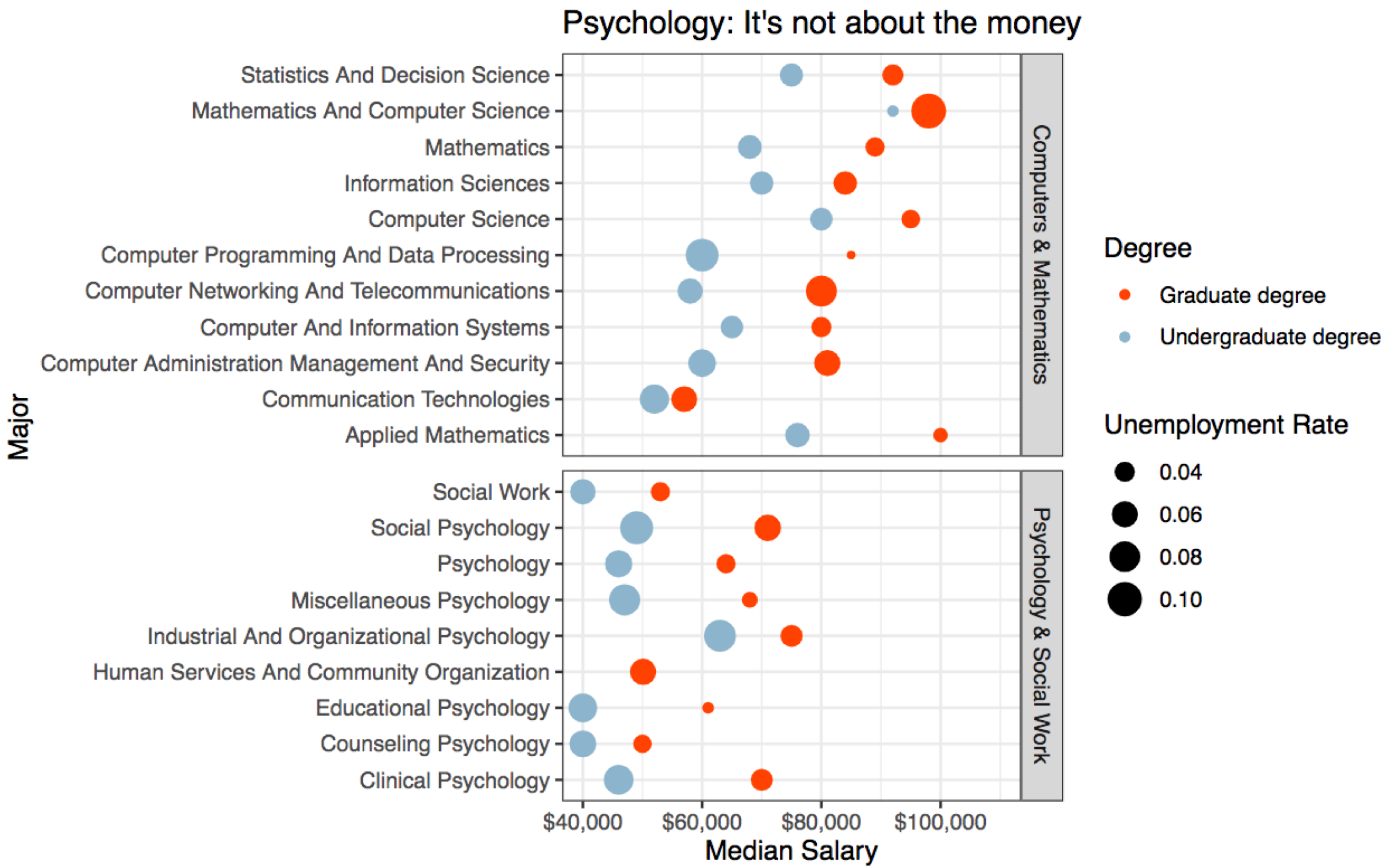


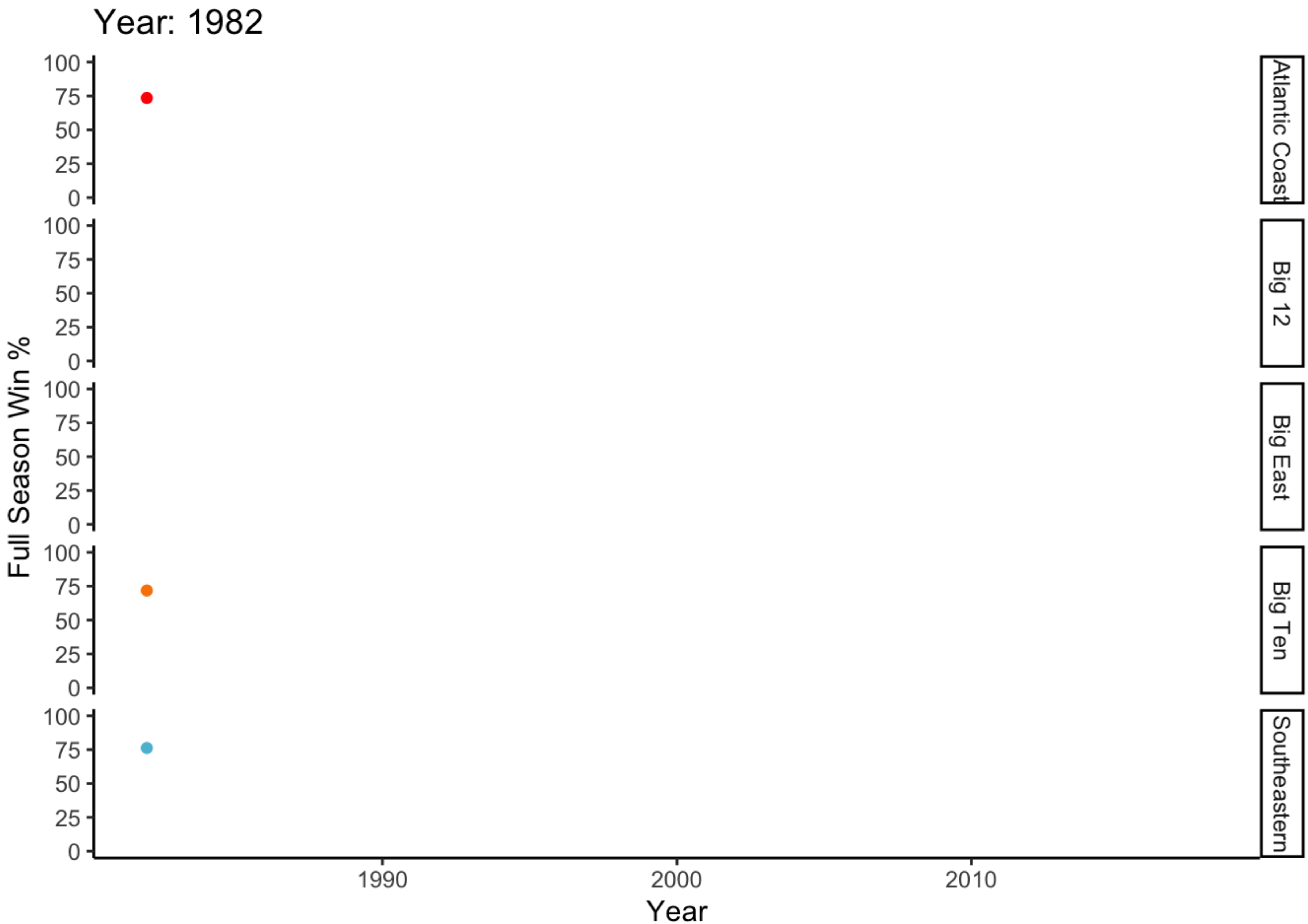
Is Technology Changing How We Check The Weather?
Adults report high likelihood of using a smartwatch to check weather
Default Weather App
17%
6%
29%
48%
Local TV News
29%
8%
30%
33%
Specific Website/App
19%
5%
30%
46%
Weather Channel
22%
8%
24%
45%
Internet Search
24%
13%
34%
29%
Newspaper
34%
6%
38%
22%
Radio
32%
16%
35%
16%
Newsletter
12%
25%
25%
38%
100%
80%
60%
40%
20%
0%
Very unlikely
Somewhat unlikely
Somewhat likely
Very likely
Data source: FiveThirtyEight, 2015



Cost and effectivess of climate change solutions
Wind Turbines (Onshore)
Insulation
Electric Vehicles
Bike Infrastructure
Atmospheric CO2 reduction (Gigatons)
Net cost (Billions USD)
sector
Buildings and Cities
Electricity Generation
Food
Land Use
Materials
Transport


A
Mean unemployment rate of college graduates by major category
Industrial Arts & Consumer Services
Arts
Computers & Mathematics
Business
Social Science
Humanities & Liberal Arts
Communications & Journalism
Psychology & Social Work
Engineering
Law & Public Policy
Agriculture & Natural Resources
Biology & Life Science
Education
Health
Physical Sciences
Interdisciplinary
Major category
0.00%
0.05%
0.10%
0.15%
Unemployment rate
Source: American Community Survey 2010–2012
B
Mean median income of college graduates by major category
Engineering
Physical Sciences
Computers & Mathematics
Business
Biology & Life Science
Health
Law & Public Policy
Social Science
Agriculture & Natural Resources
Humanities & Liberal Arts
Industrial Arts & Consumer Services
Communications & Journalism
Psychology & Social Work
Education
Arts
Interdisciplinary
Major category
$25K
$50K
$75K
$100K
$125K
Income
Source: American Community Survey 2010–2012

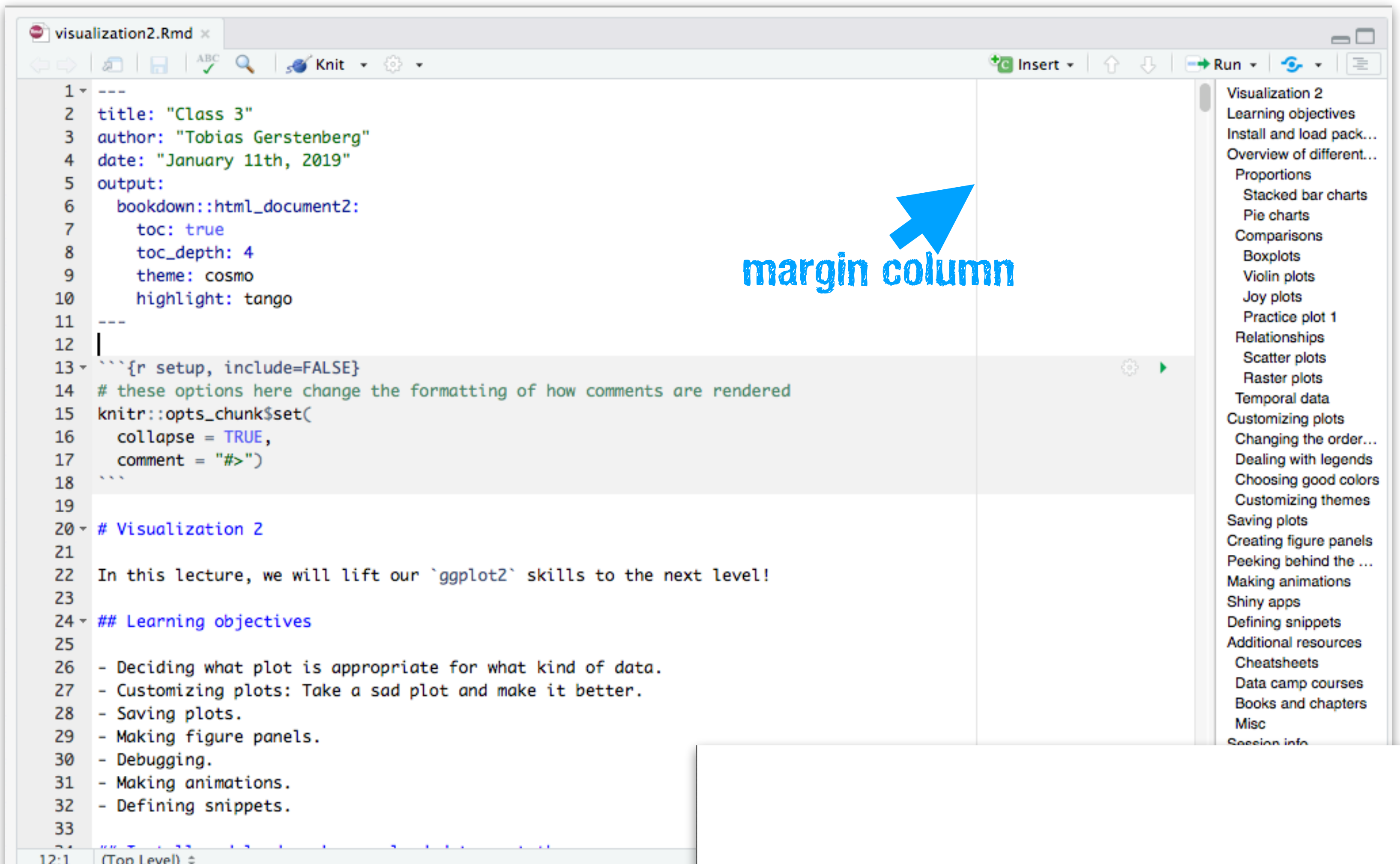

Psychology: It's not about the money
Statistics And Decision Science
Mathematics And Computer Science
Mathematics
Information Sciences
Computer Science
Computer Programming And Data Processing
Computer Networking And Telecommunications
Computer And Information Systems
Computer Administration Management And Security
Communication Technologies
Applied Mathematics
Social Work
Social Psychology
Psychology
Miscellaneous Psychology
Industrial And Organizational Psychology
Human Services And Community Organization
Educational Psychology
Counseling Psychology
Clinical Psychology
Major
$40,000
$60,000
$80,000
$100,000
Median Salary
Computers & Mathematics
Psychology & Social Work
Degree
Graduate degree
Undergraduate degree
Unemployment Rate
0.04
0.06
0.08
0.10

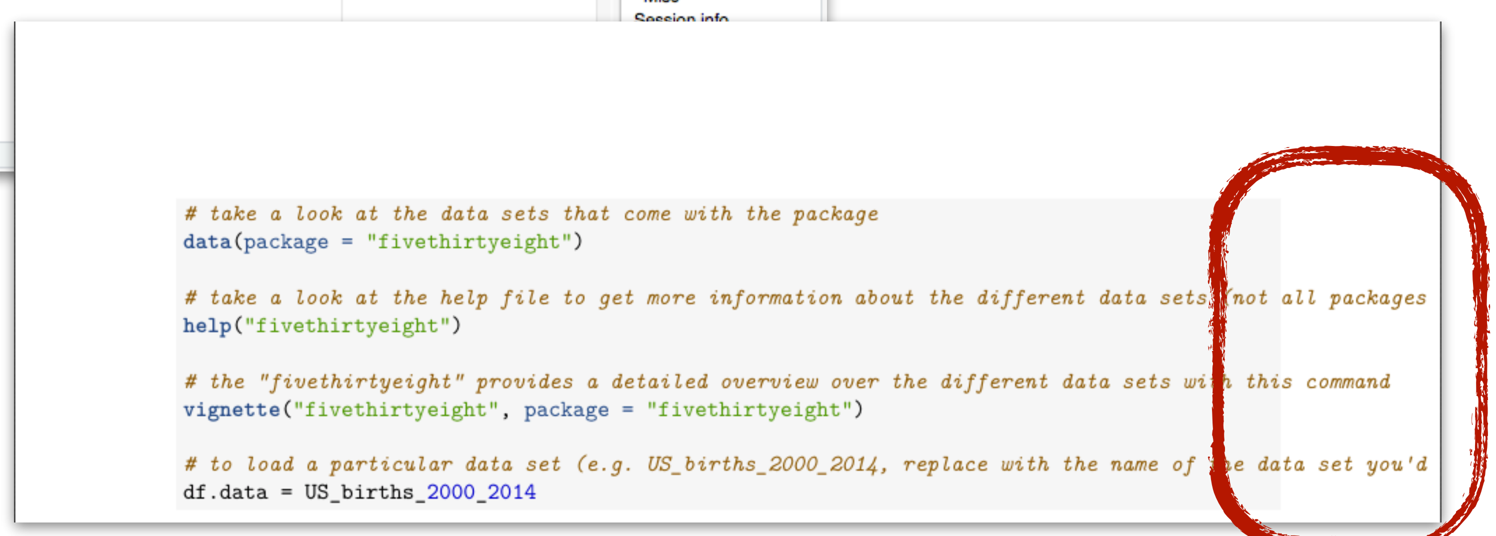


Year: 1982
Full Season Win %
100
75
50
25
0
Atlantic Coast
Big 12
Big East
Big Ten
Southeastern
1990
2000
2010
Year

some things to look out for

make sure to make your code visible!
(we can only grade what we see)





not good

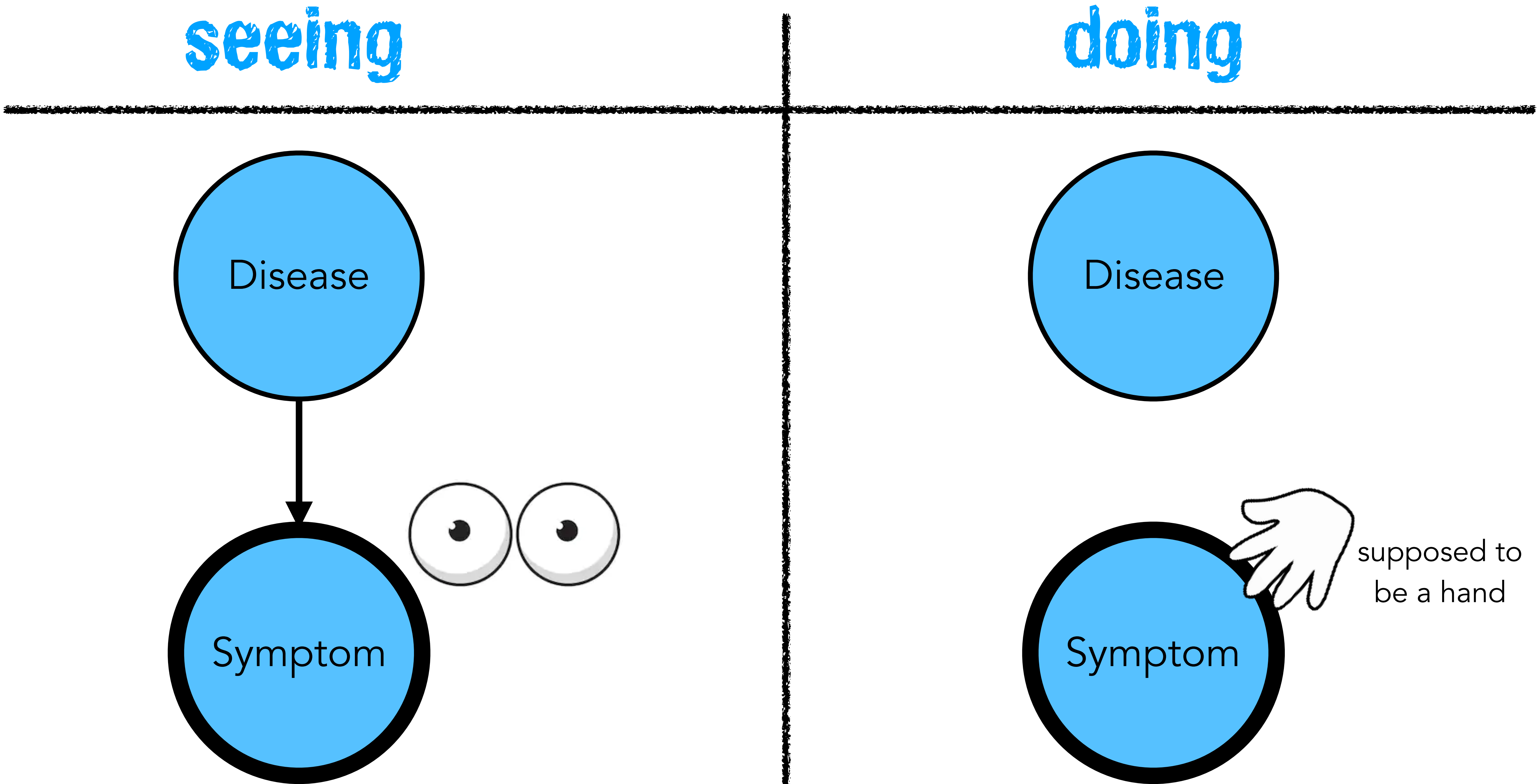
Plan for today

- Quick review of causality
- Working with probability distributions
  - `dnorm()`, `pnorm()`, `qnorm()`, `rnorm()`
  - computing probabilities
- Bayesian inference
  - analytic solution
  - via sampling
- Working with samples
  - Understanding `density()`
  - Understanding `quantile()`
  - Comparing distributions

Plan for today

- **Quick review of causality**
- Working with probability distributions
  - `dnorm()`, `pnorm()`, `qnorm()`, `rnorm()`
  - computing probabilities
- Bayesian inference
  - analytic solution
  - via sampling
- Working with samples
  - Understanding `density()`
  - Understanding `quantile()`
  - Comparing distributions

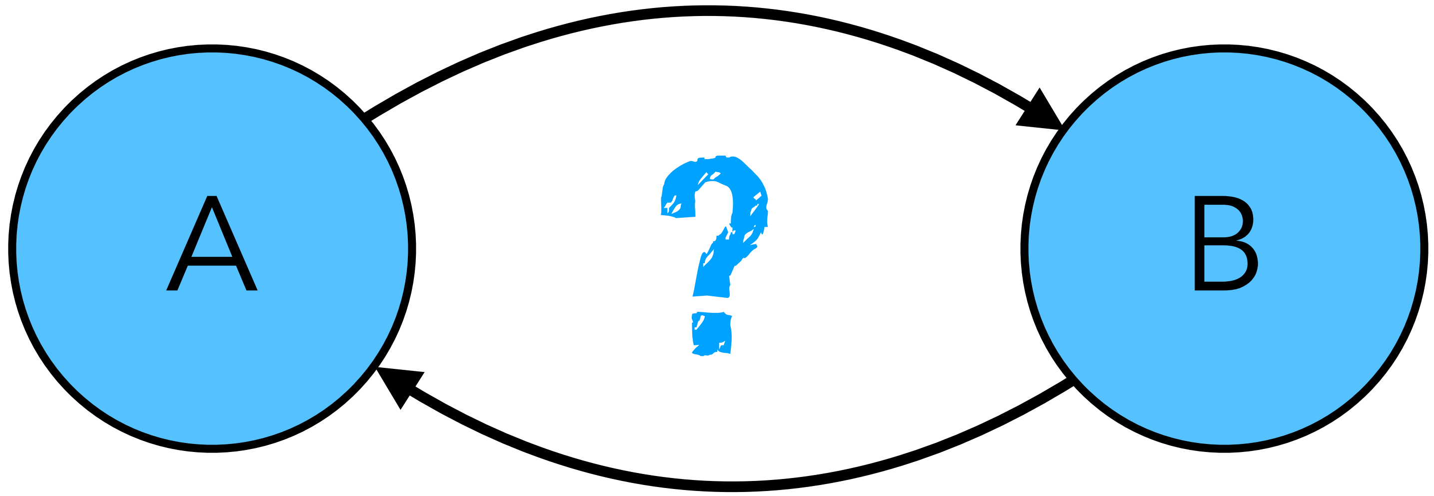
Observation vs. Intervention

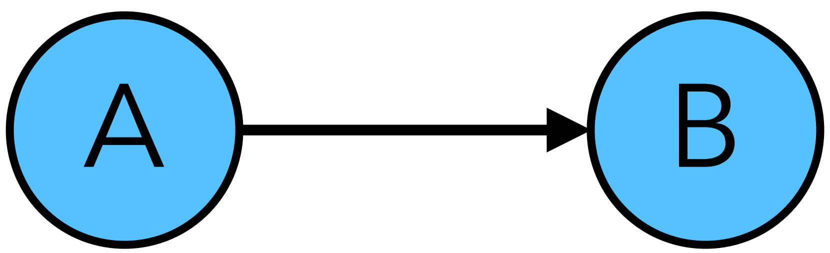


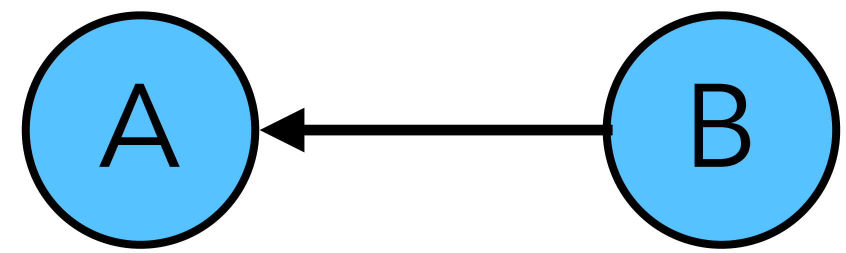
$$p(D \mid S) > p(D)$$

$$p(D \mid \mathrm{do}(S)) = p(D)$$

Inferring causal structure through intervention

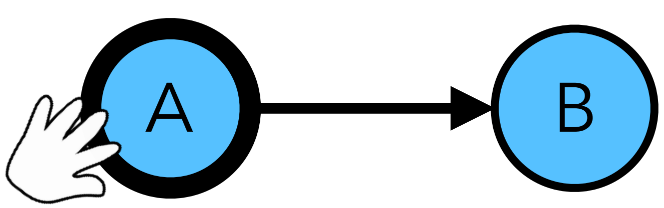






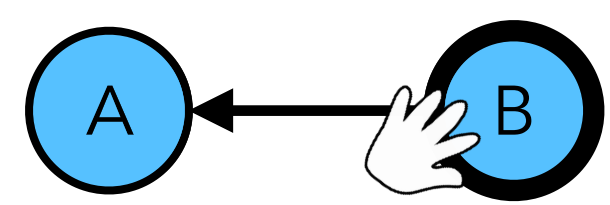
$$p(B \mid \mathrm{do}(A)) = p(B \mid A)$$

$$p(B \mid \mathrm{do}(A)) = p(B)$$



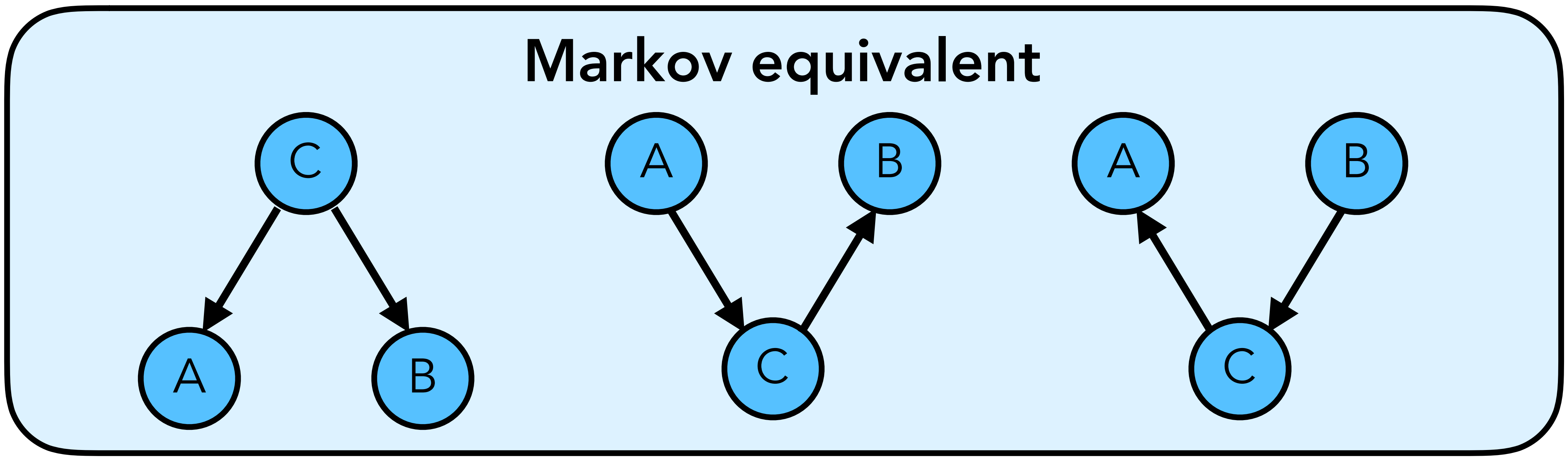
$$p(A \mid \mathrm{do}(B)) = p(A)$$

$$p(A \mid \mathrm{do}(B)) = p(A \mid B)$$



Important take home message

- correlation is not causation
- correlation (= probabilistic dependence) suggests that there is *some* causal relationship
- but we don't know which one it is

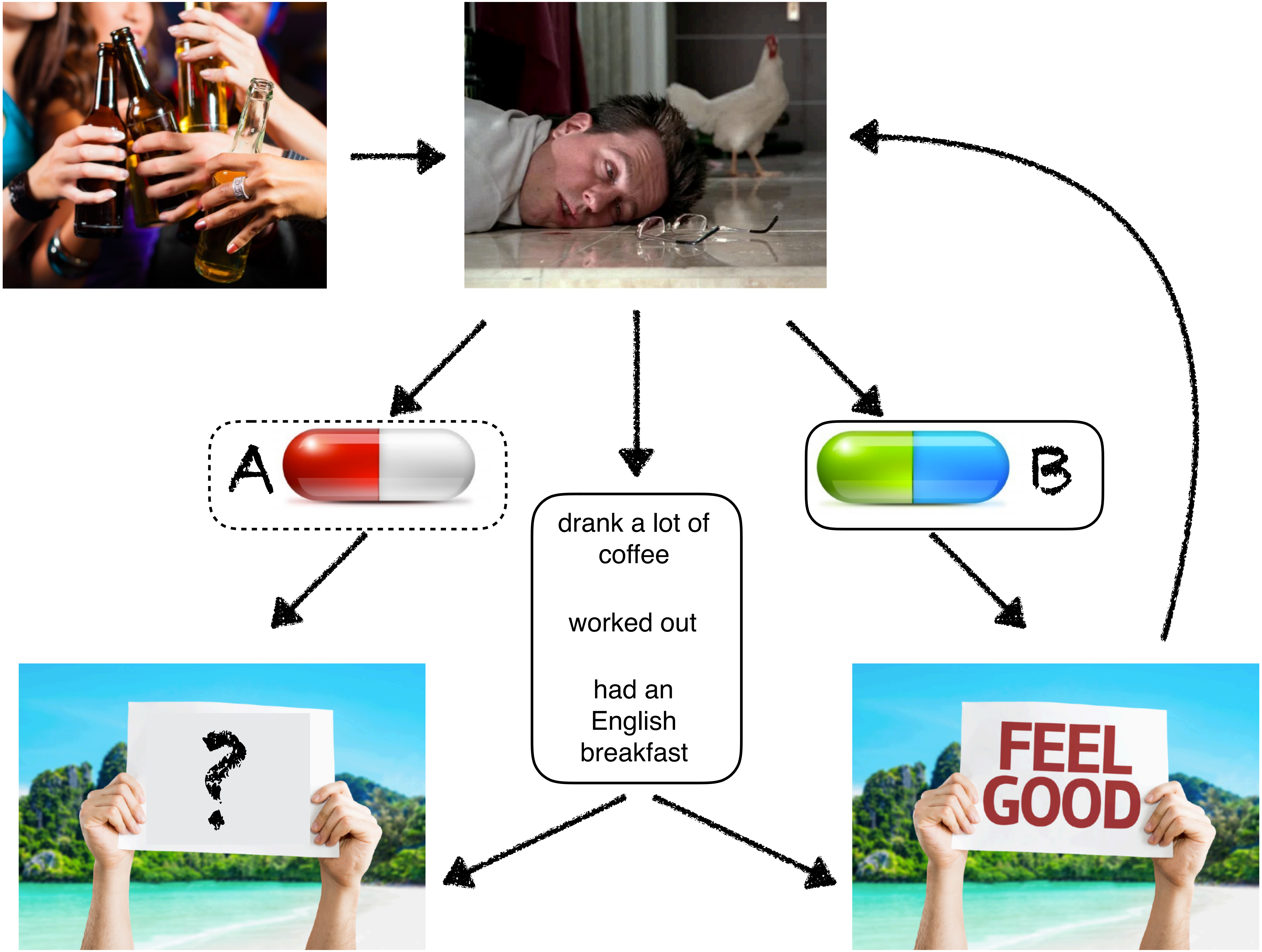


- **causal interventions** / experiments can reveal the underlying causal structure

The three layer causal hierarchy

Level (Symbol)	Typical Activity	Typical Questions	Examples
1. Association $P(y\|x)$	Seeing	What is? How would seeing X change my belief in$Y$?	What does a symptom tell me about a disease? What does a survey tell us about the election results?
2. Intervention $P(y\|do(x), z)$	Doing Intervening	What if? What if I do X?	What if I take aspirin, will my headache be cured? What if we ban cigarettes?
3. Counterfactuals $P(y_x\|x', y')$	Imagining, Retrospection	Why? Was it X that caused Y? What if I had acted differently?	Was it the aspirin that stopped my headache? Would Kennedy be alive had Oswald not shot him? What if I had not been smoking the past 2 years?

Pearl (2018) Theoretical Impediments to Machine Learning With Seven Sparks from the Causal Revolution. *Proceedings of the Eleventh International Conference on Web Search and Data Mining*


A
drank a lot of coffee
worked out
had an English breakfast
B
?
FEEL GOOD

randomized control trial

had a long afternoon nap

drank a lot of coffee

worked out

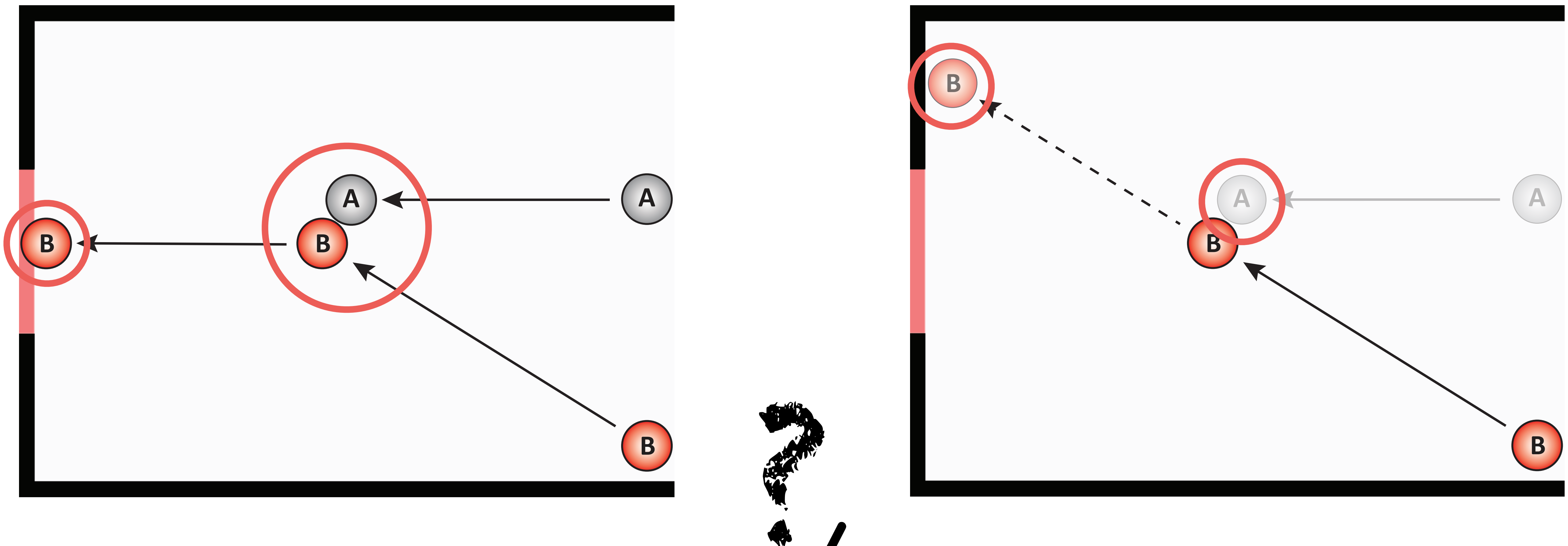
drank a lot of tea

had an English breakfast

Did A cause B to go through the gate?

Counterfactual Simulation Model

What happened? | What would have happened?



Actual situation | **Counterfactual situation**

B went through the gate $\neq$ B would have missed the gate

Gerstenberg, Goodman, Lagnado, & Tenenbaum (2012) Noisy Newtons: Unifying process and dependency accounts of causal attribution. Cognitive Science Proceedings

Gerstenberg, Goodman, Lagnado, & Tenenbaum (2014) From counterfactual simulation to causal judgment. Cognitive Science Proceedings

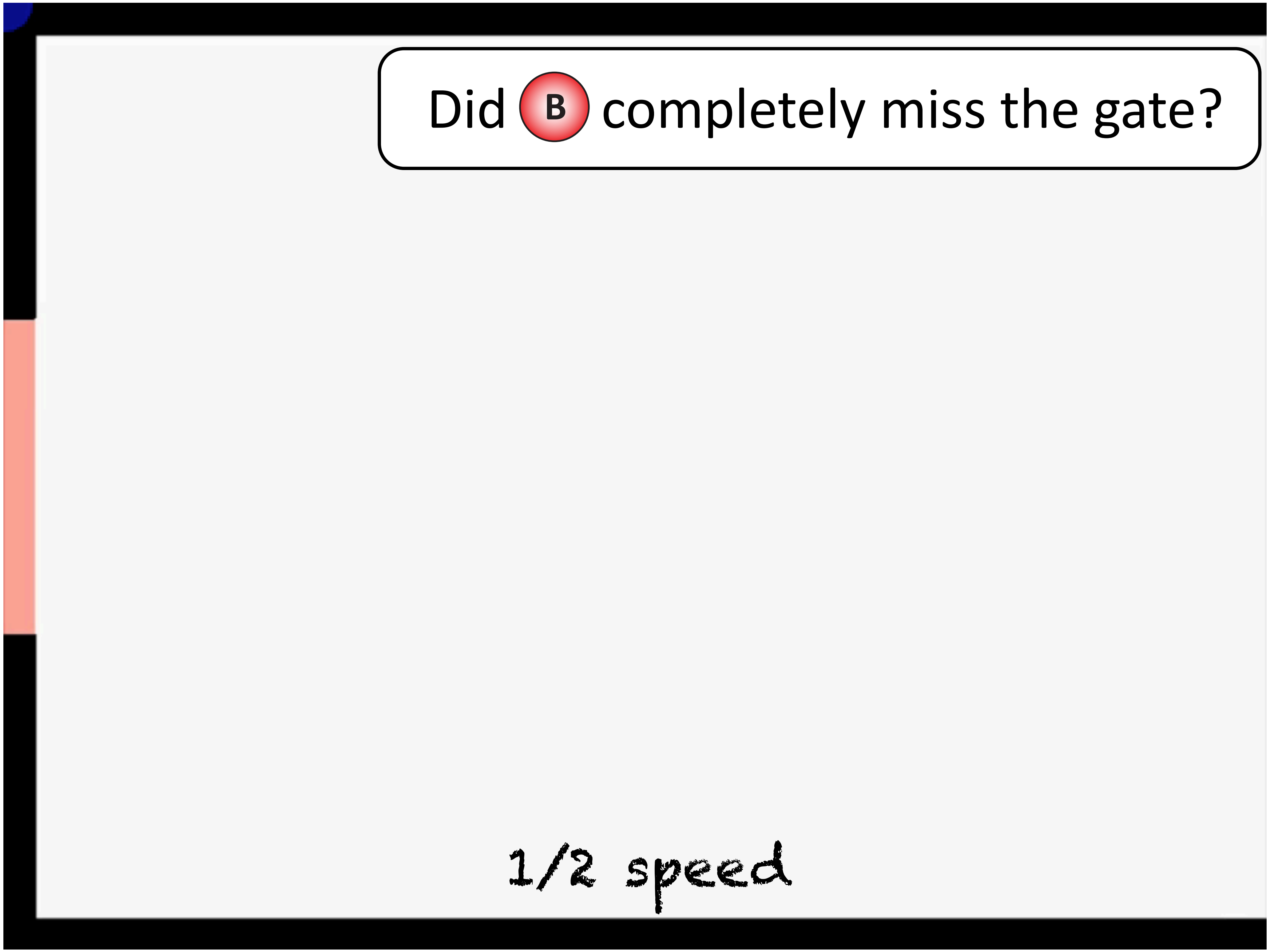
Gerstenberg, Goodman, Lagnado, & Tenenbaum (2015) How, whether, why: Causal judgments as counterfactual contrasts. Cognitive Science Proceedings

Gerstenberg & Tenenbaum (2016) Understanding ``almost'': Empirical and computational studies of near misses. Cognitive Science Proceedings

Gerstenberg & Tenenbaum (2017) Intuitive Theories. Oxford Handbook of Causal Reasoning

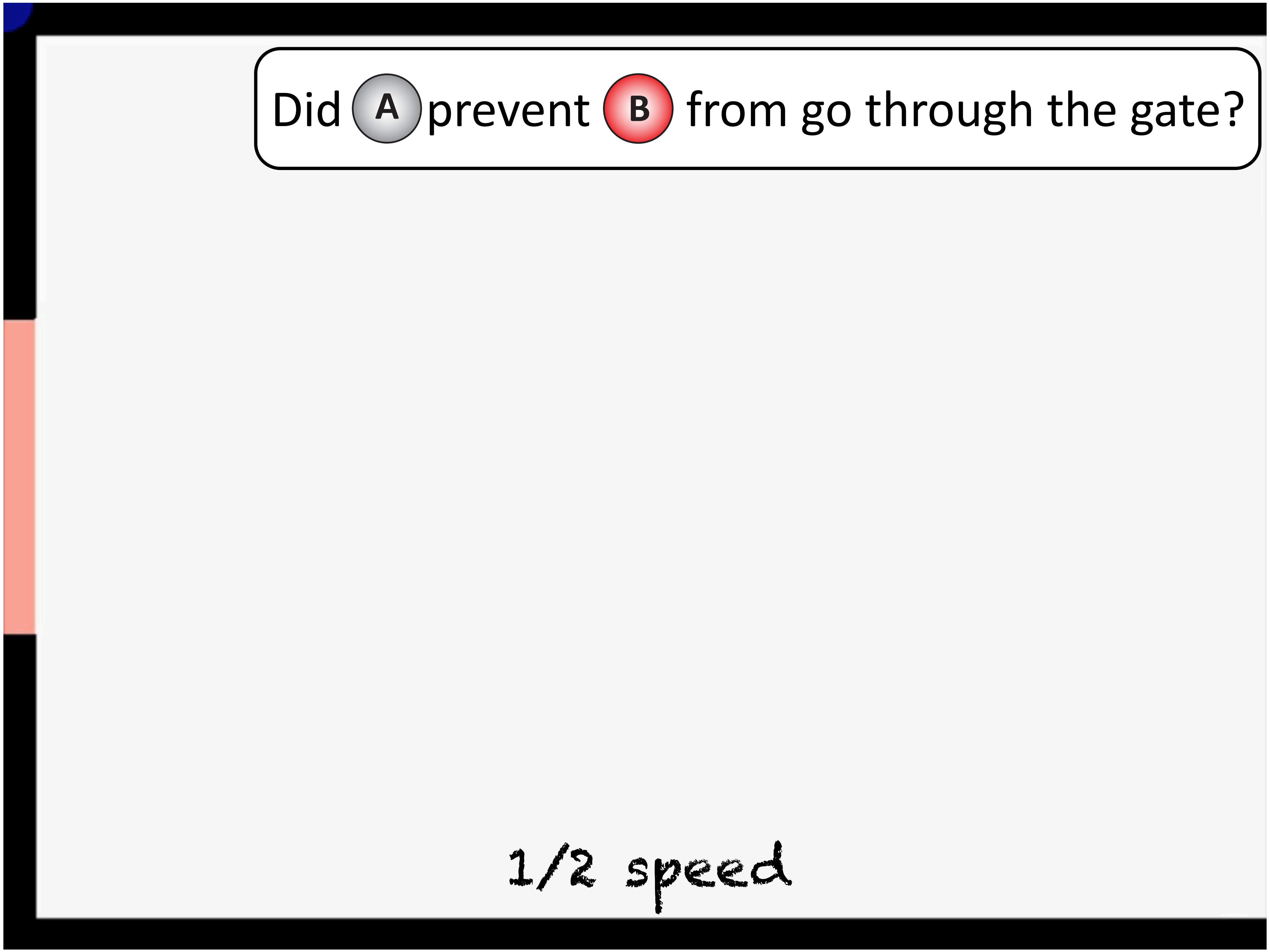
Gerstenberg, Goodman, Lagnado, & Tenenbaum (in preparation) A counterfactual simulation model of causal judgment.

Spontaneous counterfactual simulation



Gerstenberg, Peterson, Goodman, Lagnado, & Tenenbaum (2017) Eye-tracking causality. *Psychological Science*

Spontaneous counterfactual simulation



Gerstenberg, Peterson, Goodman, Lagnado, & Tenenbaum (2017) Eye-tracking causality. *Psychological Science*

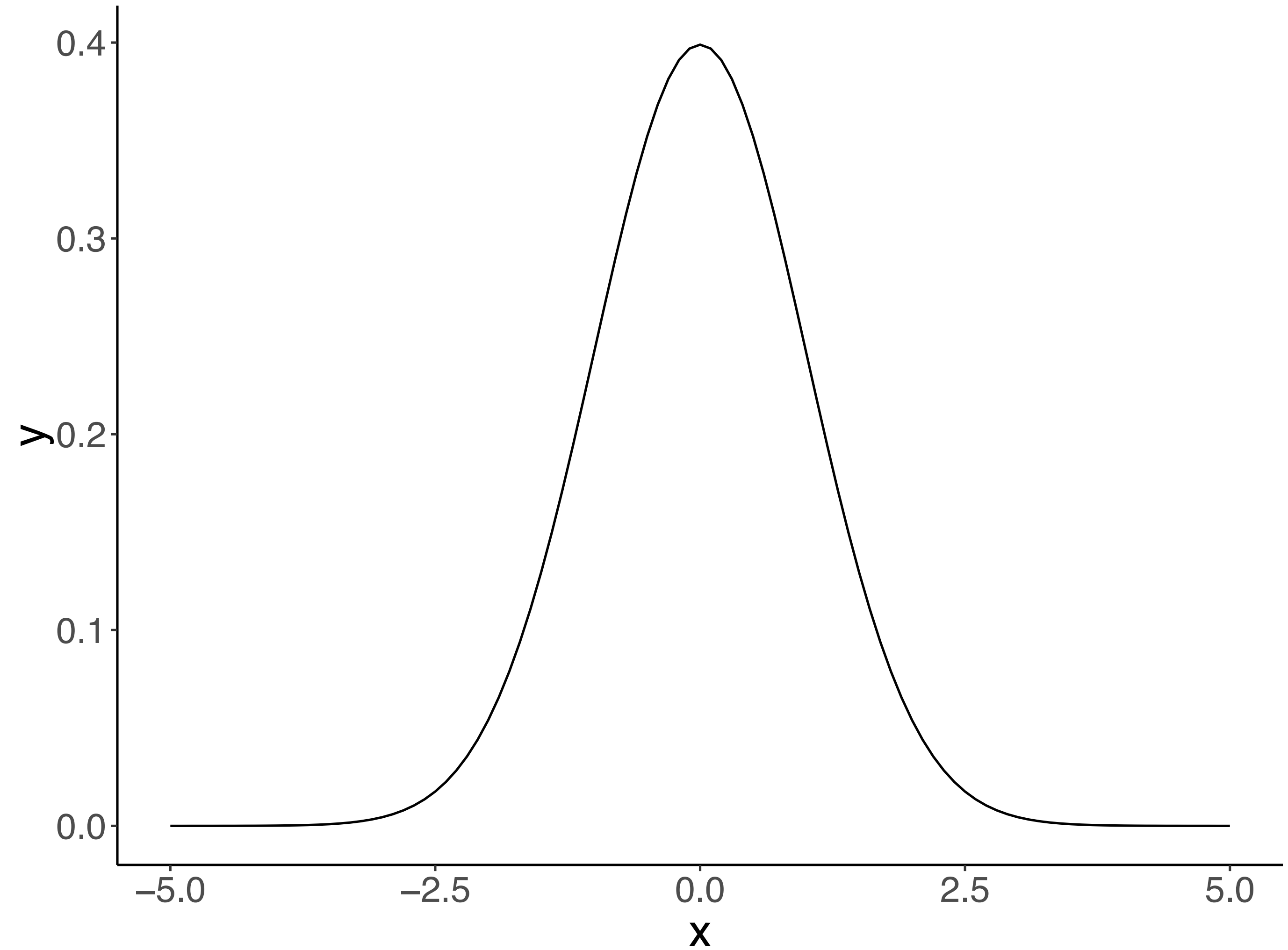
Plan for today

- Quick review of causality
- **Working with probability distributions**
  - `dnorm()`, `pnorm()`, `qnorm()`, `rnorm()`
  - computing probabilities
- Bayesian inference
  - analytic solution
  - via sampling
- Working with samples
  - Understanding `density()`
  - Understanding `quantile()`
  - Comparing distributions

Working with probability distributions

letter	description	example
`d`	for "density", the density function (probability mass function (for *discrete* variables) or probability density function (for *continuous* variables))	`dnorm()`
`p`	for "probability", the cumulative distribution function	`pnorm()`
`q`	for "quantile", the inverse cumulative distribution function	`qnorm()`
`r`	for "random", a random variable having the specified distribution	`rnorm()`

Normal distribution



mu sigma

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

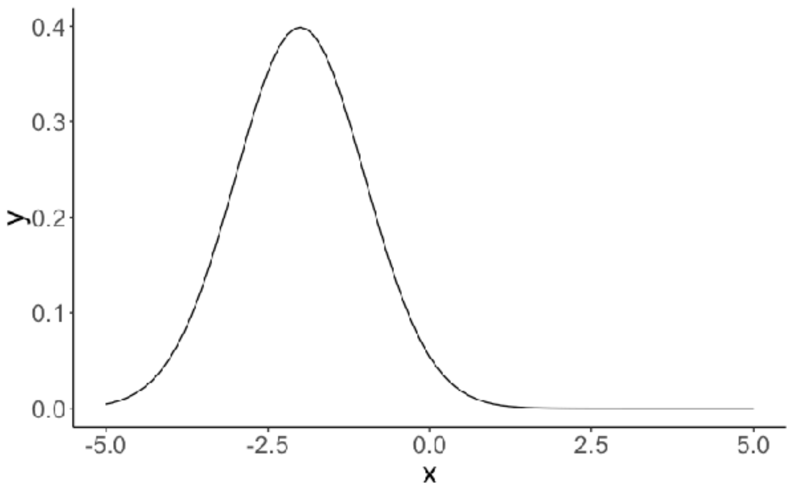
`dnorm(x, mean = 0, sd = 1)`

d = density **norm** = normal distribution

we use **Greek** letters to refer to parameters of the population (or the theoretical distributions)

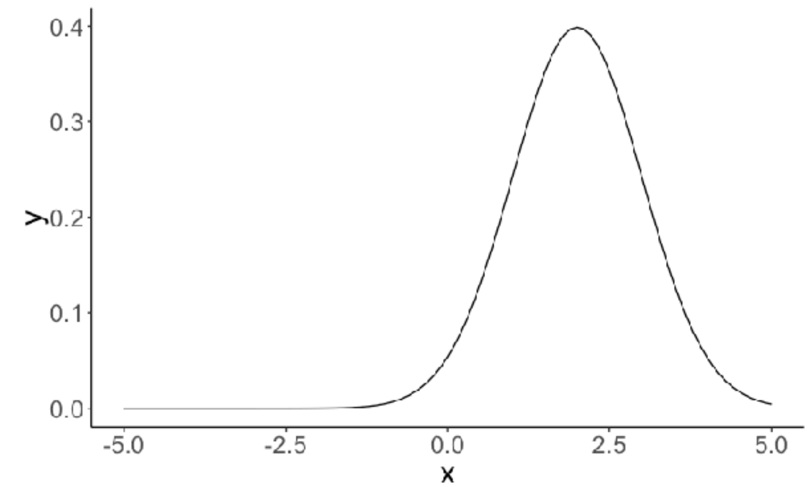
and **Roman** letters to refer to parameters in our sample from the population

Normal distribution

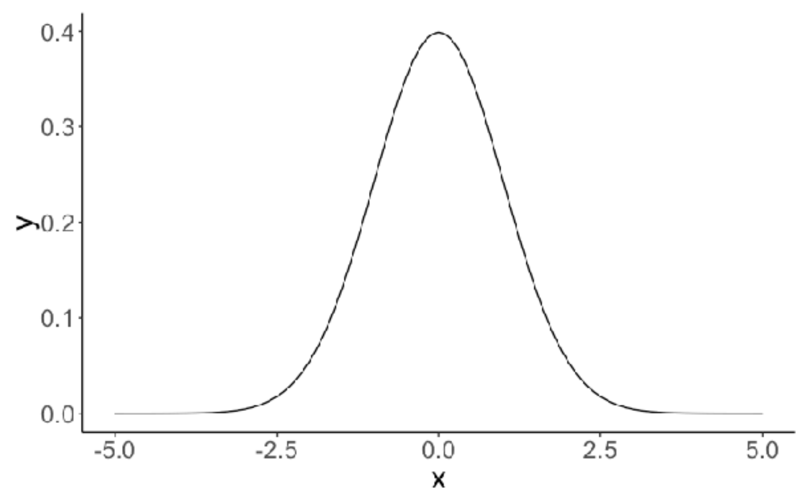


dnorm(x, mean = **-2,** sd = 1)

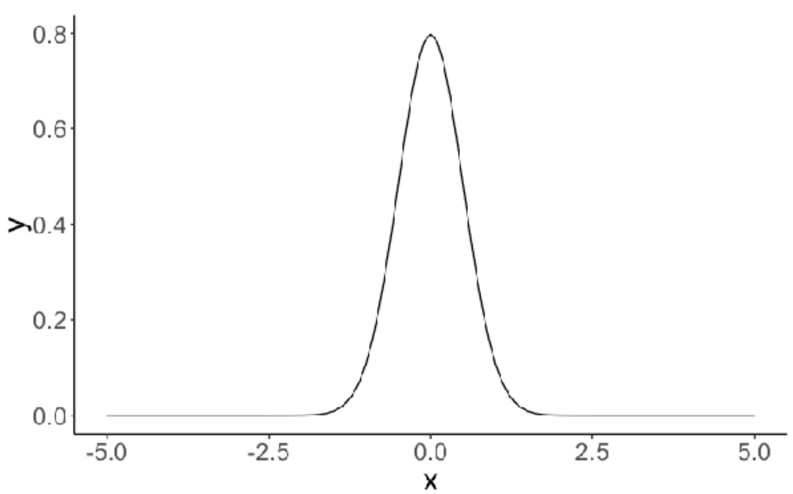
manipulating `mean`



dnorm(x, mean = **2,** sd = 1)

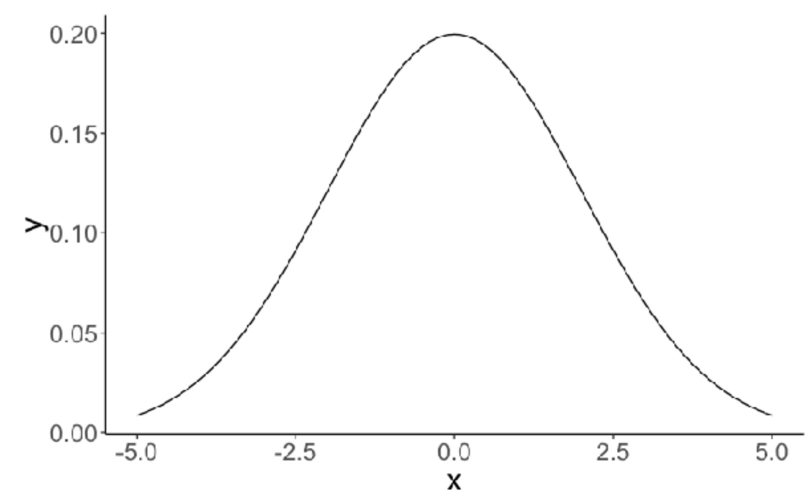


dnorm(x, mean = **0,** sd = 1)



dnorm(x, mean = 0, sd = **0.5**)

manipulating `sd`



dnorm(x, mean = 0, sd = **2**)

Plotting distributions

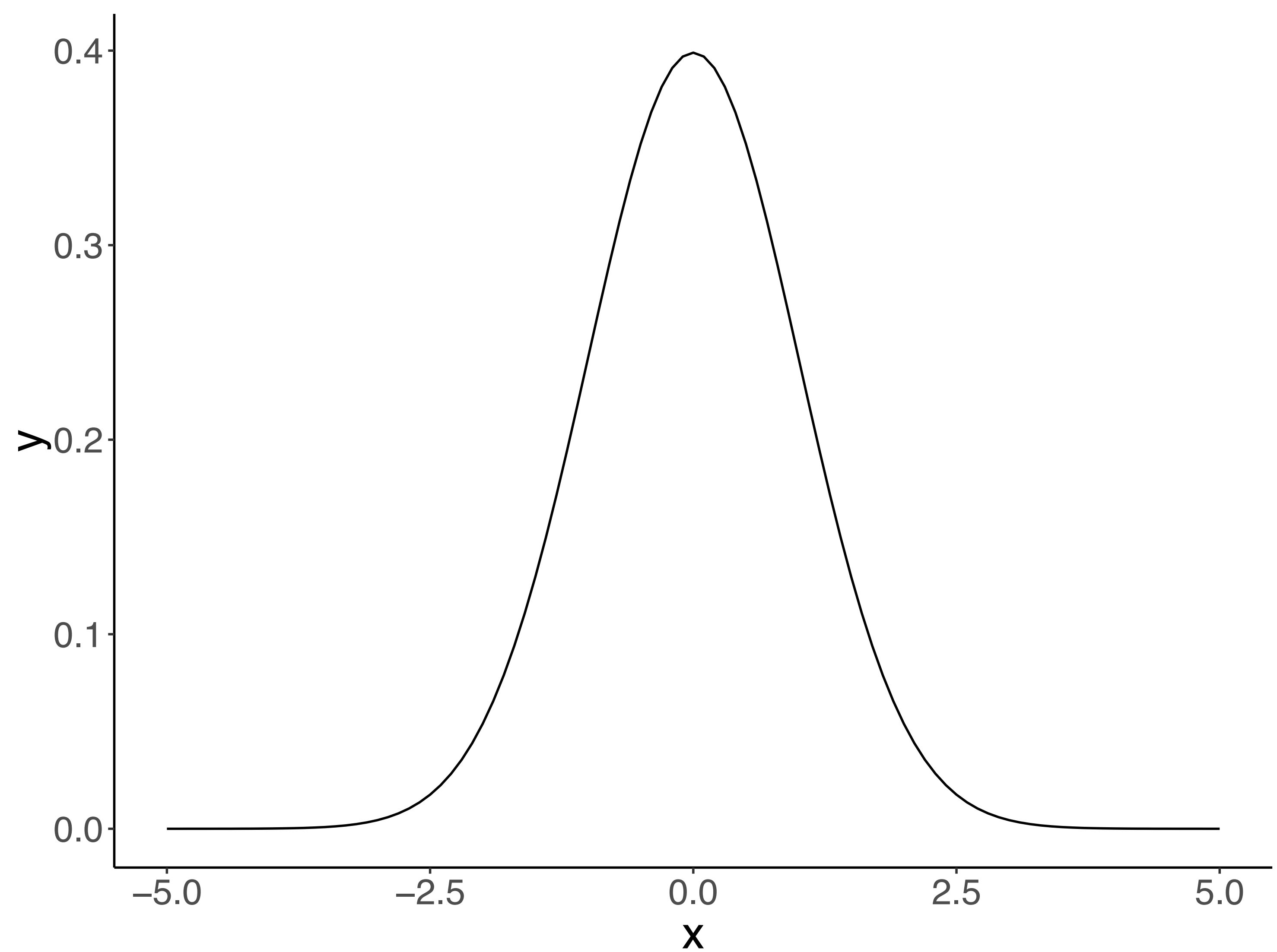
```
ggplot(data = tibble(x = c(-5, 5)),
       mapping = aes(x = x)) +
  stat_function(fun = ~ dnorm(x = .,
                              mean = 0,
                              sd = 1))
```

make data frame with minimum and maximum x-value

any parameters for the function?

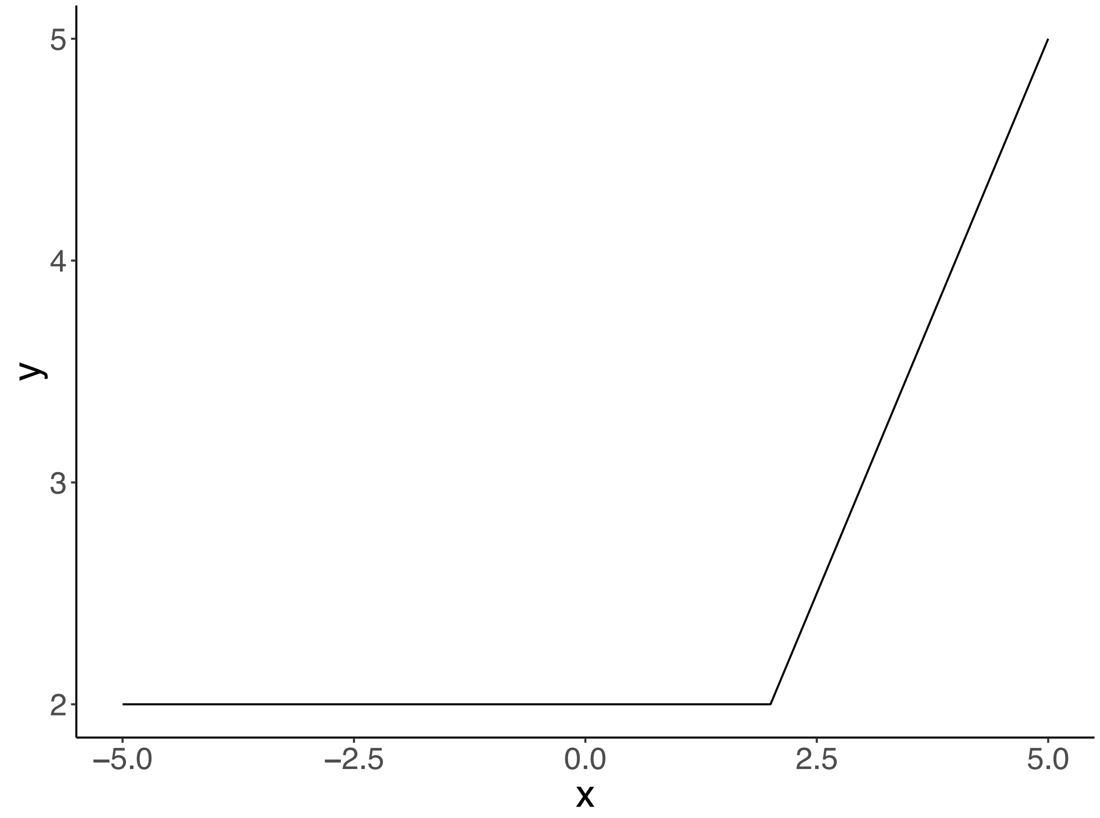
what function should be plotted?

function for plotting functions



Plotting functions

```r
# define the breakpoint function
fun.breakpoint = function(x, breakpoint){
  x[x < breakpoint] = breakpoint
  return(x)
}

# plot the function
ggplot(data = tibble(x = c(-5, 5)),
       mapping = aes(x = x)) +
  stat_function(fun = ~ fun.breakpoint(., breakpoint = 2))
```



Sampling from distributions

n = number of samples

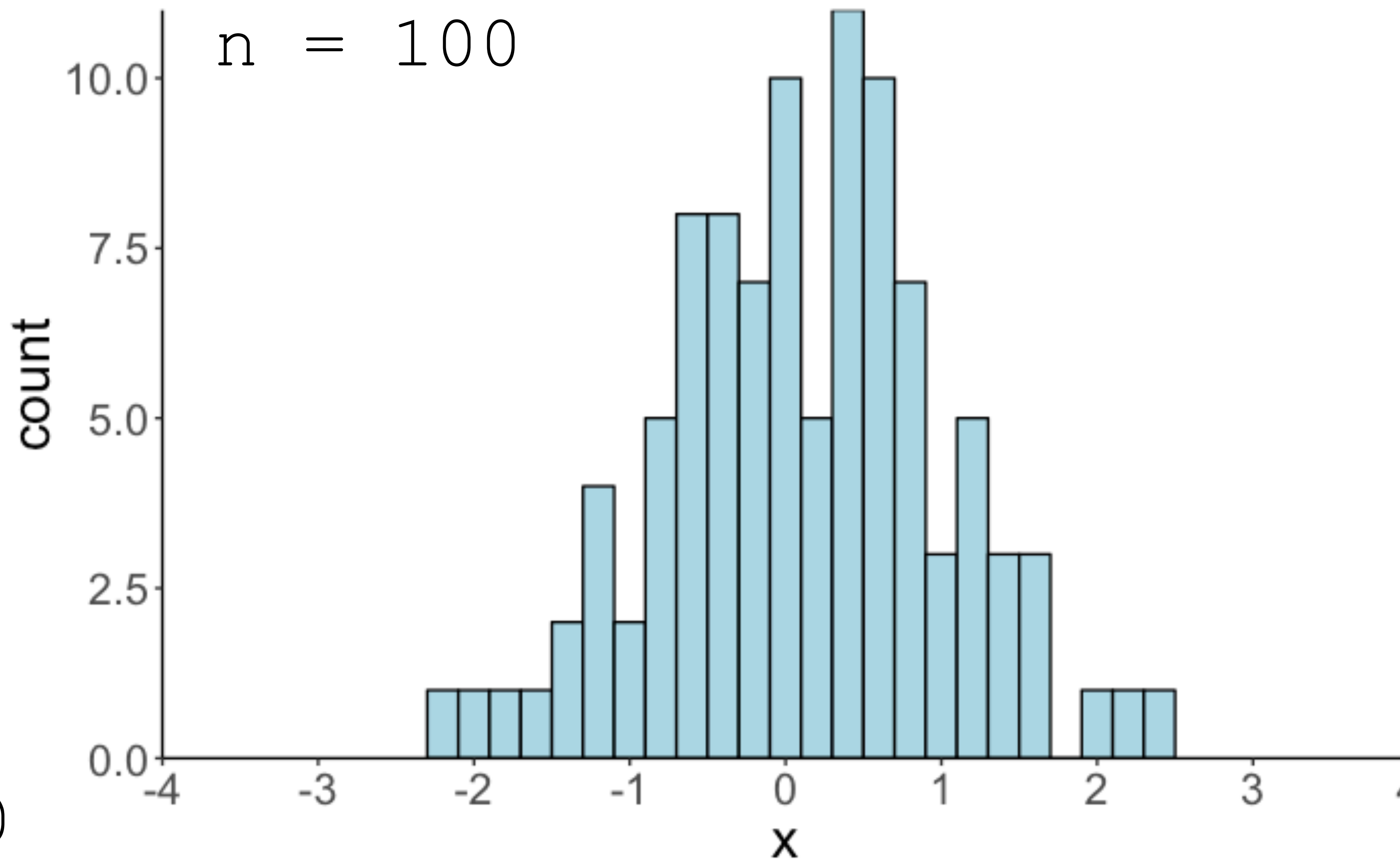
```
rnorm(n, mean = 0, sd = 1)
```

r = random samples   **norm** = normal distribution

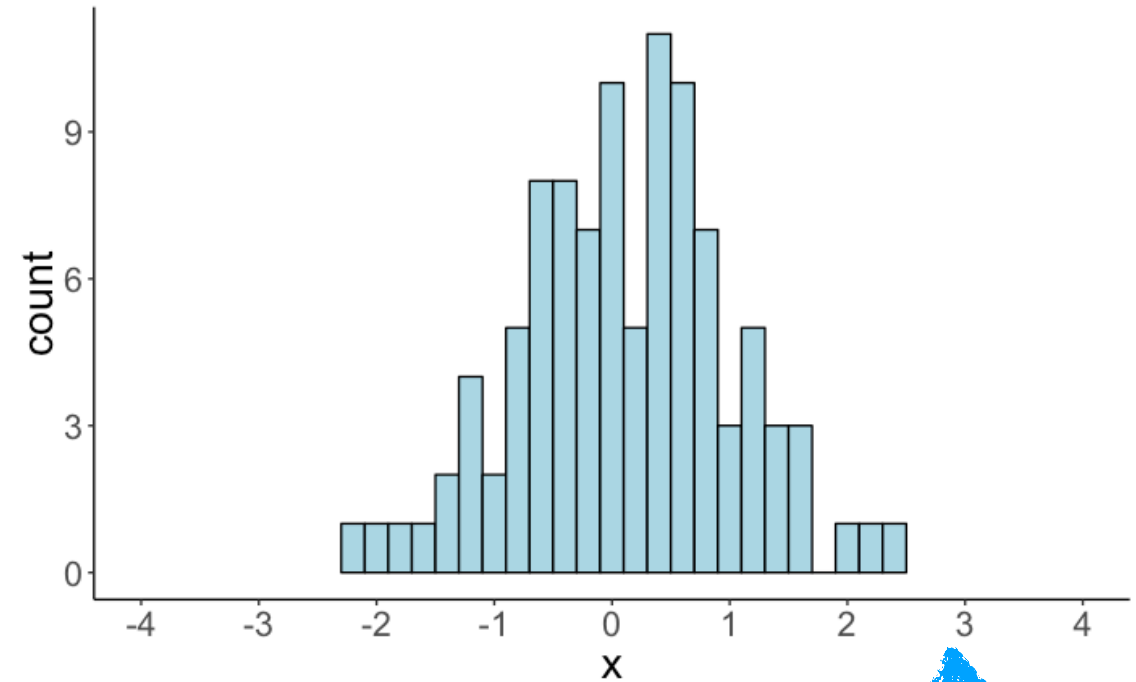
x
-0.63
0.18
-0.84
1.60
0.33
-0.82
0.49
0.74

⋮

nrow = 100



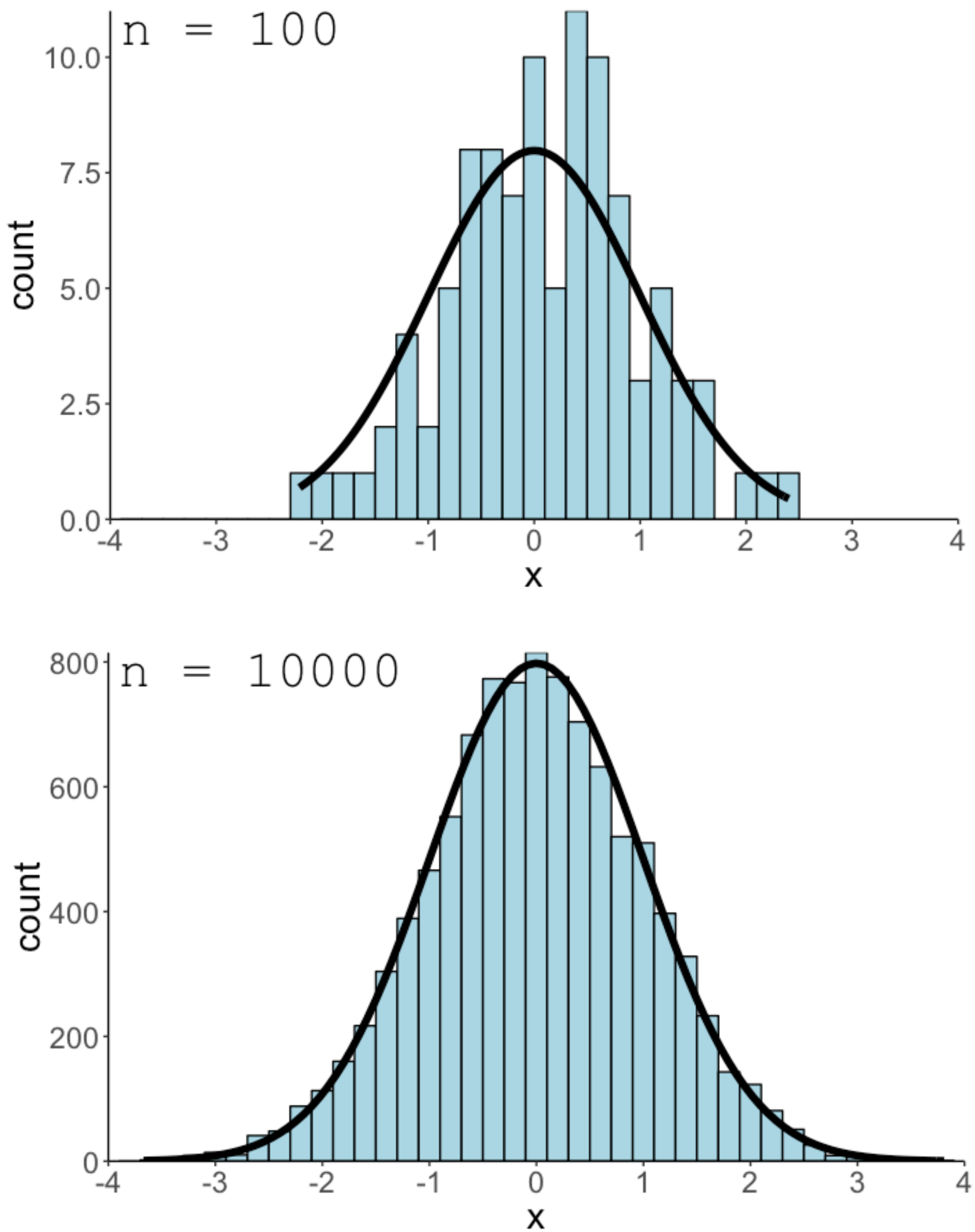
Sampling from distributions




```
# make this example reproducible
set.seed(1)

# define how many samples to draw
nsamples = 100

# make a data frame with the samples
df.plot = tibble(x = rnorm(n = nsamples, mean = 0, sd = 1))

# plot the samples using a histogram
ggplot(data = df.plot,
       mapping = aes(x = x)) +
  geom_histogram(binwidth = 0.2,
                 color = "black",
                 fill = "lightblue") +
  scale_x_continuous(breaks = -4:4, labels = -4:4) +
  coord_cartesian(xlim = c(-4, 4), expand = F)
```

expand = T

Sampling from distributions



law of large numbers

approximation to true underlying distribution improves with increased sample size

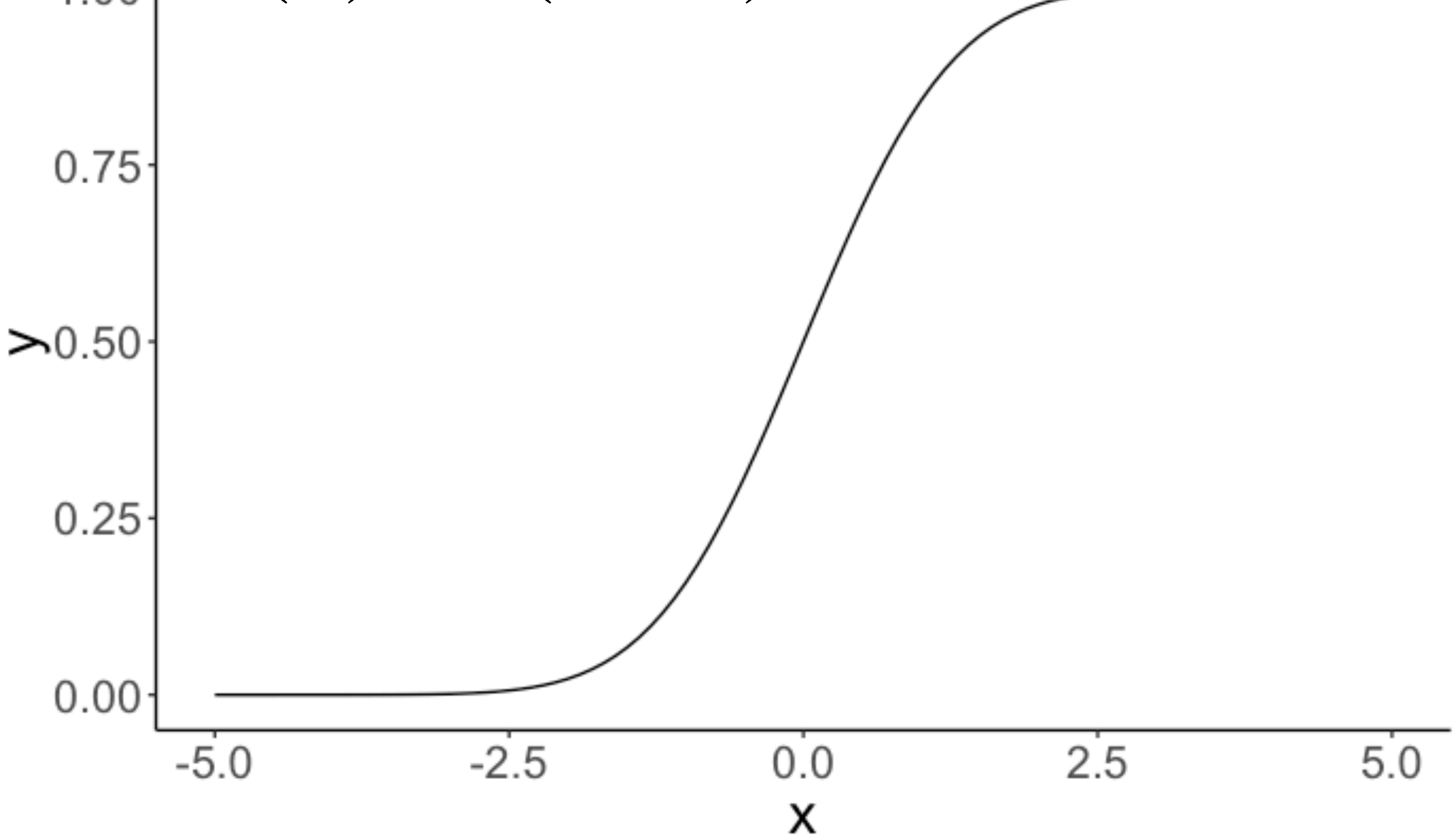
Cumulative probability distribution

```r
ggplot(data = tibble(x = c(-5, 5)),
       mapping = aes(x = x)) +
  stat_function(fun = ~ pnorm(q = .,
                              mean = 0,
                              sd = 1))
```

p = probability

cumulative distribution function

$$F(X) = P(X \leq x)$$

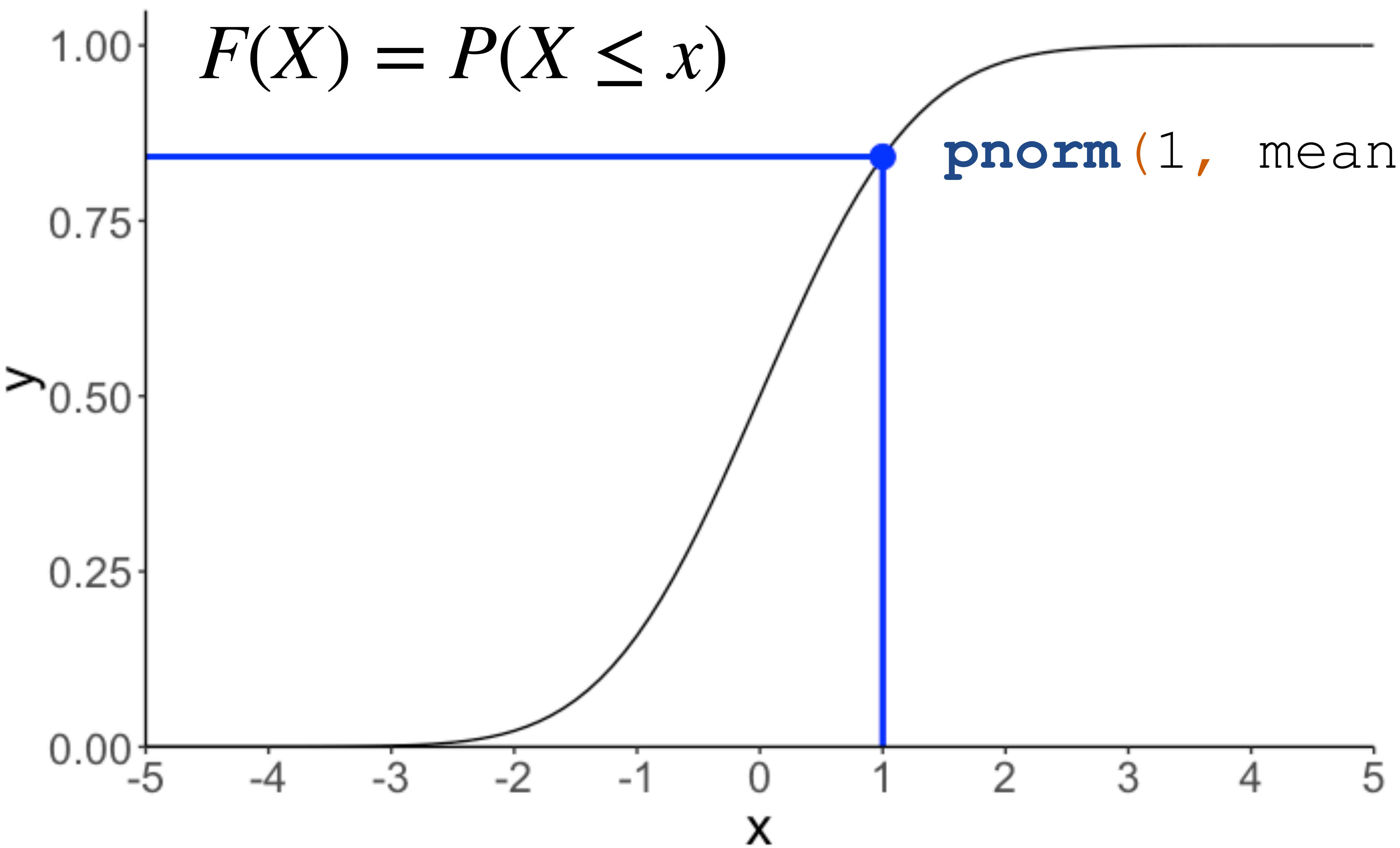


Cumulative probability distribution

```r
ggplot(data = tibble(x = c(-5, 5)),
       mapping = aes(x = x)) +
  stat_function(fun = ~ pnorm(q = .,
                              mean = 0,
                              sd = 1))
```

p = probability

cumulative distribution function



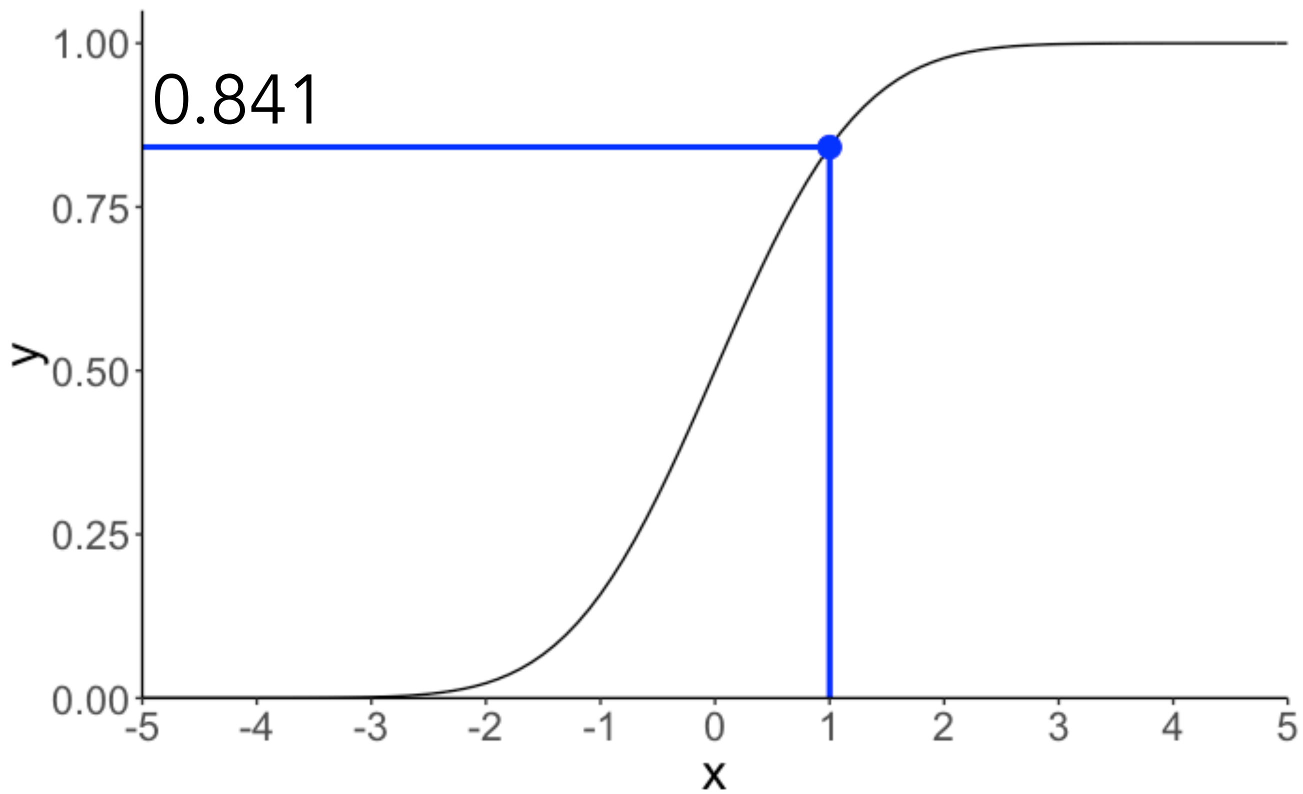
$$F(X) = P(X \leq x)$$

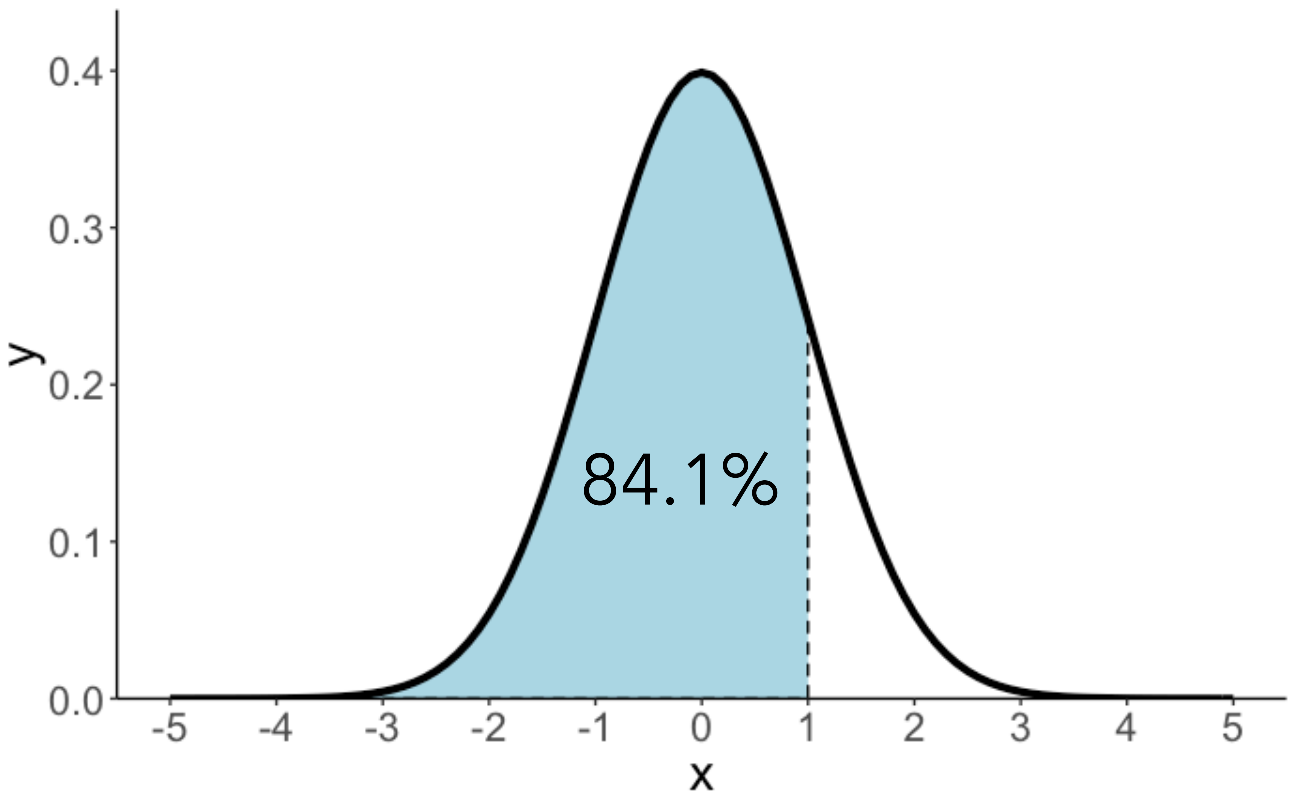
`pnorm(1, mean = 0, sd = 1)`

= 0.841

Computing probabilities

```
pnorm(1, mean = 0, sd = 1)
```





`pnorm(x)` returns the integral from $-\infty$ to `x` of the probability density function

Inverse cumulative distribution function

```r
ggplot(data = tibble(x = c(0, 1)),
       mapping = aes(x = x)) +
  stat_function(fun = ~ qnorm(p = .,
                              mean = 0,
                              sd = 1))
```

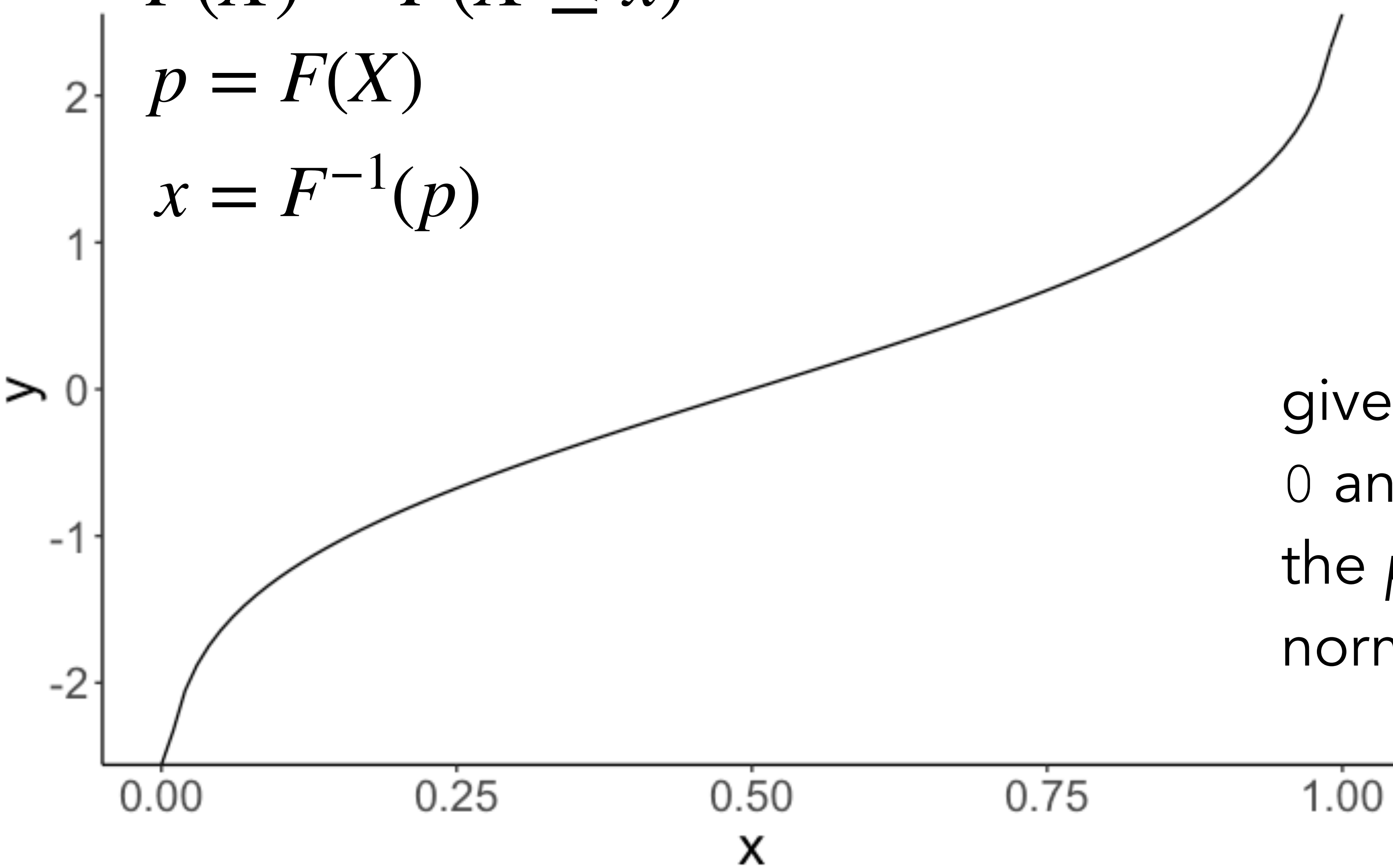
q = quantile

inverse cumulative distribution function

$$F(X) = P(X \leq x)$$

$$p = F(X)$$

$$x = F^{-1}(p)$$



given a number *p* between `0` and `1`, `qnorm` looks up the *p*-th quantile of the normal distribution

Inverse cumulative distribution function

```r
ggplot(data = tibble(x = c(0, 1)),
       mapping = aes(x = x)) +
  stat_function(fun = ~ qnorm(p = .,
                              mean = 0,
                              sd = 1))
```

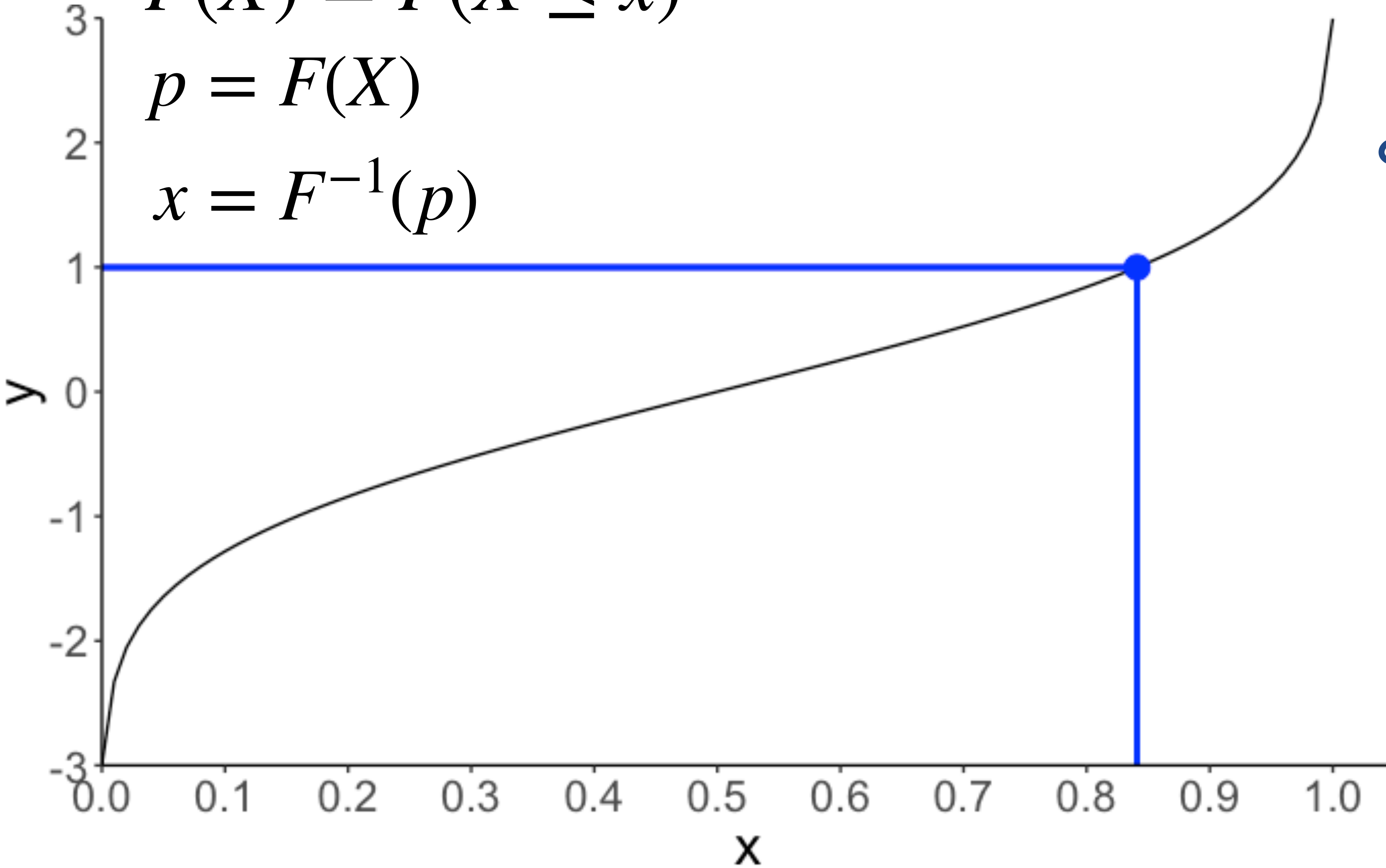
q = quantile

inverse cumulative distribution function

$$F(X) = P(X \leq x)$$

$$p = F(X)$$

$$x = F^{-1}(p)$$

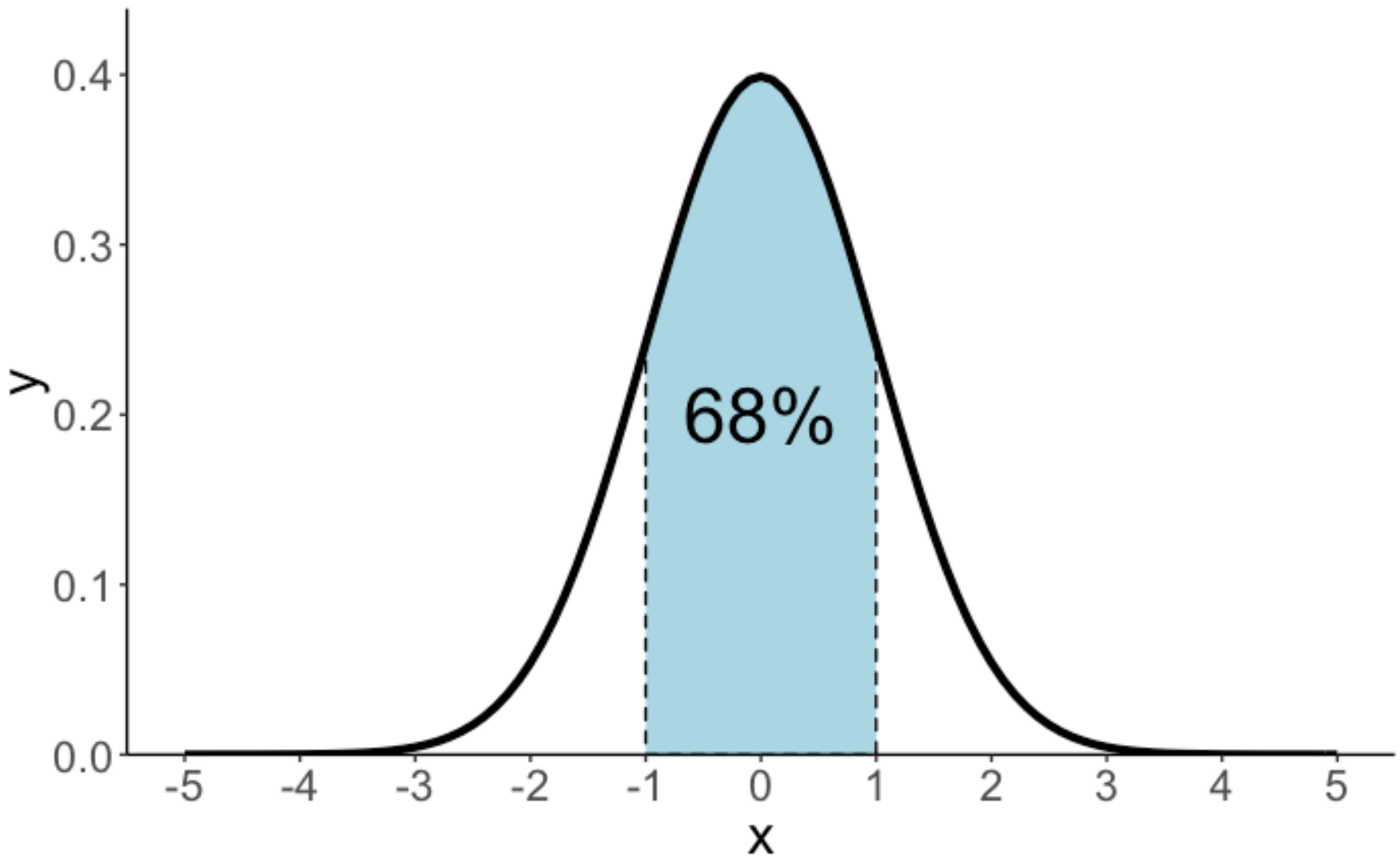



```r
qnorm(0.841,
      mean = 0,
      sd = 1)
```
≈ 1

Computing probabilities

Find the probability between two values of interest:

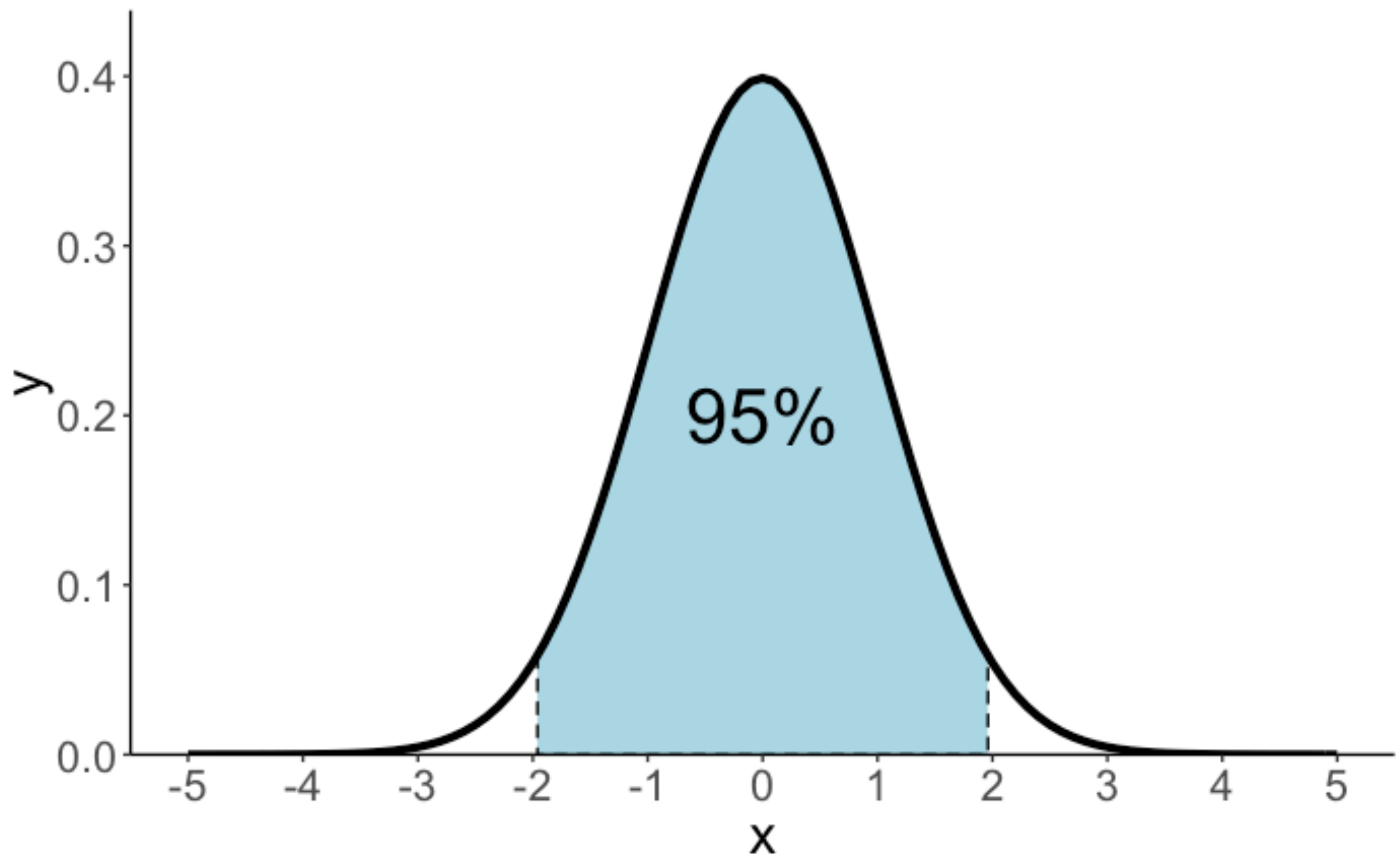
```
pnorm(1) - pnorm(-1) =
  0.84   -   0.16    = 0.68
```



Computing probabilities

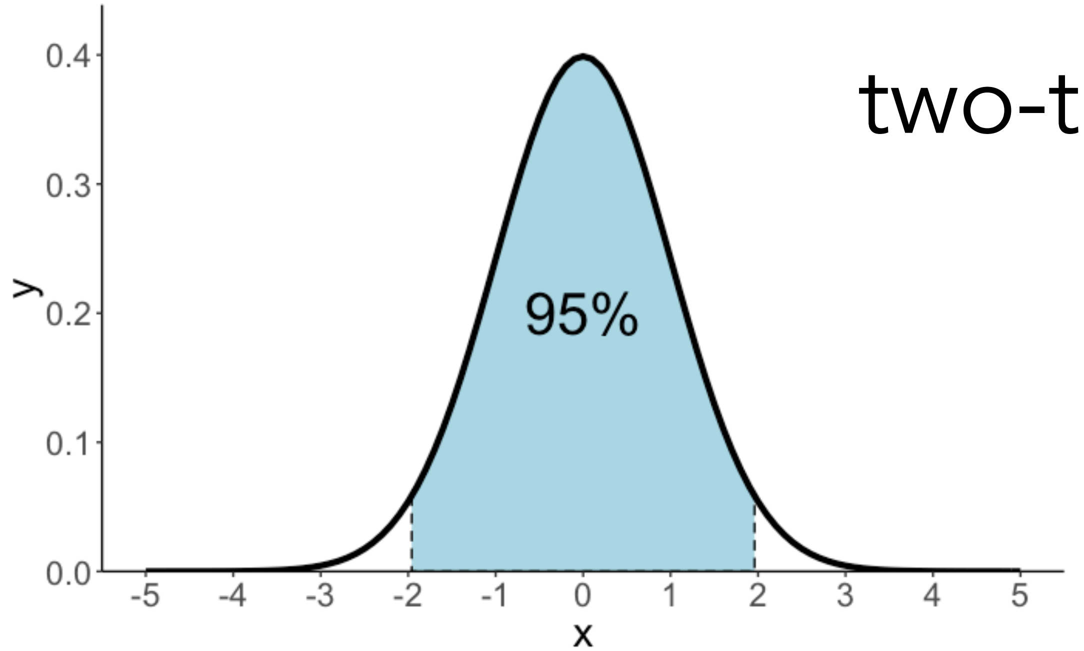
Find the lower and upper value so that 95% of the probability are contained between these values.

```
qnorm(0.025)= -1.96
qnorm(0.975)= 1.96
```



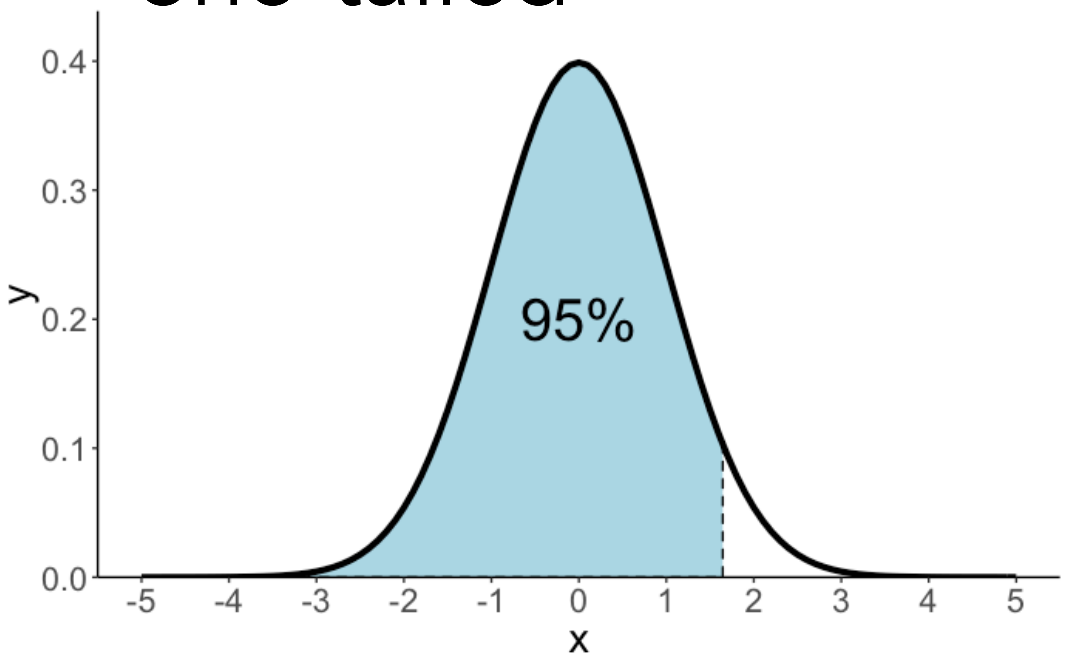
Computing probabilities

two-tailed



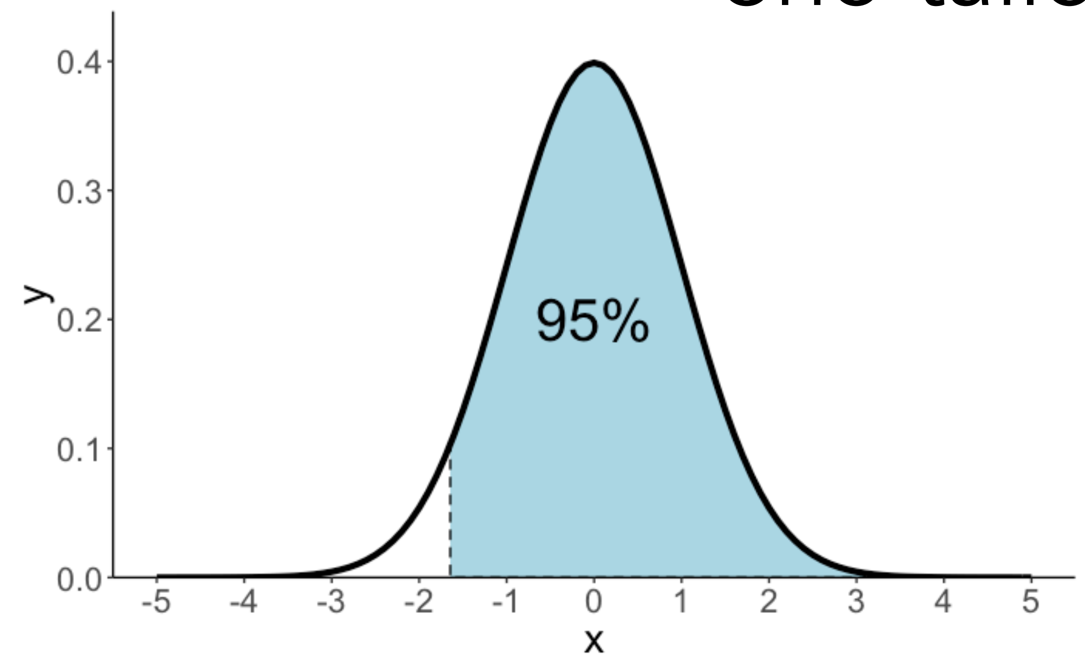

```
qnorm(0.025)= -1.96
qnorm(0.975)= 1.96
```

one-tailed




```
qnorm(0)= -Inf
qnorm(0.95)= 1.64
```

one-tailed




```
qnorm(0.05)= -1.64
qnorm(1)= Inf
```

Plan for today

- Quick review of causality
- Working with probability distributions
  - `dnorm()`, `pnorm()`, `qnorm()`, `rnorm()`
  - computing probabilities
- **Bayesian inference**
  - analytic solution
  - via sampling
- Working with samples
  - Understanding `density()`
  - Understanding `quantile()`
  - Comparing distributions

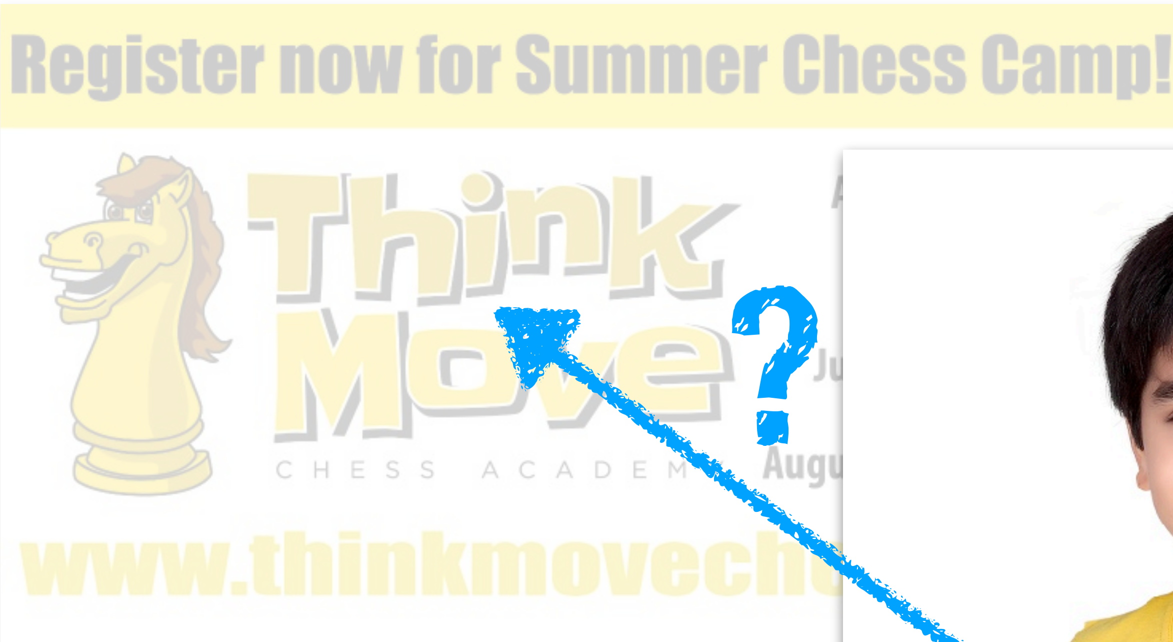
Summer camp

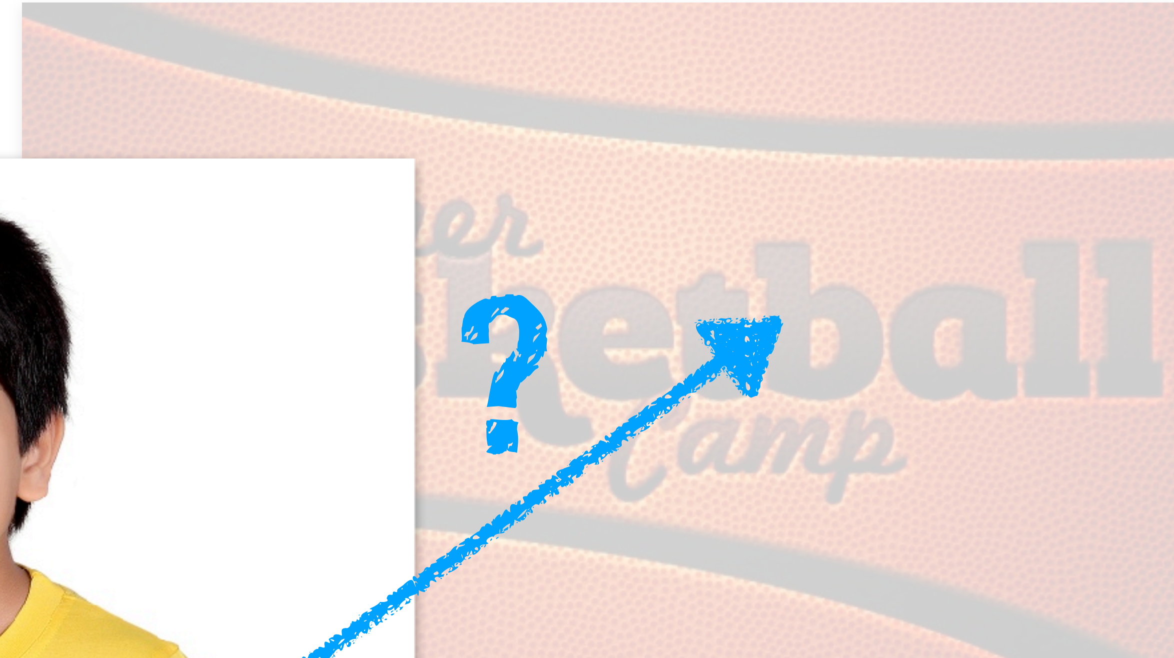
twice as many kids go to the basketball camp

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

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

Summer camp





twice as ma … etball camp

$X \sim$ Normal($\mu = 170$ … nal($\mu = 180$, $\sigma = 10$)

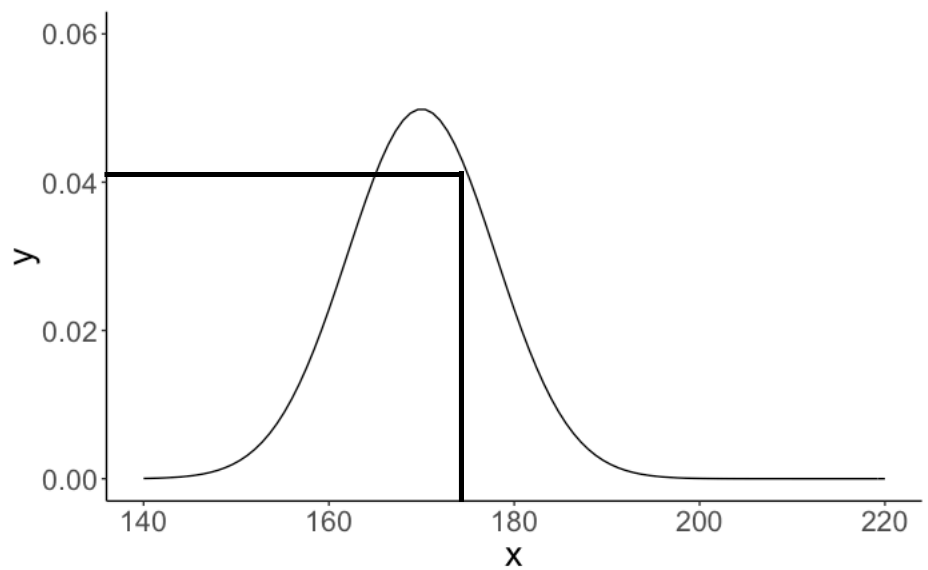
Analytic solution

Summer camp

prior

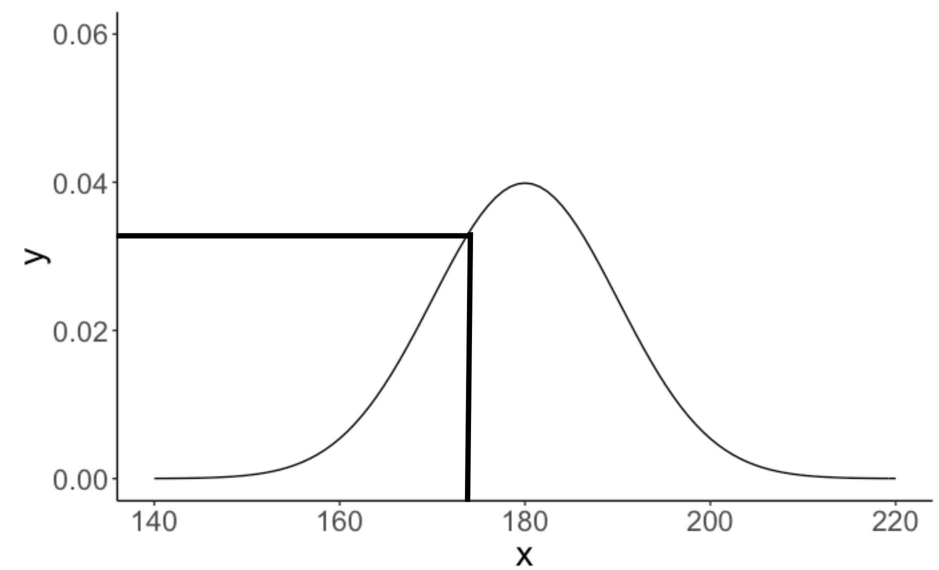
$$p(\text{chess}) = \frac{1}{3} \qquad p(\text{basketball}) = \frac{2}{3}$$

likelihood



`dnorm(175, mean = 170, sd = 8)`

= 0.041



`dnorm(175, mean = 180, sd = 10)`

= 0.035

posterior

$$p(\text{sport} = \text{basketball} \mid \text{height} = 175) = \frac{p(175 \mid \text{basketball}) \cdot p(\text{basketball})}{p(175)}$$

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

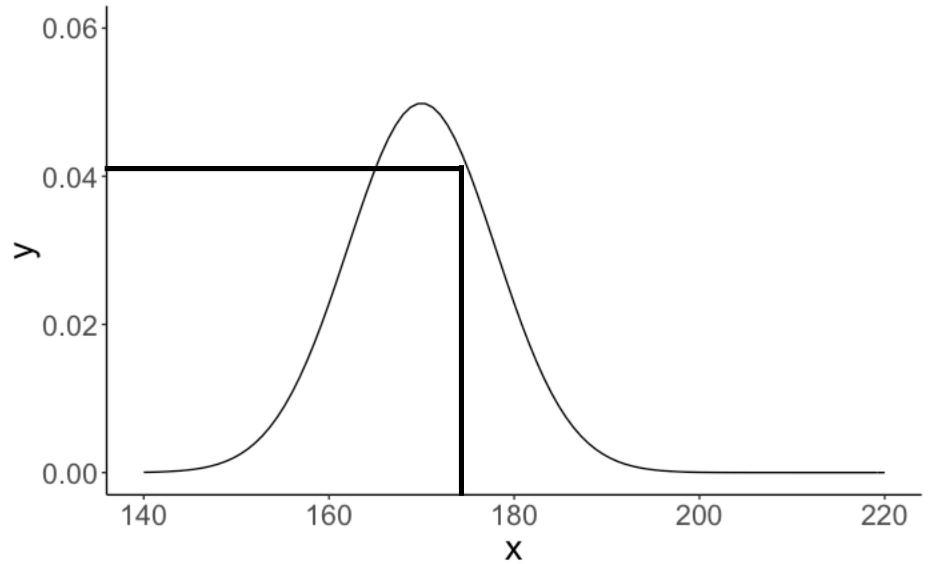
Summer camp

prior

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

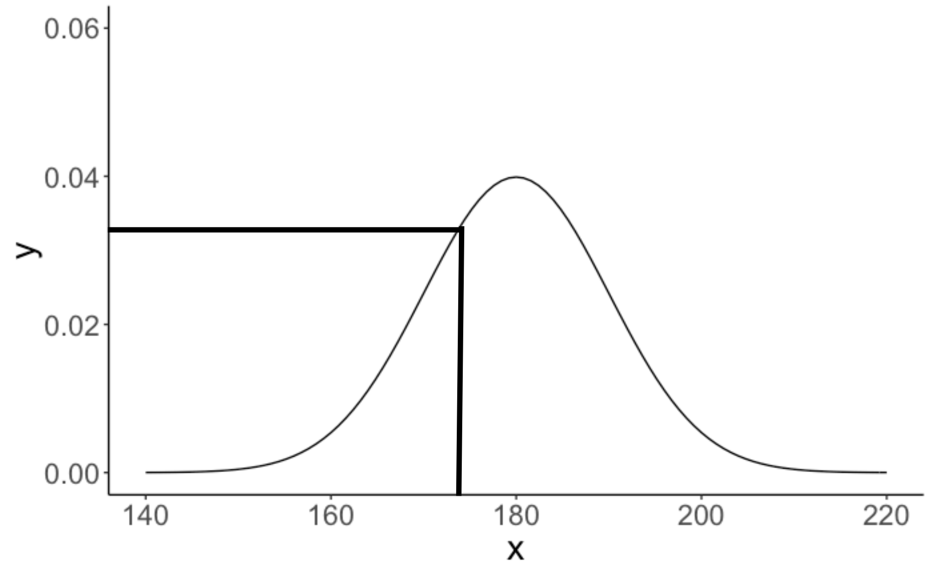
$$p(\text{basketball}) = \frac{2}{3}$$

likelihood




```
dnorm(175, mean = 170, sd = 8)
```

$= 0.041$




```
dnorm(175, mean = 180, sd = 10)
```

$= 0.035$

posterior

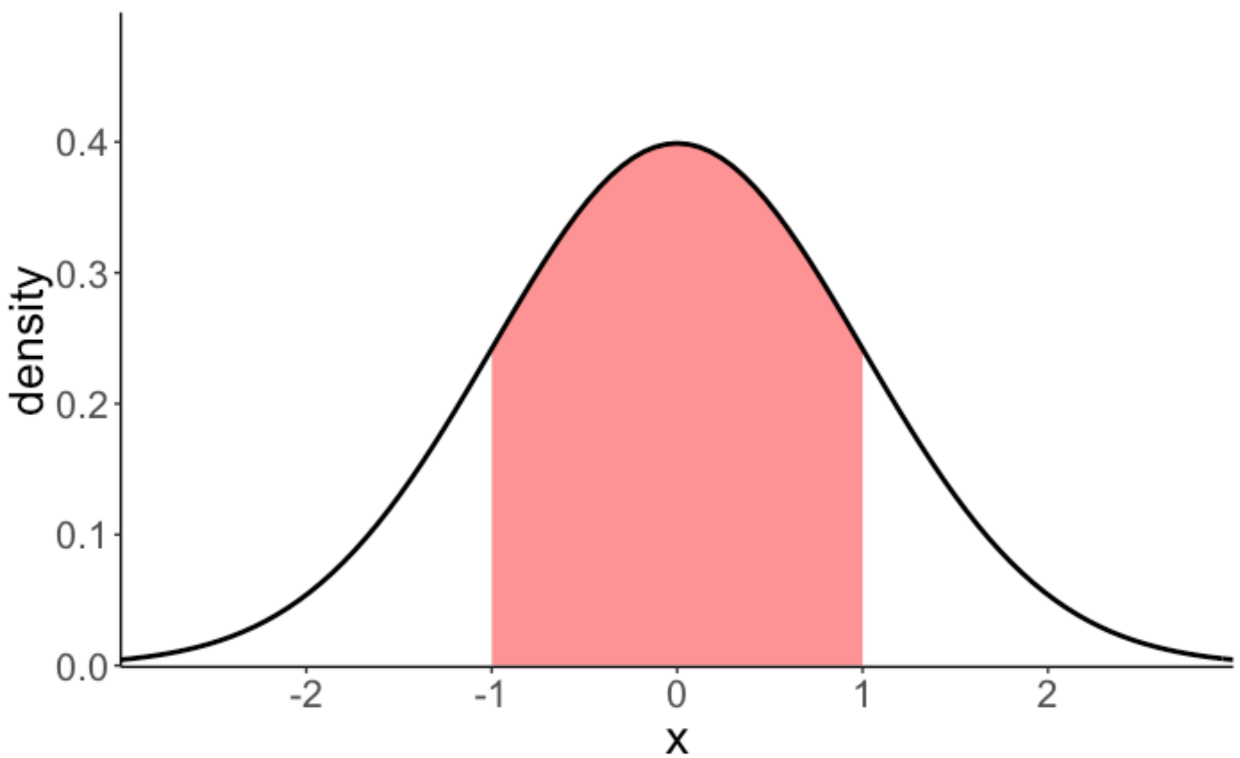
$$p(\text{basketball} \mid 175) = \frac{p(175 \mid \text{basketball}) \cdot p(\text{basketball})}{p(175 \mid \text{basketball}) \cdot p(\text{basketball}) + p(175 \mid \text{chess}) \cdot p(\text{chess})}$$

$$p(\text{basketball} \mid 175) = \frac{0.035 \cdot 2/3}{0.035 \cdot 2/3 + 0.041 \cdot 1/3} \approx 0.63$$

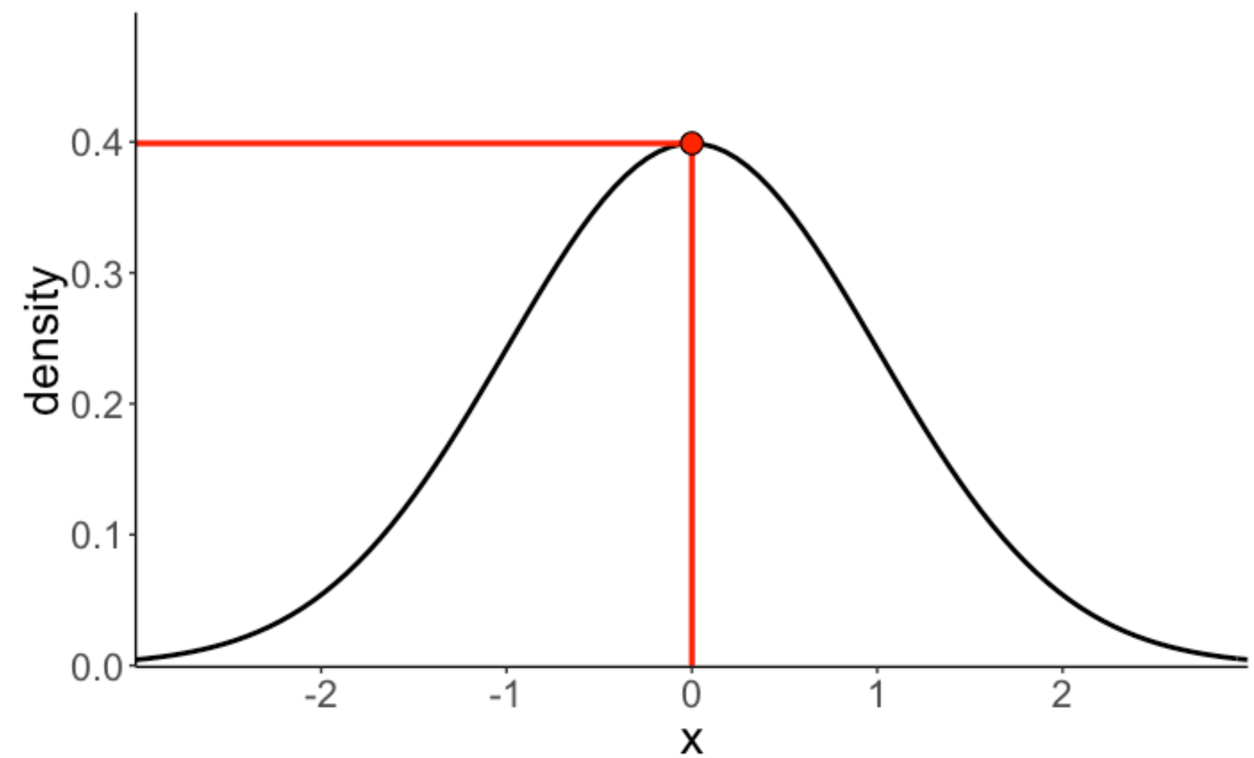
send the kid to the basketball gym!

Probability vs. likelihood

Probability

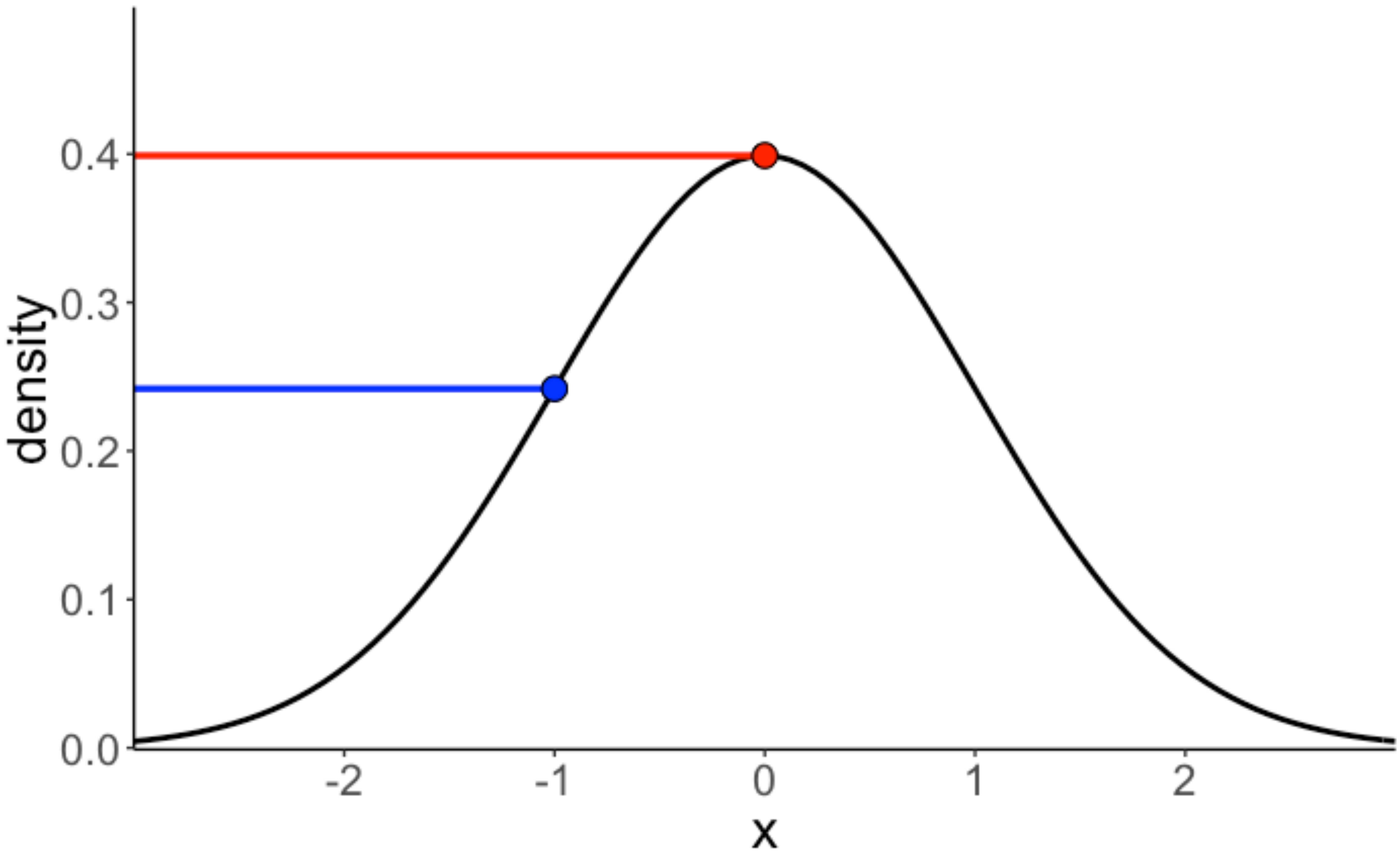


Likelihood



Probability vs. likelihood

`dnorm(0)/dnorm(-1) = 1.6487`

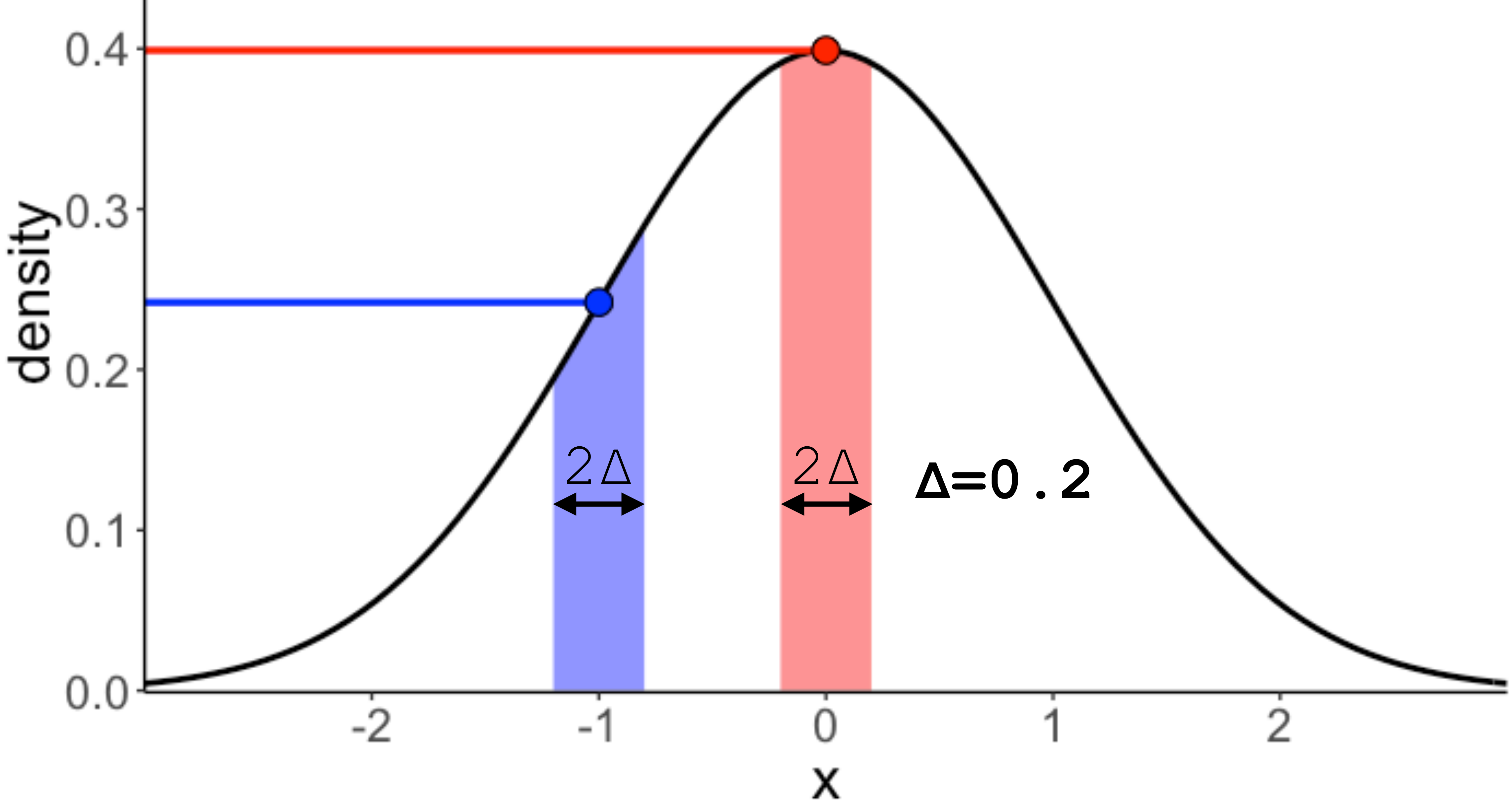


relative probability of one value vs. another

Probability vs. likelihood

`dnorm(0)/dnorm(-1) = 1.6487`

`(pnorm(0+Δ)- pnorm(0-Δ))/`
`(pnorm(-1+Δ) -(pnorm(-1-Δ) =` **`1.6378`**

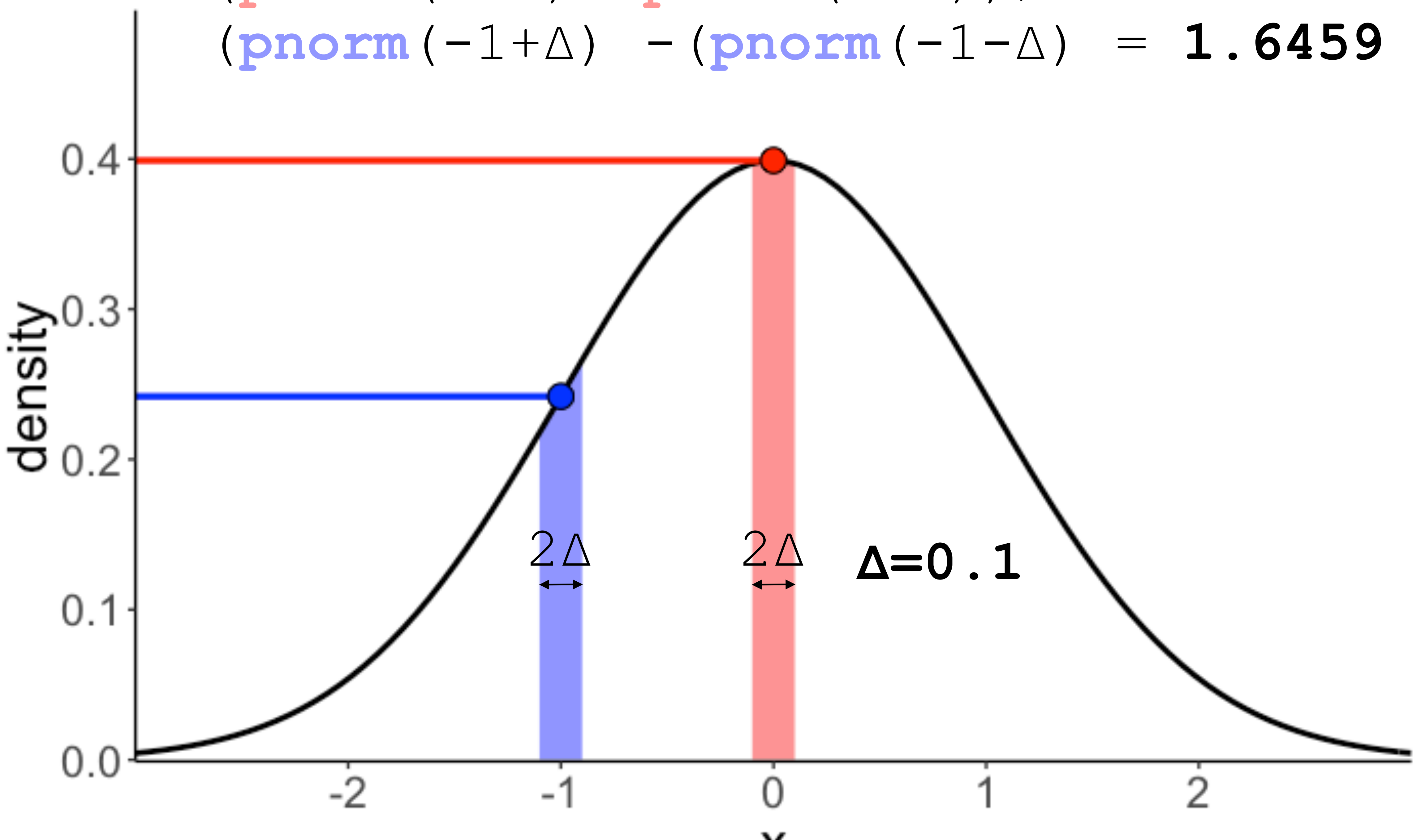


relative probability of one value vs. another

Probability vs. likelihood

`dnorm(0)/dnorm(-1) = 1.6487`

`(pnorm(0+Δ)- pnorm(0-Δ))/`
`(pnorm(-1+Δ) -(pnorm(-1-Δ) =` **`1.6459`**

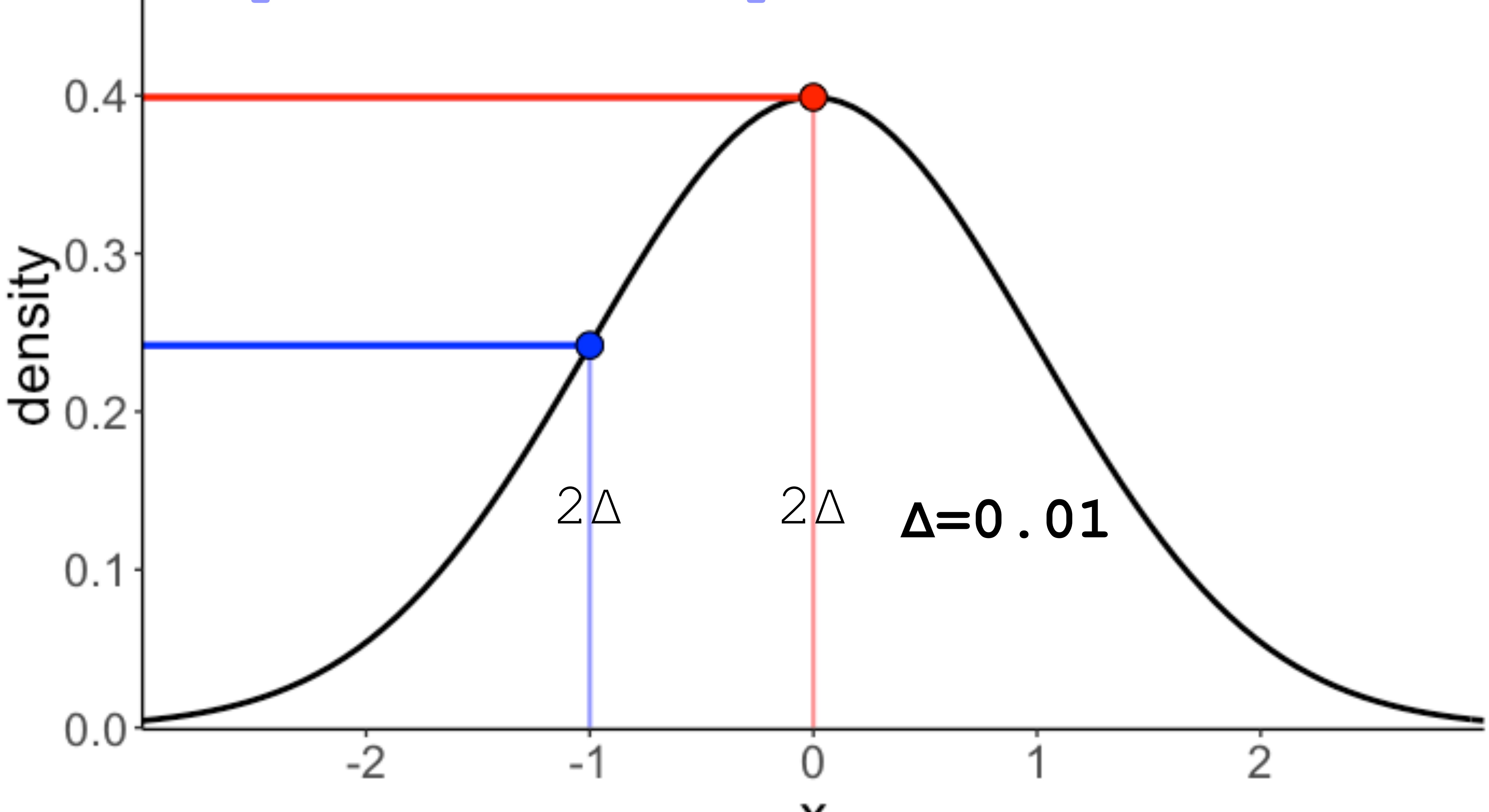


relative probability of one value vs. another

Probability vs. likelihood

```
dnorm(0)/dnorm(-1) = 1.6487
```

```
(pnorm(0+Δ)- pnorm(0-Δ))/
(pnorm(-1+Δ) -(pnorm(-1-Δ) = 1.6486
```



relative probability of one value vs. another

Sampling solution

Summer camp: Via sampling

```
df.camp = tibble(
  kid = 1:1000,
  sport = sample(c("chess", "basketball"),
                 size = 1000,
                 replace = T,
                 prob = c(1/3, 2/3))) %>%
  rowwise() %>%
  mutate(height = ifelse(test = sport == "chess",
                         yes = rnorm(., mean = 170, sd = 8),
                         no = rnorm(., mean = 180, sd = 10))) %>%
  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

```
df.camp %>%
  filter(height == 175) %>%
  count(sport) %>%
  pivot_wider(names_from = sport,
              values_from = n) %>%
  summarize(prob_basketball =
            basketball/(basketball
            + chess))
```

doesn't work!

Summer camp: Via sampling

```
df.camp = tibble(
  kid = 1:1000,
  sport = sample(c("chess", "basketball"),
                 size = 1000,
                 replace = T,
                 prob = c(1/3, 2/3))) %>%
  rowwise() %>%
  mutate(height = ifelse(test = sport == "chess",
                         yes = rnorm(., mean = 170, sd = 8),
                         no = rnorm(., mean = 180, sd = 10))) %>%
  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

```
df.camp %>%
  filter(between(height,
                 left = 174,
                 right = 176)) %>%
  count(sport) %>%
  pivot_wider(names_from = sport,
              values_from = n) %>%
  summarize(prob_basketball =
            basketball/(basketball
            + chess))
```

this works!

≈ 0.63

Plan for today

- Quick review of causality
- Working with probability distributions
  - `dnorm()`, `pnorm()`, `qnorm()`, `rnorm()`
  - computing probabilities
- Bayesian inference
  - analytic solution
  - via sampling
- **Working with samples**
  - Understanding `density()`
  - Understanding `quantile()`
  - Comparing distributions

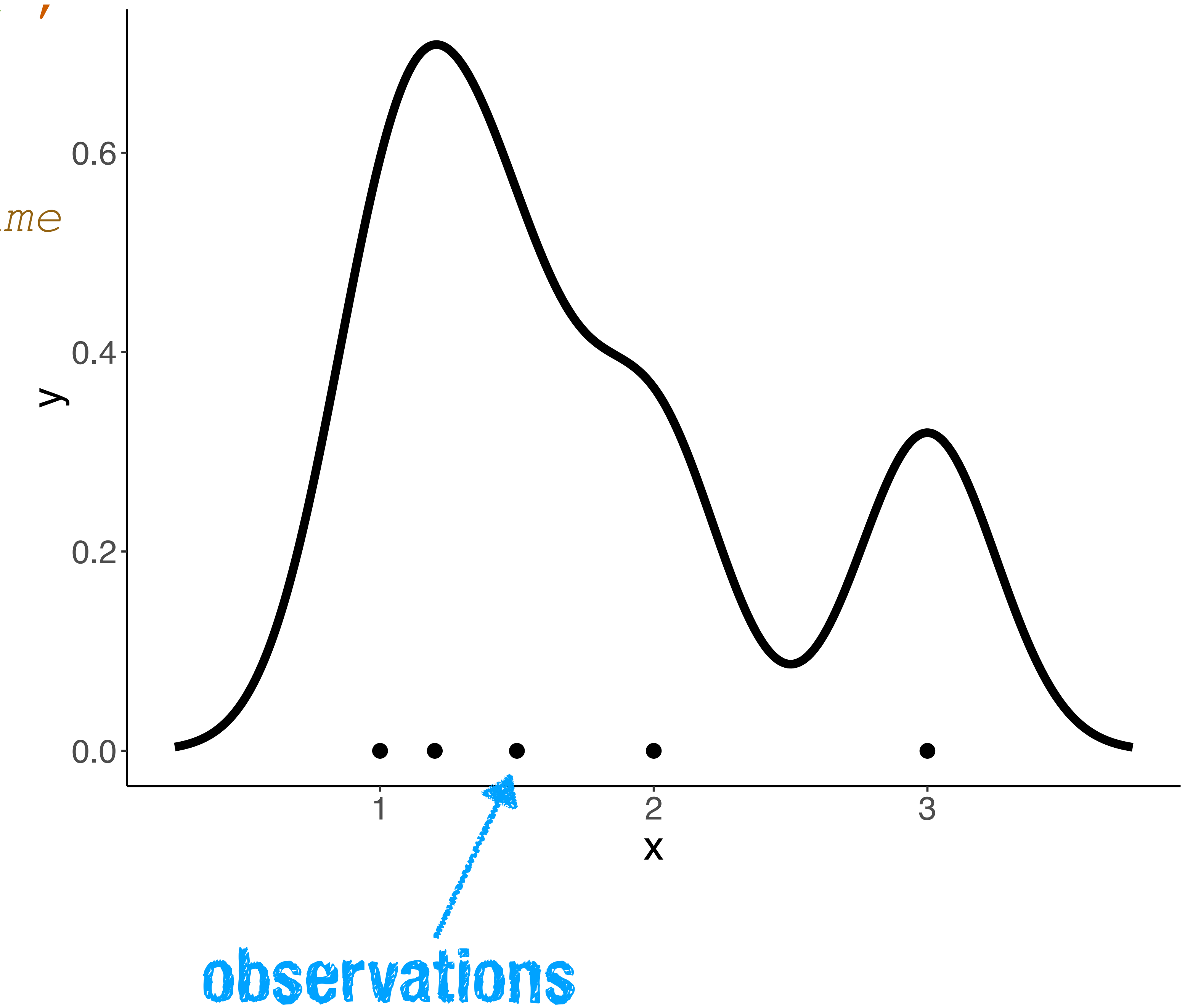
Understanding **`density()`**

Understanding `density()`

```r
# calculate density
observations = c(1, 1.2, 1.5, 2, 3)
bandwidth = 0.25
density = density(observations,
          kernel = "gaussian",
          bw = bandwidth,
          n = 512)

# save density as data frame
df.density = tibble(
  x = density$x,
  y = density$y
)
```

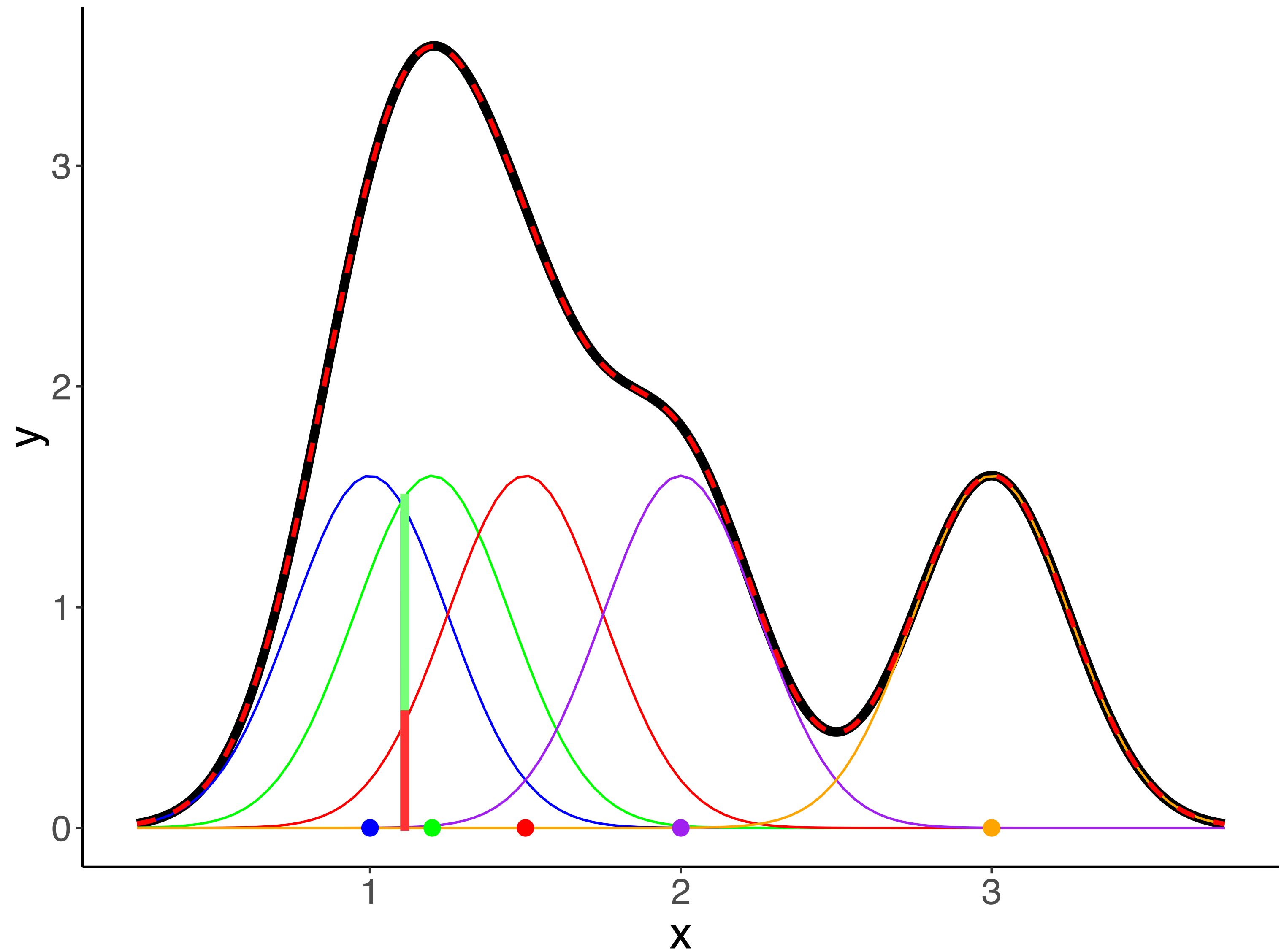
x	y
0.250	0.004
0.257	0.004
0.264	0.005
0.271	0.005
0.277	0.005
0.284	0.006

`nrow = 512`



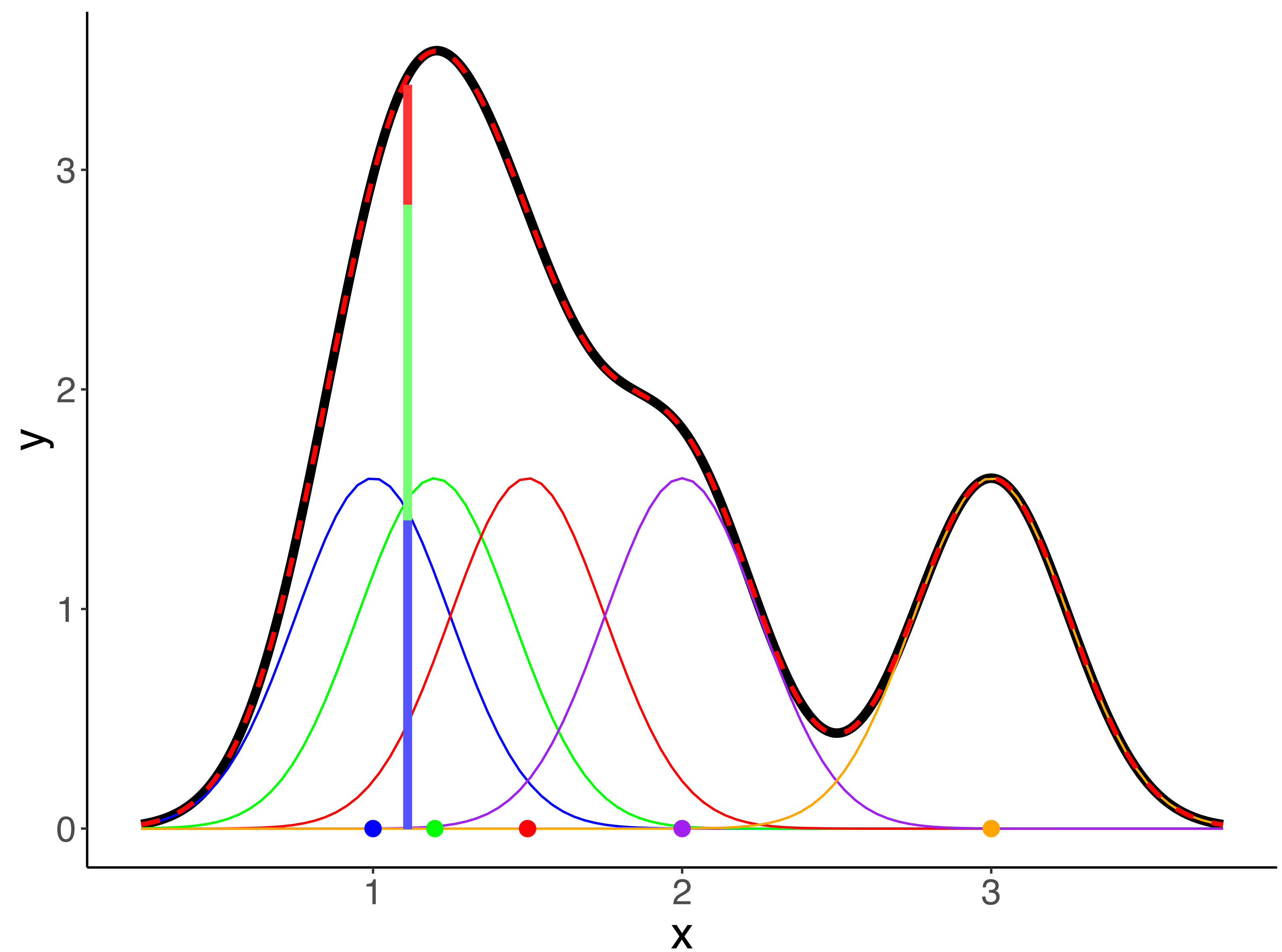
Understanding `density()`

x	y	observation_1	observation_2	observation_3	observation_4	observation_5	sum_norm
0.250	0.019	0.018	0.001	0	0	0	0.019
0.257	0.021	0.019	0.001	0	0	0	0.021
0.264	0.023	0.021	0.001	0	0	0	0.022
0.271	0.024	0.023	0.002	0	0	0	0.024
0.277	0.027	0.024	0.002	0	0	0	0.026
0.284	0.029	0.026	0.002	0	0	0	0.028

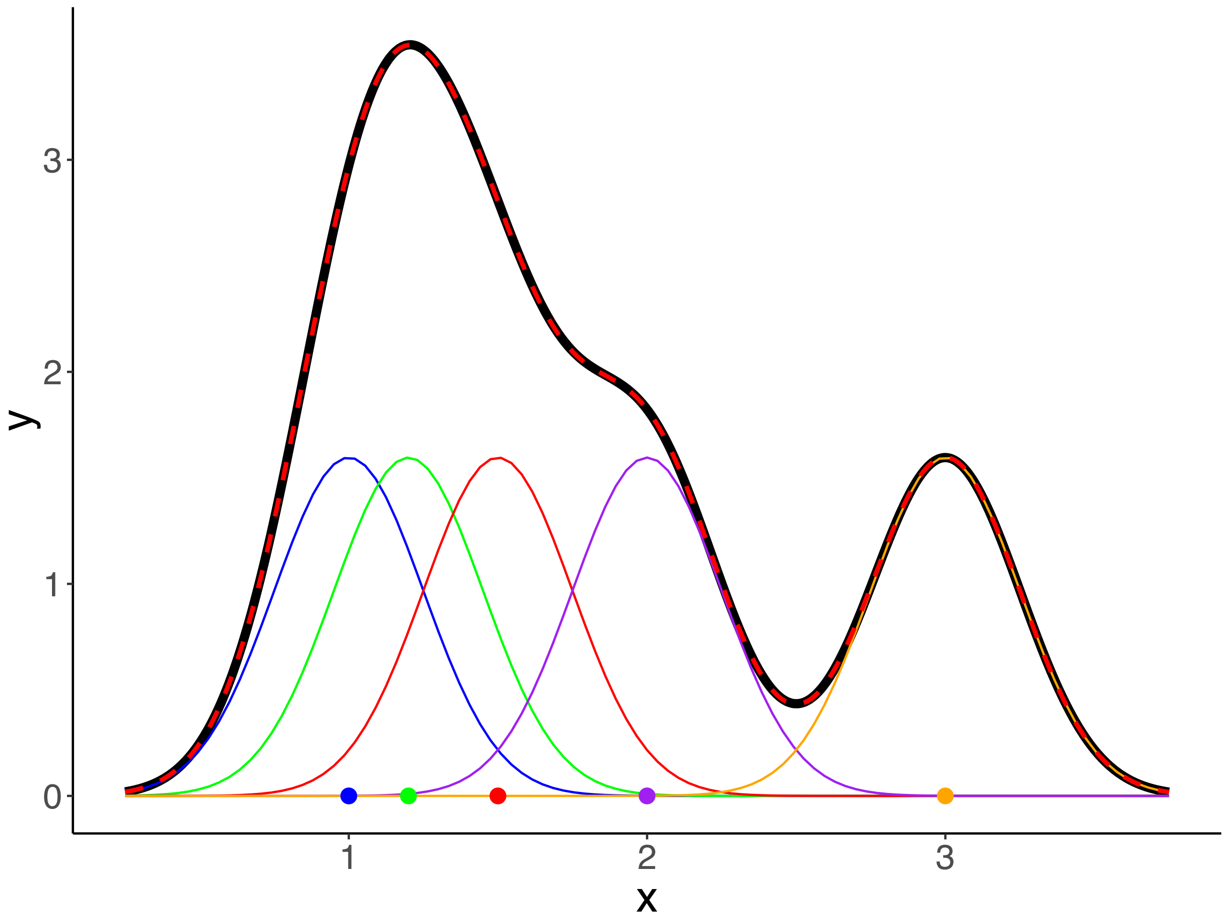


Understanding `density()`

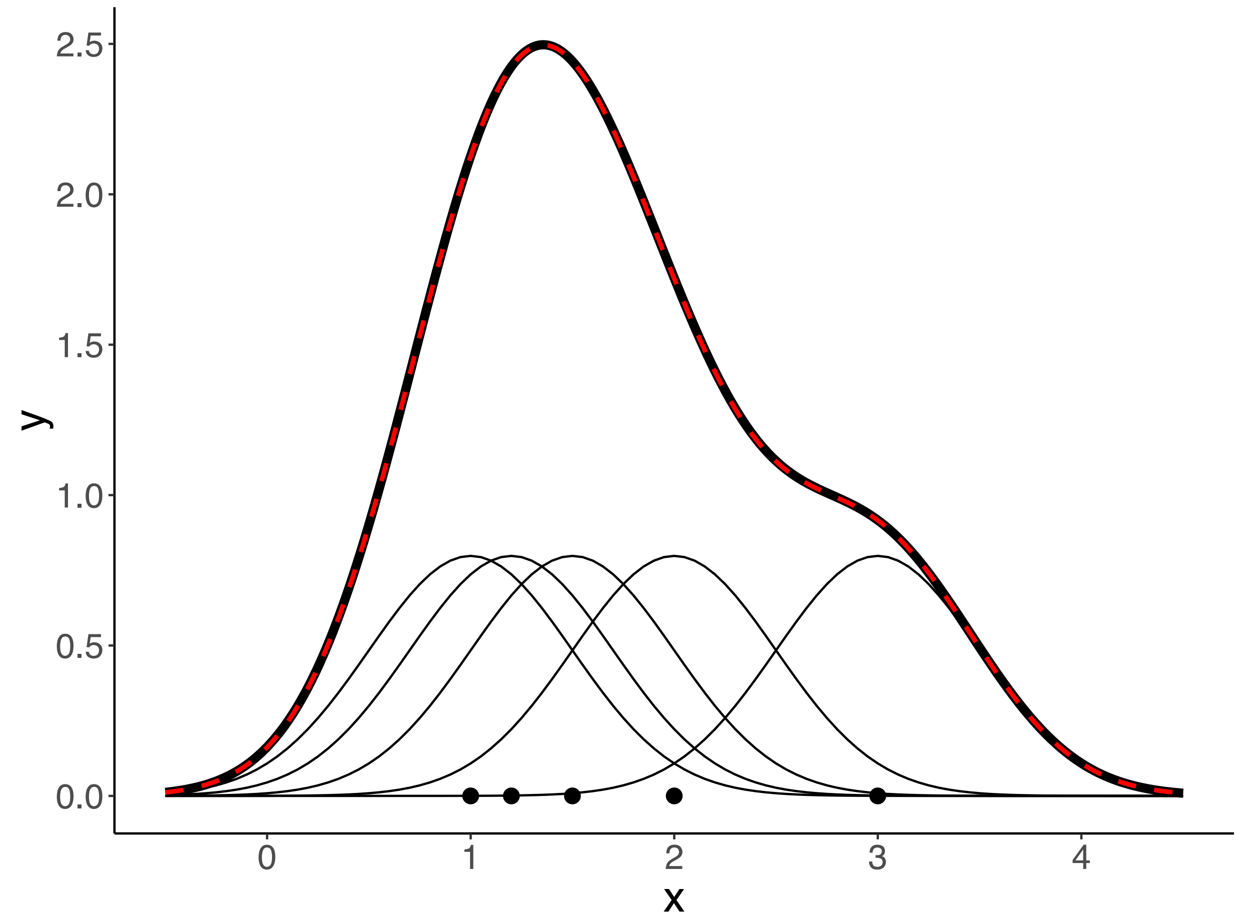
x	y	observation_1	observation_2	observation_3	observation_4	observation_5	sum_norm
0.250	0.019	0.018	0.001	0	0	0	0.019
0.257	0.021	0.019	0.001	0	0	0	0.021
0.264	0.023	0.021	0.001	0	0	0	0.022
0.271	0.024	0.023	0.002	0	0	0	0.024
0.277	0.027	0.024	0.002	0	0	0	0.026
0.284	0.029	0.026	0.002	0	0	0	0.028



Understanding `density()`



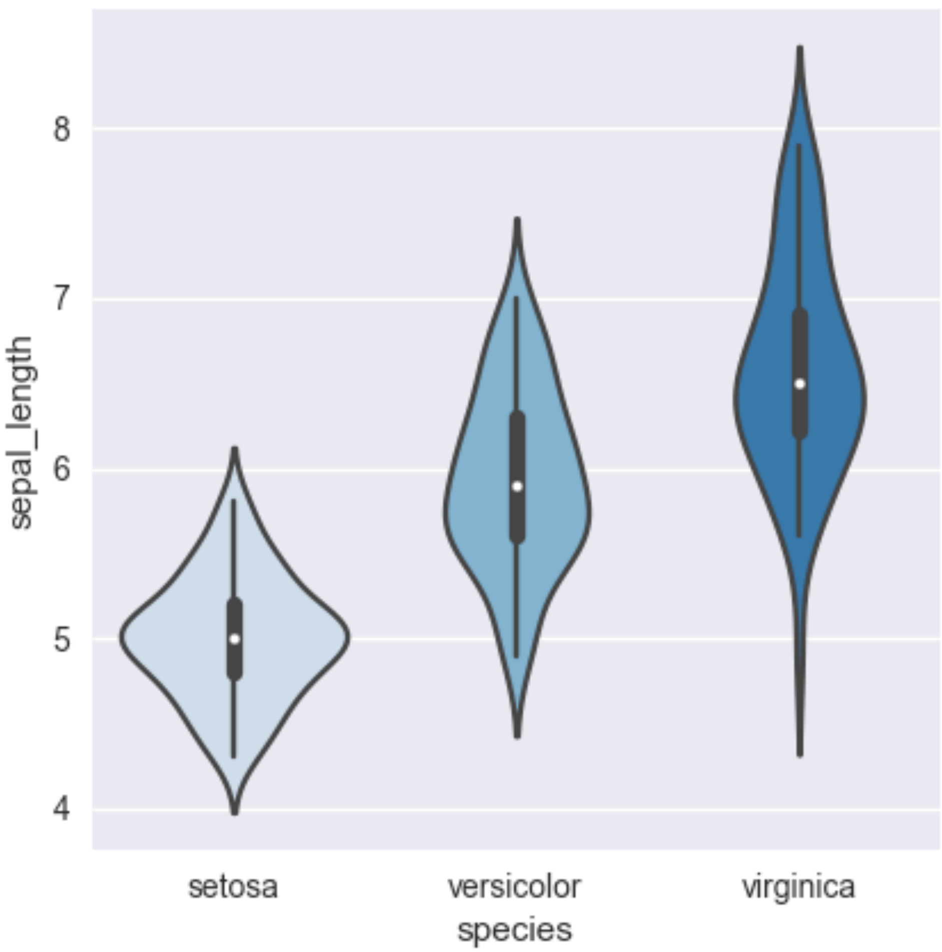
density(bw = 0.25)



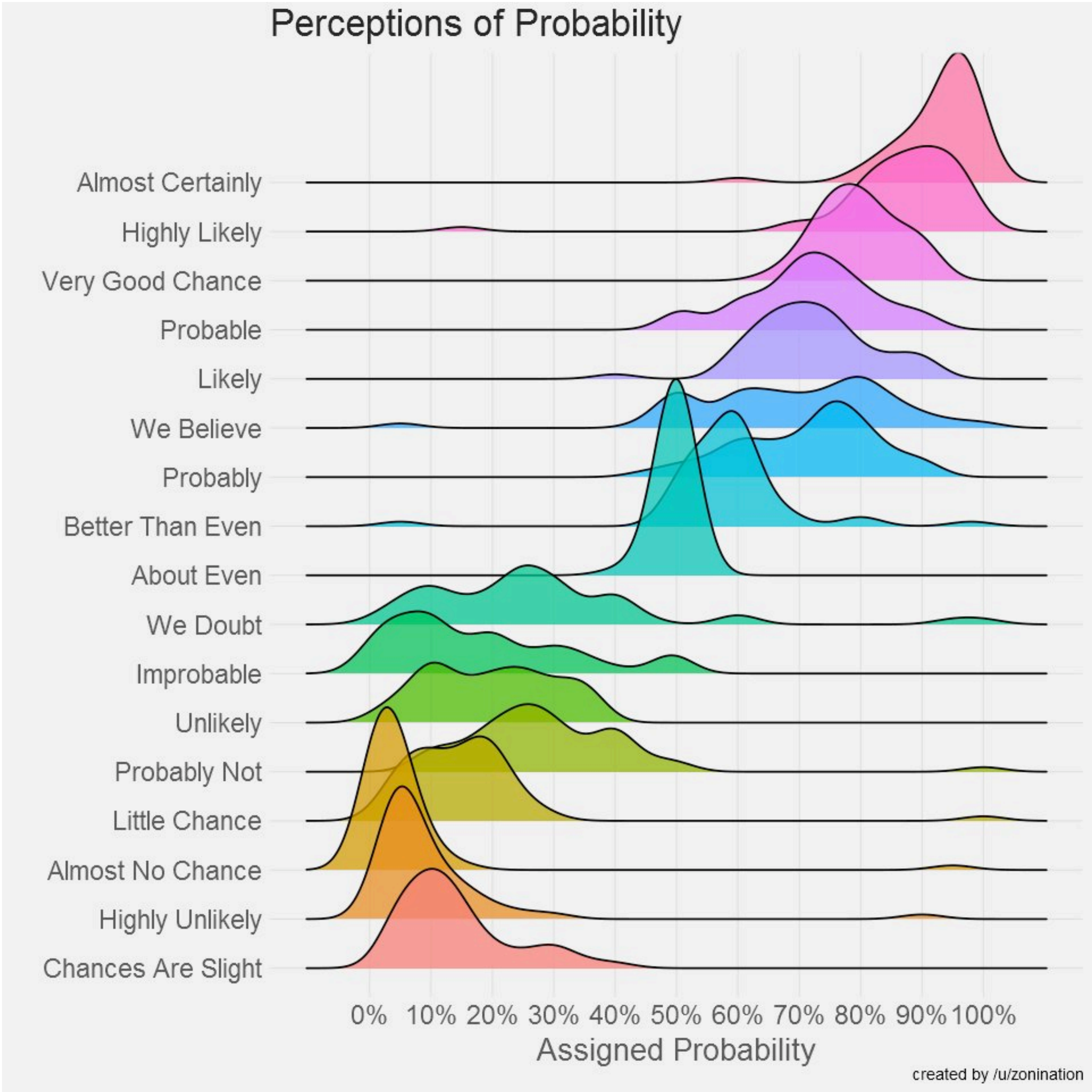
density(bw = 0.5)

Understanding `density()`

violinplot



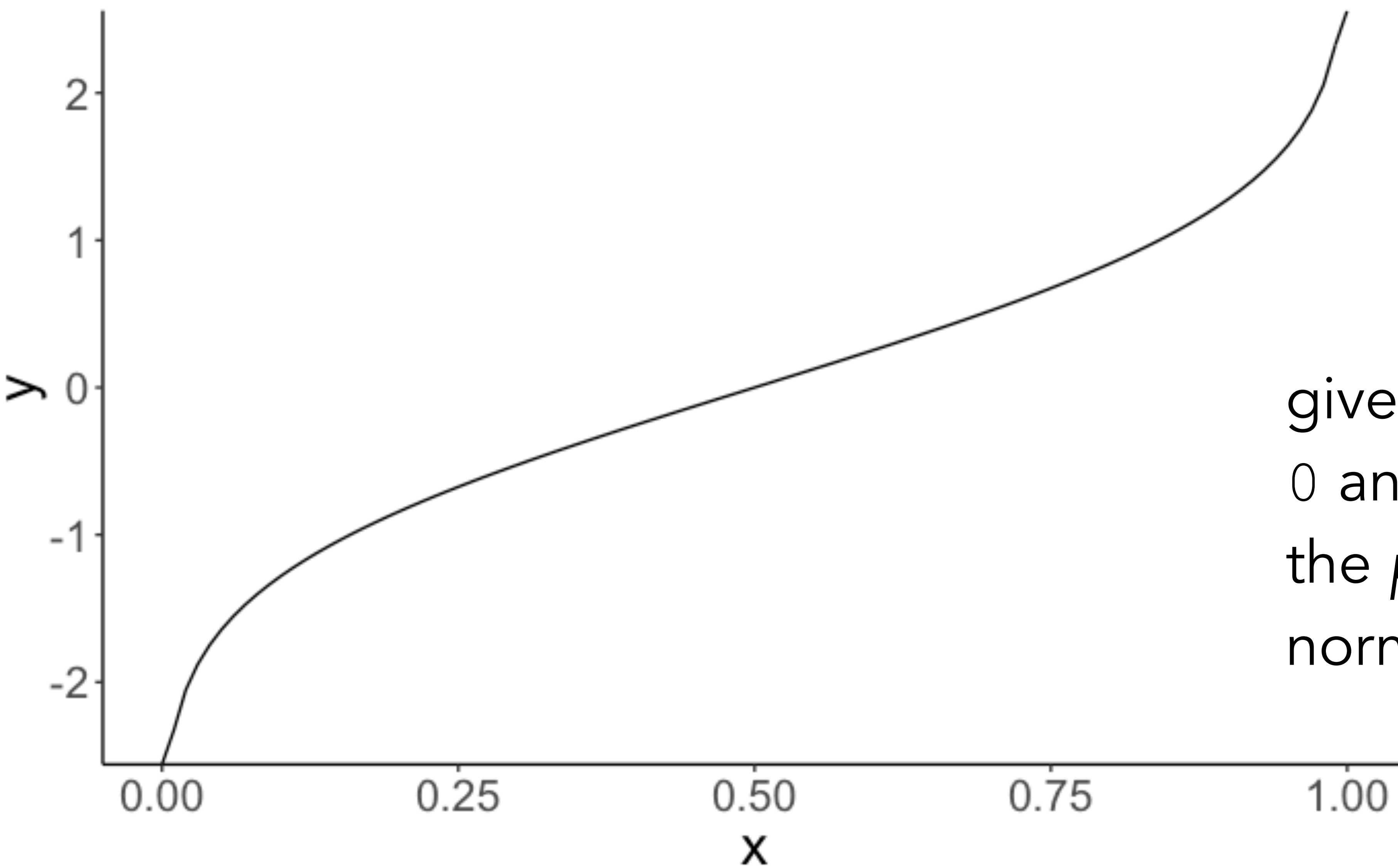
joyplot



Using **`quantile()`**

Using **`quantile()`**

- `quantile()` is for samples what `qnorm()` is for the normal distribution



given a number *p* between `0` and `1`, `qnorm` looks up the *p*-th quantile of the normal distribution

Using `quantile()`

- `quantile()` is for samples what `qnorm()` is for the normal distribution

```r
# a sample from the normal distribution
df.quantile = tibble(
  sample = 1:nsamples,
  value = rnorm(n = nsamples))
```

sample	value
1	-0.63
2	0.18
3	-0.84
4	1.60
5	0.33
6	-0.82
7	0.49
8	0.74
9	0.58
10	-0.3

median

1st quartile

3rd quartile

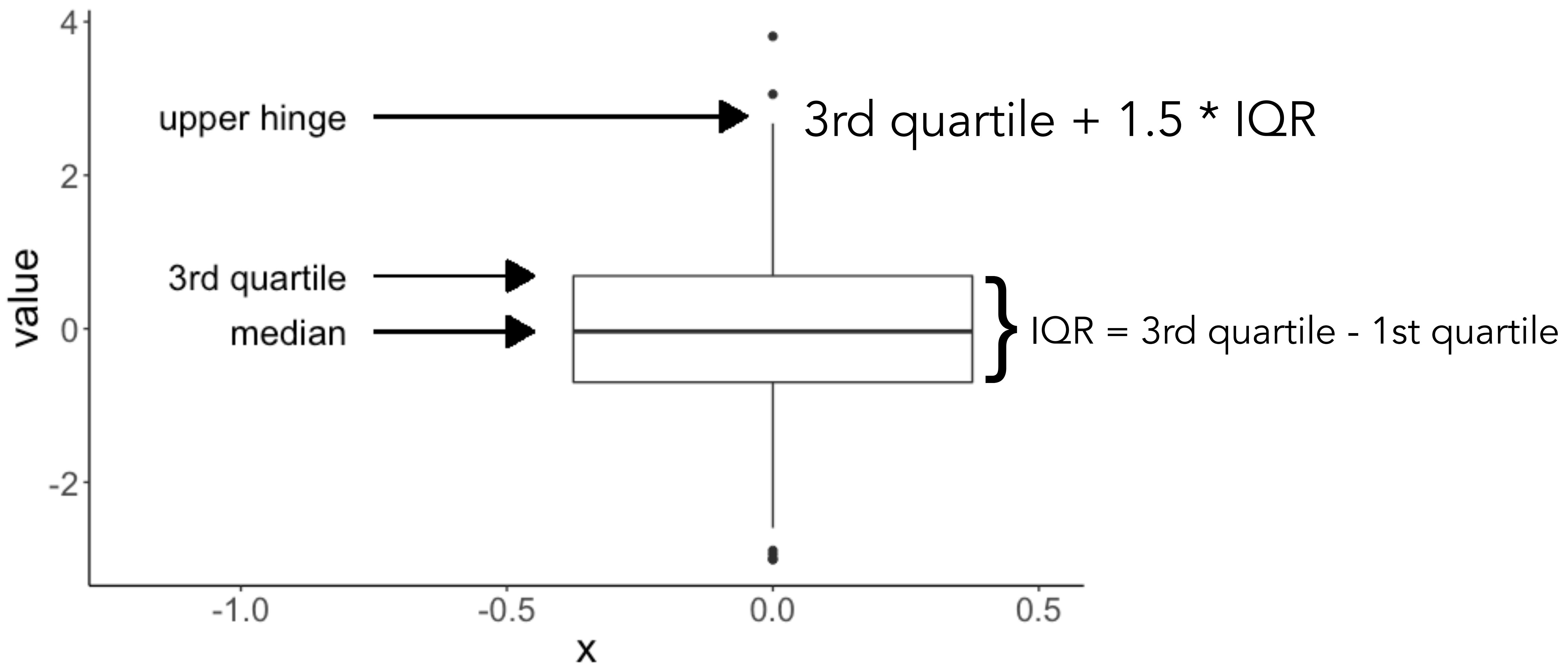
```r
quantile(df.quantile, probs = c(0.25, 0.5, 0.75))
```

Using `quantile()`

median

1st quartile 3rd quartile

```
quantile(df.quantile, probs = c(0.25, 0.5, 0.75))
```

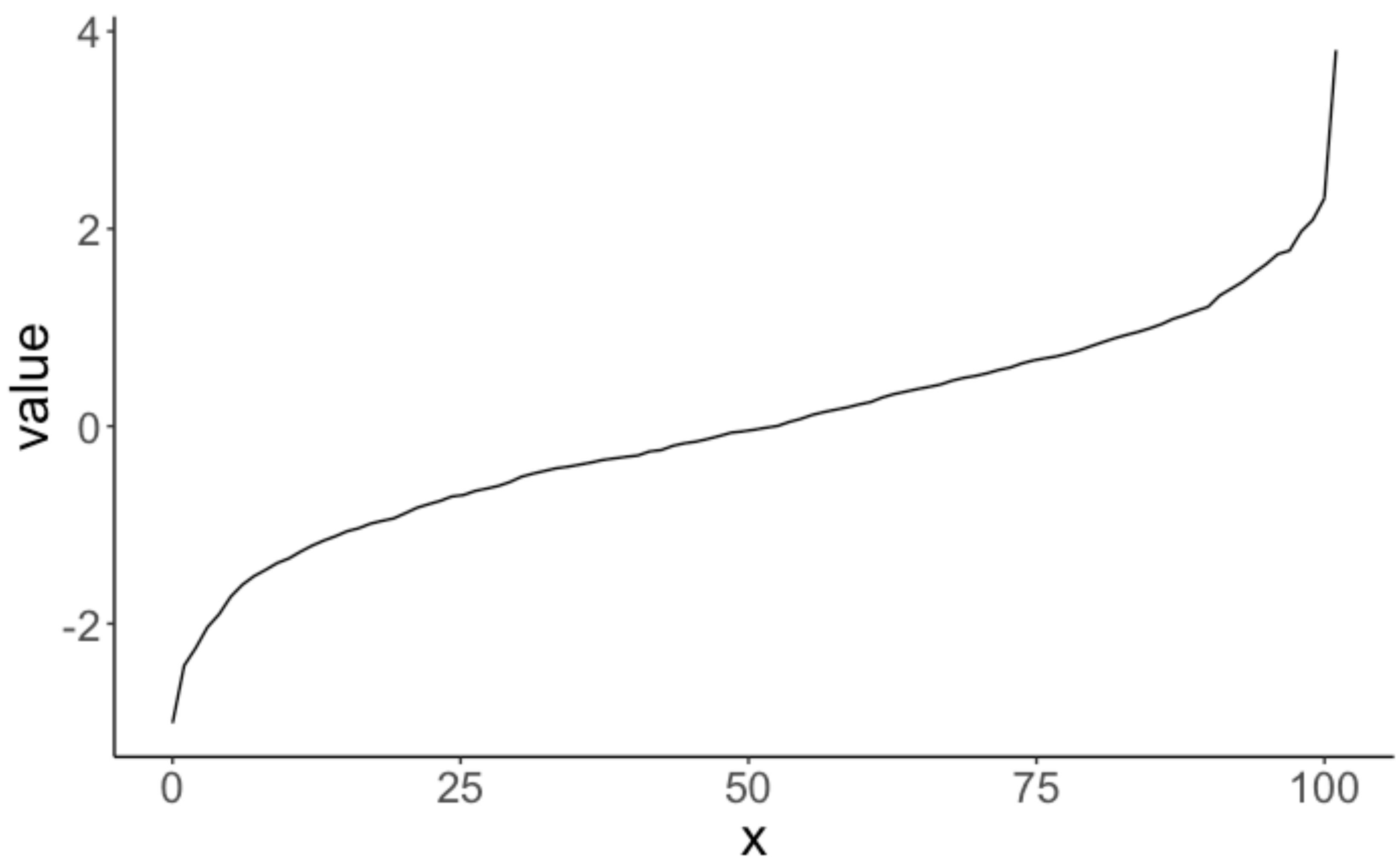


Using `quantile()`

```
df.plot = df.quantile$value %>%
  quantile(probs = seq(0, 1, 0.01)) %>%
  as_tibble() %>%
  mutate(x = seq(0, n(), length.out = n()))
```

value	x
-3.008	0.00
-2.424	1.01
-2.246	2.02
-2.036	3.03
-1.905	4.04
1.744	95.95
1.779	96.96
1.971	97.97
2.089	98.98
2.308	99.99
3.810	101.00

```
ggplot(data = df.plot,
       mapping = aes(x = x, y = value)) +
  geom_line()
```



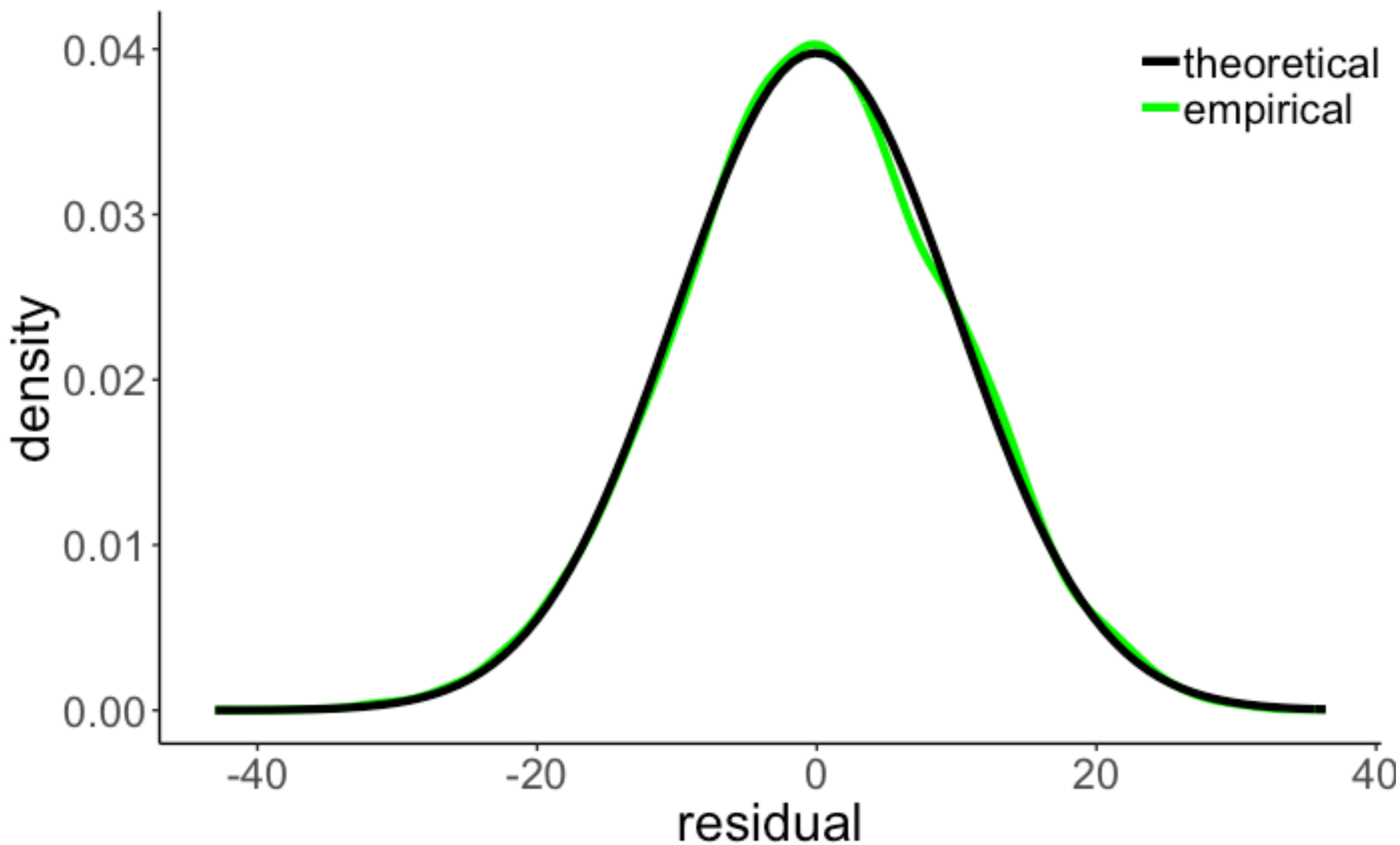
Working with samples

theoretical distribution		empirical sample
`dnorm()`	⟷	**`density()`**
`qnorm()`	⟷	**`quantile()`**

Comparing distributions

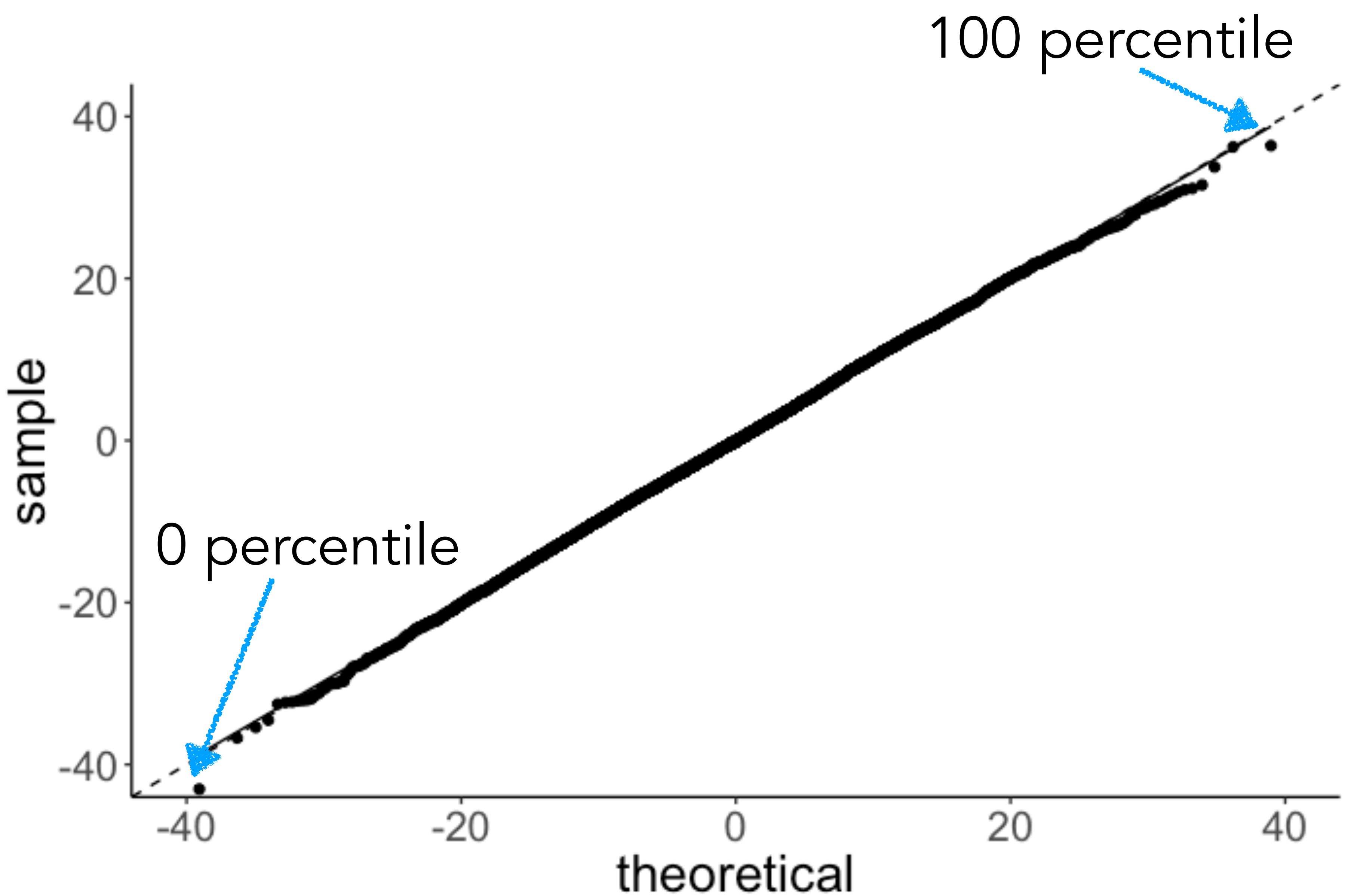
Comparing distributions

- **QQ plot** (Quantile-Quantile plot): useful for checking whether data (or errors) are normally distributed

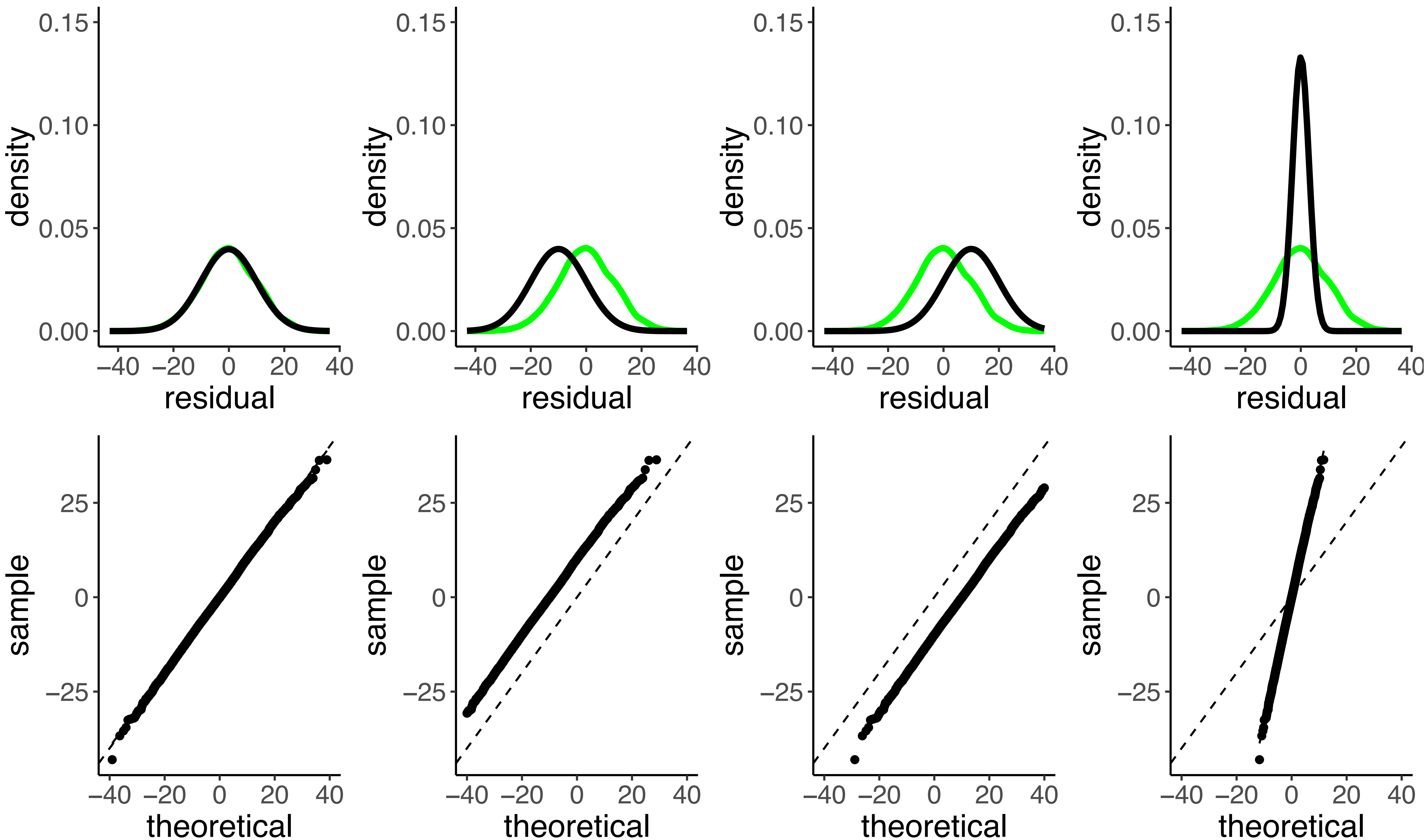


Comparing distributions

- **QQ plot** (Quantile-Quantile plot): useful for checking whether data (or errors) are normally distributed

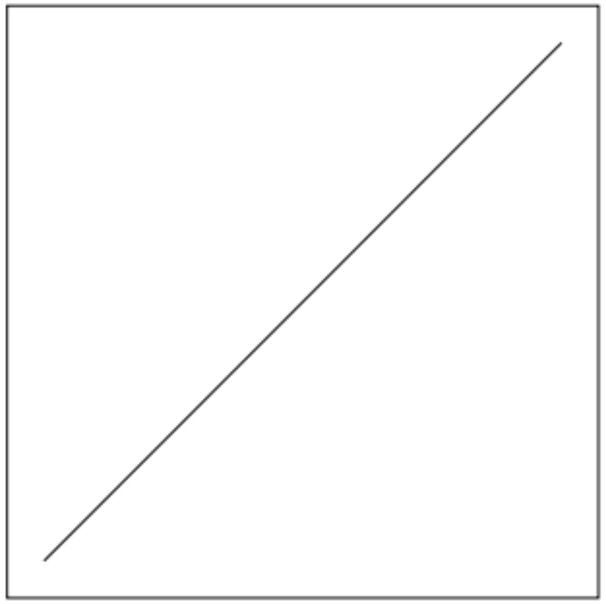


Comparing distributions



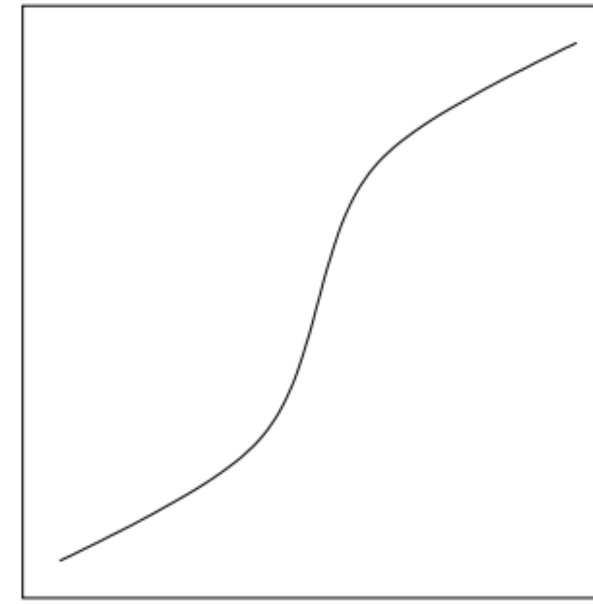
data is normally distributed if QQ plot is a line

Comparing distributions

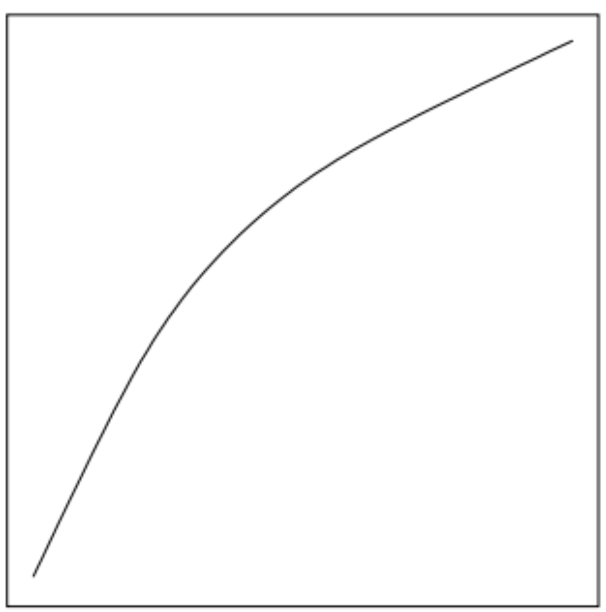


(a) Normally Distributed

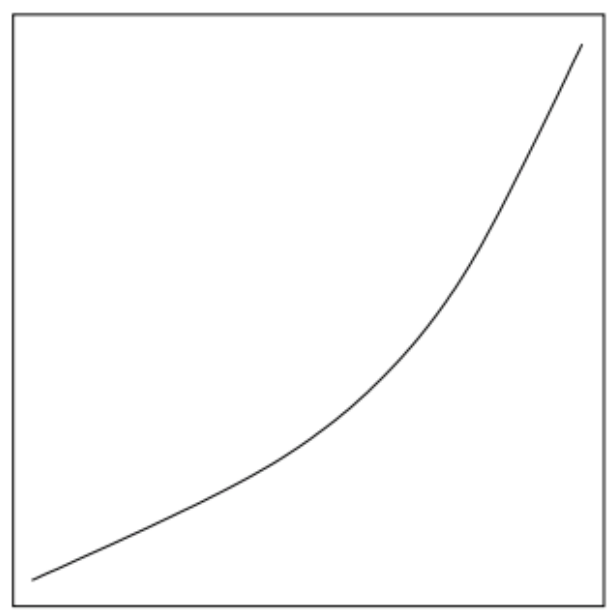
(b) Heavy Tails



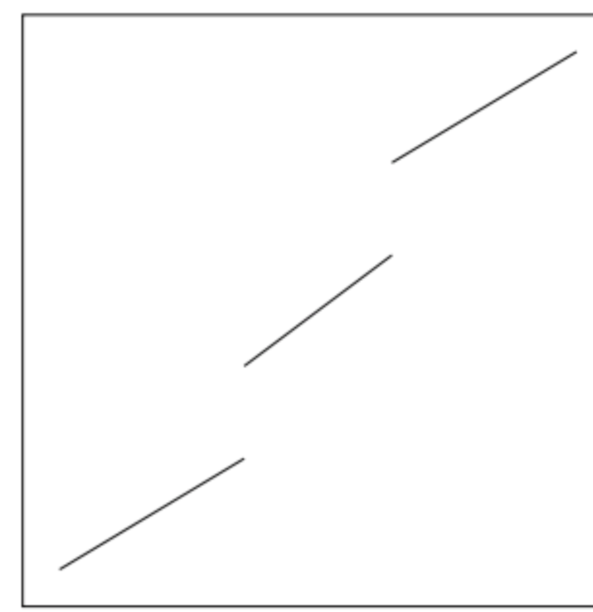
(c) Light Tails



(d) Skewed to the Left



(e) Skewed to the Right



(f) Separate Clusters

Plan for today

- **Quick review of causality**
- **Working with probability distributions**
  - `dnorm()`, `pnorm()`, `qnorm()`, `rnorm()`
  - computing probabilities
- **Bayesian inference**
  - analytic solution
  - via sampling
- **Working with samples**
  - Understanding `density()`
  - Understanding `quantile()`
  - Comparing distributions

Feedback

How was the pace of today's class?

Start the presentation to see live content. Still no live content? Install the app or get help at **PollEv.com/app**

How happy were you with today's class overall?

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

Start the presentation to see live content. Still no live content? Install the app or get help at **PollEv.com/app**

Thank you!