PSY 503: Foundations of Statistics in Psych Science

# Basics of Probability

Jason Geller, Ph.D. (he/him/his)

Princeton University

2022-09-26

# Knowledge Check

Go to **www.menti.com/alupng919mx4**

## Name

| | | |
|---|---|---|
| Nicole Horner | Sarah | Sir Branson Byers |
| Cody Dong | Henna | Jamie |
| test-retest reliability | raincloud plot | Karen Christianson |
| Jamie | | |

# Last Class

- Measurement is hard, but so important
- Make sure you understand different types of reliability:
    - Test-retest
    - Internal
    - Inter-rater
- Make sure you understand different types of validity:
    - Construct
    - Face
    - Convergent
    - Divergent

# Today

- What is Probability?
- Different ways of thinking about probability
- Rules of probability

# Probability Warm-up

1. What is probability of drawing the ace of spades from a fair deck of cards?
2. What is the probability of drawing an ace of any suit?
3. You are going to roll some dice twice. What is the chance you roll double 1s?
4. What is the chance that a live specimen of New Jersey Devil will be found?
5. Who is more likely to be a victim of a street robbery, a young man or an old lady?
6. Yesterday the whether forecaster said that there was a 30% chance of rain today, and it rained today. Was she right or wrong?

# What is Probability Theory?

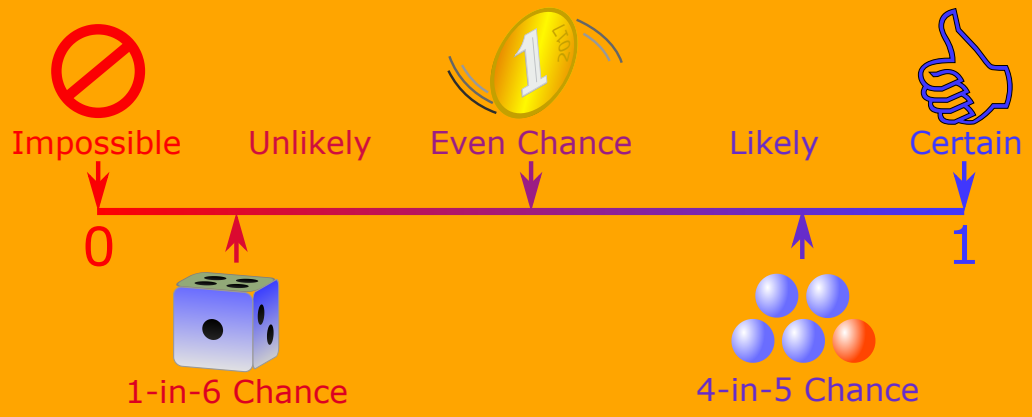
*Probability is the most important concept in modern science, especially as nobody has the slightest notion what it means.*

—Bertrand Russell, 1929

# What is Probability Theory?

Probability is the study of **random processes**

- Probability is used to characterize uncertainty/randomness



# Random Processes: Intuition

- Let's flip a fair coin

```
set.seed(973)

coinflips <- function(x) {
    flip <- rbinom(x, 1, 0.5)
    flip <- ifelse(flip==1, "Tails", "Heads")
    return(flip)
}
```

1. Can you tell me what the outcome will be?

2. If we were to flip a fair coin many many times, would you be able to tell the proportion of times that we would obtain heads?

- If answer to first question is "NO"

AND

- Answer to second question is "YES"

THEN

- You are dealing with a random process

# Definition

> Random processes are **mechanisms** that produce outcomes... from **a world/set of possible** outcomes... with some degree of **uncertainty** but with **regularity**.

# Probability Terminlogy

- **Experiment** or **Trial**:
  - Any activity that produces or observes an outcome
- **Sample space:** $\Omega$
  - The set of all possible outcomes
- **Outcome:** $\omega$
  - Possible realization of the random process
    - heads
- **Event:** $A$, $B$, $C$, etc.
  - A given outcome or set of outcomes
- **Probability**: Proportion of outcomes favoring an event

# Examples of Random Processes

- Random assignment of $N$ individuals to an experimental condition
- Random draw of a sample of $n$ individuals from a population of $N$ individuals
- Rolling a die

# Illustration: Random Assignment

- We randomly assigned an individual to a Treatment (T) vs. Control (C)
  - Sample space?
  - We could express $\Omega$ in the following ways:
  - $\Omega = \{\text{Treatment}, \text{ Control}\}$
  - $\Omega = \{\text{T}, \text{ C}\}$
- What if we assigned two individuals to Treatment (T) vs. Control (C)
  - $\Omega = \{\text{TT}, \text{ TC}, \text{ CT}, \text{ CC}\}$

# Events

- An *event* is a subset of the sample space $\Omega$ and corresponds to the realization of one or more than one outcomes $\omega$
- Let $\Omega = \{\text{TT}, \text{TC}, \text{CT}, \text{CC}\}$
- We could let $A$ be **event** that both individuals are assigned to the same experimental condition
- We could write:
  - $A = \{TT, CC\}$
- Another example?

# Notations

| Syntax | Description |
| --- | --- |
| $\Omega$ | sample space |
| $\omega$ | a possible probabilistic outcome |
| $A \cup B$ | $A$ or $B$ |
| $A \cap B$ | $A$ and $B$ |
| $A^C$ | not $A$ |
| $A_1 \cup A_2 \cup \ldots \cup A_n$ | at least one of $A_1, \ldots, A_n$ |
| $A_1 \cap A_2 \cap \ldots \cap A_n$ | all of $A_1, \ldots, A_n$ |
| $A \cap B = \emptyset$ | $A$ and $B$ are mutually exclusive |

# Practice with Events

- We randomly assign 8 participants to T vs. C
  - Possible outcome:
- $\omega = \text{TTTTCCTC}$
- Sample space:
  - Set of all possible strings of length 8 of T's and C's

# Practice with Events

- Let's randomly generate a possible outcome $\omega_j$ in R \vspace{.40cm}

```
sample(c("T", "C"),
       size = 8,
       replace = TRUE)
```

- In the background, does R draw from this sample space?
- NO: Keep in mind that R draws an outcome $\omega_j$ from $\Omega = \{T, C\}$ 8 times in a row with replacement

# Probability Warm-up

- What is probability of drawing the ace of spades from a fair deck of cards?

```
ace=1/52

ace
```

```
## [1] 0.01923077
```

- What is the probability of drawing an ace of any suit?

```
ace=4/52

ace
```

```
## [1] 0.07692308
```

- You are going to roll some dice twice. What is the chance you roll double 1s?

```
dice1s <- 1/6*1/6
dice1s
```

```
## [1] 0.02777778
```

- What is the chance that a live specimen of New Jersey Devil will be found?
  - 0%
- Who is more likely to be a victim of a street robbery, a young man or an old lady?
  - old lady
- Yesterday the whether forecaster said that there was a 30% chance of rain today, and it rained today. Was she right or wrong?
  - Depends

CAPITAL WEATHER GANG

# Top Hungarian weather service officials fired after wrong forecast

A forecast warning of intense storms prompted a decision to postpone a massive fireworks display on a key Hungarian national holiday

By Zach Rosenthal

August 23, 2022 at 1:31 p.m. EDT

# Different Ways of Thinking About Probability

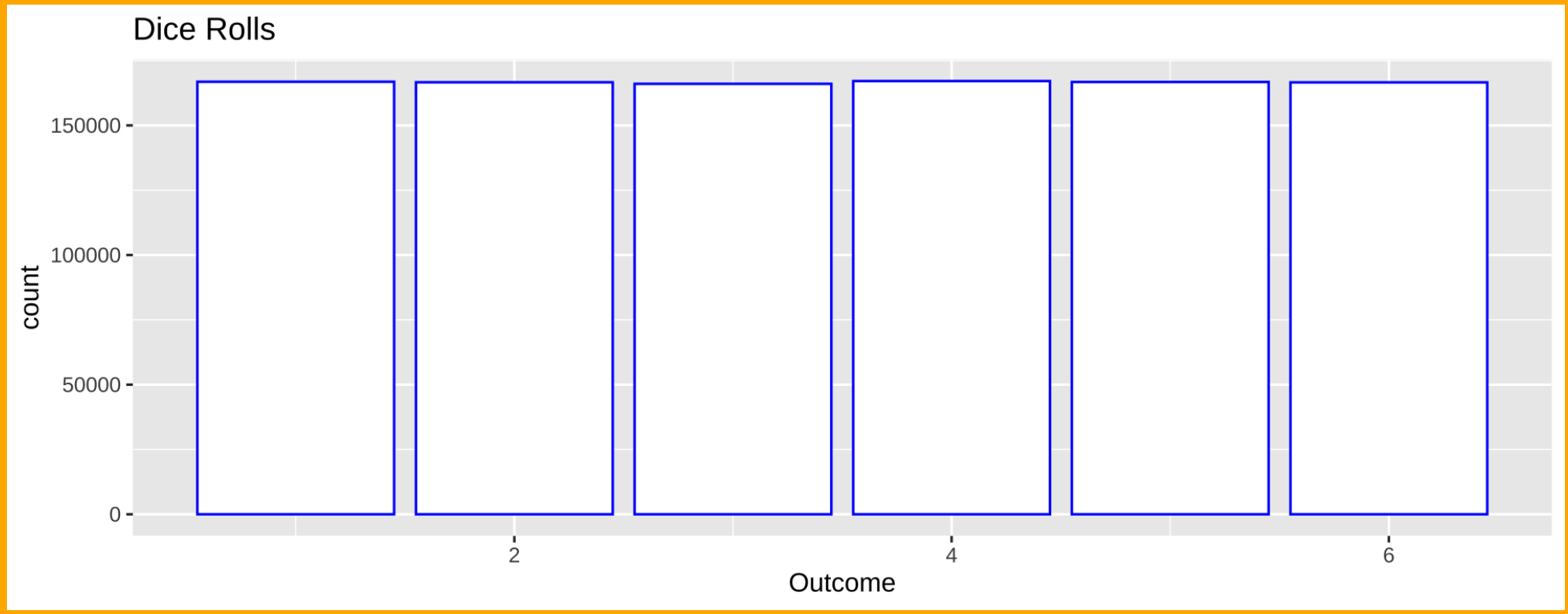
- Classic/Naive
  - **All outcomes are equally likely**

Let $A$ be an event with a finite sample space $\Omega$. The *naive probability* of $A$ is

$$P(A) = \frac{|A|}{|\Omega|}$$

in which |A| is the number of possible outcomes $\omega$ that satisfy A, and |$\Omega$| is the total number of possible outcomes $\omega$ within $\Omega$.

# Dice Rolls



# Wait, why is this naive?

- Requires $\Omega$ to be finite
- Requires each possible outcome $\omega$ to have the same weight
  - This can be misleading!

# Wait, why is this naive?

Is the assumption of equal probability realistic?

- $d_1$: Watching a horror movie
- $d_0$: Watching a neutral movie
- $Y$: Fear response measured
    - Is their an equal probability of attrition in this study?
- Online data collection

# Different Ways of Thinking About Probability

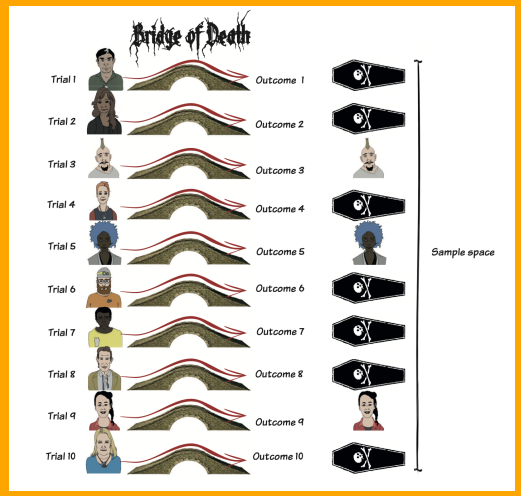
- Frequentist view
  - Past Performance
  - Relative frequency -> Proportion of times an event occurred out of all occasions it could have occurred

$$P(A) = \frac{|f|}{*N*}$$

- Where $f$ = frequency of outcome and $N$ = Total #
- Over the long-run (many repetitions) what is the probability of X event?

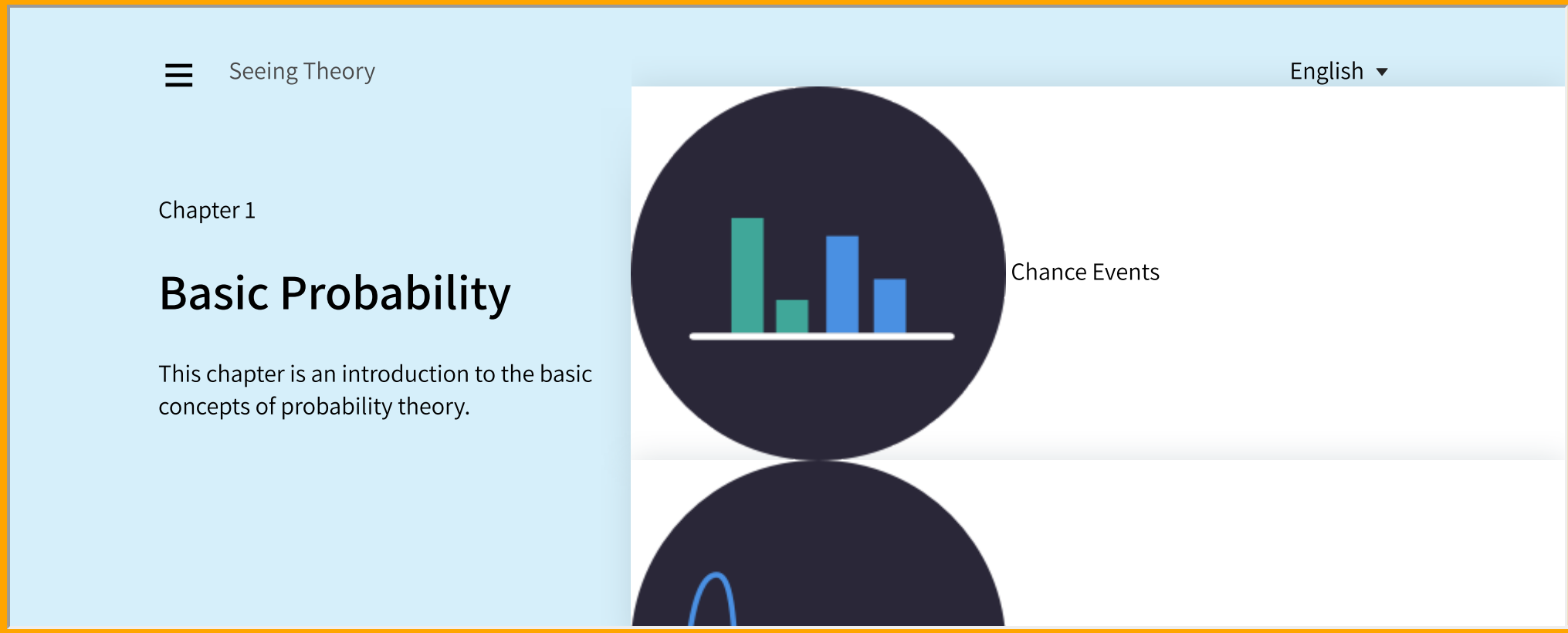
# Different Ways of Thinking About Probability

- Empirical probability
  - Should we cross the bridge?




$$P(death) = \frac{P(number of deaths)}{P(total)}$$

# Coin Flips


Seeing Theory
English
Chapter 1
Basic Probability
This chapter is an introduction to the basic concepts of probability theory.
Chance Events

# Globe Toss

# Different Ways of Thinking About Probability

- Bayesian (Personal belief)
  - In what realistic setting would we actually perform the same experiment infinite times?
  - Many probability questions concern the outcome of a singular trial rather than hypothetical repeated trials, and decision makers with the same information may differ

# Rules of Probability

# Probability Rules

- Probabilities take values between 0 and 1 (inclusive)
  - For some event $A$:

$$0 \leq P(A) \leq 1$$

- Probability cannot be negative
- Probability cannot be greater than 1

# Probability Rule # 2

- Since $\Omega$ is the entire sample space,

$$P(\Omega) = 1$$

- e.g.,If you belong to one of three political parties then the sum of P(R), P(D) and P(I) = 1

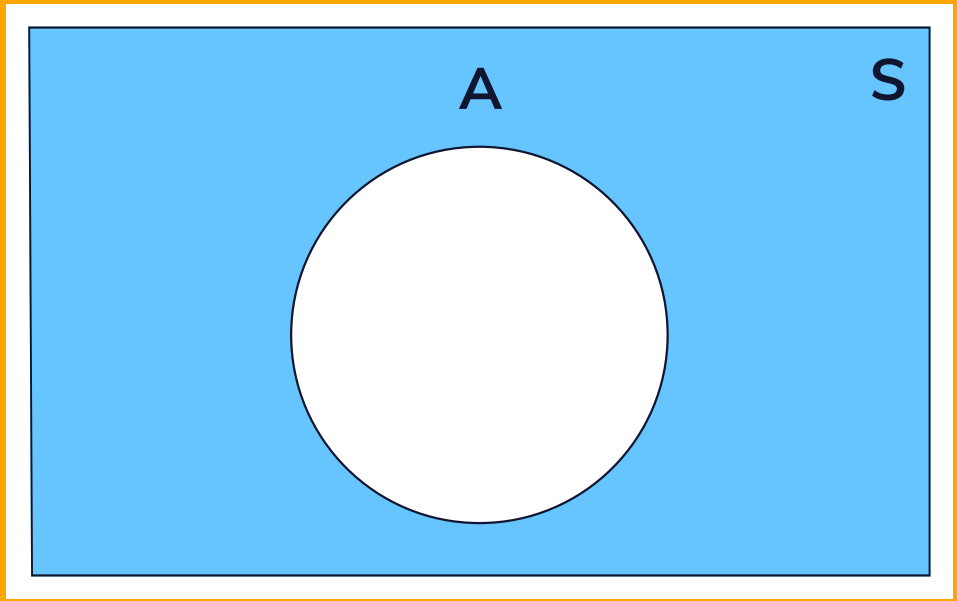
# Probability Rule #3 (Subtraction)

- Complement
  - By definition

$$P(A) + P(A^c) = 1$$

  - This implies

$$P(A^c) = 1 - P(A)$$

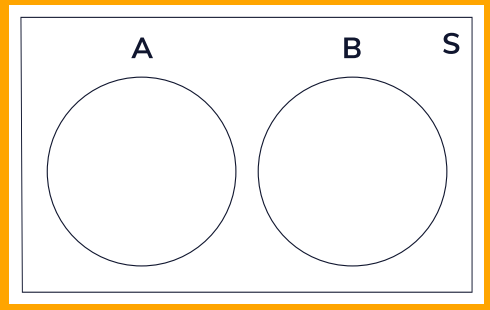


# Probability Rule # 4 (Addition)

- Addition Rule: If A and B are two events in a probability experiment, then the probability that either one of the events will occur is:

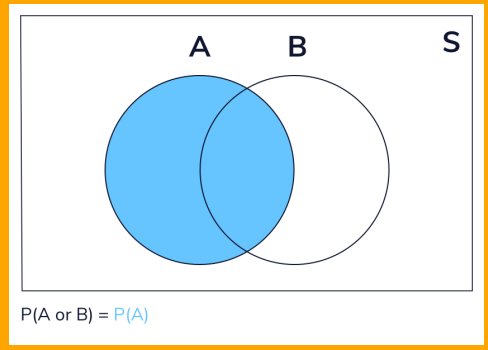
- Mutually Exclusive

P(A or B) = P(A) + P(B)




- Non-Mutually Exclusive

P(A or B) = P(A) + P(B)-(A and B)



# Practice

| Color | Count |
|---|---|
| Brown | 13 |
| Red | 13 |
| Yellow | 14 |
| Green | 16 |
| Orange | 20 |
| Blue | 24 |

*p*(blue or green)

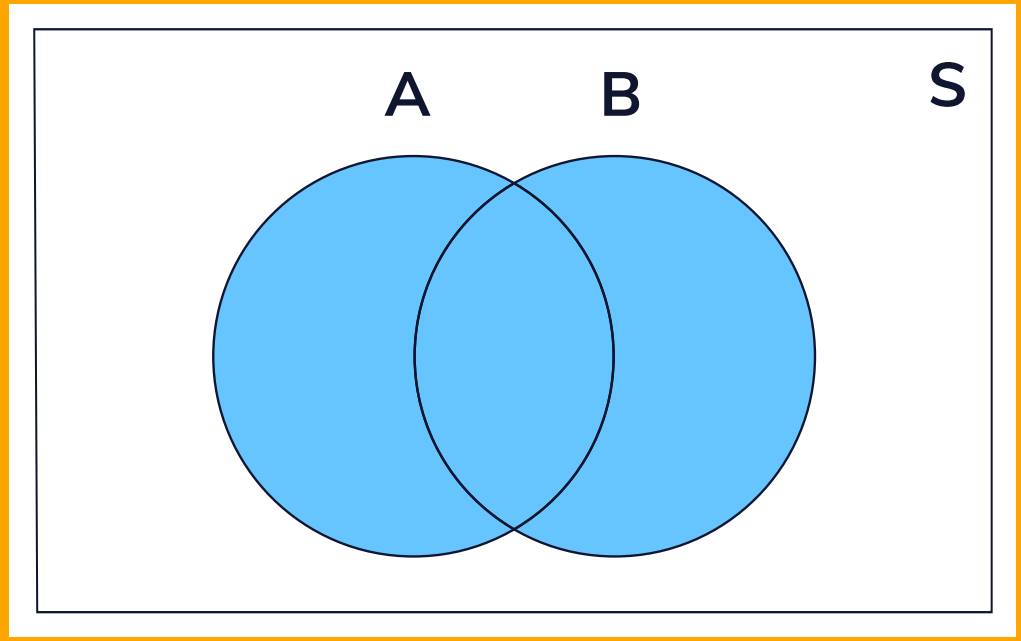
| Color | Count |
|---|---|
| Brown | 13 |
| Red | 13 |
| Yellow | 14 |
| Green | 16 |
| Orange | 20 |
| Blue | 24 |

$p$(blue or green)

```
## [1] 0.4
```

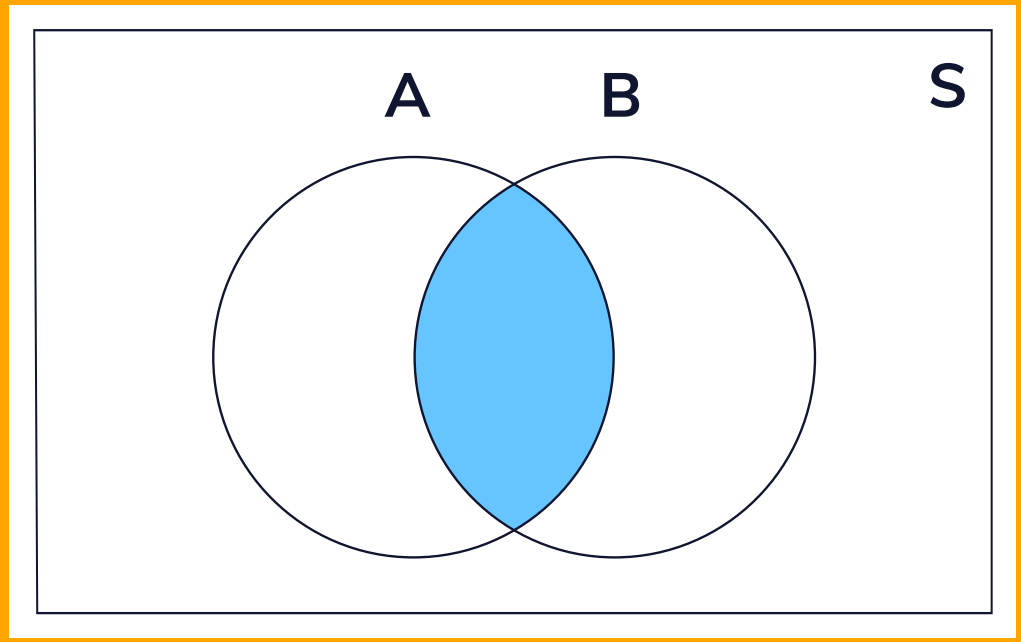
# Union

> The union of two sets encompasses any element that exists in either one or both of them. We can represent this visually as a venn diagram as shown.



# Intersection

> The intersection between two sets encompasses any element that exists in BOTH sets and is often written out as




- Joint probability

# Multiplication Rule

- The multiplication rule is used to find the probability of two events, $A$ and $B$, happening simultaneously.

Dependent:

$$P(AandB) = P(A) * P(B|A)$$

Independent:

$$P(AandB) = P(A) * P(B)$$

# Independent Events

- $A$ and $B$ are independent if the occurrence of $A$ does not influence the occurrence of $B$, and if the occurrence of $B$ does not influence the occurrence of $A$.

If two events $A$ and $B$ are independent, knowing that $A$ occurred does not inform the chances that $B$ occurred. We have:

$$P(A|B) = P(A)$$

$$P(B|A) = P(B)$$

# M&Ms

What is the $p$(blue and blue)?

```
24/100*24/100
```

```
## [1] 0.0576
```

# Knowledge Check

Go to **www.menti.com/al6kt5xojfoa**

Name

6

# Practice with grant proposal

You are about to send a grant proposal to an organization. While you read about the grant, you realize that your grant proposal will be sent to 5 different referees, who can be either social or cognitive psychologists. Imagine that for each grant proposal, the committee flips a coin five times and assigns the proposal to a social psychologist every time the flip returns heads, and to a cognitive psychologist every time the flip returns tails.

Assume an infinite pool of social and cognitive psychologists. What are the chances that your grant proposal is assigned to 5 cognitive psychologists?

# Practice with grant proposal

Let $C_i$ be the event that your grant proposal is assigned to a cognitive psychologist. Since the events are independent from each other, we have:

$$
\begin{aligned}
P(C_1 \cap C_2 \cap C_3 \cap C_4 \cap C_5) &= P(C_1) \times P(C_2) \times P(C_3) \times P(C_4) \times P(C_5) \\
&= (\frac{1}{2})^5 \\
&= \frac{1}{32}
\end{aligned}
$$

# Today

- More fun with probability
  - Conditional probability
  - Bayes' Rule
- Probability and Statistics
  - Probability density function (PDF)
  - Cumulative distribution function (CDF)
- Computing conditional probabilities from data

# Conditional Probablity

- The likelihood of an event or outcome occurring, based on the occurrence of a previous event or outcome

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

- $p$(B|A) -> Conditional probability
- $p$(A ∩ B)-> Joint probability
- $p$(A) -> Marginal probability

# Conditional Probability

- Marginal probability: Probability of single event occurring independent of other events
- Joint probability: Intersection (overlap) of A and B
- Conditional probability: Likelihood that an outcome randomly sampled from the subset with $B$ has $A$ (i.e., conditional is opposed to marginal)
    - We would say “B given A” or B conditional on A”

# Conditional Probability Practice

A math teacher gave her class two tests. 25% of the class passed both tests and 42% of the class passed the first test. What percent of those who passed the first test also passed the second test?

$$p(second \mid first)$$

- $p(A \cap B)$: .25
- $p(A)$: .42

$$p(second \mid first) = .6$$

# Conditional Probability Practice

I just got accepted to graduate school. The acceptance rate is 30%. Not everyone gets funding if they have been accepted (only 13% do). What is the probability I receive funding given that I was accepted?

$$p(funding|accepted)$$

- $p$(A ∩ B): .13
- $p$(A): .3

$$p(funding \mid accepted) = .43$$

# Bayes' Rule

- Reversing a conditional probability allows us to find $P(A|B)$ if we know $P(B|A)$:
  - Bayes' rule:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

$$P(A|B) = \frac{P(B|A) * P(A)}{P(B|A) * P(A) + P(B|\neg A) * P(\neg A)}$$

# Bayes' Rule

- Allows us to update the probability of an event $A$ based on the occurrence of another event
- $P(A)$ is called the *prior probability*
- $P(A|B)$ is called the *posterior probability*

# Monty Hall Problem

# Monty Hall

| Behind Door #1 | Behind Door #2 | Behind Door #3 | Outcome if Stick with Door #1 | Outcome if Switch to the Door Offered |
|---|---|---|---|---|
| **Car** | Goat | Goat | **Car** | Goat |
| Goat | **Car** | Goat | Goat | **Car** |
| Goat | Goat | **Car** | Goat | **Car** |

- The winning strategy is to switch, but how is this possible?
  - Our intuition tells us our chance of winning the car increases from 1/3 to 1/2 when there are only two doors to choose from
  - In reality, our chance of winning the car remains 1/3 if we stick with our original choice, but increases to 2/3 if we switch

# Monty Hall Simulations

```
monty <- function() {
    prize <- sample(x = 1:3, size = 1, replace = TRUE)
    choice <- sample(x = 1:3, size = 1, replace = TRUE)
    monty <- sample(x = c(1:3)[-c(choice, prize)], size = 1, replace = TRUE)
    return(ifelse(prize != choice, yes = "Switch", no = "Stick"))
}

monty
```

```
## function() {
##     prize <- sample(x = 1:3, size = 1, replace = TRUE)
##     choice <- sample(x = 1:3, size = 1, replace = TRUE)
##     monty <- sample(x = c(1:3)[-c(choice, prize)], size = 1, replace = TRUE)
##     return(ifelse(prize != choice, yes = "Switch", no = "Stick"))
## }
```

```r
run <- rep(NA, 100000)

for (i in 1:100000) {
    run[i] <- monty()
}

prop.table(table(run))
```

```
## run
##   Stick  Switch
## 0.33393 0.66607
```

```r
## strategy
##   Stick  Switch
## 0.33147 0.66853
```

# Illustration: Bayes' Rule

Doctors recommend getting a PSA test after 50 to screen for prostate cancer

- If you tested positive for prostate cancer, what is the chance you actually have it?

    1. 80% of the people who test positive have prostate cancer $(\text{sensitivity} = P(\text{positive test}|\text{ disease}))$

    2. 70% of the people who have a negative test do not have cancer $(\text{specificity} = P(\text{negative test}|\text{no disease}))$

    3. 5% of individuals over 60 have prostate cancer

# Illustration: Prostate Cancer

$$P(\text{cancer}|\text{test}) = \frac{P(\text{test}|\text{cancer}) * P(\text{cancer})}{P(\text{test}|\text{cancer}) * P(\text{cancer}) + P(\text{test}|\neg\text{cancer}) * P(\neg\text{cancer})}$$

$$= \frac{0.8 * 0.058}{0.8 * 0.058 + 0.3 * 0.942} =$$

0.14

# Class Activity

Suppose there is a disease outbreak in an enclosed population. It is turning folks into zombies.

- Your friend tested positive. How likely is it that they are a zombie?

    1. 99% of the people who test positive have Zombie Virus
    $(\text{sensitivity} = P(\text{positive test}|\text{ zombie}))$

    2. 85% of the people who have a negative test do not have Zombie virus
    $(\text{specificity} = P(\text{negative test}|\text{not zombie}))$

    3. 15% of individuals are zombies

03:00

# Bayes' Rule: Zombie Outbreak

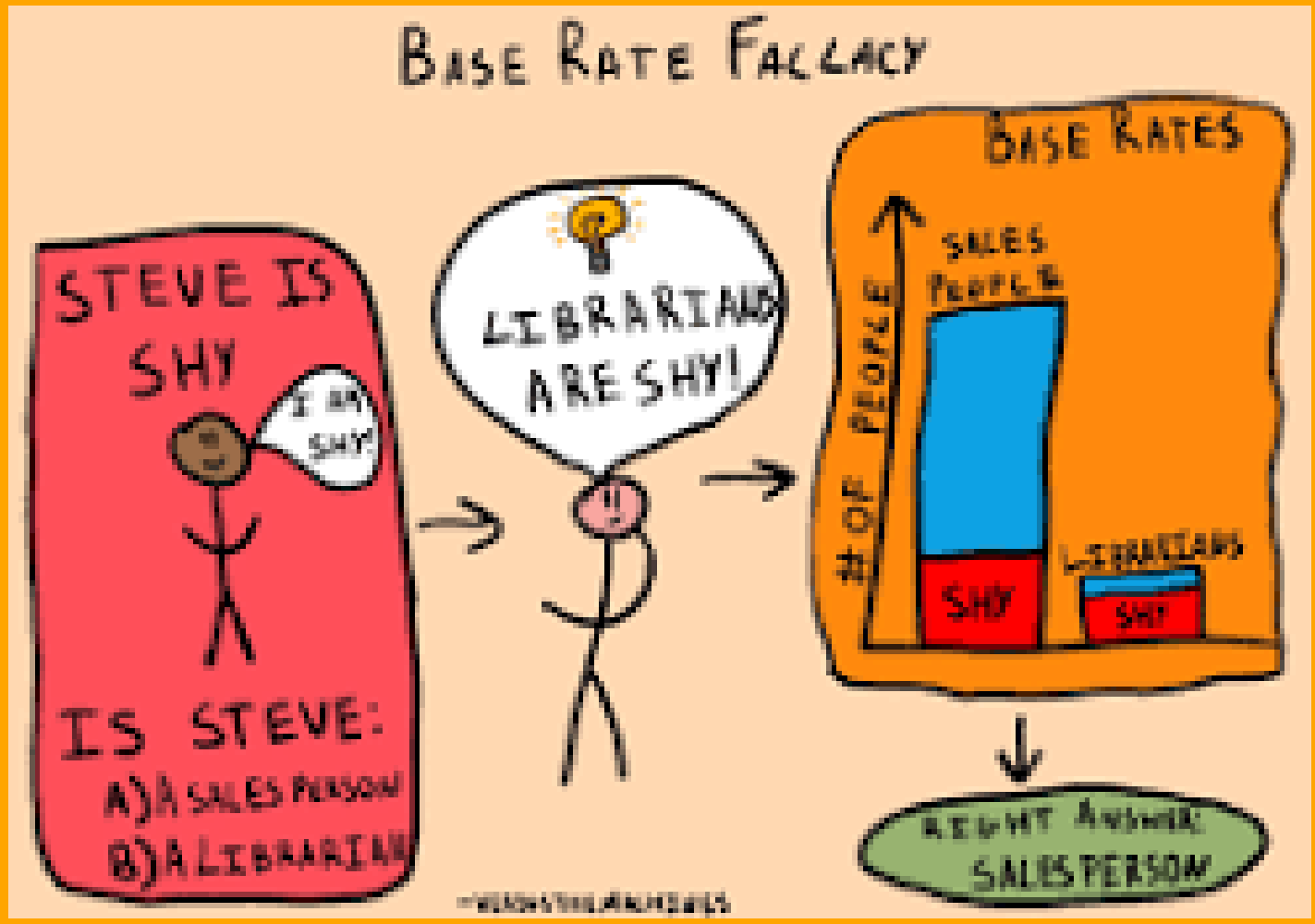
$$P(\text{zombie}|\text{test}) = \frac{P(\text{test}|\text{zombie}) * P(\text{zombie})}{P(\text{test}|\text{zombie}) * P(\text{zombie}) + P(\text{test}|\neg\text{zombie}) * P(\neg\text{zombie})}$$

$$= \frac{0.99 * 0.15}{0.99 * 0.15 + 0.15 * 0.85} = 0.53$$

# Lessons from Bayes' rule

- Based on the results of this test, the probability that your friend actually is a zombie is .54
    - That's a 54% chance of being zombie
        - What would you do?
- Bayes' rule often yields counter-intuitive results!
- Importance of base rates

# Base-rate neglect



# Probability Theory vs. Statistical Inference

# Probability Theory

- For any given random phenomenon, probability theory is a set of tools that assume prior knowledge of:
  - The sample space
  - The probability of a set of events defined on that sample space
- Allows you to find the probability of any other possible event from that sample space

# Problem

- We usually don't know the probability model
- OK, we can find the probability of every outcome in the sample space by observing many many repetitions
    - BUT most random phenomena cannot be repeated again, again, and again
- We generally need to infer the probability of each possible outcome using information on a few realizations of the random phenomenon of interest

# Probability and Statistics

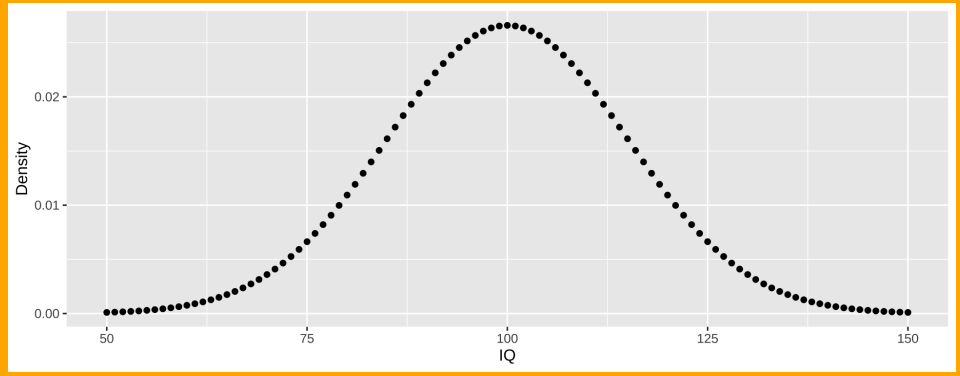
- By knowing your population makeup, you have a better idea of the probability of obtaining certain samples.
- Probability links population with samples.
- Inferential statistics rely on this connection when they use sample data as the basis for making conclusions about populations.

# Probability Distributions

- Probability density function (PDF)
  - Indicates the probability of observing a measurement with specific value

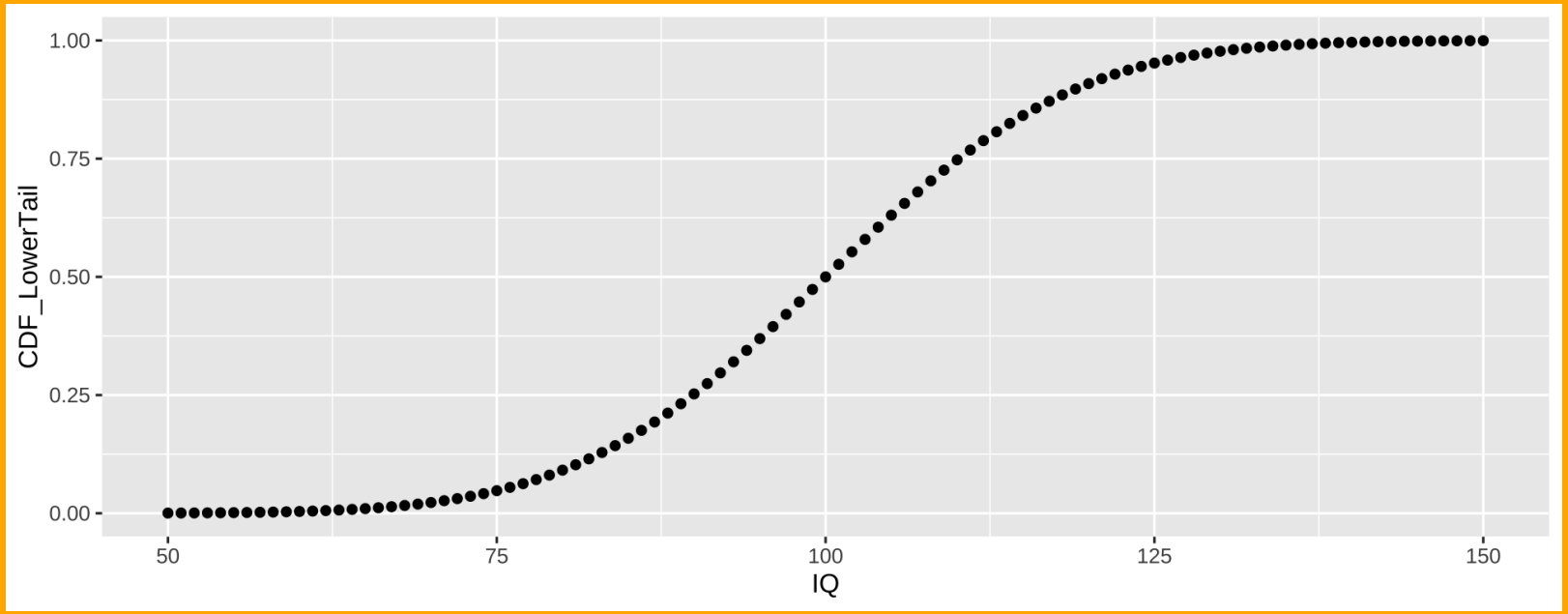
$$f(x|\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

  - e.g., Where does an IQ of 140 lie?



# Probability Distributions

- Cumulative distribution function (CDF)
  - $X <= x$ (less than or equal)
  - E.g., Is IQ less than or equal given value



# R

- PDF

```r
dnorm(x,            # X-axis values (grid)
      mean = 0,     # Integer or vector representing the mean/s
      sd = 1,       # Integer or vector representing the standard deviation/s
)
```

- CDF

```r
pnorm()
```

# In-Class Analysis

# Data

- Florida voter registration data

```
library(here)

voter=read.csv(here::here("static","slides","05-Probability","data", "florida-voters.csv"))
voter <- na.omit(voter)
voter %>%
  glimpse()
```

```
## Rows: 9,113
## Columns: 6
## $ surname <chr> "PIEDRA", "LYNCH", "LATHROP", "HUMMEL", "CHRISTISON", "HOMAN",…
## $ county  <int> 115, 115, 115, 115, 115, 115, 115, 1, 1, 115, 115, 115, 115, 1…
## $ VTD     <int> 66, 13, 80, 8, 55, 84, 48, 41, 39, 26, 45, 11, 48, 88, 25, 82,…
## $ age     <int> 58, 51, 54, 77, 49, 77, 34, 56, 60, 44, 45, 80, 83, 55, 33, 63…
## $ gender  <chr> "f", "m", "m", "f", "m", "f", "f", "f", "m", "m", "f", "m", "f…
## $ race    <chr> "white", "white", "white", "white", "white", "white", "white",…
```

# Data: Setup

```
library(kableExtra)

head(voter) %>%
kable(align = "cccccc")%>%
   kable_material_dark()
```

| | surname | county | VTD | age | gender | race |
|---|---|---|---|---|---|---|
| 1 | PIEDRA | 115 | 66 | 58 | f | white |
| 2 | LYNCH | 115 | 13 | 51 | m | white |
| 4 | LATHROP | 115 | 80 | 54 | m | white |
| 5 | HUMMEL | 115 | 8 | 77 | f | white |
| 6 | CHRISTISON | 115 | 55 | 49 | m | white |

# Marginal Probabilties

- What are these again?
    - The probability of an event irrespective of the outcomes

```
marg.race <- voter %>%
  count(race)%>%
  mutate(prop=prop.table(n))
```

| race | n | prop |
|---|---|---|
| asian | 175 | 0.0192033 |
| black | 1194 | 0.1310216 |
| hispanic | 1192 | 0.1308022 |
| native | 29 | 0.0031823 |
| other | 310 | 0.0340173 |
| white | 6213 | 0.6817733 |

# Gender

```
marg.gender <- voter %>%
  group_by(gender) %>%
  summarise(n=n())%>%
  mutate(freq=n/sum(n))

marg.gender %>%
  kable(align = "cccccc")%>%
   kable_material_dark()
```

| gender | n | freq |
|---|---|---|
| f | 4883 | 0.5358279 |
| m | 4230 | 0.4641721 |

# Conditional Probability

$$P(black|male) =$$

| race | gender | n | prop |
|---|---|---|---|
| asian | m | 92 | 0.0217494 |
| black | m | 516 | 0.1219858 |
| hispanic | m | 526 | 0.1243499 |
| native | m | 12 | 0.0028369 |
| other | m | 152 | 0.0359338 |
| white | m | 2932 | 0.6931442 |

# Joint Probability

$$P(black \cap male)$$

```
library(janitor)
joint <- voter %>%
  select(race, gender) %>%
  group_by(race, gender) %>%
  count(race, gender) %>%
  ungroup() %>%
  mutate(total=sum(n), prop=n/total)
```

| race | gender | n | total | prop |
|---|---|---|---|---|
| asian | f | 83 | 9113 | 0.0091079 |
| asian | m | 92 | 9113 | 0.0100955 |
| black | f | 678 | 9113 | 0.0743992 |
| black | m | 516 | 9113 | 0.0566224 |
| hispanic | f | 666 | 9113 | 0.0730824 |
| hispanic | m | 526 | 9113 | 0.0577197 |
| native | f | 17 | 9113 | 0.0018655 |
| native | m | 12 | 9113 | 0.0013168 |
| other | f | 158 | 9113 | 0.0173379 |

# Data: Independance

$$P(black \cap male)$$

$$p(black \cap male) = p(black) \times p(male)$$

```
#Are race and gender independent? Recall that two #events are independent if and only if, for example:

marg.race <- voter %>%
 group_by(race)%>%
  tabyl(race)

marg.gender <- voter %>%
   group_by(gender)%>%
  tabyl(gender)
```

| race | n | percent |
|---|---|---|
| asian | 175 | 0.0192033 |
| black | 1194 | 0.1310216 |
| hispanic | 1192 | 0.1308022 |
| native | 29 | 0.0031823 |
| other | 310 | 0.0340173 |
| white | 6213 | 0.6817733 |

```
marg.gender %>%
kable() %>%
 kable_material_dark()
```

| gender | n | percent |
|---|---|---|
| f | 4883 | 0.5358279 |
| m | 4230 | 0.4641721 |

```
0.13*0.464
```

```
## [1] 0.06032
```

```
voter %>%
 select(race, gender) %>%
 group_by(race, gender) %>%
 count(race, gender) %>%
 ungroup() %>%
 mutate(total=sum(n), prop=n/total) %>%
 kable(align = "cccccc") %>%
 kable_material_dark()
```

| race | gender | n | total | prop |
|---|---|---|---|---|
| asian | f | 83 | 9113 | 0.0091079 |
| asian | m | 92 | 9113 | 0.0100955 |
| black | f | 678 | 9113 | 0.0743992 |
| black | m | 516 | 9113 | 0.0566224 |
| hispanic | f | 666 | 9113 | 0.0730824 |

# Your Turn

1. What is the conditional probability: $(P(black|female))$

```
library(tidyverse)
data <- read_csv("https://raw.githubusercontent.com/jgeller112/psy503-psych_stats/master/static/slides/
```

```
cond_racegender <- voter  %>%
  dplyr::filter(gender=="f")%>%
  dplyr::count(race, gender) %>%
  dplyr::mutate(prop=n / sum(n))%>%
  kable(align = "cccccc")%>%
   kable_material_dark()

cond_racegender
```

| race | gender | n | prop |
|---|---|---|---|
| asian | f | 83 | 0.0169977 |
| black | f | 678 | 0.1388491 |
| hispanic | f | 666 | 0.1363916 |
| native | f | 17 | 0.0034815 |
| other | f | 158 | 0.0323572 |