

Probability 1

Lecture 2

STA 371G

1. Announcements and logistics
2. Probability rules
3. Getting our hands dirty with R

Announcements

- You can find lecture slides, data, etc on Canvas
- Practice quiz tomorrow night at 6:30 PM (15 minutes, doesn't count towards your grade!)



TEXAS CONVERGENT

SPRING 2021 RECRUITMENT



Build Team and Incubator applications
go live 1/21 6:00 CDT





INFO SESSIONS

Tuesday, Jan 26th, 5:00 - 6:00 PM CDT
Wednesday, Jan 27th, 5:00 - 6:00 PM CDT

VIRTUAL ORG FAIR

Feb 1st, 11:00 AM - 3:00 PM CDT

OFFICE HOURS

Feb 1st, 2:00 PM - 6:00 PM CDT
Feb 4th, 12:00 PM - 4:00 PM CDT



APPLY NOW!

www.txconvergent.com/build

Applications due on
Feb 5th, 11:59 PM CDT

1. Announcements and logistics
2. Probability rules
3. Getting our hands dirty with R

Probability rules

1. The chance of an event happening is between 0% and 100%, i.e. $0 \leq P(E) \leq 1$ for any event E .
2. The probabilities for all possible outcomes put together add up to 1.
3. The probability that something doesn't happen is 100% minus the probability that it does happen, i.e. $P(E^c) = 1 - P(E)$.

Disjoint events

Definition

Two events A and B are **disjoint** if they never both occur, i.e., if $P(A \text{ and } B) = 0$.

Disjoint events

Definition

Two events A and B are **disjoint** if they never both occur, i.e., if $P(A \text{ and } B) = 0$.

If we consider two UT students selected at random,

- “Has a 3.0 GPA” and “Has a 4.0 GPA” are disjoint
- “Has a major in McCombs” and “Has a major in CNS” are not disjoint (you can do both!)

Let's say we pick a *Jersey Shore: Family Vacation* cast member at random.





Mike



Ronnie



Vinny



Pauly



J-Woww



Snooki



Deena

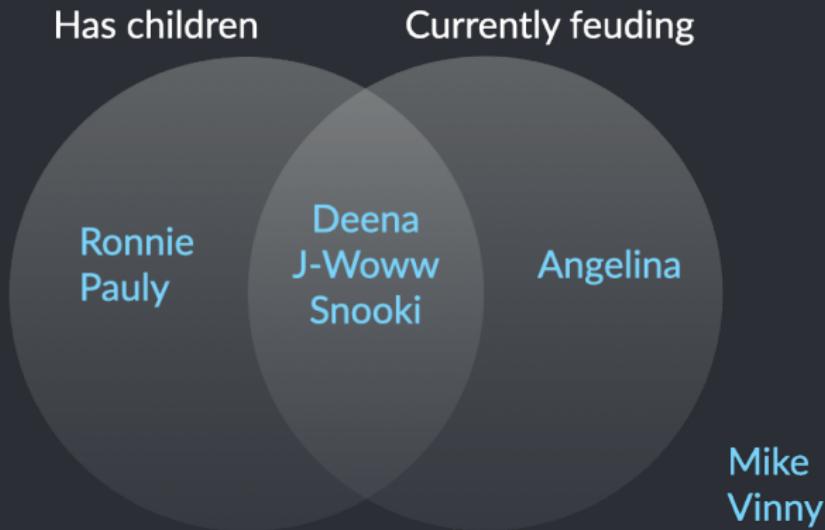


Angelina

Probability rules

1. The chance of an event happening is between 0% and 100%, i.e.
 $0 \leq P(E) \leq 1$ for any event E .
2. The probabilities for all possible outcomes put together add up to 1.
3. The probability that something doesn't happen is 100% minus the probability that it does happen, i.e. $P(E^c) = 1 - P(E)$.
4. If A and B are disjoint, $P(A \text{ or } B) = P(A) + P(B)$.

What if the events are not disjoint?



$$P(C) = \frac{5}{8} \quad P(F) = \frac{4}{8} \quad P(C \text{ or } F) = \frac{6}{8} \neq P(C) + P(F)$$

What if the events are not disjoint?



We are double-counting Deena, J-Woww, and Snooki!

$$P(C \text{ or } F) = P(C) + P(F) - P(C \text{ and } F) = \frac{5}{8} + \frac{4}{8} - \frac{3}{8} = \frac{6}{8}$$

Probability rules

1. The chance of an event happening is between 0% and 100%, i.e.
 $0 \leq P(E) \leq 1$ for any event E .
2. The probabilities for all possible outcomes put together add up to 1.
3. The probability that something doesn't happen is 100% minus the probability that it does happen, i.e. $P(E^c) = 1 - P(E)$.
4. $P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$.

Independent events

Definition

Two events A and B are **independent** if knowing that one event happened tells you nothing about whether the other happened.

Independent events

Definition

Two events A and B are **independent** if knowing that one event happened tells you nothing about whether the other happened.

If we consider two UT students selected at random,

- “Has blue eyes” and “Has a 4.0 GPA” are independent
- “Got a 1600 on their SAT” and “Has a 4.0 GPA” are not independent

Probability rules

1. The chance of an event happening is between 0% and 100%, i.e.
 $0 \leq P(E) \leq 1$ for any event E .
2. The probabilities for all possible outcomes put together add up to 1.
3. The probability that something doesn't happen is 100% minus the probability that it does happen, i.e. $P(E^c) = 1 - P(E)$.
4. $P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$.
5. If A and B are independent, $P(A \text{ and } B) = P(A)P(B)$.

What if the events are not independent?



We are double-counting Deena, J-Woww, and Snooki!

$$P(C) = \frac{5}{8} \quad P(F) = \frac{4}{8} \quad P(C \text{ and } F) = \frac{3}{8} \neq P(C)P(F)$$

1. Announcements and logistics
2. Probability rules
3. Getting our hands dirty with R

A quick overview of R

- Data window, console window, environment window, plot/help window
- Using R as a calculator
- Using the console window
- Creating R scripts (why and how)
- Adding comments to scripts
- Importing data from a spreadsheet
- Loading RData files

Summarizing categorical data

Use the `table` function to create a contingency (frequency) table for categorical data:

```
table(titanic$Sex, titanic$Survived)
```

	No	Yes
female	71	217
male	372	96

Use the `addmargins` function to add “margins” (the row, column, and overall totals):

```
addmargins(table(titanic$Sex, titanic$Survived))
```

	No	Yes	Sum
female	71	217	288
male	372	96	468
Sum	443	313	756

Apply `prop.table` to a table to tally up percentages by rows or columns:

```
# By rows
```

```
prop.table(table(titanic$Sex, titanic$Survived), 1)
```

	No	Yes
female	0.2465278	0.7534722
male	0.7948718	0.2051282

```
# By columns
```

```
prop.table(table(titanic$Sex, titanic$Survived), 2)
```

	No	Yes
female	0.1602709	0.6932907
male	0.8397291	0.3067093

Getting more help with R

Additional resources for learning R are available on Canvas (see also the STA 309 R help pages if you are new to R).