Session 4: Probability (part 1)

Stats 60/Psych 10
Ismael Lemhadri
Summer 2020

What is a “probability”?

- Informal:
  - A number between zero and one that denotes how likely some event is to occur
    - close to zero: not very likely
    - close to one: pretty likely
- More formally
  - First, some definitions…

Sample space

- The collection of all possible basic outcomes of an experiment
- Coin flip:
  - {H,T}
- Roll of six-sided die
  - {1,2,3,4,5,6}
- Count of how many Facebook friends a person has
  - {0,1,2,3,…}

Events

- An event is a subset of the sample space
- We will focus on “elementary events”
  - exactly one possible outcome
- Examples
  - coin flip == H
  - die roll == 4
  - # of FB friends == 324

The algebra of sets

- An event is just a set of things
  - The "algebra of sets" tells us what we can do with sets
- The "null set" is empty: {}
- The "union" of two sets (∪) includes any element that appears in either
  - $\{A,B,C\} \cup \{B,C,D\} = \{A,B,C,D\}$
  - like a logical "or"
- The "intersection" of two sets (∩) includes only elements that appear in both sets
  - $\{A,B,C\} \cap \{B,C,D\} = \{B,C\}$
  - like a logical "and"

Working with sets in R

```
setA <- c(1,2,3)
setB <- c(2,3,4,5)

union(setA,setB)
## [1] 1 2 3 4 5

intersect(setA,setB)
## [1] 2 3
```

What is a probability?

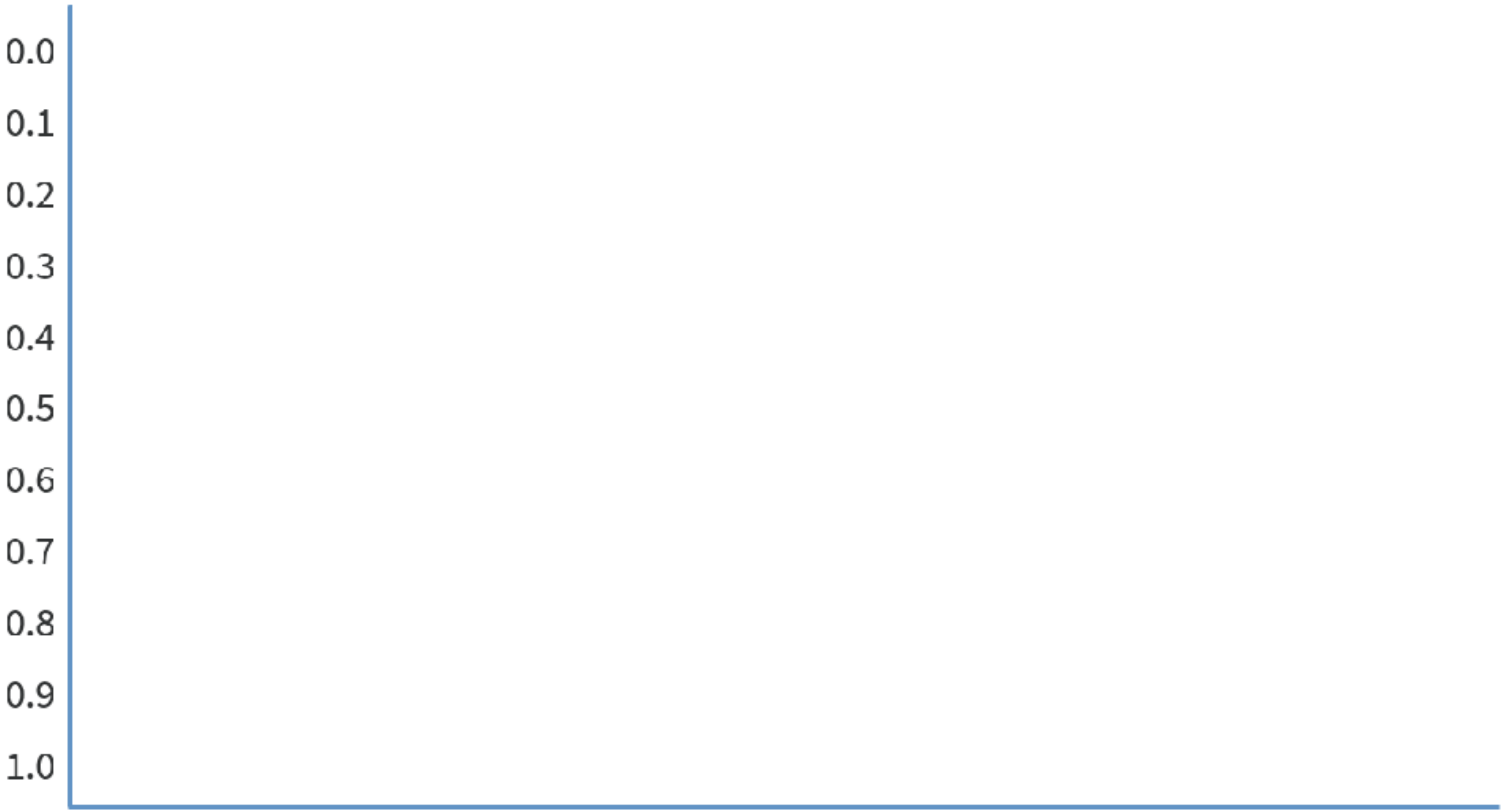
- A probability of an outcome X_i — denoted $P(X_i)$ — must have particular characteristics (known as *axioms*)
- Probability cannot be negative
  - $P(X_i) \geq 0$
- The total probability of all outcomes in the sample space is 1
  - $P(X_0) + P(X_1) + \ldots P(X_N) = 1$
  - this means that $P(X_i) \leq 1$

Andrei Kolmogrov

How do we obtain the probability of an event?

- Personal opinion
  - Sometimes this is the only way!

What do you think is the probability that Donald Trump would have beaten Bernie Sanders if Sanders had been the Democratic nominee in 2016?



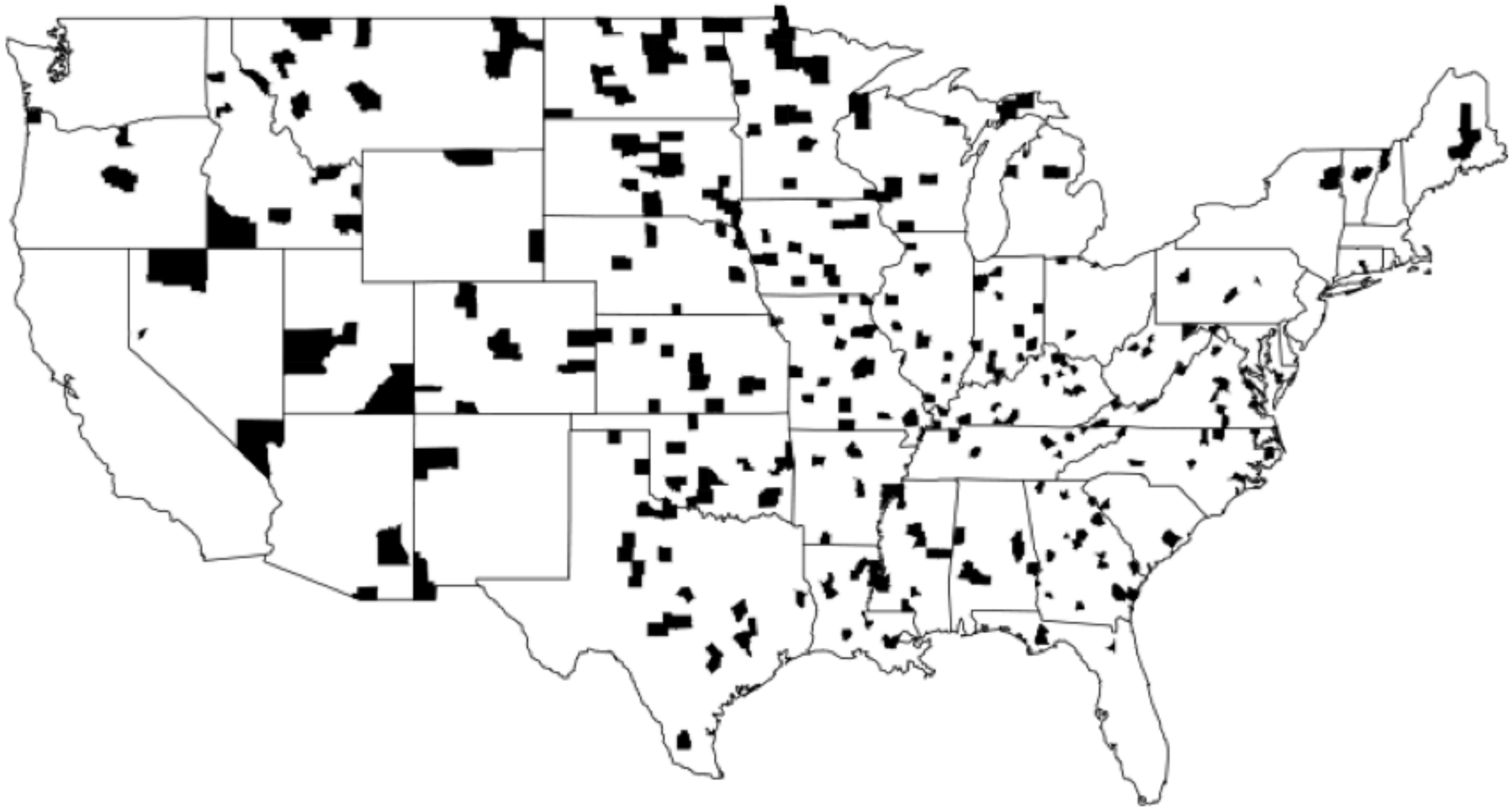
How do we obtain the probability of an event?

- Empirical probability (aka relative frequency)
  - What is the probability of rain in San Francisco?
    - Obtain National Weather Service data from downtown SF weather station for each day in 2017 (from https://www.ncdc.noaa.gov/)

nRainyDays	nDaysMeasured
73	365

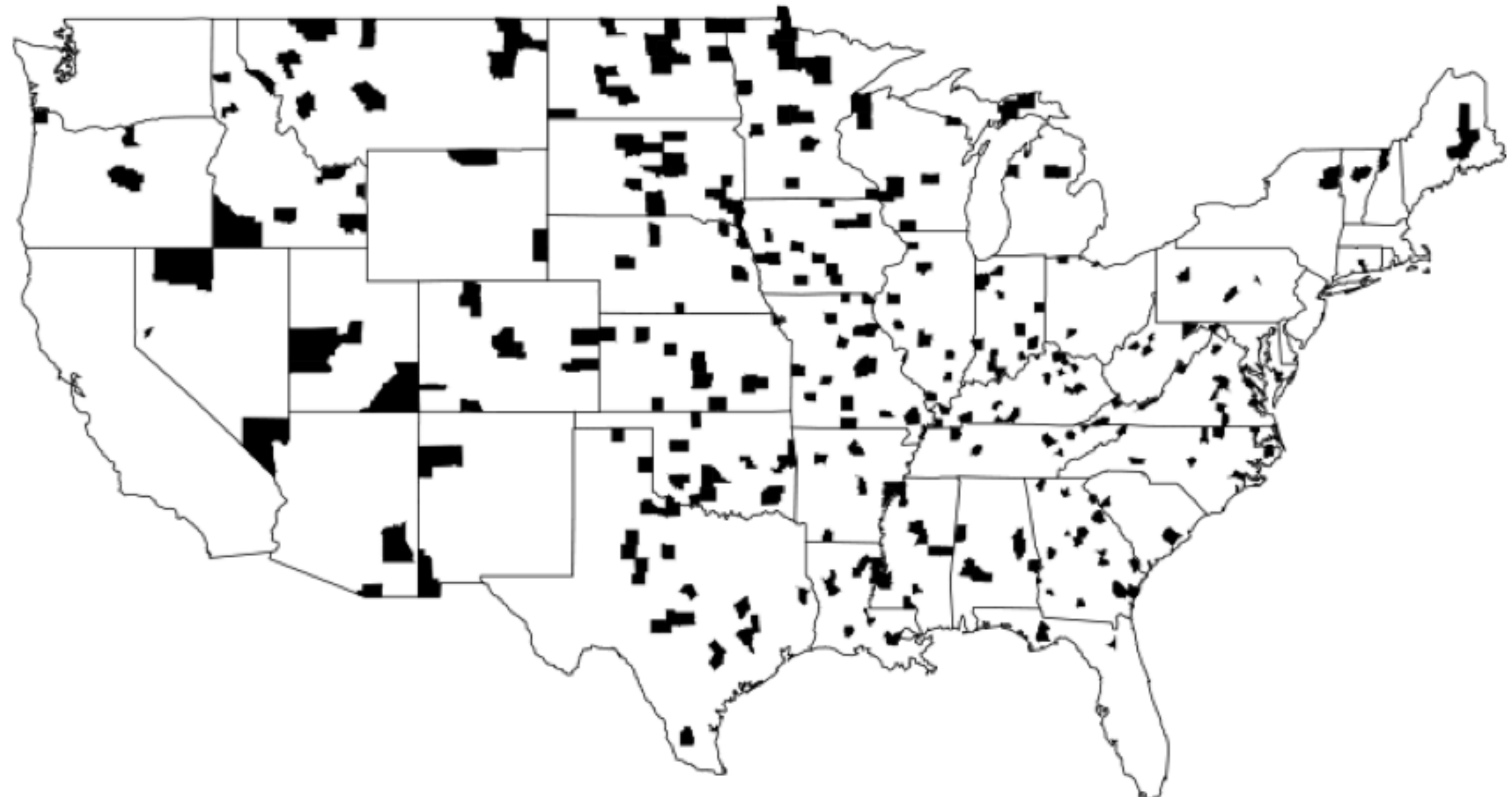
$$P(\text{rain in SF}) = \frac{\text{number of rainy days}}{\text{number of days measured}} = \frac{73}{365} = 0.2$$

Counties with highest kidney cancer rates

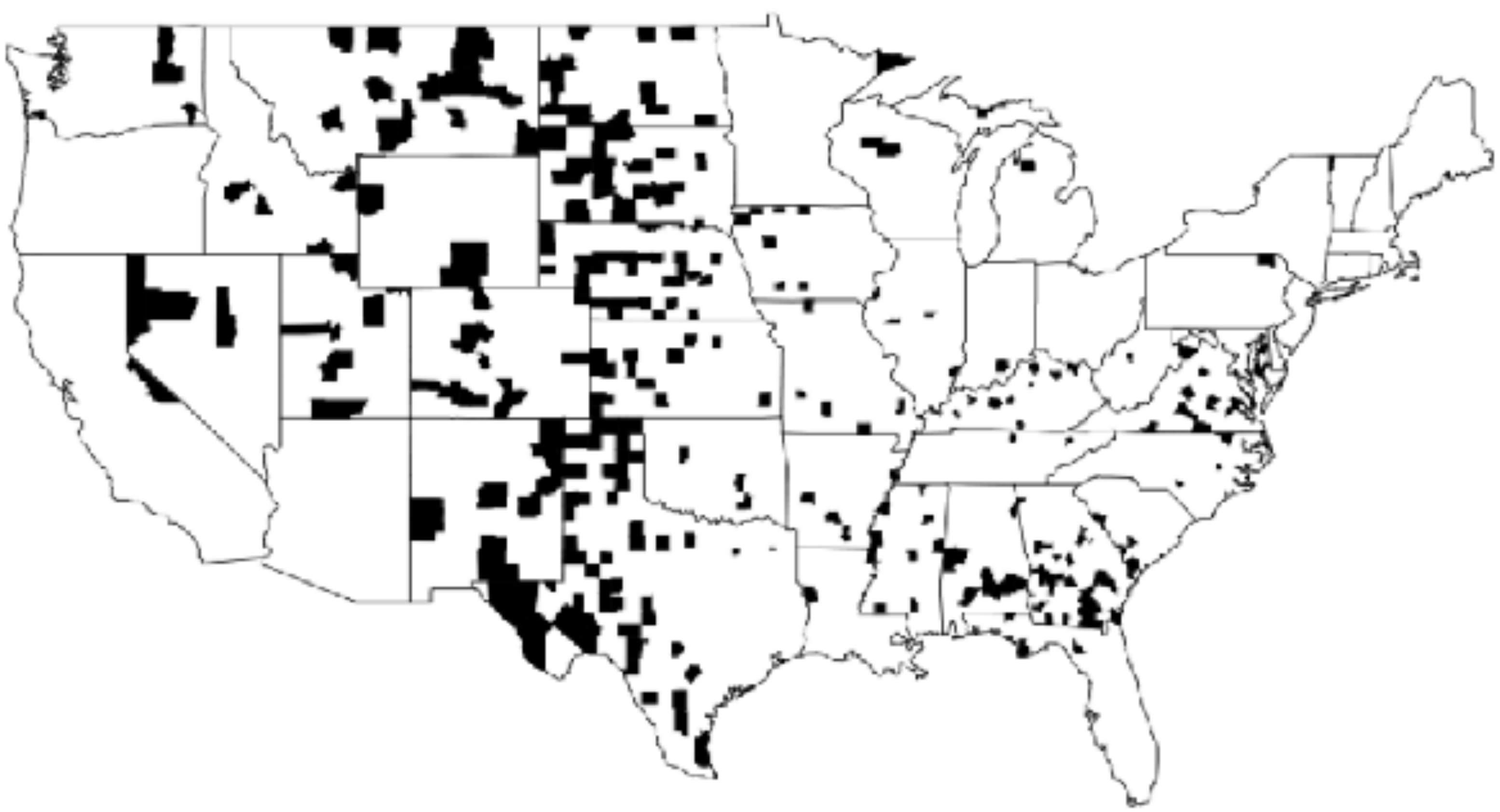


- What do you notice?
- What do you think might be causing this?

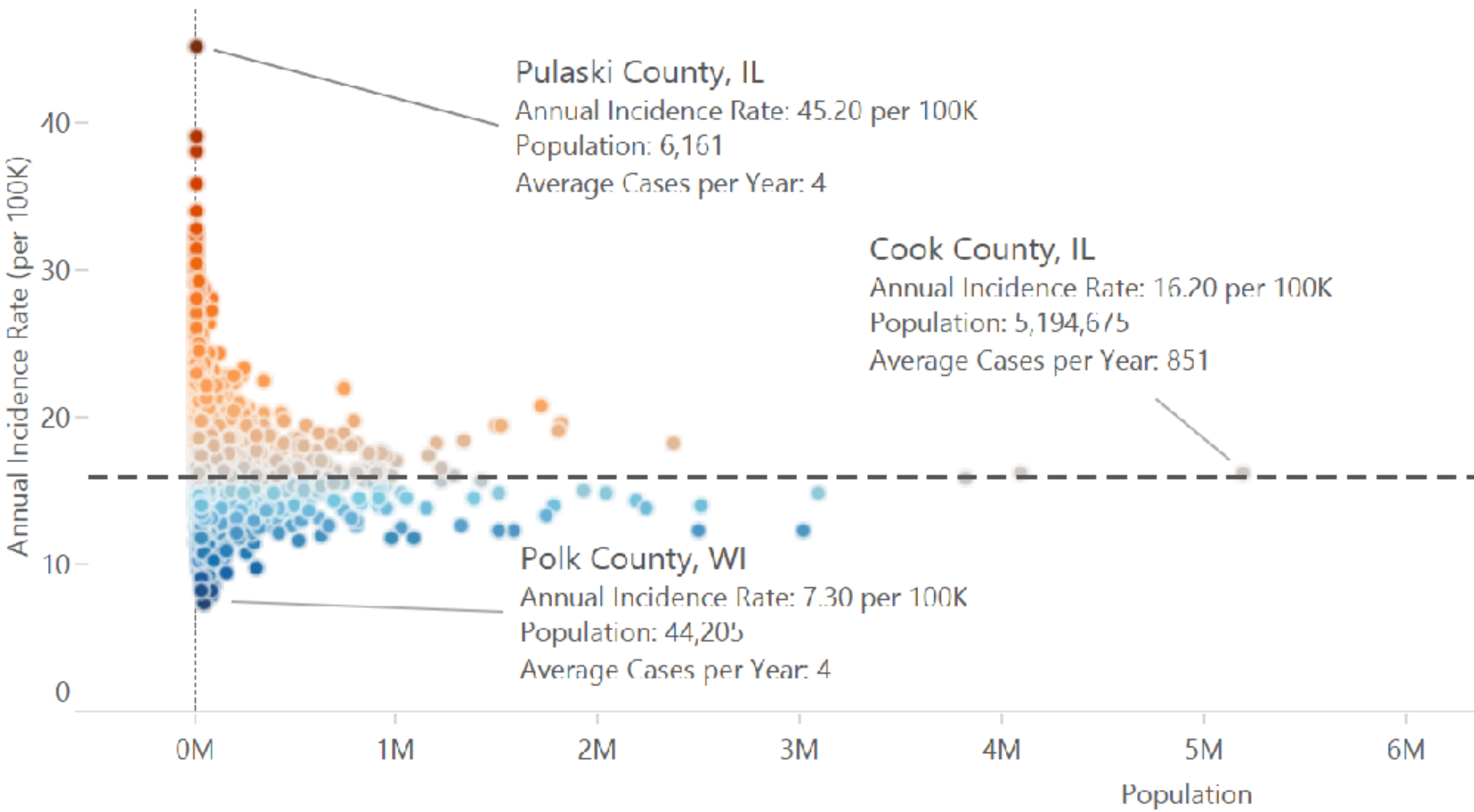
Counties with *highest* kidney cancer rates



Counties with *lowest* kidney cancer rates



from Andrew Gelman



http://dataremixed.com/2015/01/avoiding-data-pitfalls-part-2/

A real life example: Dec 12, 2017

Doug Jones   Roy Moore

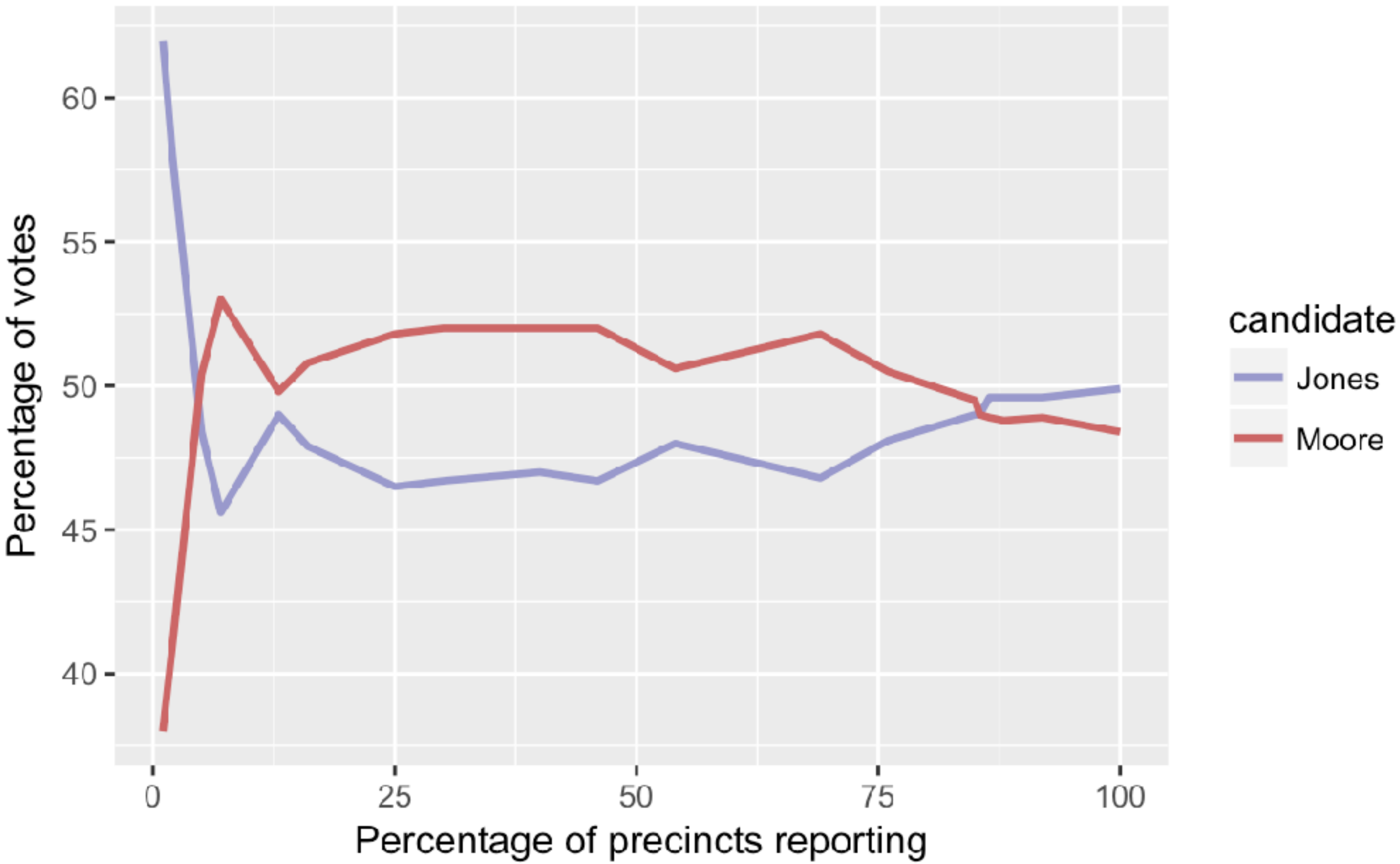
The New York Times

POLITICS

Once a Long Shot, Democrat Doug Jones Wins Alabama Senate Race

By ALEXANDER BURNS and JONATHAN MARTIN DEC. 12, 2017

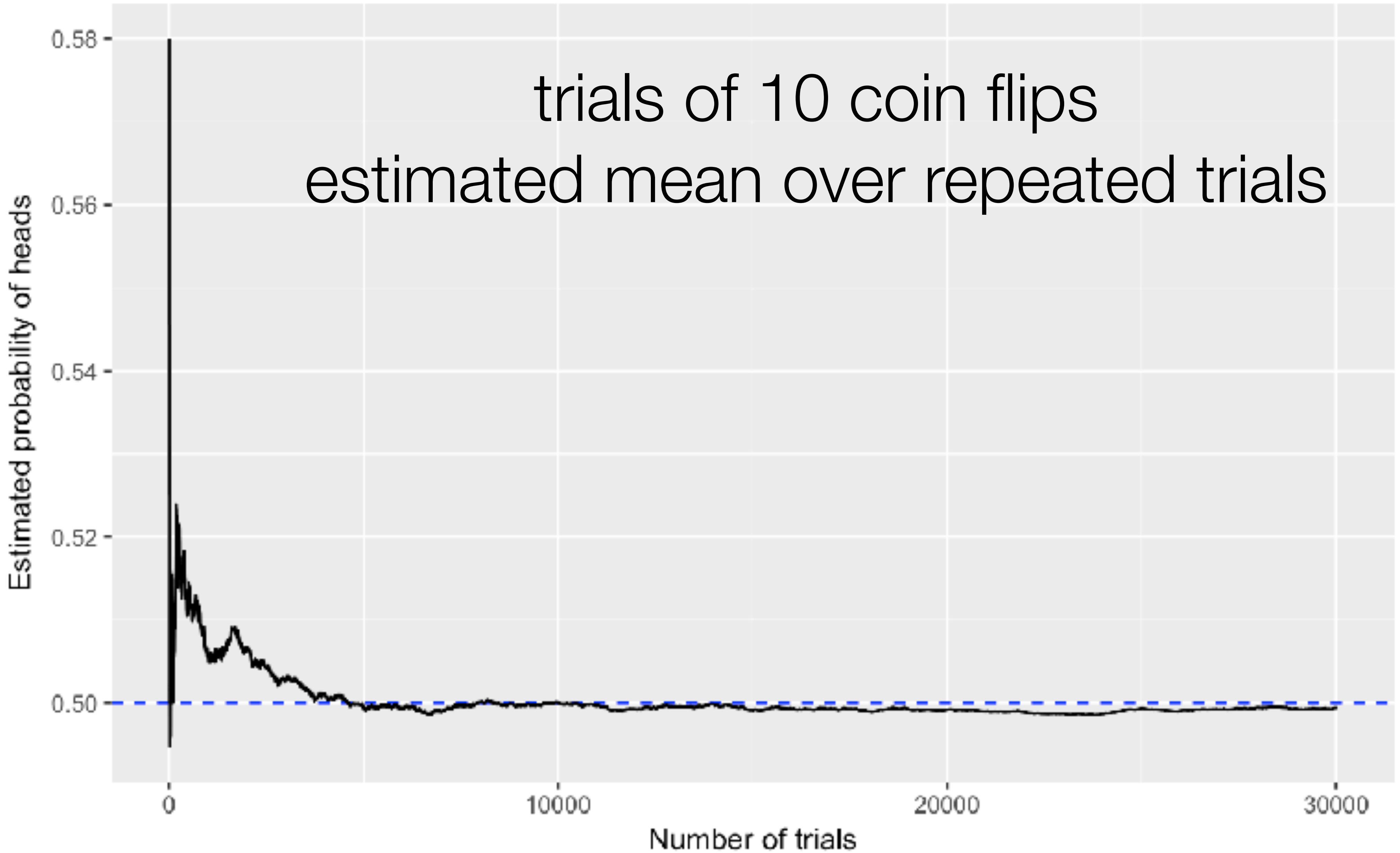
1997



Percentage of votes
60
55
50
45
40
0
25
50
75
100
Percentage of precincts reporting
candidate
Jones
Moore

The law of large numbers

- The average empirical probability converges on the true expected value as the sample size increases



The “law of small numbers”

Psychological Bulletin
1971, Vol. 76, No. 2, 105–110

BELIEF IN THE LAW OF SMALL NUMBERS

AMOS TVERSKY AND DANIEL KAHNEMAN [1]

Hebrew University of Jerusalem

People have erroneous intuitions about the laws of chance. In particular, they regard a sample randomly drawn from a population as highly representative, that is, similar to the population in all essential characteristics. The prevalence of the belief and its unfortunate consequences for psychological research are illustrated by the responses of professional psychologists to a questionnaire concerning research decisions.

People underestimate the variability in statistical estimates

How do we obtain the probability of an event?

- Classical probability
  - Arose initially from study of gambling

Chevalier de Méré
1654

I keep losing money at dice - can you help me Blaise?

Blaise Pascal

de Méré's dilemma

From Gonick, The Cartoon Guide to Statistics

Classical probability

- Assume that all elementary events in the sample space are equally likely

$$P(outcome_i) = \frac{1}{\text{number of possible outcomes}}$$

Sample space for a single six-sided die: {1,2,3,4,5,6}

P(1) = P(2) … = 1/6 = 0.17

Probability for complex events

- To compute the probability of a complex event, add together the probabilities of the elementary events

Example: roll two six-sided dice

Sample space: {11,12,13,14,15,21,22,23…}

$$P(11) = P(12) = P(13) = P(14) \ldots = 1/36$$

$$P(11 \cup 12) = P(11) + P(12) = 2/36$$

Graphical view

$$P(11 \cup 12) = P(11) + P(12) = 2/36$$

Die 2

Die 1						
	11	12	13	14	15	16
	21	22	23	24	25	26
	31	32	33	34	35	36
	41	42	43	44	45	46
	51	52	53	54	55	56
	61	62	63	64	65	66

Let's say that we draw two cards from a 52 card deck (no jokers), and we record the suit (D=diamond,H=heart,C=club,S=spade} of each card. What is the sample space for these events?

{D,H,C,S}

{DD,DH,DC,DS,HD,HH,HC,HS,CD,CH,CC,CS,SS,SC,SH,SD}

{DD,DH,DC,DS,HH,HC,HS,CC,CS,SS}

What is the probability of drawing a club on each of the two draws?

What is the probability of drawing a pair (SS,CC,HH,or DD)?

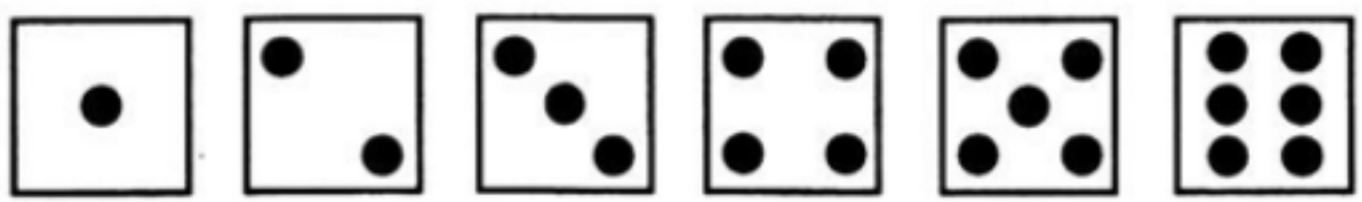
de Méré's reasoning

Chance of six on a roll of one die	$\frac{1}{6}$
Chance of at least one six on four rolls	$4 * \frac{1}{6} = \frac{4}{6} = \frac{2}{3}$
Chance of a double-six on a roll of two die	$\frac{1}{36}$
Chance of at least one double-six on 24 rolls of two dice	$24 * \frac{1}{36} = \frac{24}{36} = \frac{2}{3}$

If this was true, then why did he win money on the first bet and lose money on the second bet?

Rules of probability

- How do we work with probabilities?
- Rule of subtraction:
  - $P(\bar{A}) = 1 - P(A)$



$$P(1 \cup 2) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$$

$$P(\overline{1 \cup 2}) = 1 - P(1 \cup 2) = \frac{2}{3}$$

Rules of probability

- Rule of addition
  - subtract the intersection to avoid double-counting

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

P(1X)=6/36

P(X1)=6/36

P(11)=1/36

Die 2

Die 1

11	12	13	14	15	16
21	22	23	24	25	26
31	32	33	34	35	36
41	42	43	44	45	46
51	52	53	54	55	56
61	62	63	64	65	66

Die 2

Die 1

11	12	13	14	15	16
21	22	23	24	25	26
31	32	33	34	35	36
41	42	43	44	45	46
51	52	53	54	55	56
61	62	63	64	65	66

Die 2

Die 1

11	12	13	14	15	16
21	22	23	24	25	26
31	32	33	34	35	36
41	42	43	44	45	46
51	52	53	54	55	56
61	62	63	64	65	66

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

P(1 on either roll)=P(1X) + P(X1) - P(11) = 11/36

Rules of probability

- Special rule of multiplication for independent events
  - We will define “independent” in the next lecture - just assume that dice rolls are independent for now

$$P(A \cap B) = P(A) * P(B) \text{ iff A and B are independent}$$

Example: two rolls of a single die

P(11)=P(1) * P(1) = 1/36

P(1X)=6/36 P(X1)=6/36 P(11)=1/36

Die 2 (columns), Die 1 (rows)

11	12	13	14	15	16
21	22	23	24	25	26
31	32	33	34	35	36
41	42	43	44	45	46
51	52	53	54	55	56
61	62	63	64	65	66

Die 2 (columns), Die 1 (rows)

11	12	13	14	15	16
21	22	23	24	25	26
31	32	33	34	35	36
41	42	43	44	45	46
51	52	53	54	55	56
61	62	63	64	65	66

Die 2 (columns), Die 1 (rows)

11	12	13	14	15	16
21	22	23	24	25	26
31	32	33	34	35	36
41	42	43	44	45	46
51	52	53	54	55	56
61	62	63	64	65	66

$$P(A \cap B) = P(A) * P(B)$$

P(11)=P(1X) * P(X1) = 1/36

(assuming outcomes on Die 1 and 2 are independent)

Question time (~5 mins)

- Please break into groups of 3 and come up with a question about a point so far in the lecture that is unclear
- Be ready to present your question if called upon

de Méré's error

Chance of six on a roll of one die	$\frac{1}{6}$
Chance of at least one six on four rolls	$4 * \frac{1}{6} = \frac{4}{6} = \frac{2}{3}$
Chance of a double-six on a roll of two die	$\frac{1}{36}$
Chance of at least one double-six on 24 rolls of two dice	$24 * \frac{1}{36} = \frac{24}{36} = \frac{2}{3}$

needed to multiply rather than add probabilities

$$P(A \cap B) =$$

$$P(A) * P(B)$$

Blaise Pascal

Pascal's solution to de Méré's first problem: Chance of at least one six on four rolls

$$P(\text{no six in four rolls}) = \frac{5}{6} \times \frac{5}{6} \times \frac{5}{6} \times \frac{5}{6} = \left(\frac{5}{6}\right)^4 = 0.482253$$

$$\begin{aligned} P(\text{at least one six in four rolls}) &= 1 - P(\text{no six in four rolls}) \\ &= 1 - 0.482253 \\ &= 0.517747 \end{aligned}$$

Blaise Pascal

Pascal's solution to de Méré's second problem:
Chance of at least one double-six on 24 rolls of two dice

$$P(\text{no double six in 24 rolls}) = \left(\frac{35}{36}\right)^{24} = 0.5086$$

$$\begin{aligned} P(\text{at least one double six in 24 rolls}) &= 1 - P(\text{no double six in 24 rolls}) \\ &= 1 - 0.5086 \\ &= 0.4914 \end{aligned}$$

Another example: The Birthday Problem

$$P(birthday_A = birthday_B) = \frac{1}{365}$$

$$P(birthday_A \neq birthday_B) = 1 - \frac{1}{365}$$

$$P(\text{no matches for } birthday_A \text{ in n people}) = (1 - \frac{1}{365})^{n-1}$$

$$\text{expected number of people with no match} = n * (1 - \frac{1}{365})^{n-1}$$

$$\text{expected number of people with a match} = n - n * (1 - \frac{1}{365})^{n-1}$$

Probability distributions

- A probability distribution describes the probability of any score occurring
- Example: binomial distribution
  - Probability of k successes out of n trials, when probability of success on a single trial is p
  - in R: `dbinom(k,n,p)`

$$P(k;n,p) = P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$$

Example

$$P(k; n, p) = P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

- On Jan 20 2018, Steph Curry hit only 2 out of 4 free throws vs. Houston
- How likely is this, given that his overall percentage is 91%?
  - k=2
  - n=4
  - p=0.91

$$P(k;n,p) = P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$$

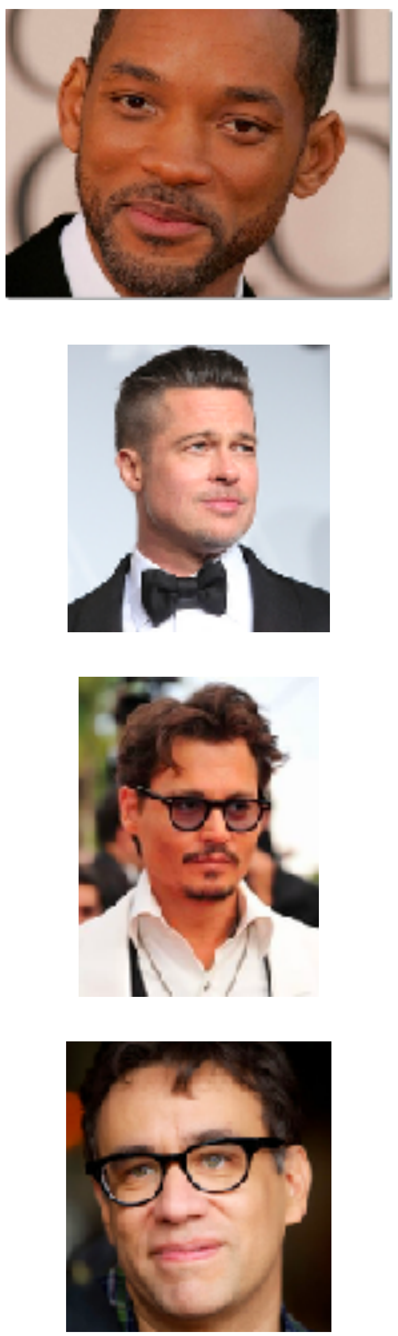
"n choose k" - how many ways are there to choose k items out of n possibilities?

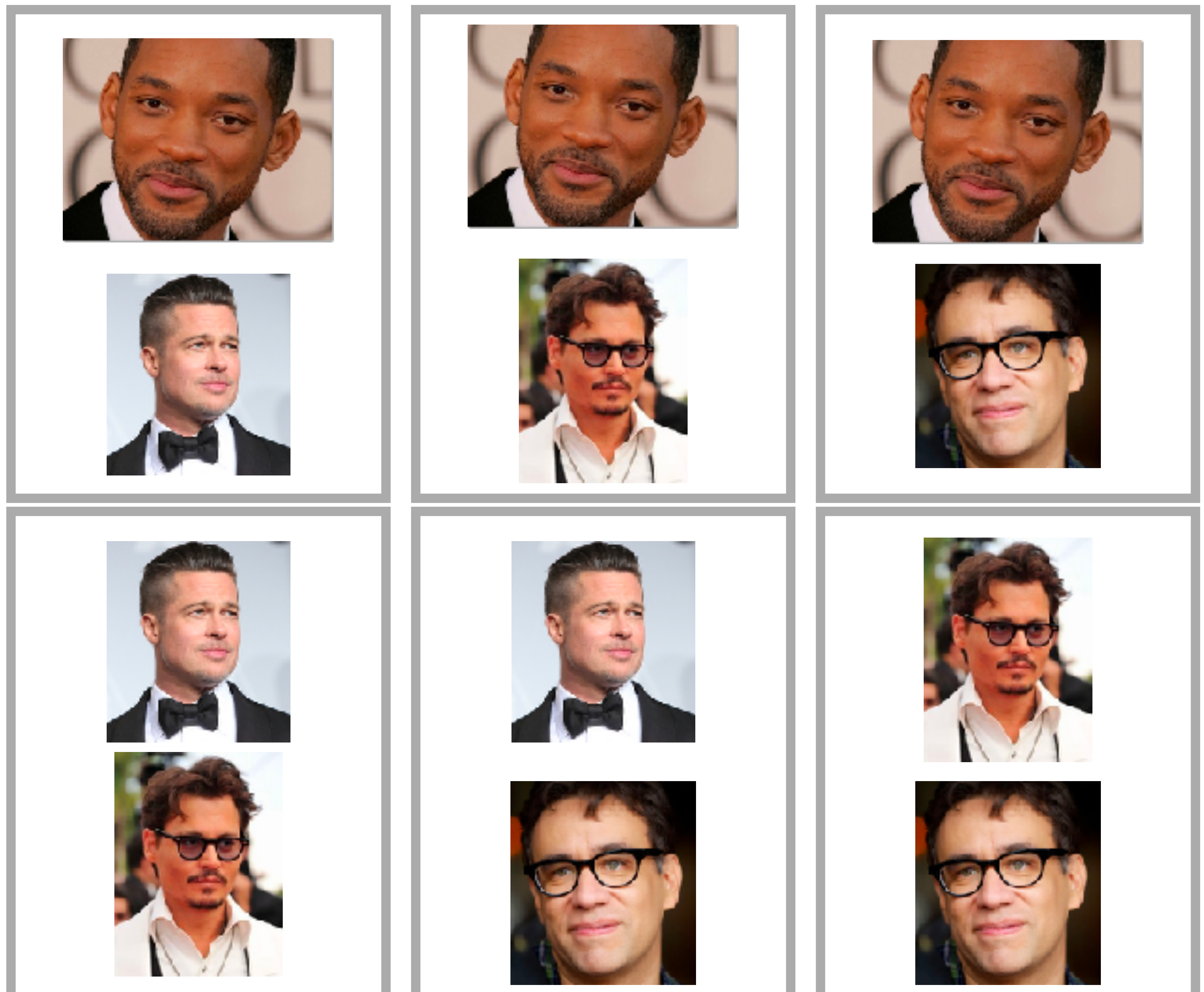
"n factorial"

$$\binom{n}{k} = \frac{n!}{(n-k)!k!}$$

$$n! = n * (n-1) * (n-2) * \ldots * 3 * 2 * 1$$

$$\binom{n}{k} = \frac{n!}{(n-k)!k!} \qquad \binom{4}{2} = \frac{4*3*2*1}{(4-2)!2!} = \frac{24}{2*2} = 6$$





Example: Steph Curry's free throws

$$P(k; n, p) = P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

$$P(2; 4, 0.91) = P(X = 2) = \binom{4}{2} 0.91^2 (1 - 0.91)^{4-2}$$

$$P(2; 4, 0.91) = P(X = 2) = 6 * 0.91^2 (0.09)^2 = 0.040$$

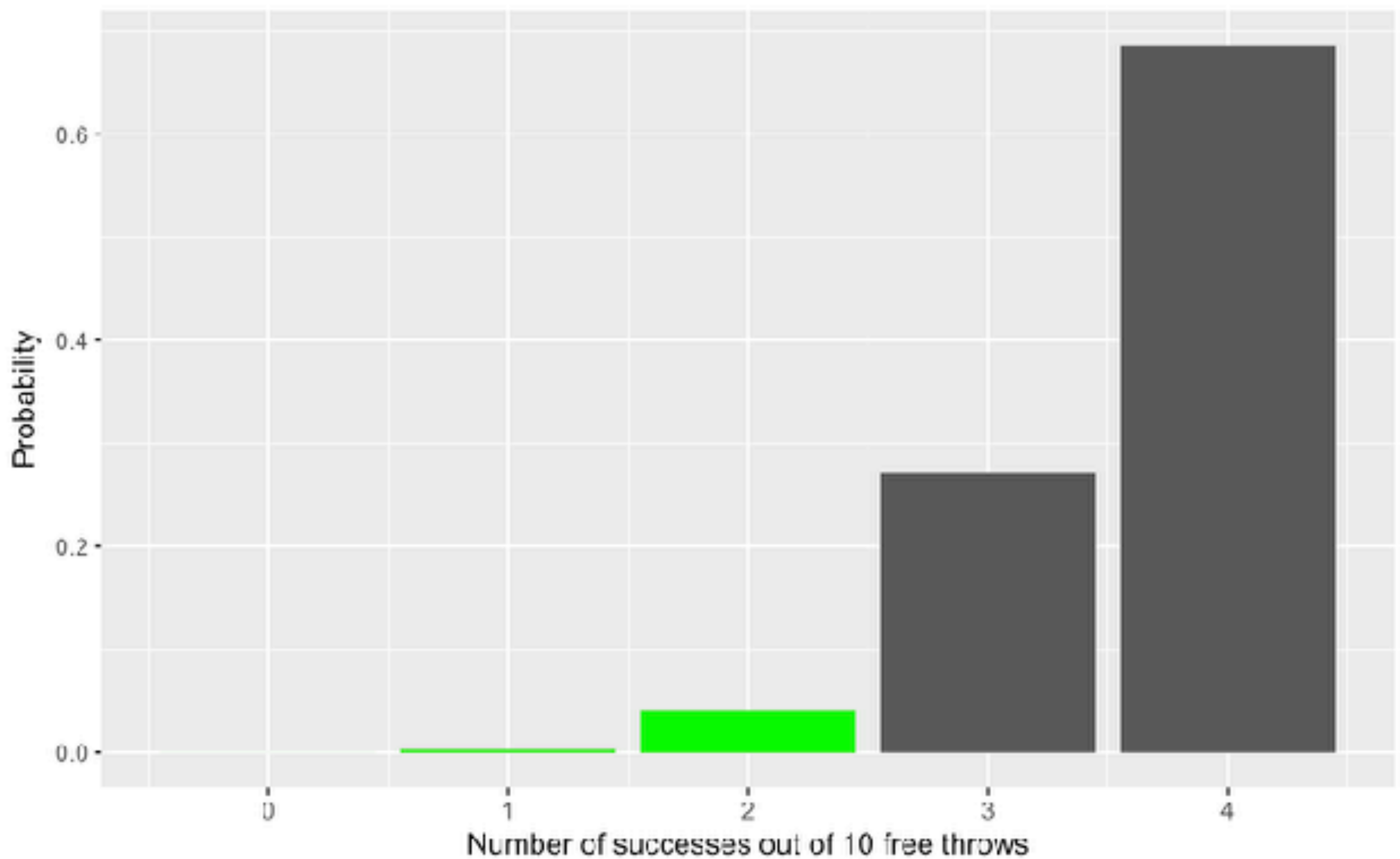
```
pFreeThrows=dbinom(seq(0,4),4,0.91)
data.frame(numSuccesses=seq(0,4),
           probability=pFreeThrows)
```

numSuccesses	probability
0	0.00006561
1	0.00265356
2	0.04024566
3	0.27128556
4	0.68574961

Computing tail probabilities

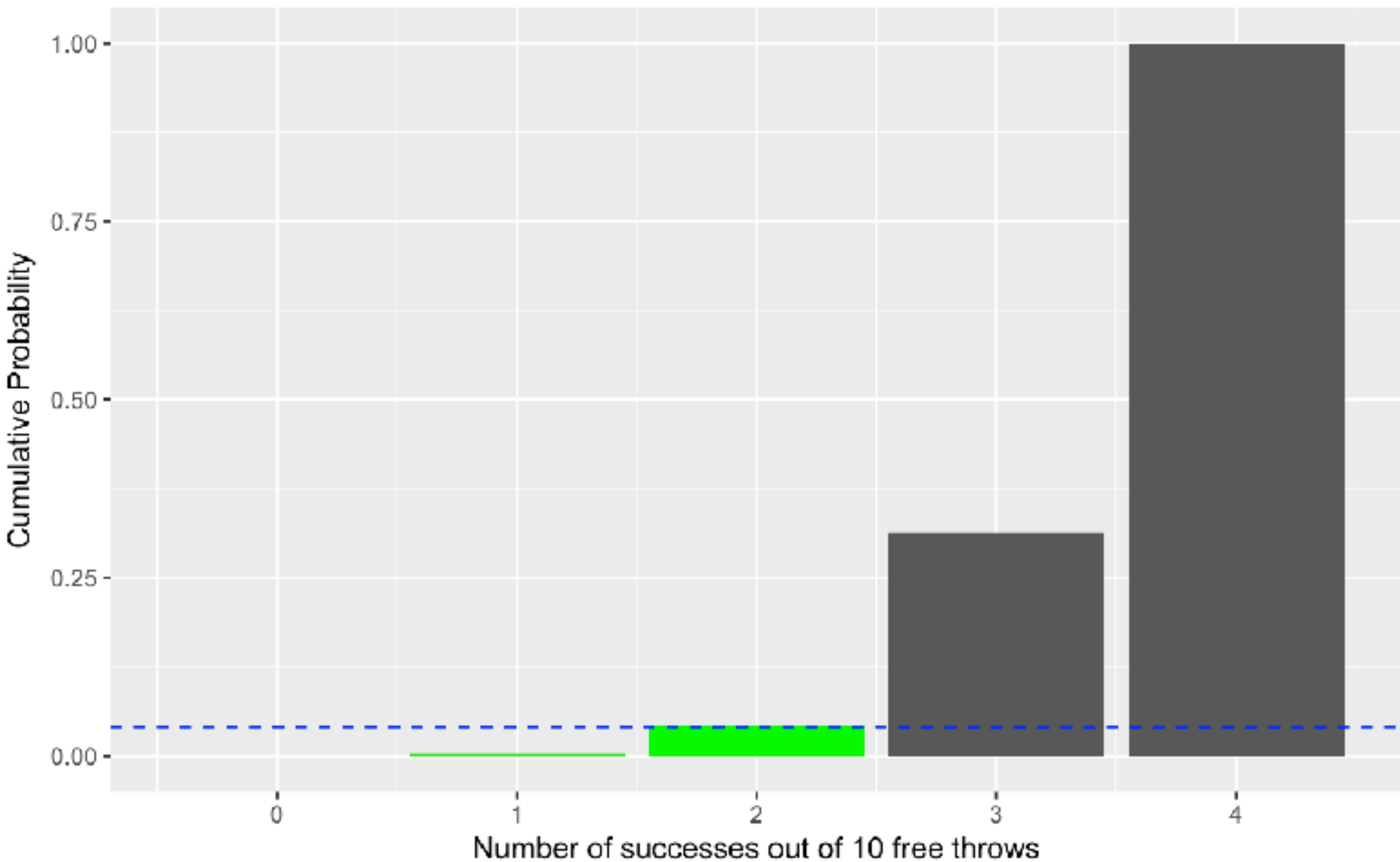
- What is the probability of 2 or fewer successes on 10 throws?

$$P(k \leq 2) = 6e\text{-}5 + .002 + .040 = .043$$



Computing tail probabilities using the cumulative

$P(k \leq 2) = \text{pbinom}(2,4,0.91) = .043$



Summary

- Probabilities are numbers between zero and one that express the likelihood of some event
- We can compute probabilities from data or from theory
- Probability distributions describe the likelihood of various outcomes