Multiple events, conditioning, and independence

DSE 210

People's probability judgements

Experiment by Kahneman-Tversky. Subjects were told:

> *Linda is 31, single, outspoken, and very bright. She majored in philosophy in college. As a student, she was deeply concerned with racial discrimination and other social issues, and participated in anti-nuclear demonstrations.*

They were then asked to rank three possibilities:

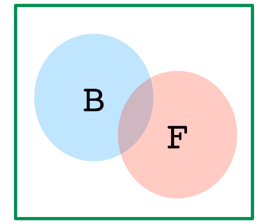
(a) Linda is active in the feminist movement.

(b) Linda is a bank teller.

(c) Linda is a bank teller and is active in the feminist movement.

Over 85% respondents chose (a) > (c) > (b).

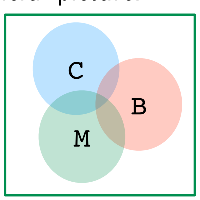
But:
$\Pr(\text{bank teller, feminist}) \leq \Pr(\text{bank teller})$.



In a city, 60% of people have a car, 20% of people have a bike, and 10% of people have a motorcycle. Anyone without at least one of these walks to work. What is the minimum fraction of people who walk to work?

Let $\Omega = \{\text{people in the town}\}$. Let
$C = \{\text{has car}\}, B = \{\text{has bike}\}, M = \{\text{has motorcyle}\}, W = \{\text{walks}\}$.

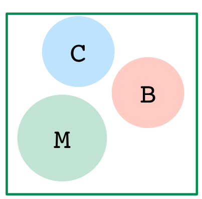
General picture:



$$\Pr(W) \geq 1 - \Pr(C \cup B \cup M)$$

$$\begin{aligned}\Pr(C \cup B \cup M) &\leq \Pr(C) + \Pr(B) + \Pr(M) \\ &= 0.6 + 0.2 + 0.1 = 0.9\end{aligned}$$

and thus $\Pr(W) \geq 0.1$.



Complements and unions

The complement of an event.

Let Ω be a sample space and $E \subset \Omega$ an event.
Write E^c for the event that E does not occur, that is, $E^c = \Omega \setminus E$.

$$\Pr(E^c) = 1 - \Pr(E).$$

The union bound.

For any events $E_1, \ldots, E_k$:

$$\Pr(E_1 \cup \cdots \cup E_k) \leq \Pr(E_1) + \cdots + \Pr(E_k).$$

This inequality is exact when the events are disjoint.

Coupon-collector problem

Each cereal box has one of k action figures. How many boxes do you need to buy so that you are likely to get all k figures?

Say we buy n boxes.
Let A_i be the event that the ith action figure is *not* obtained.

$$\begin{aligned}\Pr(A_i) &= \Pr(\text{not in 1st box}) \cdot \Pr(\text{not in 2nd box}) \cdots \Pr(\text{not in } n\text{th box}) \\ &= \left(1 - \frac{1}{k}\right)^n \le e^{-n/k}\end{aligned}$$

By union bound, the probability of missing some figure is

$$\Pr(A_1 \cup \cdots \cup A_k) \le \Pr(A_1) + \cdots + \Pr(A_k) \le ke^{-n/k}.$$

Setting $n \ge k \ln 2k$ makes this $\le 1/2$.

Therefore: enough to buy $O(k \log k)$ cereal boxes.

Pregnancy test.

The following data is obtained on a pregnancy test:

$\Omega = \{\text{women who use the test}\}$
$P = \{\text{women using the test who are actually pregnant}\}$
$T = \{\text{women for whom the test comes out positive}\}$

Suppose $T \subset P$ and $\Pr(P) = 0.4$ and $\Pr(T) = 0.3$.

Suppose the test comes out positive. What is the chance of pregnancy?
Exactly 1.

Suppose the test comes out negative. What is the chance of pregnancy?

$$\Pr(P|T^c) = \frac{\Pr(P \cap T^c)}{\Pr(T^c)} = \frac{\Pr(P) - \Pr(T)}{1 - \Pr(T)} = \frac{0.1}{0.7} = \frac{1}{7}.$$

Conditional probability

You meet a stranger at a bar. What is the chance he votes Republican?

Just use the average for your town: 0.5, say.

Now suppose you find out he plays tennis.

Sample space $\Omega = \{\text{all people in your town}\}$
Two events of interest:
$R = \{\text{votes Republican}\}$
$T = \{\text{plays tennis}\}$
What is $\Pr(R|T)$?

Formula for conditional probability:

$$\Pr(R|T) = \frac{\Pr(R \cap T)}{\Pr(T)} = \frac{\#\text{ people who vote Republican and play tennis}}{\#\text{ people who play tennis}}.$$

Rolls of a die.

You roll a die twice. What is the probability that the sum is ≥ 10:

If the first roll is 6?

$$\Pr(\text{sum} \ge 10 | \text{first} = 6) = \Pr(\text{second} \ge 4) = \frac{1}{2}.$$

If the first roll is ≥ 3?

$$\begin{aligned}\Pr(\text{sum} \ge 10 | \text{first} \ge 3) &= \frac{\Pr(\text{sum} \ge 10, \text{first} \ge 3)}{\Pr(\text{first} \ge 3)} \\ &= \frac{\Pr(\{(4,6),(5,5),(5,6),(6,4),(6,5),(6,6)\})}{2/3} = \frac{1}{4}.\end{aligned}$$

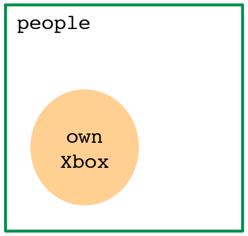
If the first roll is < 6?

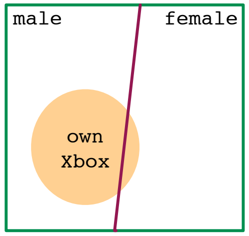
$$\begin{aligned}\Pr(\text{sum} \ge 10 | \text{first} < 6) &= \frac{\Pr(\text{sum} \ge 10, \text{first} < 6)}{\Pr(\text{first} < 6)} \\ &= \frac{\Pr(\{(5,5),(5,6),(4,6)\})}{5/6} = \frac{1}{10}.\end{aligned}$$

Summation rule

Breaking down a probability into disjoint pieces.

Example: what fraction of people own an Xbox?





$$\begin{aligned}&\Pr(\text{own Xbox})\\&= \Pr(\text{own Xbox, male}) + \Pr(\text{own Xbox, female})\\&= \Pr(\text{own Xbox}|\text{male})\Pr(\text{male}) + \Pr(\text{own Xbox}|\text{female})\Pr(\text{female})\end{aligned}$$

Suppose events $A_1, \ldots, A_k$ are disjoint and $A_1 \cup \cdots \cup A_k = \Omega$: that is, one of these events must occur. Then for any other event E,

$$\begin{aligned}\Pr(E) &= \Pr(E, A_1) + \Pr(E, A_2) + \cdots + \Pr(E, A_k)\\&= \Pr(E|A_1)\Pr(A_1) + \Pr(E|A_2)\Pr(A_2) + \cdots + \Pr(E|A_k)\Pr(A_k)\end{aligned}$$

The Monty Hall game

Three doors: one has a treasure chest behind it and the other two have goats. You pick a door and indicate it to Monty. He opens one of the other two doors to reveal a goat. Now, should you stick to your initial choice, or switch to the other unopened door?

You should switch.

First argument:

$\Pr(\text{initial choice has treasure}) = 1/3$
No matter what Monty does, he can't change this fact. So
$\Pr(\text{other unopened door has treasure}) = 2/3$

Second argument:

$$\begin{aligned}&\Pr(\text{treasure in other door})\\&= \Pr(\text{treasure in other door}|\text{initial choice correct})\Pr(\text{initial choice correct}) +\\&\quad \Pr(\text{treasure in other door}|\text{initial choice wrong})\Pr(\text{initial choice wrong})\\&= 0 \cdot \frac{1}{3} + 1 \cdot \frac{2}{3} = \frac{2}{3}.\end{aligned}$$

Sex bias in graduate admissions

In 1969, there were 12673 applicants for graduate study at Berkeley. 44% of the male applicants were accepted, and 35% of the female applicants.

Define:

- $\Omega = \{\text{all applicants}\}$
- $M = \{\text{male applicants}\}$
- What is M^c? $M^c = \{\text{female applicants}\}$
- $A = \{\text{accepted applicants}\}$

What do the percentages 44% and 35% correspond to?

$\Pr(A|M) = 0.44$ and $\Pr(A|M^c) = 0.35$.

The administration found, however, that in every department, the accept rate for female applicants was at least as high as the accept rate for male applicants. How could this be?

Bayes' rule

Pearl: You wake up in the middle of the night to the shrill sound of your burglar alarm. What is the chance that a burglary has been attempted?

The facts:

- There is a 95% chance that an attempted burglary will trigger the alarm.
$$\Pr(\text{alarm}|\text{burglary}) = 0.95$$
- There is a 1% chance of a false alarm.
$$\Pr(\text{alarm}|\text{no burglary}) = 0.01$$
- Based on local crime statistics, there is a 1-in-10,000 chance that a given house will be burglarized on a given night.
$$\Pr(\text{burglary}) = 10^{-4}$$

We need to compute

$$\Pr(\text{burglary}|\text{alarm}) = \frac{\Pr(\text{burglary, alarm})}{\Pr(\text{alarm})} = \frac{\Pr(\text{alarm}|\text{burglary})\Pr(\text{burglary})}{\Pr(\text{alarm})}$$

We need to compute

$$\Pr(\text{burglary}|\text{alarm}) = \frac{\Pr(\text{alarm}|\text{burglary})\Pr(\text{burglary})}{\Pr(\text{alarm})}$$

Now,

$$\Pr(\text{alarm}) = \Pr(\text{alarm}|\text{burglary})\Pr(\text{burglary}) + \Pr(\text{alarm}|\text{no burglary})\Pr(\text{no burglary})$$

Therefore,

$$\Pr(\text{burglary}|\text{alarm}) = \frac{0.95 \times 10^{-4}}{0.95 \times 10^{-4} + 0.01 \times (1 - 10^{-4})} \approx 0.00941$$

The alarm increases one's belief in a burglary hundredfold, from $1/10000$ to roughly $1/100$.

Bayes' rule:

$$\Pr(H|E) = \frac{\Pr(E|H)}{\Pr(E)}\Pr(H).$$

The three prisoners

Three prisoners – A, B, C – are in a jail one night and one of them (they don't know whom) will be declared guilty and executed in the morning. Racked by worry, prisoner A calls the prison guard and begs to be told whether he is the unlucky one. The guard is not allowed to tell him – but he can say only that B will be declared innocent. Now A thinks to himself, "previously my chance of being executed was $1/3$, and now, because of an innocuous inquiry, it seems to have gone up to $1/2$. How can this be?"

Analyze using these events:
G_A = the event that A will be declared guilty
I_B = the event that the guard, when prompted, will declare B innocent

$$\Pr(G_A|I_B) = \frac{\Pr(I_B|G_A)\Pr(G_A)}{\Pr(I_B)} = \frac{\frac{1}{2}\cdot\frac{1}{3}}{\frac{1}{2}} = \frac{1}{3}$$

ndependence

Two events A, B are **independent** if the probability of B occurring is the same whether or not A occurs.

Example: toss two coins.
$A = \{\text{first coin is heads}\}$
$B = \{\text{second coin is heads}\}$

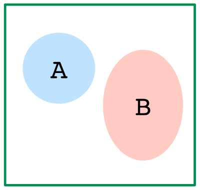
Formally, we say A, B are independent if $\Pr(A \cap B) = \Pr(A)\Pr(B)$.

The independence of A and B implies:

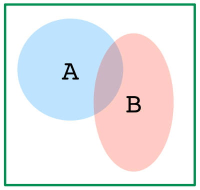
- $\Pr(A|B) = \Pr(A)$
- $\Pr(B|A) = \Pr(B)$
- $\Pr(A|B^c) = \Pr(A)$

Examples: independent or not?

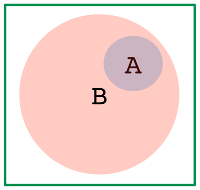
1. You have two children.
   $A = \{\text{first child is a boy}\}$, $B = \{\text{second child is a girl}\}$.
   Independent.
2. You throw two dice.
   $A = \{\text{first is a six}\}$, $B = \{\text{sum} > 10\}$.
   Not independent.
3. You get dealt two cards at random from a deck of 52.
   $A = \{\text{first is a heart}\}$, $B = \{\text{second is a club}\}$.
   Not independent: $\Pr(A) = \Pr(B) = 1/4$, but
   $$\Pr(A \cap B) = \frac{1}{4}\cdot\frac{13}{51} > \Pr(A)\Pr(B).$$
4. You are dealt two cards.
   $A = \{\text{first is a heart}\}$, $B = \{\text{second is a 10}\}$.
   Independent: $\Pr(A) = 1/4, \Pr(B) = 1/13$, and
   $$\Pr(A \cap B) = \frac{12}{52}\cdot\frac{4}{51} + \frac{1}{52}\cdot\frac{3}{51} = \frac{1}{52} = \Pr(A)\Pr(B).$$



Not independent



Possibly independent



Not independent