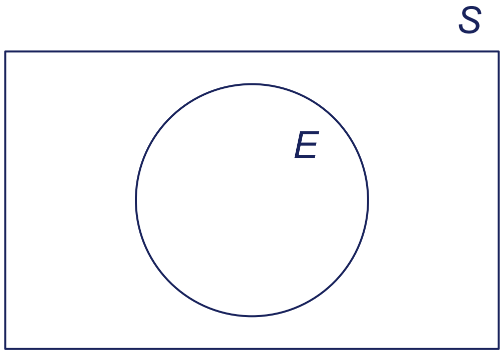
MIT PRIMES Circle

Probability Theory: Why You Are Falsely Convicted and Lonely

By Elena Baskakova and Alice He
Mentor Jeremy Smithline
May 21, 2022

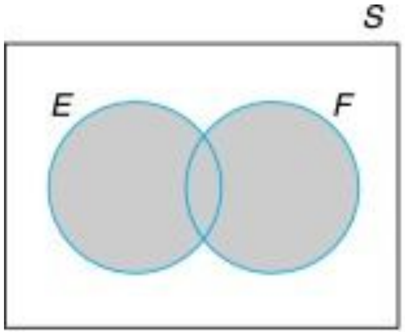
Definitions

- **Sample Space *(S)*** - the set of all possible outcomes
- **Event *(E)*** - any subset of outcomes within the sample space
- ***P(E)*** - the probability that the outcome of the experiment is contained in *E*

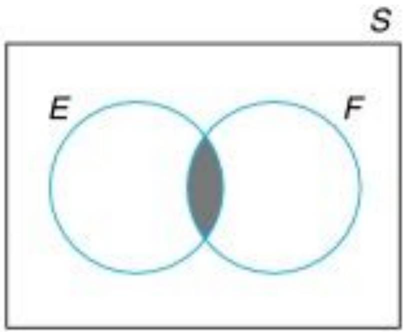


Definitions

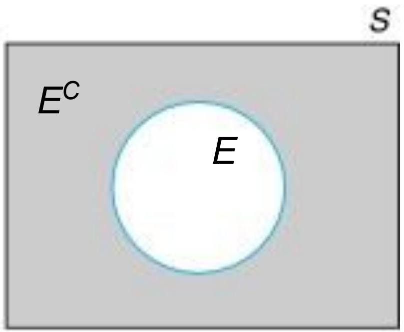
- **Union ($E \cup F$)** - the set of outcomes contained in either *E, F,* or both
- **Intersection ($E \cap F$ or EF)** - the set of outcomes contained in both *E* and *F*
- **Complement (E^C)** - the set of all outcomes in the sample space *S* that are *not* contained in E
  - *E^C occurs if and only if E does not!*
  - $E + E^C = S$



union ($E \cup F$)



intersection ($E \cap F$)



complement (E^C)

independence & dependence

given events E and F, does knowing that one has already occurred affect the probability of the other one occurring?

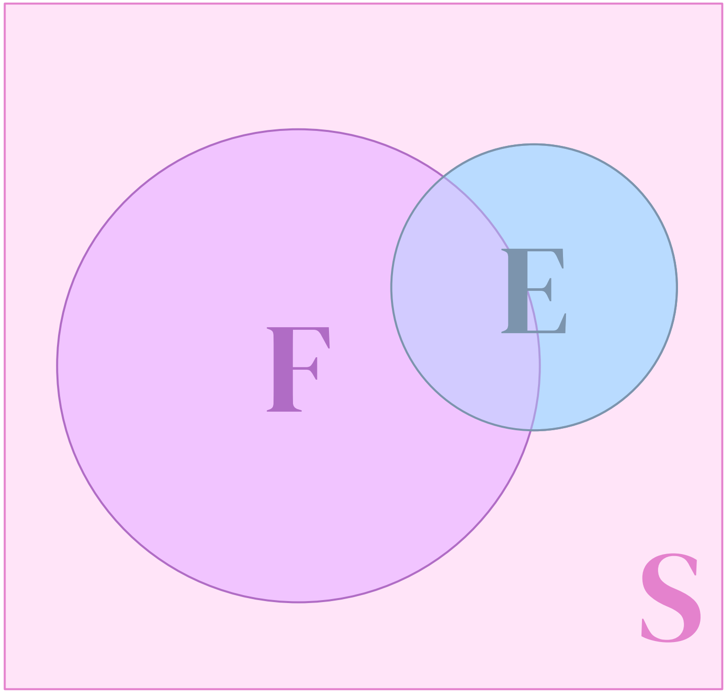
in general,

two events are said to be **independent** if $P(E)P(F) = P(EF)$, and **dependent** if this equation does not hold.

Conditional Probability

$P(E|F)$ denotes the probability that event E occurs given that F occurred.

$$P(E|F) = \frac{P(E \cap F)}{P(F)}.$$



Bayes' Formula

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

Bayes's formula describes the probability of a event based on some prior knowledge about the conditions in which it occurs.

Prosecutor's Fallacy

Associative evidence: evidence about 'matches'

Incidence rate: rarity of these factors occurring in general population

The fallacy: "If the probability of having all the factors is very low, and the defendant has them, then they're very probably guilty"

People vs. Collins

Eyewitness description:

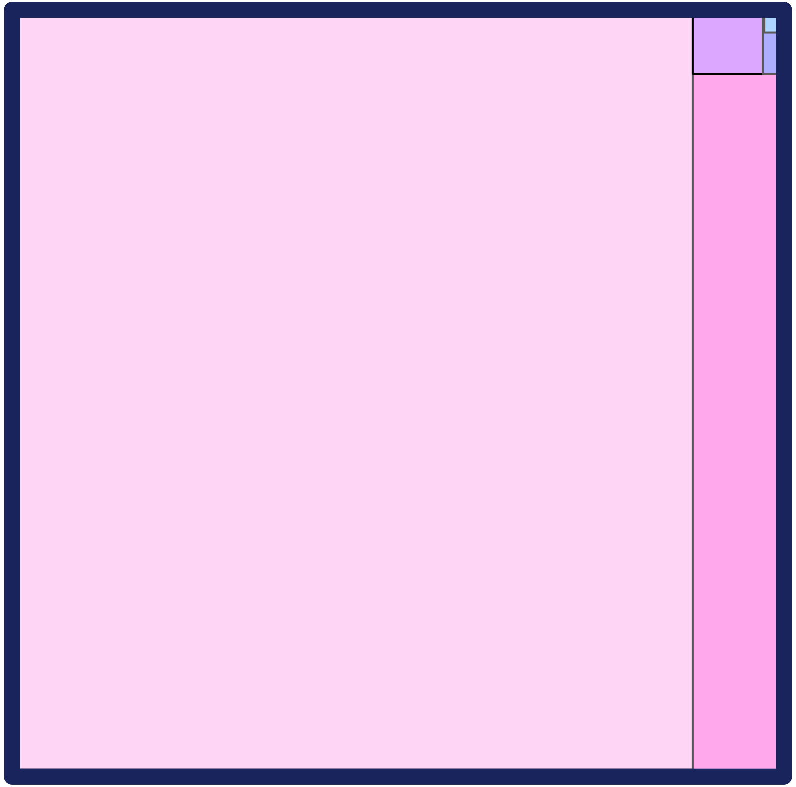
- Young White woman
- About 145 pounds, ordinary build
- Wearing something dark
- Dark blond hair in a ponytail
- Yellow car
- Black man, who had a mustache and beard

Mathematician's Testimony

$$\frac{1}{10} \times \frac{1}{10} \times \frac{1}{5} \times \frac{1}{200} = \frac{1}{100,000}$$

... times some other things gives

1 in 12,000,000



“Something like one in a billion”

What went wrong?

- Assumed independence
- Answering the wrong question!

The Question Answered:

what is the probability that a random innocent couple matches description?

$$P(\text{Match the Description}|\text{Innocent})$$

The Real Question:

what is the probability that a couple that matches the description is innocent?

$$P(\text{Innocent}|\text{Match the Description})$$

Using *Bayes' theorem*, we are looking for:

$$\frac{P(\text{Innocent})P(\text{Match the description}|\text{Innocent})}{P(\text{Match the Description})}.$$

Data assumed

– 5 million couples in California in 1964

– 1 in 1 million couples matched the description

– 1 guilty couple that matches

	Guilty	Not Guilty
Match	1	4
Don't Match	0	4,999,995

$$\frac{P(\text{Innocent})P(\text{Match the description}|\text{Innocent})}{P(\text{Match the Description})}.$$

$P(\text{Innocent})$ .... $\frac{4,999,999}{5,000,000}$,

$P(\text{Match the description} \mid \text{Innocent})$ .... $\frac{4}{4,999,999}$,

$P(\text{Match the description})$.. $\frac{5}{5,000,000}$.

$$\text{P(Innocent|Match)} = \frac{4,999,999}{5,000,000} \cdot \frac{4}{4,999,999} \cdot \frac{5,000,000}{5} = \frac{4}{5}.$$

$$\frac{4}{5} \neq \frac{1}{12,000,000}!$$

random variables

any function of the outcome of an experiment that takes on values with defined probabilities

expected value *E[X]*

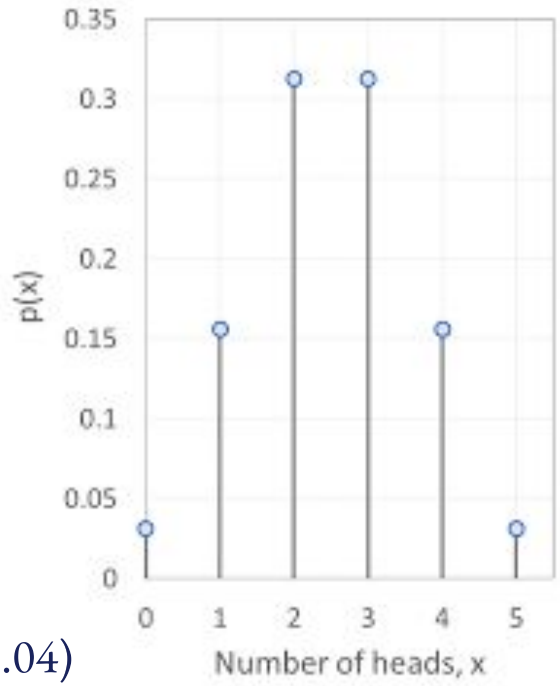
the weighted average of all possible values of *X*, with the weight of each value being the probability that *X* assumes it

$$E[X] = \sum_{x:p(x)>0} xp(x)$$

x = the value that X takes

$p(x)$ = the probability that X takes on the value x

graph at right: $E[X] = 0(0.04) + 1(0.16) + 2(0.31) \ldots 5(0.04)$



expected value *E[X]*

Find *E[X]*, where *X* is the outcome of one roll of a fair die.

- *X* is a discrete random variable with possible values 1, 2, 3, 4, 5, and 6
- fair die → probability of rolling any number is ⅙

$$E[X] = 1\left(\frac{1}{6}\right) + 2\left(\frac{1}{6}\right) + 3\left(\frac{1}{6}\right) + 4\left(\frac{1}{6}\right) + 5\left(\frac{1}{6}\right) + 6\left(\frac{1}{6}\right) = \frac{7}{2}$$

same idea applies to the

expectation of a function of a random variable, E[f(X)]

$$E[f(X)] = \sum_i f(x_i)p(x_i)$$

variance

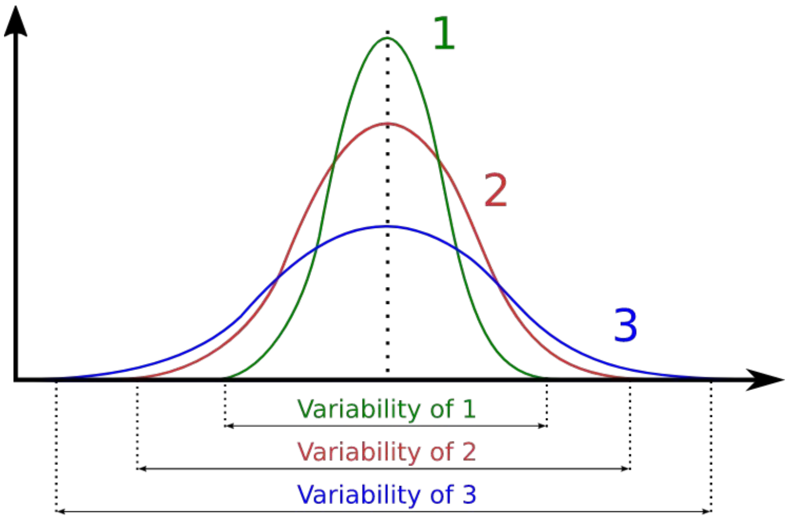
tells us about the *spread* of the possible values of X

$$\text{Var}(X) = E[X^2] - (E[X])^2$$

since variance is nonnegative,

$$\text{Var}(X) \geq 0 \quad \longrightarrow \quad E[X^2] - (E[X])^2 \geq 0$$

$$E[X^2] \geq (E[X])^2$$

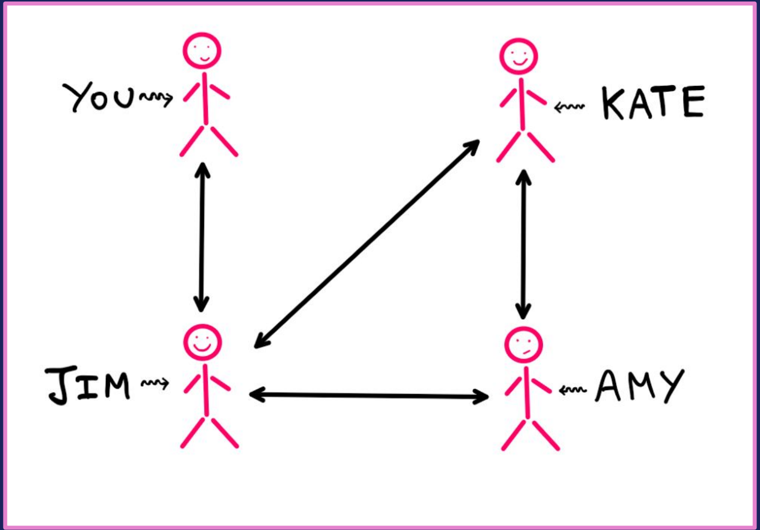


the friendship paradox

the *friendship paradox* states that on average, your friends have more friends than you do. is this true, and if so, why?

the friendship paradox

- *n* students at a school, each person numbered 1, 2, 3 . . . *n*
- *f(i)* = number of friends of person *i*
- $t = \sum_{i=1}^{n} f(i)$ = total number of one-way friendships at school



8 one-way friendships

if we choose a random individual X . . .

- $E[f(X)]$ = expected number of friends of X

$$E[f(X)] = \sum_{i=1}^{n} f(i)P\{X = i\}$$
$$= \sum_{i=1}^{n} f(i) \cdot \frac{1}{n}$$
$$= \frac{t}{n}$$

- **now, everyone writes down all their friends, one name per sheet of paper**
  - a person with $f(i)$ friends will use $f(i)$ sheets of paper
  - there will be $f(i)$ sheets of paper with person i's name
  - total of t sheets of paper

the friendship paradox

Let Y be the name on a randomly chosen sheet of paper, and $E[f(Y)]$ be the expected number of friends of that person.

- person i's name appears on $f(i)$ out of t sheets of paper

$$P\{Y = i\} = \frac{f(i)}{t},$$

where $i = 1, 2, \ldots n$

- thus, expected number of friends of Y:

$$E[f(Y)] = \sum_{i=1}^{n} f(i) P\{Y = i\}$$

$$= \sum_{i=1}^{n} f(i) \cdot \frac{f(i)}{t}$$

$$= \sum_{i=1}^{n} \frac{f^2(i)}{t}$$

to summarize:

$$E[f(X)] = \frac{t}{n}$$

$$E[f(Y)] = \sum_{i=1}^{n} \frac{f^2(i)}{t}$$

hello variance!

$E[f^2(X)]$ = expectation of the square of the number of friends of X

$$E[f^2(X)] = \sum_{i=1}^{n} f^2(i) P\{X = i\}$$
$$= \sum_{i=1}^{n} f^2(i) \cdot \frac{1}{n}$$
$$= \sum_{i=1}^{n} \frac{f^2(i)}{n}$$

thus, we have:

$$\frac{E[f^2(X)]}{E[f(X)]} = \frac{\sum_{i=1}^{n} f^2(i)}{t}$$

the friendship paradox

The Power of Variance:

$$E[f(Y)] = \frac{E[f^2(X)]}{E[f(X)]} \geq E[f(X)] \quad \longrightarrow \quad \boxed{E[f(Y)] \geq E[f(X)]}$$

average # of friends of random friend ≥ average # of friends of a random individual

The Intuitive Reasoning

the math, explained in less mathematical terms

- X = randomly chosen individual
  - **equally likely** to be any of the n students
- Y = random friend selected from the sheets of paper
  - probability that person i is picked is proportional to number of slips with their name
  - ***Y* is biased to those with more friends**
- naturally then, $E[Y] \geq E[X]$

the tl; dr

people with more friends are more likely to be your friend, and similarly, you are less likely to be friends with someone who has very few friends.

thus, on average, your friends tend to have more friends than you do.

many thanks to...

our mentor Jeremy

the PRIMES Circle program

our families

Sheldon Ross, for his book *A First Course in Probability*