PSY 201: Statistics in Psychology

Lecture 13

Probability

Coincidences are rarely interesting.

Greg Francis

Purdue University

Fall 2019

PROBABILITY

number between 0 and 1
probability of event A is written as

$$P(A)$$

if

$$P(A) = 1.0$$

it indicates with certainty that event A will happen
if

$$P(A) = 0$$

it indicates with certainty that event A will **not** happen

EVERYDAY EVENTS

- people often have misconceptions about the way probabilities interact
- things that seem rare may not actually be
- interesting to analyze the probability of events that seem unusual
  - Julius Ceasar
  - Hitting streaks
  - Predictive dreams
  - Shared birthdays
  - Con games with cards

JULIUS CAESAR

- Some 2000 years ago (or so) Julius Caesar is said to have gasped "You too, Brutus? Then I die." as his friend stabbed him to death
- What are the chances that you just inhaled a molecule that came out of his mouth?
- Surprisingly good! Almost 0.99.
- Assumes
  - Caesar's dying breath contained about $A = 2.2 \times 10^{22}$ molecules
  - Those molecules are free and distributed around the globe evenly.
  - Your inward breath contained about $B = 2.2 \times 10^{22}$ molecules
  - The atmosphere contains about $N = 10^{44}$ molecules

JULIUS CAESAR

- If there are N molecules and Caesar exhaled A of them, then the probability that any given molecule you inhale is from Caesar is

$$P(\text{m from C}) = \frac{A}{N} = 2.2^{-23}$$

- which is very small!
- So the probability that any given molecule you inhale is **not** from Caesar is the complement:

$$P(\text{m not from C}) = 1 - \frac{A}{N} = 1 - 2.2^{-23}$$

JULIUS CAESAR

- So the probability of inhaling B molecules that are not from Caesar is

$$P(\text{breath not from C}) = \left(1 - \frac{A}{N}\right)^{B} \approx 0.01$$

- So the probability of your breath containing a molecule from Caesar is approximately 1-0.01 = 0.99!

HITTING STREAKS

- Pete Rose set a National League record with 44 consecutive games with a safe hit
- this is impressive, but is it rare?
- Rose batted around 0.300 (had a safe hit 30% of the time)
- so, assuming 4 at bats per game, the probability of **not** getting a hit during a game is

$$P(\text{no hit}) = (1-0.3)^4 = 0.24$$

- So the probability of getting at least one hit is 1-0.24 = 0.76.

HITTING STREAKS

- Still, the probability of getting hits in any given sequence of 44 games is

$$P(\text{44 streak}) = (0.76)^{44} = 0.000005699$$

- and the probability of not getting a streak is

$$P(\text{not 44 streak}) = 1 - (0.76)^{44} = 0.999994301$$

HITTING STREAKS

- But there are 162 games in a season, so there are 118 sets of 44 consecutive games.
- Thus, the probability of not getting a streak of hits in at least 44 consecutive games out of a 162 game season is:

$$P(\text{no streak}) = (0.999994)^{118} = 0.999327$$

- so the probability of a 44-game streak is

$$P(\text{streak}) = 1 - (0.999994)^{118} = 0.000672$$

- (includes the possibility of streaks of more than 44 games)
- Still very rare!

HITTING STREAKS

- But how many players have been in the Major Leagues at any given time? (say 30 that bat like Rose)
- the probability that **every** player will **not** get a streak of at least 44 games in a given year is

$$P(\text{no streak}) = (0.9993)^{30} = 0.9800$$

- So probability that at least one player gets such a streak is

$$1.0 - 0.980027651 = 0.019972349$$

- still small!

HITTING STREAKS

- And how many years has baseball been played? (say 100)
- the probability that **every** year everyone will **not** get a streak of at least 44 games in a given year is

$$P(\text{no streak}) = (0.9800)^{100} = 0.1329$$

- So probability that at least one player on some year gets such a streak is

$$1.0 - 0.132994269 = 0.867005731$$

- which is pretty good odds!
- Thus, we can expect that Rose's streak will be broken eventually (unless pitchers become much better)

PREDICTIVE DREAMS

- ever dream something and had it come true?
- Many people take this occurence as evidence of extrasensory perception and "other worlds". But it's actually not that uncommon from a probabilitistic point of view
- suppose that the probability that a night's dream matches some later event in life is 1 in 10000

$$P(\text{predictive dream}) = 0.0001$$

- Then the chance that a dream is non-predictive is

$$P(\text{non predictive dream}) = 1 - 0.0001 = 0.9999$$

- assume that dream predictiveness is independent

PREDICTIVE DREAMS

- With 365 days a year, the probability that all 365 nights have non-predictive dreams is

$$P(\text{non predictive}) = (0.9999)^{365} = 0.96415$$

- so the probability that an individual has a predictive dream during a year is

$$P(\text{predictive}) = 1.0 - 0.96415 = 0.03585$$

- or about 3.6% of people have a predictive dream during a year
- considering that there are billions of people, this corresponds to millions of dreams (and lots of people talk about them!)

PREDICTIVE DREAMS

- but what about for an individual?
- over a span of 20 years, the probability that **all** your dreams are non predictive is

$$P(\text{non predictive}) = (0.96415)^{20} = 0.481$$

- which means that the probability of having a predictive dream is

$$P(\text{predictive}) = 1.0 - 0.481 = 0.519$$

- better than 50% chance!
- It might be unusual to **not** have had a predictive dream!

SHARED BIRTHDAYS

- ever been amazed to find that a group of people has two members with a shared birthday?
- you shouldn't be; it is not much of a coincidence
- Consider that a year has 366 days (counting February 29)
- to be **certain** that a group of people has a common birthday you would need a group of size 367
- what if we were willing to be just 50% certain of a shared birthday? How big would the group need to be?
- the surprising answer is 23

SHARED BIRTHDAYS

- what is the probability that a group of 23 people have no shared birthdays?
- how many ways to have birthdates from 23 people?

$$(366)^{23} = 9.1214727 \times 10^{58}$$

- How many ways to have 23 birthdates with no shared birthdays?

$$366 \times 365 \times 364 \times ... \times 344 = 4.5030611 \times 10^{58}$$

SHARED BIRTHDAYS

- probability of no shared birthdays is number of ways to have no shared birthdays divided by number of ways to have birthdays

$$P(\text{no shared}) = \frac{4.5030 \times 10^{58}}{9.1214 \times 10^{58}} = 0.4936$$

- so the probability of at least one shared birthday is

$$P(\text{shared}) = 1.0 - 0.4936 = 0.5063$$

- just about 50%
- Test it!

CON GAMES

- Here is a game that is played on the streets of some cities
- A man has 3 cards
  - Card 1: Black on both sides.
  - Card 2: Red on both sides.
  - Card 3: Black on one side and red on the other.
- He drops the cards in a hat, turns around and asks you to pick a card. Then he asks you to show him only one side of the card.
- Suppose you show him a red side. Now the man knows that the card cannot be Card 1 (black on both sides) and the card in your hand must be either Card 2 or Card 3.
- He offers you a bet of even money that he can guess the card. Is this a fair bet?

CON GAMES

- It might seem that this is fair. After all, the card in your hand is either Card 2 or Card 3. He has a 50% chance of guessing correctly, right?
- No.

CON GAMES

- Given that you have shown him one red side he knows that what you have shown is either:
  - first side of card 2
  - second side of card 2
  - red side of card 3
- Thus, of the possibilities, two are consistent with his guess of Card 2, and only one is consistent with your option of Card 3. He wins two-thirds of the time.

CONCLUSIONS

- probability
- apply to lots of situations
- coincidences are not as interesting as you might expect

NEXT TIME

- Decision making from noisy data
- Signal detection

Is that your phone?