Lecture 17: Continuing Probability.

Events, Conditional Probability, Independence, Bayes' Rule

Probability Basics Review

Setup:

- Random Experiment.
  Flip a fair coin twice.
- Probability Space.
  - **Sample Space:** Set of outcomes, Ω.
    $\Omega = \{HH, HT, TH, TT\}$
    (Note: Not $\Omega = \{H, T\}$ with two picks!)
  - **Probability:** $Pr[\omega]$ for all $\omega \in \Omega$.
    $Pr[HH] = \cdots = Pr[TT] = 1/4$
    1. $0 \le Pr[\omega] \le 1$.
    2. $\sum_{\omega \in \Omega} Pr[\omega] = 1$.
  - **Events.**
    Event $A \subseteq \Omega$, $Pr[A] = \sum_{\omega \in \Omega} Pr[\omega]$.

Probability is Additive

Theorem

(a) If events A and B are disjoint, i.e., $A \cap B = \emptyset$, then

$$Pr[A \cup B] = Pr[A] + Pr[B].$$

(b) If events $A_1, \ldots, A_n$ are pairwise disjoint,
i.e., $A_k \cap A_m = \emptyset, \forall k \ne m$, then

$$Pr[A_1 \cup \cdots \cup A_n] = Pr[A_1] + \cdots + Pr[A_n].$$

Proof:

(a) $Pr[A \cup B] = \sum_{\omega \in A \cup B} Pr[\omega]$
$= \sum_{\omega \in A} Pr[\omega] + \sum_{\omega in B} Pr[\omega]$ since $A \cap B = \emptyset$. $\quad = Pr[A] + Pr[B]$

(b) Either induction, or argue over sample points.

Consequences of Additivity

Theorem

(a) $Pr[A \cup B] = Pr[A] + Pr[B] - Pr[A \cap B]$;
(inclusion-exclusion property)

(b) $Pr[A_1 \cup \cdots \cup A_n] \le Pr[A_1] + \cdots + Pr[A_n]$;
(union bound)

(c) If $A_1, \ldots A_N$ are a partition of Ω, i.e.,
pairwise disjoint and $\cup_{m=1}^{N} A_m = \Omega$, then

$$Pr[B] = Pr[B \cap A_1] + \cdots + Pr[B \cap A_N].$$

(law of total probability)

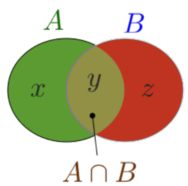
Proof:

(b) follows from the fact that every $\omega \in A_1 \cup \cdots A_n$ is included at least once in the right hand side.

Proofs for (a) and (c)? Next...

Inclusion/Exclusion

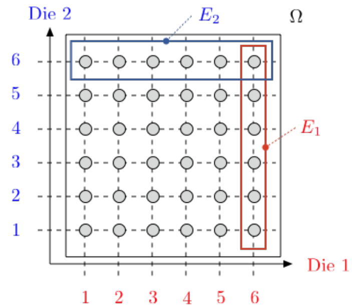
$$Pr[A \cup B] = Pr[A] + Pr[B] - Pr[A \cap B]$$



$$Pr[A] = x + y$$
$$Pr[B] = y + z$$
$$Pr[A \cap B] = y$$
$$Pr[A \cup B] = x + y + z$$

Another view. Any $\omega \in A \cup B$ is in $A \cap \overline{B}$, $A \cup B$, or $\overline{A} \cap B$. So, add it up.

Roll a Red and a Blue Die.



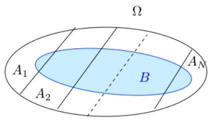
$$|E_1 \cup E_2| = |E_1| + |E_2| - |E_1 \cap E_2|$$

E_1 = 'Red die shows 6'; E_2 = 'Blue die shows 6'
$E_1 \cup E_2$ = 'At least one die shows 6'

$Pr[E_1] = \frac{6}{36}, Pr[E_2] = \frac{6}{36}, Pr[E_1 \cup E_2] = \frac{11}{36}$.

Total probability

Assume that Ω is the union of the disjoint sets $A_1, \ldots, A_N$.



Then,

$$Pr[B] = Pr[A_1 \cap B] + \cdots + Pr[A_N \cap B].$$

Indeed, B is the union of the disjoint sets $A_n \cap B$ for $n = 1, \ldots, N$.

In "math": $\omega \in B$ is in exactly one of $A_i \cap B$.

Adding up probability of them, get $Pr[\omega]$ in sum.

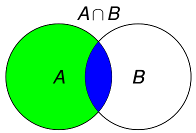
..Did I say...

Add it up.

Conditional Probability.

Definition: The **conditional probability** of B given A is

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

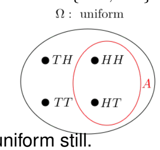


In A!
In B?
Must be in $A \cap B$.

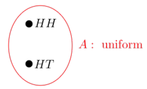
$$Pr[B|A] = \frac{Pr[A \cap B]}{Pr[A]}.$$

Conditional probability: example.

Two coin flips. First flip is heads. Probability of two heads?
$\Omega = \{HH, HT, TH, TT\}$; Uniform probability space.
Event A = first flip is heads: $A = \{HH, HT\}$.



New sample space: A; uniform still.



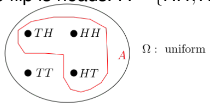
Event B = two heads.

The probability of two heads if the first flip is heads.
The probability of B given A is $1/2$.

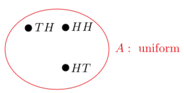
A similar example.

Two coin flips. At least one of the flips is heads.
→ Probability of two heads?

$\Omega = \{HH, HT, TH, TT\}$; uniform.
Event A = at least one flip is heads. $A = \{HH, HT, TH\}$.



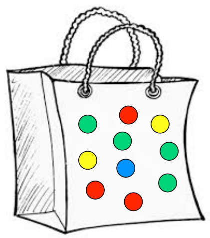
New sample space: A; uniform still.



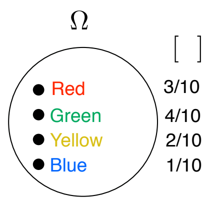
Event B = two heads.

The probability of two heads if at least one flip is heads.
The probability of B given A is $1/3$.

Conditional Probability: A non-uniform example



Physical experiment



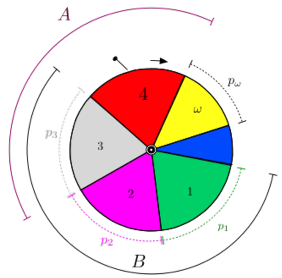
Probability model

Ω = {Red, Green, Yellow, Blue}

$$Pr[\text{Red}|\text{Red or Green}] = \frac{3}{7} = \frac{Pr[\text{Red} \cap (\text{Red or Green})]}{Pr[\text{Red or Green}]}$$

Another non-uniform example

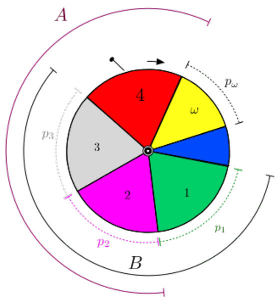
Consider $\Omega = \{1, 2, \ldots, N\}$ with $Pr[n] = p_n$.
Let $A = \{3, 4\}, B = \{1, 2, 3\}$.



$$Pr[A|B] = \frac{p_3}{p_1 + p_2 + p_3} = \frac{Pr[A \cap B]}{Pr[B]}.$$

Yet another non-uniform example

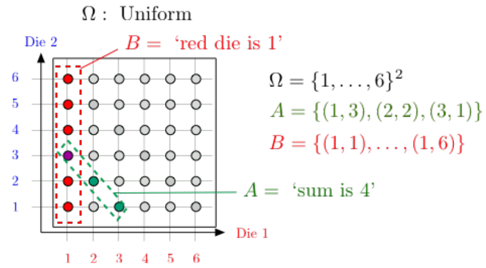
Consider $\Omega = \{1,2,\ldots,N\}$ with $Pr[n] = p_n$.
Let $A = \{2,3,4\}$, $B = \{1,2,3\}$.



$$Pr[A|B] = \frac{p_2 + p_3}{p_1 + p_2 + p_3} = \frac{Pr[A \cap B]}{Pr[B]}.$$

More fun with conditional probability.

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

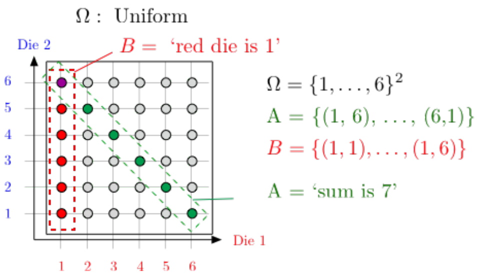


$Pr[B|A] = \frac{|B \cap A|}{|A|} = \frac{1}{3}$; versus $Pr[B] = 1/6$.

B is more likely given A.

Yet more fun with conditional probability.

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



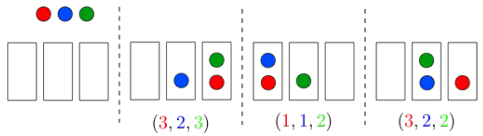
$Pr[B|A] = \frac{|B \cap A|}{|A|} = \frac{1}{6}$; versus $Pr[B] = \frac{1}{6}$.

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

Emptiness..

Suppose I toss 3 balls into 3 bins.
A ="1st bin empty"; B ="2nd bin empty." What is $Pr[A|B]$?

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



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

$Pr[B] = Pr[\{(a,b,c) \mid a,b,c \in \{1,3\}] = Pr[\{1,3\}^3] = \frac{8}{27}$

$Pr[A \cap B] = Pr[(3,3,3)] = \frac{1}{27}$

$Pr[A|B] = \frac{Pr[A \cap B]}{Pr[B]} = \frac{(1/27)}{(8/27)} = 1/8$; vs. $Pr[A] = \frac{8}{27}$.

A is less likely given B: If second bin is empty the first is more likely to have balls in it.

Gambler's fallacy.

Flip a fair coin 51 times.
A = "first 50 flips are heads"
B = "the 51st is heads"
$Pr[B|A]$?

$A = \{HH \cdots HT, HH \cdots HH\}$
$B \cap A = \{HH \cdots HH\}$

Uniform probability space.

$Pr[B|A] = \frac{|B \cap A|}{|A|} = \frac{1}{2}$.

Same as $Pr[B]$.

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

Product Rule

Recall the definition:

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

Hence,

$$Pr[A \cap B] = Pr[A]Pr[B|A].$$

Consequently,

$$\begin{aligned} Pr[A \cap B \cap C] &= Pr[(A \cap B) \cap C] \\ &= Pr[A \cap B]Pr[C|A \cap B] \\ &= Pr[A]Pr[B|A]Pr[C|A \cap B]. \end{aligned}$$

Product Rule

Theorem Product Rule
Let $A_1, A_2, \dots, A_n$ be events. Then

$$Pr[A_1 \cap \cdots \cap A_n] = Pr[A_1]Pr[A_2|A_1] \cdots Pr[A_n|A_1 \cap \cdots \cap A_{n-1}].$$

Proof: By induction.
Assume the result is true for n. (It holds for $n = 2$.) Then,

$$\begin{aligned} & Pr[A_1 \cap \cdots \cap A_n \cap A_{n+1}] \\ & = Pr[A_1 \cap \cdots \cap A_n] Pr[A_{n+1}|A_1 \cap \cdots \cap A_n] \\ & = Pr[A_1]Pr[A_2|A_1] \cdots Pr[A_n|A_1 \cap \cdots \cap A_{n-1}] Pr[A_{n+1}|A_1 \cap \cdots \cap A_n], \end{aligned}$$

so that the result holds for $n+1$. □

Correlation

An example.
Random experiment: Pick a person at random.
Event A: the person has lung cancer.
Event B: the person is a heavy smoker.

Fact:

$$Pr[A|B] = 1.17 \times Pr[A].$$

Conclusion:

- Smoking increases the probability of lung cancer by 17%.
- Smoking causes lung cancer.

Correlation

Event A: the person has lung cancer. Event B: the person is a heavy smoker. $Pr[A|B] = 1.17 \times Pr[A]$.

A second look.

Note that

$$\begin{aligned} Pr[A|B] = 1.17 \times Pr[A] & \Leftrightarrow \frac{Pr[A \cap B]}{Pr[B]} = 1.17 \times Pr[A] \\ & \Leftrightarrow Pr[A \cap B] = 1.17 \times Pr[A]Pr[B] \\ & \Leftrightarrow Pr[B|A] = 1.17 \times Pr[B]. \end{aligned}$$

Conclusion:

- Lung cancer increases the probability of smoking by 17%.
- Lung cancer causes smoking. Really?

Causality vs. Correlation

Events A and B are **positively correlated** if

$$Pr[A \cap B] > Pr[A]Pr[B].$$

(E.g., smoking and lung cancer.)

A and B being positively correlated does not mean that A causes B or that B causes A.

Other examples:

- Tesla owners are more likely to be rich. That does not mean that poor people should buy a Tesla to get rich.
- People who go to the opera are more likely to have a good career. That does not mean that going to the opera will improve your career.
- Rabbits eat more carrots and do not wear glasses. Are carrots good for eyesight?

Proving Causality

Proving causality is generally difficult. One has to eliminate external causes of correlation and be able to test the cause/effect relationship (e.g., randomized clinical trials).

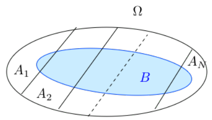
Some difficulties:

- A and B may be positively correlated because they have a common cause. (E.g., being a rabbit.)
- If B precedes A, then B is more likely to be the cause. (E.g., smoking.) However, they could have a common cause that induces B before A. (E.g., smart, CS70, Tesla.)

More about such questions later. For fun, check "N. Taleb: Fooled by randomness."

Total probability with Conditional Probability.

Assume that Ω is the union of the disjoint sets $A_1, \dots, A_N$.



Then,

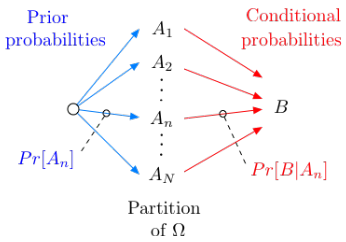
$$Pr[B] = Pr[A_1 \cap B] + \cdots + Pr[A_N \cap B].$$

Indeed, B is the union of the disjoint sets $A_n \cap B$ for $n = 1, \dots, N$. Thus,

$$Pr[B] = Pr[A_1]Pr[B|A_1] + \cdots + Pr[A_N]Pr[B|A_N].$$

Total probability

Assume that Ω is the union of the disjoint sets $A_1, \ldots, A_N$.



$Pr[B] = Pr[A_1]Pr[B|A_1] + \cdots + Pr[A_N]Pr[B|A_N]$.

Simple Bayes Rule.

$Pr[A|B] = \frac{Pr[A \cap B]}{Pr[B]}$, $Pr[B|A] = \frac{Pr[A \cap B]}{Pr[A]}$.

$Pr[A \cap B] = Pr[A|B]Pr[B] = Pr[B|A]Pr[A]$.

Bayes Rule: $Pr[A|B] = \frac{Pr[B|A]Pr[A]}{Pr[B]}$.

Lecture basically ended here.