Conditional Probability and Independence

CS70: Discrete Mathematics and Probability Theory

UC Berkeley – Summer 2026

Lecture 17

Ref: Note 14

Lecture 17: More Basic Probability

Today: High likelihood of covering more probability theory

Many important topics:

- Conditional Probability and Bayesian Inference
- Total Probability
- Bayes' Rule
- Independent Events and Event Intersection
- Intersection for Non-Independent Events
- Event Union and Inclusion-Exclusion

And lots of examples...

Note: Two days for Lecture 17 (and Note 14)

Probability Basics: Quick Review Quiz

What is a probability space?

- A set and a function on the elements. Yes
- The values of the function are positive integers. No
- The values of the function are real numbers. Yes
- An element of the set is an outcome. Yes
- There is an experiment associated with a probability space. Yes
- The values in the set are integers. No

A probability space is a set Ω of outcomes and a function $\mathbb{P} : \Omega \rightarrow \mathbb{R}$ such that

- (Non-negativity): $0 \leq \mathbb{P}[\omega] \leq 1$ for all $\omega \in \Omega$
- (Sum to 1): $\sum_{\omega \in \Omega} \mathbb{P}[\omega] = 1$

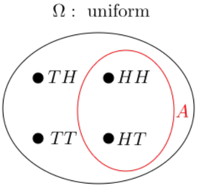
Conditional Probability: The Idea

Knowledge is power

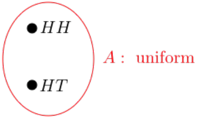
Two coin flips. Uniform space: $\Omega = \{HH, HT, TH, TT\}$

Event $B =$ both heads. $\implies \mathbb{P}[B] = 1/4$

Event $A =$ first flip is heads.



What if you know A happened? New sample space A ; still uniform.



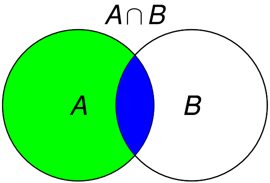
The probability of two heads *given that* the first flip is heads?

“Probability of B given A.” Notation: $\mathbb{P}[B \mid A] = 1/2$

Conditional Probability: The Definition

Definition: For events A and B, with $\mathbb{P}[A] > 0$, the **conditional probability** of B given A is

$$\mathbb{P}[B|A] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[A]}$$



In A!

In B?

Must be in $A \cap B$.

$$\mathbb{P}[B|A] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[A]}.$$

Brief Side Track: Gambler's Fallacy

Prev example: Probability that 2nd flip is heads given that 1st is heads... $= 1/2$
$\Rightarrow$ *Next outcome isn't affected by previous*

Gambler's Fallacy: "I've been losing so long I'm due a win!"
$\Rightarrow$ Reality: "losing so long" is completely irrelevant

Taken to an extreme: Flip a fair coin 51 times
A = "first 50 flips are heads"
B = "the 51st is heads"

What is $\mathbb{P}[B|A]$?
$A = \{\mathtt{HH\cdots HT}, \mathtt{HH\cdots HH}\}$ (*only two possibilities!*)
$B \cap A = \{\mathtt{HH\cdots HH}\}$ (only one of the two ends in $\mathtt{H}$)

Uniform probability space: $\mathbb{P}[B|A] = \frac{|B \cap A|}{|A|} = \frac{1}{2}$ Same as $\mathbb{P}[B]$

The likelihood of 51st heads does not depend on the previous flips.

Concept we'll see later: *independence*

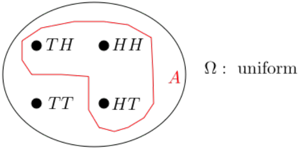
Conditional Probability: Different Coin Problem

Two coin flips – uniform distribution of outcomes.

$\Omega = \{HH, HT, TH, TT\}$; $\quad \mathbb{P}[\omega] = 1/4$ for all $\omega \in \Omega$

Events:

$A =$ at least one flip is heads $\qquad B =$ both flips are heads



$\mathbb{P}[A] = 3/4 \qquad \mathbb{P}[B] = 1/4$

What if we know that A happened? What about $\mathbb{P}[B|A]$?

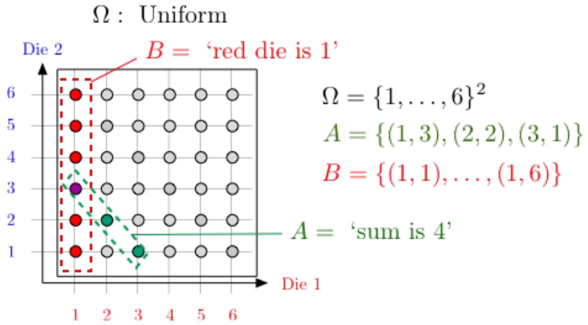
$\Rightarrow$ One possibility out of three equally likely outcomes, so $\mathbb{P}[B|A] = 1/3$

The math: $\mathbb{P}[B|A] = \frac{\mathbb{P}[B \cap A]}{\mathbb{P}[A]} = \frac{1/4}{3/4} = \frac{1}{3}$

Another puzzle: *If a person has two kids, and you know one is a boy, what is the probability that the other is a boy?*

Conditional Probability: An Example with Dice

Toss a red and a blue die, sum is 4 — What is probability that red is 1?



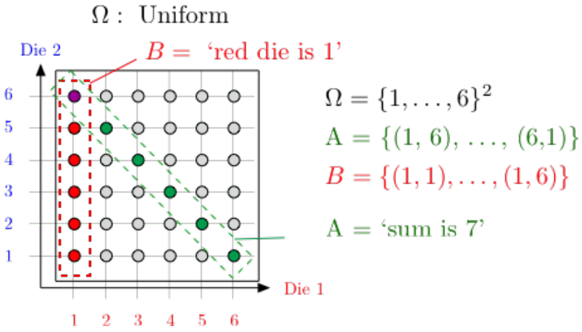
$\mathbb{P}[B|A] = \frac{|B \cap A|}{|A|} = \frac{|\{(1,3)\}|}{|\{(1,3),(2,2),(3,1)\}|} = \frac{1}{3}$

With no knowledge about sum: $\mathbb{P}[B] = \frac{|B|}{|\Omega|} = \frac{6}{36} = \frac{1}{6}$

$\Rightarrow$ *B* is more likely given *A*.

Conditional Probability: Dice with Different Sum

Toss a red and a blue die, sum is 7 — What is probability that red is 1?



$\mathbb{P}[B|A] = \frac{|B \cap A|}{|A|} = \frac{1}{6}$

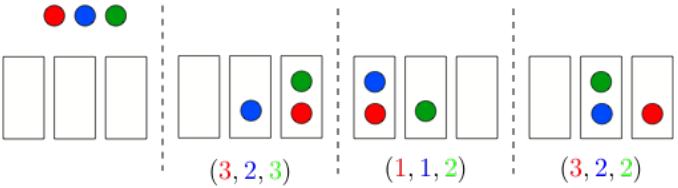
With no knowledge about sum: $\mathbb{P}[B] = \frac{|B|}{|\Omega|} = \frac{6}{36} = \frac{1}{6}$

$\Rightarrow$ Observing A does not change your mind about the likelihood of B.

Conditional Probability: Balls and Bins

Toss 3 balls into 3 bins. Events: A =“1st bin empty” ; B =“2nd bin empty.”

$$\Omega = \{1, 2, 3\}^3$$



$$\omega = (\text{bin of red ball}, \text{bin of blue ball}, \text{bin of green ball})$$

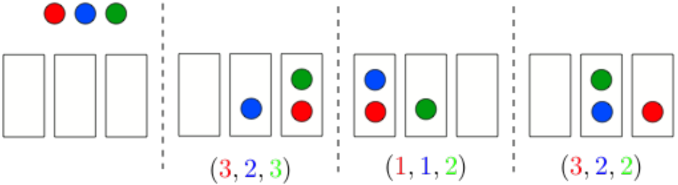
What is $\mathbb{P}[A \mid B]$?

(A) 1/27
(B) 8/27
(C) 1/8
(D) 0
(E) 2

Conditional Probability: Balls and Bins

Toss 3 balls into 3 bins. Events: A =“1st bin empty” ; B =“2nd bin empty.”

$$\Omega = \{1,2,3\}^3$$



$$\omega = (\text{bin of red ball}, \text{bin of blue ball}, \text{bin of green ball})$$

What is $\mathbb{P}[A \mid B]$?

$\mathbb{P}[A \cap B]$ (all balls in bin 3) $= \mathbb{P}[\{(3,3,3)\}] = \frac{1}{|\Omega|} = \frac{1}{27}$

$\mathbb{P}[B]$ (only bins 1 and 3) $= \mathbb{P}[\{(a,b,c) \mid a,b,c \in \{1,3\}] = \mathbb{P}[\{1,3\}^3] = \frac{2^3}{|\Omega|} = \frac{8}{27}$

$\therefore \quad \mathbb{P}[A|B] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]} = \frac{1/27}{8/27} = \frac{1}{8}$ – compare to $\mathbb{P}[A] = \frac{8}{27}$

A is less likely given B: Second bin is empty $\Longrightarrow$ first is less likely to be empty.

Conditional Probability: A Non-Uniform Example

Experiment: Roll a single weighted 6-sided die.

Probability space:

Roll (ω)	1	2	3	4	5	6
$\mathbb{P}[\omega]$	1/2	1/10	1/10	1/10	1/10	1/10

Is this a valid probability space? Yes!

Events: $A =$ "rolled a 1" ; $B =$ "rolled a 2"

If we know we *didn't* roll a 1, what's the probability we rolled a 2?
$\Rightarrow$ Mathematically, we want $\mathbb{P}[B|\overline{A}]$

$\mathbb{P}[B \cap \overline{A}] =$ "rolled a 2 and not a 1"
Question: Does "not a 1" change anything? No!
So: $\mathbb{P}[B \cap \overline{A}] = \mathbb{P}[B] = 1/10$

So: $\mathbb{P}[B|\overline{A}] = \frac{\mathbb{P}[B \cap \overline{A}]}{\mathbb{P}[\overline{A}]} = \frac{1/10}{1/2} = \frac{1}{5}$

Consider: How does this compare to the same conditional probability with a fair die?

Conditional Probability: Counting Cards

Dealing from a standard, shuffled deck of cards.

Events:

$A =$ "first card is a 10-point card (10, J, Q, K)" *4 cards in each of 4 suits*

$B =$ "second card is a 10-point card"

$\mathbb{P}[A] = \mathbb{P}[B] = \frac{16}{52} \approx 0.3077$

$\mathbb{P}[A \cap B] = \frac{16 \cdot 15}{52 \cdot 51}$ and $\mathbb{P}[\overline{A} \cap B] = \frac{36 \cdot 16}{52 \cdot 51}$

$\mathbb{P}[B|A] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[A]} = \frac{(16 \cdot 15)/(52 \cdot 51)}{16/52} = \frac{15}{51} \approx 0.2941$

$\mathbb{P}[B|\overline{A}] = \frac{\mathbb{P}[\overline{A} \cap B]}{\mathbb{P}[\overline{A}]} = \frac{(36 \cdot 16)/(52 \cdot 51)}{36/52} = \frac{16}{51} \approx 0.3137$

Knowledge of first card affects probability of "big card" on second

$\Rightarrow$ This is a super-simplified version of "card counting" (blackjack)

Question: Does "I've gotten low cards for a long time, so I'm due a big card" violate the gambler's fallacy?

No! Finite set of cards used without replacement, not *de novo* games.

A Historical Note...

Can you know enough to tilt blackjack odds in your favor (instead of house)?

We Asked a Founder of the MIT Blackjack Team How to Get Rich Beating Casinos

You too can win a fortune counting cards. Probably.

Image via okan akdeniz/Getty Images

For most people, gambling is a fast way to lose money while slurping down free (well, sort of free) drinks. But for a few years in the 1980s, Bill Kaplan and a few associates figured out a way to clean house and rake in cash, and thus, the legend of the MIT Blackjack Team was born. Helmed by Kaplan (who didn't actually go to MIT), the MIT team used card counting and a strict system to carve out an advantage against the dealer that they ruthlessly exploited. With the Encore Boston Harbor Casino throwing open its doors, I wondered: Do you have to be an MIT-level genius to pull off a scheme like this, or can any schlub with a will to learn and dollar signs in his eyes make do? Kaplan was kind enough to hop on the phone and break down just what it takes to beat the house.

Fascinating story, but coordination between multiple players is key...

Modern card-counting mitigations include multiple decks, frequent reshuffling, ...

Recall the Shopping Center Puzzle

Shopping center 1:
3 Indian restaurants:
2 Moroccan restaurants:

Shopping center 2:
1 Indian restaurant:
1 Moroccan restaurants:

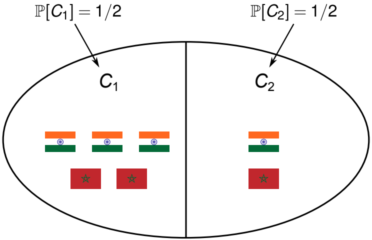
Joe picks a center randomly and then a restaurant at that center randomly

$\Rightarrow$ *In this case, "randomly" means "uniformly" (all outcomes equally likely)*

Joe comes back and says he ate at a Moroccan restaurant

What was the probability that he went to shopping center 1?

Recall the Shopping Center Puzzle



View sample space

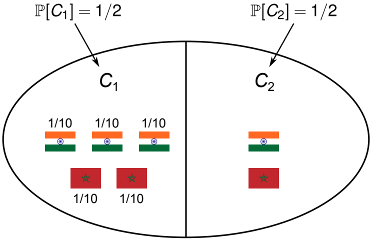
Events: $C_k =$ Joe chose Center k

Joe chooses Center randomly first, so $\mathbb{P}[C_1] = \mathbb{P}[C_2] = 1/2$

Joe chooses a restaurant uniformly next

What is the probability of each restaurant in C_1?

Recall the Shopping Center Puzzle

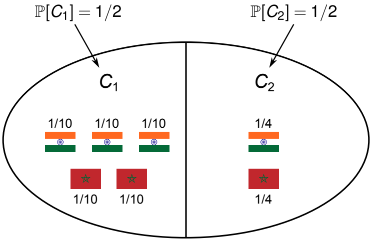


The $1/2$ probability in event C_1 is spread out evenly between 5 restaurants

$\Rightarrow$ Each probability is $1/10$

What is the probability of each restaurant in C_2?

Recall the Shopping Center Puzzle



Joe said he ate at a Moroccan restaurant ( – event M)

⇒ *What's the probability he went to Center 1?*

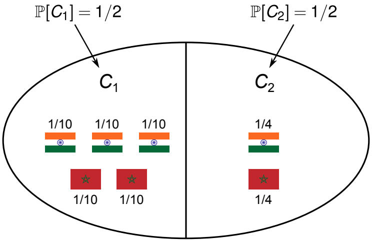
⇒ This is $\mathbb{P}[C_1 \mid M] = \frac{\mathbb{P}[C_1 \cap M]}{\mathbb{P}[M]}$

$\mathbb{P}[C_1 \cap M] = 1/10 + 1/10 = 2/10$

$\mathbb{P}[M] = 1/10 + 1/10 + 1/4 = 9/20$

$$\mathbb{P}[C_1 \mid M] = \frac{2/10}{9/20} = \frac{4/20}{9/20} = \frac{4}{9}$$

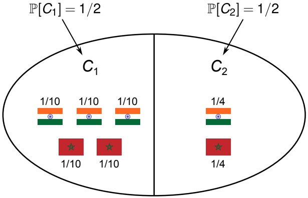
Recall the Shopping Center Puzzle

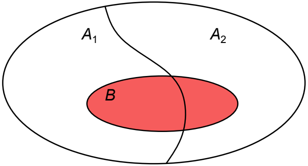


$$\mathbb{P}[C_1 \mid M] = \frac{2/10}{9/20} = \frac{4/20}{9/20} = \frac{4}{9}$$

- Considered an experiment with a sequence of choices, and calculated the prob of *unseen initial choice* given final output. *This is very useful!*
- This is the “balls and bins” problem in Note 14 – re-framing with shopping centers due to Josh Hug.

Generalizing what we just did....





C_1 and C_2 disjoint and equally likely

Restaurants uniform in center

Calculated:

$$\mathbb{P}[M] = 1/10 + 1/10 + 1/4 = 9/20$$

$$\mathbb{P}[C_1 \mid M] = \frac{\mathbb{P}[C_1 \cap M]}{\mathbb{P}[M]} = \frac{\mathbb{P}[M \mid C_1] \cdot \mathbb{P}[C_1]}{\mathbb{P}[M]}$$

Events A_1 and A_2 **partition** space

Know $\mathbb{P}[B \mid A_1]$ and $\mathbb{P}[B \mid A_2]$

Calculate (we'll see why later...)

$$\mathbb{P}[B] = \mathbb{P}[B \mid A_1] \cdot \mathbb{P}[A_1] + \mathbb{P}[B \mid A_2] \cdot \mathbb{P}[A_2]$$

$$\mathbb{P}[A_1 \mid B] = \frac{\mathbb{P}[A_1 \cap B]}{\mathbb{P}[B]} = \frac{\mathbb{P}[B \mid A_1] \cdot \mathbb{P}[A_1]}{\mathbb{P}[B]}$$

Bayesian Inference: *How does knowledge of B affect probability of A_1 ?*

Two General Rules

Total Probability Rule: For any events $A_1, A_2, \ldots, A_n$ that partition the sample space Ω,

$$\mathbb{P}[B] = \sum_{j=1}^{n} \mathbb{P}[B \mid A_j] \cdot \mathbb{P}[A_j] \,.$$

Note: From conditional probability formula, this is just $\sum_{j=1}^{n} \mathbb{P}[B \cap A_j]$

Bayes' Rule: For any events A and B with non-zero probabilities,

$$\mathbb{P}[A \mid B] = \frac{\mathbb{P}[B \mid A] \cdot \mathbb{P}[A]}{\mathbb{P}[B]} \,.$$

Note that if $B_1, \ldots, B_n$ partition Ω, then we can use the Total Probability Rule to calculate the denominator so that for any B_i:

$$\mathbb{P}[A_i \mid B] = \frac{\mathbb{P}[B \mid A_i] \cdot \mathbb{P}[A_i]}{\sum_{j=1}^{n} \mathbb{P}[B \mid A_j] \cdot \mathbb{P}[A_j]} \,.$$

Total Probability Example

Playing Dungeons and Dragons

Roll an 8-sided die (a “D8”)...

- Roll 1 – 2: Encounter a Goblin (event G), killed with probability 2/20
- Roll 3 – 6: Encounter an Ogre (event O), killed with probability 8/20
- Roll 7 – 8: Encounter Kraken (event K), killed with probability 19/20

Events G, O, and K partition Ω. Event D is “you die.” What do we know?

- $\mathbb{P}[G] = 2/8$; $\mathbb{P}[D|G] = 2/20$
- $\mathbb{P}[O] = 4/8$; $\mathbb{P}[D|O] = 8/20$
- $\mathbb{P}[K] = 2/8$; $\mathbb{P}[D|K] = 19/20$

What is the probability that you are killed?

$$\begin{aligned}\mathbb{P}[D] &= \mathbb{P}[D|G]\cdot\mathbb{P}[G] + \mathbb{P}[D|O]\cdot\mathbb{P}[O] + \mathbb{P}[D|K]\cdot\mathbb{P}[K] \\ &= \frac{2}{20}\cdot\frac{2}{8} + \frac{8}{20}\cdot\frac{4}{8} + \frac{19}{20}\cdot\frac{2}{8} = \frac{4+32+38}{160} = \frac{74}{160}\end{aligned}$$

Beware the Kraken!

Recall the COVID Test Puzzle

Company BigPharma developed a new test for COVID-19. Tests show that:

- Given to a person with COVID-19, says “positive” 90% of the time
- Given to a person without COVID-19, says “positive” 20% of the time
  - $\Rightarrow$ These are called *false positives*
- 5% of the population has COVID-19

Question: If you test “positive”, what is the chance you are actually infected?

Spoiler: Not nearly as high as you might think...

COVID Test: What Information Do You Need?

Want to know: Given a positive test result, what is probability infected?

Events: $P =$ positive test result ; $C =$ has COVID

We know:

- Given to a person with COVID-19, says "positive" 90% of the time
  - $\Rightarrow$ So says "negative" 10% of the time – false negatives
  - $\Rightarrow$ $?\mathbb{P}[P|C] = 9/10$
- Given to a person without COVID-19, says "negative" 80% of the time
  - $\Rightarrow$ So says "positive" 20% of the time – false positives
  - $\Rightarrow$ $?\mathbb{P}[P|\overline{C}] = 2/10$
- 5% of the population has COVID-19
  - $\Rightarrow$ $?\mathbb{P}[C] = 1/20$

We want: $\mathbb{P}[C|P]$

Since C and $\overline{C}$ partition Ω, we have everything we need
(*and we need everything we're given!*)

COVID Test: What Information Do You Need?

Events: $P =$ positive test result ; $C =$ has COVID

We know:

- Given to a person with COVID-19, says "positive" 90% of the time
  $\Rightarrow \quad \mathbb{P}[P|C] = 9/10$
- Given to a person without COVID-19, says "negative" 80% of the time
  $\Rightarrow \quad \mathbb{P}[P|\overline{C}] = 2/10$
- 5% of the population has COVID-19
  $\Rightarrow \quad \mathbb{P}[C] = 1/20$

We want: $\mathbb{P}[C|P] = \frac{\mathbb{P}[P|C] \cdot \mathbb{P}[C]}{\mathbb{P}[P]}$ (Bayes' Rule)

Numerator (all known): $\mathbb{P}[P|C] \cdot \mathbb{P}[C] = (9/10) \cdot (1/20) = 9/200$

Denominator (Total Probability):

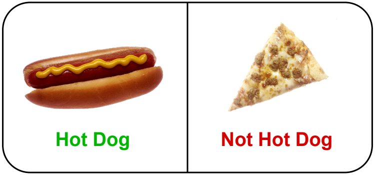
$$\mathbb{P}[P] = \mathbb{P}[P|C] \cdot \mathbb{P}[C] + \mathbb{P}[P|\overline{C}] \cdot \mathbb{P}[\overline{C}] = (9/10) \cdot (1/20) + (2/10) \cdot (19/20) = 47/200$$

So: $\mathbb{P}[C|P] = \frac{9/200}{47/200} = \frac{9}{47} \approx 0.1915$ (or $\approx 19\%$)

Machine Learning Classifiers

You want to test an image classifier that identifies hot dogs

Ref Silicon Valley, S4E4: `https://www.youtube.com/watch?v=ACmydtFDTGs`



Events: $C =$ classified as hot dog ; $H =$ is a hot dog

Want: $\mathbb{P}[H \mid C]$

1. Run tests with lots of pictures of hot dogs to estimate $\mathbb{P}[C \mid H]$
2. Run tests with lots of pictures of not hot dogs to estimate $\mathbb{P}[C \mid \overline{H}]$
3. Estimate how many picture of hot dogs there are $\mathbb{P}[H]$

Same exact problem as COVID tests!

Independence

Idea: Events A and B are independent if probability of A doesn't change based on B

$\Rightarrow$ In other words, $\mathbb{P}[A|B] = \mathbb{P}[A]$

Formal definition is a little different...

Definition: Events A and B in the same probability space are **independent** if $\mathbb{P}[A \cap B] = \mathbb{P}[A] \cdot \mathbb{P}[B]$

Mostly the same...

$$\mathbb{P}[A \cap B] = \mathbb{P}[A] \cdot \mathbb{P}[B] \implies \mathbb{P}[A|B] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]} = \frac{\mathbb{P}[A] \cdot \mathbb{P}[B]}{\mathbb{P}[B]} = \mathbb{P}[A]$$

So why just "*Mostly*"?

$\Rightarrow$ What if $\mathbb{P}[B] = 0$?

Remember: Conditional probability $\mathbb{P}[A|B]$ is only defined for $\mathbb{P}[B] > 0$

Independence: Example

Independent Events in a Uniform Probability Space

Two independent flips of a fair coin – H_i = "heads on flip i

$\mathbb{P}[H_1 \cap H_2] = \mathbb{P}[H_1] \cdot \mathbb{P}[H_2] = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$

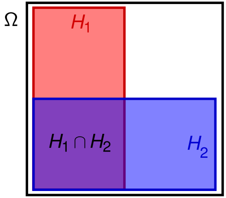
Same for $\mathbb{P}[H_1 \cap \overline{H_2}]$, $\mathbb{P}[\overline{H_1} \cap H_2]$, *and* $\mathbb{P}[\overline{H_1} \cap \overline{H_2}]$

In general:

Uniform experiments over sample space Ω_1, for all $\omega \in \Omega_1$: $\mathbb{P}[\omega] = \frac{1}{|\Omega_1|}$

Do it twice: $\Omega_2 = \Omega_1 \times \Omega_1$, and for all $\omega \in \Omega_2$: $\mathbb{P}[\omega] = \frac{1}{|\Omega_2|} = \frac{1}{|\Omega_1|^2}$

What do independent events in a uniform sample space look like?



Proportion of H_1 inside H_2 same as proportion of H_1 in Ω

Proportion of H_2 inside H_1 same as proportion of H_2 in Ω

Independence: Example

Independent Events in a Non-Uniform Probability Space

Two independent flips of a *biased* coin, with $\mathbb{P}[H_i] = \frac{2}{3}$

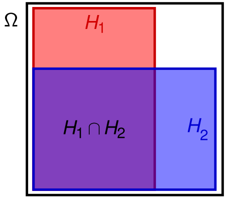
Result HH: $\mathbb{P}[H_1 \cap H_2] = \mathbb{P}[H_1] \cdot \mathbb{P}[H_2] = \frac{2}{3} \cdot \frac{2}{3} = \frac{4}{9}$

Result HT: $\mathbb{P}[H_1 \cap \overline{H_2}] = \frac{2}{3} \cdot \frac{1}{3} = \frac{2}{9}$

Result TH: $\mathbb{P}[\overline{H_1} \cap H_2] = \frac{1}{3} \cdot \frac{2}{3} = \frac{2}{9}$

Result TT: $\mathbb{P}[\overline{H_1} \cap \overline{H_2}] = \frac{1}{3} \cdot \frac{1}{3} = \frac{1}{9}$

What do independent events in a non-uniform sample space look like?



Proportion of H_1 inside H_2 same as proportion of H_1 in Ω

Proportion of H_2 inside H_1 same as proportion of H_2 in Ω

Independence with More Events

Definition: A set of events $E_1, E_2, \ldots, E_n$ is pairwise independent if for every $i \neq j$, we have $\mathbb{P}[E_i \cap E_j] = \mathbb{P}[E_i] \cdot \mathbb{P}[E_j]$.

Definition: A set of events $E_1, E_2, \ldots, E_n$ is mutually independent if for every subset $I \subseteq \{1, \ldots, n\}$ with size $|I| \geq 2$,

$$\mathbb{P}\left[\bigcap_{i \in I} E_i\right] = \prod_{i \in I} \mathbb{P}[E_i].$$

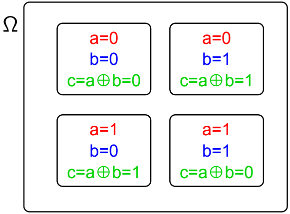
A set of events that is mutually independent is also pairwise independent. The converse is *not* necessarily true though.

Note: You can also define k-wise independent if needed.

Points in (t, n)-threshold secret sharing are t-wise independent.

Independence with More Events: Example

Consider: Two random uniform bits a and b, and a computed bit $c = a \oplus b$.



Each outcome $\omega \in \Omega$ has values for a, b, and the computed value for c.

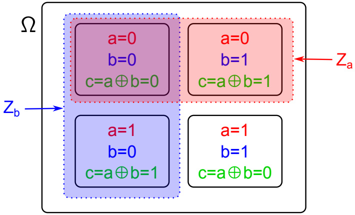
$|\Omega| = 4$ uniform so $\mathbb{P}[\omega] = 1/4$ for every outcome

Events for every bit being zero: $Z_a =$"a is 0" $Z_b =$"b is 0" $Z_c =$"c is 0"

Question: Where are events on the probability space diagram?

Independence with More Events: Example

Consider: Two random uniform bits a and b, and a computed bit $c = a \oplus b$.



Events: $Z_a =$"a is 0" $Z_b =$"b is 0" $Z_c =$"c is 0"

$\mathbb{P}[Z_a] = 1/2 \quad \mathbb{P}[Z_b] = 1/2 \quad \mathbb{P}[Z_c] = 1/2$

$$\left.\begin{array}{l} \mathbb{P}[Z_a \cap Z_b] = 1/4 = \mathbb{P}[Z_a] \cdot \mathbb{P}[Z_b] \\ \mathbb{P}[Z_b \cap Z_c] = 1/4 = \mathbb{P}[Z_b] \cdot \mathbb{P}[Z_c] \\ \mathbb{P}[Z_a \cap Z_c] = 1/4 = \mathbb{P}[Z_a] \cdot \mathbb{P}[Z_c] \end{array}\right\} \quad \text{Pairwise independent}$$

$$\left.\begin{array}{l} \mathbb{P}[Z_a \cap Z_b \cap Z_c] = 1/4 \\ \mathbb{P}[Z_a] \cdot \mathbb{P}[Z_b] \cdot \mathbb{P}[Z_c] = 1/8 \end{array}\right\} \quad \text{NOT 3-wise independent}$$

Non-Independent Intersection

Product Rule for 3 Events

Recall the definition for two events A and B (with $\mathbb{P}[A] > 0$):

$$\mathbb{P}[B|A] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[A]}.$$

Hence,

$$\mathbb{P}[A \cap B] = \mathbb{P}[A] \cdot \mathbb{P}[B|A].$$

Consider three events: E_1, E_2, and E_3:

$$\mathbb{P}[E_1 \cap E_2 \cap E_3] = \mathbb{P}[(E_1 \cap E_2) \cap E_3]$$

Now use the above formula with $A = E_1 \cap E_2$ and $B = E_3$:

$$\mathbb{P}[(E_1 \cap E_2) \cap E_3] = \mathbb{P}[E_1 \cap E_2] \cdot \mathbb{P}[E_3 | E_1 \cap E_2]$$

And finally apply again for first probability (with $A = E_1$ and $B = E_2$):

$$\mathbb{P}[E_1 \cap E_2 \cap E_3] = \mathbb{P}[E_1] \cdot \mathbb{P}[E_2|E_1] \cdot \mathbb{P}[E_3|E_1 \cap E_2]$$

This is the product rule for three events.

Think about: what if E_1, E_2, and E_3 are mutually independent?

Non-Independent Intersection

Product Rule for n Events

Theorem (Product Rule): Let $E_1, E_2, \ldots, E_n$ be events. Then

$$\mathbb{P}[E_1 \cap \cdots \cap E_n] = \mathbb{P}[E_1] \cdot \mathbb{P}[E_2 | E_1] \cdot \mathbb{P}[E_3 | E_1 \cap E_2] \cdots \mathbb{P}[E_n | E_1 \cap \cdots \cap E_{n-1}].$$

Proof: We prove this by induction on n.

Base Case ($n = 2$): Follows from conditional probability definition.

Induction Hypothesis: Assume the theorem holds for $n = k$.

Inductive Step: We prove the holds at $n = k + 1$:

$$\mathbb{P}[E_1 \cap \cdots \cap E_k \cap E_{k+1}] = \mathbb{P}[(E_1 \cap \cdots \cap E_k) \cap E_{k+1}]$$

By conditional probability definition:

$$= \mathbb{P}[E_1 \cap \cdots \cap E_k] \cdot \mathbb{P}[E_{k+1} | E_1 \cap \cdots \cap E_k]$$

Using the induction hypothesis:

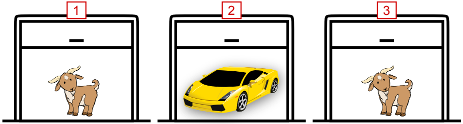
$$= (\mathbb{P}[E_1] \cdot \mathbb{P}[E_2 | E_1] \cdots \mathbb{P}[E_k | E_1 \cap \cdots \cap E_{k-1}]) \cdot \mathbb{P}[E_{k+1} | E_1 \cap \cdots \cap E_k],$$

so that the result holds at $n = k + 1$. □

Monty Hall... Using Conditional Probabilities

Events:

C_1 = "you choose door 1"
H_3 = "host opens door 3"
P_2 = "prize is behind door 2"



Deciding whether to switch, want $\mathbb{P}[P_2|C_1 \cap H_3] = \frac{C_1 \cap P_2 \cap H_3}{C_1 \cap H_3}$

For $\mathbb{P}[C_1 \cap P_2 \cap H_3]$: Are these events independent?

$\Rightarrow$ No! Host won't open contestant's door *or* the door with the prize

So... $\mathbb{P}[H_3|C_1 \cap P_2] = 1$

Non-independent events – use the product rule:

$$\mathbb{P}[C_1 \cap P_2 \cap H_3] = \mathbb{P}[C_1] \cdot \mathbb{P}[P_2|C_1] \cdot \mathbb{P}[H_3|C_1 \cap P_2] = \frac{1}{3} \cdot \frac{1}{3} \cdot 1 = \frac{1}{9}$$

What about $\mathbb{P}[C_1 \cap H_3] = \mathbb{P}[H_3|C_1] \cdot \mathbb{P}[C_1]$?

$\Rightarrow$ $\mathbb{P}[H_3|C_1] = \frac{1}{2}$ (2 or 3 opened, symmetric) so $\mathbb{P}[C_1 \cap H_3] = \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6}$

Conclusion: $\mathbb{P}[P_2|C_1 \cap H_3] = \frac{C_1 \cap P_2 \cap H_3}{C_1 \cap H_3} = \frac{1/9}{1/6} = \frac{2}{3}$

- $\Rightarrow$ So $2/3$ probability of winning if you switch in this situation
- $\Rightarrow$ Generalizing: Always choose one and host opens different
- $\Rightarrow$ Names (1 and 3) are arbitrary $\Rightarrow$ always prob prize is behind other

Union of Events

Problem: Roll a six-sided die until you get a 1 or 2 (“success”)

Event $S_i =$ “success on roll i”

$\mathbb{P}[S_1] = \frac{1}{3} \quad \mathbb{P}[S_2] = \frac{1}{3} \quad \mathbb{P}[S_3] = \frac{1}{3} \quad \mathbb{P}[S_4] = \frac{1}{3} \quad \dots$

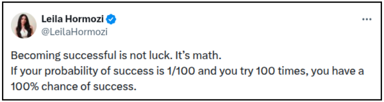
Success on first roll: $\mathbb{P}[S_1] = \frac{1}{3}$

Success in first two rolls? $\mathbb{P}[S_1 \cup S_2] = \frac{1}{3} + \frac{1}{3} = \frac{2}{3}$? No!

Success in first three rolls? $\mathbb{P}[S_1 \cup S_2 \cup S_3] = \frac{1}{3} + \frac{1}{3} + \frac{1}{3} = 1$?

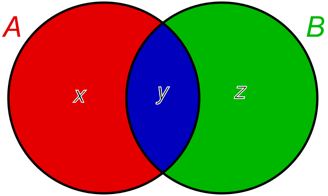
Take it one more... $\mathbb{P}[S_1 \cup S_2 \cup S_3 \cup S_4] = \frac{1}{3} + \frac{1}{3} + \frac{1}{3} + \frac{1}{3} = \frac{4}{3}$? ... What?

Why you shouldn’t learn math from “motivational speakers”:



Probability is really $\approx 1 - 1/e \approx 0.632$...

Event Union: Inclusion-Exclusion – 2 Events



$\mathbb{P}[A] = x + y$
$\mathbb{P}[B] = y + z$
$\mathbb{P}[A \cap B] = y$
$\mathbb{P}[A \cup B] = x + y + z$

If we just added up probabilities?

$$\mathbb{P}[A] + \mathbb{P}[B] = (x + y) + (y + z) = x + 2y + z$$

Added in y outcomes twice! So subtract out to correct....

$$\mathbb{P}[A \cup B] = x + y + z = \mathbb{P}[A] + \mathbb{P}[B] - \mathbb{P}[A \cap B]$$

This is the simple (two event) version of “inclusion-exclusion”

Inclusion-Exclusion Example

Problem: Roll a six-sided die until you get a 1 or 2 ("success")

Event $S_i =$ "success on roll i"

$\mathbb{P}[S_1] = \frac{1}{3} \quad \mathbb{P}[S_2] = \frac{1}{3} \quad \mathbb{P}[S_3] = \frac{1}{3} \quad \mathbb{P}[S_4] = \frac{1}{3} \quad \ldots$

Success on first roll: $\mathbb{P}[S_1] = \frac{1}{3}$?

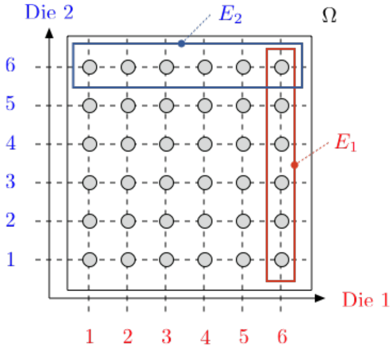
Success in first two rolls? $\mathbb{P}[S_1 \cup S_2] = \frac{1}{3} + \frac{1}{3} = \frac{2}{3}$? No!

Correct: $\mathbb{P}[S_1 \cup S_2] = \mathbb{P}[S_1] + \mathbb{P}[S_2] - \mathbb{P}[S_1 \cap S_2]$

Rolls are independent, so $\mathbb{P}[S_1 \cap S_2] = \mathbb{P}[S_1] \cdot \mathbb{P}[S_2] = \frac{1}{3} \cdot \frac{1}{3} = \frac{1}{9}$

So: $\mathbb{P}[S_1 \cup S_2] = \mathbb{P}[S_1] + \mathbb{P}[S_2] - \mathbb{P}[S_1 \cap S_2] = \frac{1}{3} + \frac{1}{3} - \frac{1}{9} = \frac{5}{9}$

Inclusion-Exclusion Example: Dice – Roll a Six



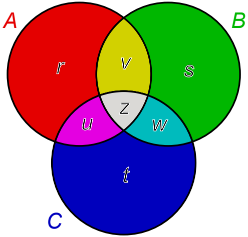
$$|E_1 \cup E_2| = |E_1| + |E_2| - |E_1 \cap E_2|$$

Events: $E_1 =$ Red die shows 6 $\qquad E_2 =$ Blue die shows 6

$E_1 \cup E_2 =$ At least one die shows 6

$$\mathbb{P}[E_1] = \frac{6}{36}, \quad \mathbb{P}[E_2] = \frac{6}{36}, \quad \mathbb{P}[E_1 \cap E_2] = \frac{1}{36}, \quad \mathbb{P}[E_1 \cup E_2] = \frac{6}{36} + \frac{6}{36} - \frac{1}{36} = \frac{11}{36}$$

Event Union: Inclusion-Exclusion – 3 Events



$\mathbb{P}[A] = r+u+v+z$
$\mathbb{P}[B] = s+v+w+z$
$\mathbb{P}[C] = t+u+w+z$
$\mathbb{P}[A\cap B] = v+z$
$\mathbb{P}[A\cap C] = u+z$
$\mathbb{P}[B\cap C] = w+z$
$\mathbb{P}[A\cap B\cap C] = z$

Adding: $$\begin{aligned}\mathbb{P}[A]+\mathbb{P}[B]+\mathbb{P}[C] &= (r+u+v+z)+(s+v+w+z)+(t+u+w+z) \\ &= r+s+t+2u+2v+2w+3z\end{aligned}$$

Oops... counted overlapping parts multiple times – take out intersections:

$$r+s+t+2u+2v+2w+3z-\mathbb{P}[A\cap B]-\mathbb{P}[A\cap C]-\mathbb{P}[B\cap C]$$

$$=r+s+t+2u+2v+2w+3z-(v+z)-(u+z)-(w+z)=r+s+t+u+v+w$$

Oops... now completely removed z, so add it back in:

$$=r+s+t+u+v+w+\mathbb{P}[A\cap B\cap C]$$

$$\Longrightarrow \mathbb{P}[A\cup B\cup C]=\mathbb{P}[A]+\mathbb{P}[B]+\mathbb{P}[C]-\mathbb{P}[A\cap B]-\mathbb{P}[A\cap C]-\mathbb{P}[B\cap C]+\mathbb{P}[A\cap B\cap C]$$

What a mess.... but there's a pattern!

Recall XOR Example...

Consider: Two random uniform bits a and b, and a computed bit $c = a \oplus b$.

Events: Z_a ="a is 0" Z_b ="b is 0" Z_c ="c is 0"

$\mathbb{P}[Z_a] = 1/2 \quad \mathbb{P}[Z_b] = 1/2 \quad \mathbb{P}[Z_c] = 1/2$

$$\left.\begin{array}{l} \mathbb{P}[Z_a \cap Z_b] = 1/4 = \mathbb{P}[Z_a] \cdot \mathbb{P}[Z_b] \\ \mathbb{P}[Z_b \cap Z_c] = 1/4 = \mathbb{P}[Z_b] \cdot \mathbb{P}[Z_c] \\ \mathbb{P}[Z_a \cap Z_c] = 1/4 = \mathbb{P}[Z_a] \cdot \mathbb{P}[Z_c] \end{array}\right\} \quad \text{Pairwise independent}$$

$$\left.\begin{array}{l} \mathbb{P}[Z_a \cap Z_b \cap Z_c] = 1/4 \\ \mathbb{P}[Z_a] \cdot \mathbb{P}[Z_b] \cdot \mathbb{P}[Z_c] = 1/8 \end{array}\right\} \quad \text{NOT 3-wise independent}$$

$\mathbb{P}[Z_a \cup Z_b \cup Z_c] = ?$

$\mathbb{P}[Z_a] + \mathbb{P}[Z_b] + \mathbb{P}[Z_c] = \frac{1}{2} + \frac{1}{2} + \frac{1}{2} = \frac{3}{2}$ obviously too big!

$\mathbb{P}[Z_a] + \mathbb{P}[Z_b] + \mathbb{P}[Z_c] - \mathbb{P}[Z_a \cap Z_b] - \mathbb{P}[Z_b \cap Z_c] - \mathbb{P}[Z_a \cap Z_c] = \frac{3}{2} - \frac{1}{4} - \frac{1}{4} - \frac{1}{4} = \frac{3}{4}$

$\mathbb{P}[Z_a] + \mathbb{P}[Z_b] + \mathbb{P}[Z_c] - \mathbb{P}[Z_a \cap Z_b] - \mathbb{P}[Z_b \cap Z_c] - \mathbb{P}[Z_a \cap Z_c] + \mathbb{P}[Z_a \cap Z_b \cap Z_c] = \frac{3}{4} + \frac{1}{4} = 1$

Recall XOR Example...

Consider: Two random uniform bits a and b, and a computed bit $c = a \oplus b$.

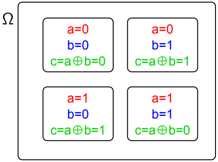
Events: $Z_a =$"a is 0" $Z_b =$"b is 0" $Z_c =$"c is 0"

$\mathbb{P}[Z_a] = 1/2$ $\mathbb{P}[Z_b] = 1/2$ $\mathbb{P}[Z_c] = 1/2$

$$\left.\begin{array}{l} \mathbb{P}[Z_a \cap Z_b] = 1/4 = \mathbb{P}[Z_a] \cdot \mathbb{P}[Z_b] \\ \mathbb{P}[Z_b \cap Z_c] = 1/4 = \mathbb{P}[Z_b] \cdot \mathbb{P}[Z_c] \\ \mathbb{P}[Z_a \cap Z_c] = 1/4 = \mathbb{P}[Z_a] \cdot \mathbb{P}[Z_c] \end{array}\right\} \quad \text{Pairwise independent}$$

$$\left.\begin{array}{l} \mathbb{P}[Z_a \cap Z_b \cap Z_c] = 1/4 \\ \mathbb{P}[Z_a] \cdot \mathbb{P}[Z_b] \cdot \mathbb{P}[Z_c] = 1/8 \end{array}\right\} \quad \text{NOT 3-wise independent}$$

$\mathbb{P}[Z_a \cup Z_b \cup Z_c] = 1$ – can this be right?



Every outcome has $a = 0$, $b = 0$, or $c = 0$

Event Union: Inclusion-Exclusion – *n* Events

Theorem (Inclusion-Exclusion): Let $E_1, \ldots, E_n$ be events in some probability space. Then

$$\mathbb{P}[E_1 \cup \cdots \cup E_n] = \sum_{k=1}^{n} (-1)^{k-1} \sum_{S \subseteq \{1,\ldots,n\} : |S|=k} \mathbb{P}[\cap_{i \in S} E_i].$$

Think it through:

- Iterates through subsets of size 1, 2, ...
- Alternates sign (add, subtract, add, subtract, ..)
- Starting (size 1) with positive (add)

Inclusion-Exclusion: Special Cases and Bounds

Two cases where the full inclusion-exclusion formula isn't needed.

We can't just add event probabilities, because this counts some outcomes multiple times. So the sum is too high – it makes a good upper-bound though!

Union Bound (or Boole's Inequality): Let $E_1, \ldots, E_n$ be events in some probability space. Then

$$\mathbb{P}[\bigcup_{i=1}^{n} E_i] \leq \sum_{i=1}^{n} \mathbb{P}[E_i].$$

If events don't overlap, this is a tight bound (nothing to subtract out!)...

Mutually Exclusive Events: Let $E_1, \ldots, E_n$ be *mutually exclusive* events in some probability space, so $E_i \cap E_j = \varnothing$ for all $i \neq j$. Then

$$\mathbb{P}[\bigcup_{i=1}^{n} E_i] = \sum_{i=1}^{n} \mathbb{P}[E_i].$$

Summary

Topics:

- Conditional Probability and Bayesian Inference

  $\mathbb{P}[A|B] = \frac{\mathbb{P}[A\cap B]}{\mathbb{P}[B]}$

- Total Probability

  $\mathbb{P}[A] = \mathbb{P}[A|B]\cdot\mathbb{P}[B] + \mathbb{P}[A|\overline{B}]\cdot\mathbb{P}[\overline{B}]$

- Bayes' Rule

  $\mathbb{P}[A|B] = \frac{\mathbb{P}[B|A]\cdot\mathbb{P}[A]}{\mathbb{P}[B]}$

- Independent Events and Event Intersection

  Independent A and B $\iff$ $\mathbb{P}[A\cap B] = \mathbb{P}[A]\cdot\mathbb{P}[B]$

- Intersection for Non-Independent Events (Product Rule)

  $\mathbb{P}[A\cap B\cap C] = \mathbb{P}[A]\cdot\mathbb{P}[B|A]\cdot\mathbb{P}[C|A\cap B]$

- Event Union and Inclusion-Exclusion

  $\mathbb{P}[A\cup B] = \mathbb{P}[A] + \mathbb{P}[B] - \mathbb{P}[A\cap B]$

Note: Formulas are only simplest forms, as reminders