Conditional Probability

- Many times, we want to consider the chance that an event A occurs under the assumption that B occurs.

- There are three possibilities:

  (1) The occurrence of B might increase the chances of A, so that A and B are positively correlated.

  (2) The occurrence of B might decrease the chances of A, so that A and B are negatively correlated.

  (3) The occurrence of B might not change the chances of A, so that A and B are independent.

Ex: Suppose we flip a fair coin 3 times.

Our sample space is

$$\Omega = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$$

and each outcome $\omega \in \Omega$ is equally likely.

Let B be the event that the first coin flip is heads. (Recall: this means $B = \{HHH, HHT, HTH, HTT\}$.)

Also, define events A_1, A_2, A_3 as follows:

A_1 = majority of the flips are heads

A_2 = majority of the flips are tails

A_3 = odd number of the flips are heads.

Then A_1 and B are positively correlated

A_2 and B are negatively correlated

A_3 and B are independent.

def The <u>probability of an event A given that an event B occurs</u>, written $\Pr(A|B)$ is defined as

$$\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)}.$$

(assuming $\Pr(B) > 0$.)

<u>Think</u>: $\Pr(A|B)$ means, "If I restrict myself to considering outcomes $\omega \in \Omega$ where B occurs (i.e. $\omega \in B$), what are the odds that A also occurs?"

We divide by $\Pr(B)$ to renormalize so that B and $\Pr(\cdot|B)$ give a new probability space.

<u>Ex</u>:
- $\Pr(B|B) = \frac{\Pr(B \cap B)}{\Pr(B)} = \frac{\Pr(B)}{\Pr(B)} = 1$
- $\Pr(A|B) + \Pr(\bar{A}|B) = \frac{\Pr(A \cap B)}{\Pr(B)} + \frac{\Pr(\bar{A} \cap B)}{\Pr(B)}$

$$= \frac{\Pr(B)}{\Pr(B)} = 1$$

Let us return to our 3-coin example.

- $Pr(B) = \frac{|B|}{|\Omega|} = \frac{4}{8} = \frac{1}{2}$
- $Pr(A_1) = \frac{|A_1|}{|\Omega|} = \frac{4}{8} = \frac{1}{2}$
- $Pr(A_2) = \frac{|A_2|}{|\Omega|} = \frac{4}{8} = \frac{1}{2}$
- $Pr(A_3) = \frac{|A_3|}{|\Omega|} = \frac{4}{8} = \frac{1}{2}$

(1) $Pr(A_1 | B) = \frac{Pr(A_1 \cap B)}{Pr(B)} = \frac{|A_1 \cap B|/8}{1/2}$

$= \frac{3/8}{1/2} = \frac{6}{8} = 3/4$

Because $Pr(A_1|B) = 3/4 > \frac{1}{2} = Pr(A_1)$, we see that A_1 and B are positively correlated.

(2) $Pr(A_2 | B) = 1 - Pr(\overline{A_2} | B)$

$= 1 - Pr(A_1 | B) = 1 - 3/4 = 1/4$

so $Pr(A_2|B) = 1/4 < Pr(A_2)$ and A_2 and B are negatively correlated.

(3) $Pr(A_3 | B) = \frac{Pr(A_3 \cap B)}{Pr(B)} = \frac{|A_3 \cap B|/8}{1/2} =$

$= \frac{2/8}{1/2} = \frac{4}{8} = \frac{1}{2}$

so $Pr(A_3 | B) = \frac{1}{2} = Pr(A_3)$ and A_3 and B are independent.

<u>Note:</u> $Pr(A|B) = Pr(A) \iff \frac{Pr(A \cap B)}{Pr(B)} = Pr(A)$

$$\iff Pr(A \cap B) = Pr(A) \cdot Pr(B)$$

$$\iff \frac{Pr(A \cap B)}{Pr(A)} = Pr(B)$$

$$\iff Pr(B|A) = Pr(B)$$

<u>def</u> Two events A and B are <u>independent</u> if $Pr(A \cap B) = Pr(A) \cdot Pr(B)$.

A collection of events $\mathcal{A}$ is <u>mutually independent</u> if, for each finite subset

$$\{A_1, A_2, \dots, A_n\} \subseteq \mathcal{A}$$

we have

$$Pr(A_1 \cap A_2 \cap \cdots \cap A_n) = Pr(A_1) \cdot Pr(A_2) \cdot \cdots \cdot Pr(A_n)$$

A collection of events $\mathcal{A}$ is <u>pairwise independent</u>

if each pair $\{A_1, A_2\} \subseteq \mathcal{A}$ of events in $\mathcal{A}$ are independent.

Warning:

- $\mathcal{A}$ mutually independent $\Longrightarrow$ $\mathcal{A}$ pairwise independent is true, but
- $\mathcal{A}$ pairwise independent $\Rrightarrow$ $\mathcal{A}$ mutually independent is $\boxed{\text{false}}$!

Ex: Suppose we flip a fair coin twice:

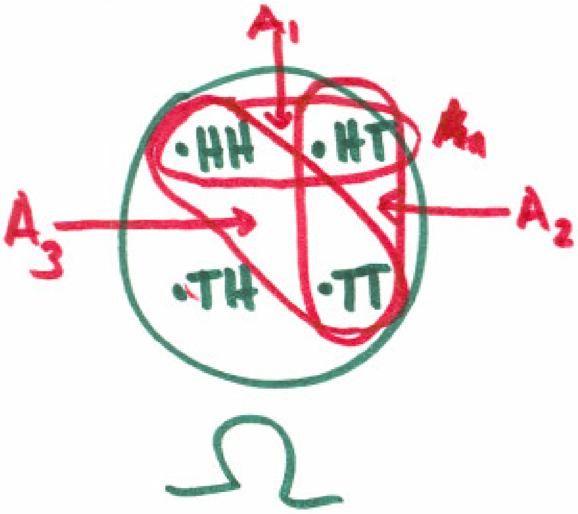
$$\Omega = \{HH, HT, TH, TT\}$$

and let

A_1 = first flip is H

A_2 = second flip is T

A_3 = both flips the same.



Although $\mathcal{A} = \{A_1, A_2, A_3\}$ is pairwise independent, $\mathcal{A}$ is not mutually independent.

Ex: By our claim, A_1 and A_2 are independent.

$$Pr(A_1 \cap A_2) = \frac{|A_1 \cap A_2|}{|\Omega|} = \frac{1}{4}$$

$$Pr(A_1) \cdot Pr(A_2) = \frac{|A_1|}{|\Omega|} \cdot \frac{|A_2|}{|\Omega|} = \frac{2}{4} \cdot \frac{2}{4} = \frac{1}{4} \checkmark$$

Note: $\mathcal{A}$ is not mutually independent because the product rule does not hold for A_1, A_2, A_3:

$$Pr(A_1 \cap A_2 \cap A_3) = \frac{|A_1 \cap A_2 \cap A_3|}{|\Omega|} = \frac{0}{4} = 0$$

$$Pr(A_1) \cdot Pr(A_2) \cdot Pr(A_3) = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8}$$

An Important Equality

- What is $\Pr(A_1 \cap A_2 \cap \cdots \cap A_n)$?

- If $\{A_1, \ldots, A_n\}$ are mutually independent, then

$$\Pr(A_1 \cap A_2 \cap \cdots \cap A_n) = \Pr(A_1)\cdot \Pr(A_2) \cdot \cdots \cdot \Pr(A_n).$$

- If $A_1, \ldots, A_n$ are not mutually independent, we can still sometimes compute $\Pr(A_1 \cap \cdots \cap A_n)$ with this formula:

Prop ~~Let~~ If $A_1, A_2, \ldots, A_n$ be events, then

$$\boxed{\begin{aligned}\Pr(A_1 \cap A_2 \cap \cdots \cap A_n) = {} & \Pr(A_1)\cdot \Pr(A_2 \mid A_1) \cdot \Pr(A_3 \mid A_1 \cap A_2) \\ & \cdot \Pr(A_4 \mid A_1 \cap A_2 \cap A_3) \cdot \cdots \cdot \Pr(A_n \mid A_1 \cap \cdots \cap A_{n-1})\end{aligned}}$$

Pf: $\Pr(A_1)\cdot \Pr(A_2 \mid A_1) \cdot \cdots \cdot \Pr(A_n \mid A_1 \cap \cdots \cap A_{n-1})$

$$= \Pr(A_1) \cdot \frac{\Pr(A_2 \cap A_1)}{\Pr(A_1)} \cdot \frac{\Pr(A_3 \cap A_2 \cap A_1)}{\Pr(A_1 \cap A_2)} \cdot \cdots \cdot \frac{\Pr(A_n \cap A_1 \cap \cdots \cap A_{n-1})}{\Pr(A_1 \cap \cdots \cap A_{n-1})}$$

$$= \Pr(A_1 \cap A_2 \cap \cdots \cap A_n)$$

Application: Suppose r balls are assigned to n bins. The balls are assigned

(·) <u>uniformly</u>: a ball is equally likely to be in any of the bins, and

(·) <u>independently</u>: the location of a ball is not influenced by the locations of other balls.

- Let A be the event that each ~~no~~ bin contains at most one ball. What is $\Pr(A)$?
- Note: if $r > n$, the pigeonhole principle tells us $\Pr(A) = 0$ because $A = \emptyset$.
- What if $r \leq n$?
- ~~Let~~

- For each $1 \le k \le r$, let A_k be the event that <u>the first k balls</u> are placed into bins so that each bin contains at most one ball.

- Note: $A = A_n \subseteq A_{n-1} \subseteq A_{n-2} \subseteq \cdots \subseteq A_2 \subseteq A_1 = \Omega$

  so $A_1 \cap A_2 \cap \cdots \cap A_k = A_k$.

- Therefore

$$\Pr(A) = \Pr(A_1 \cap \cdots \cap A_n)$$

$$= \Pr(A_1) \cdot \Pr(A_2 \mid A_1) \cdot \Pr(A_3 \mid A_1 \cap A_2) \cdot \cdots \cdot \Pr(A_n \mid A_1 \cap \cdots \cap A_{n-1})$$

$$= \Pr(A_1) \cdot \Pr(A_2 \mid A_1) \cdot \Pr(A_3 \mid A_2) \cdot \cdots \cdot \Pr(A_n \mid A_{n-1})$$

- What is $\Pr(A_k \mid A_{k-1})$? We could try to evaluate this directly from the definition, but that will put us right back where we started.

- Instead, think: If A_{k-1} occurs, (i.e. the first $k-1$ balls land in different bins), what are the chances that A_k occurs (i.e the kth ball lands in an unoccupied bin)?

- Well, the location of the kth ball is independent of the first $k-1$ balls, and it is equally likely to go in any bin. There are $n-(k-1) = n-k+1$ unoccupied bins, so the chance that the kth ball lands in one of the unoccupied bins is $\frac{n-k+1}{n}$.

- Therefore $Pr(A_k \mid A_{k-1}) = \frac{n-k+1}{n}$ and

$$Pr(A) = Pr(A_1) \cdot Pr(A_2 \mid A_1) \cdot Pr(A_3 \mid A_2) \cdot \cdots \cdot Pr(A_r \mid A_{r-1})$$

$$= 1 \cdot \frac{n-1}{n} \cdot \frac{n-2}{n} \cdot \cdots \cdot \frac{n-r+1}{n}$$

- This is a special case of what is called the <u>birthday</u> "<u>paradox</u>":

How many people do we need to gather before it is more likely than not that two have the same birthday?

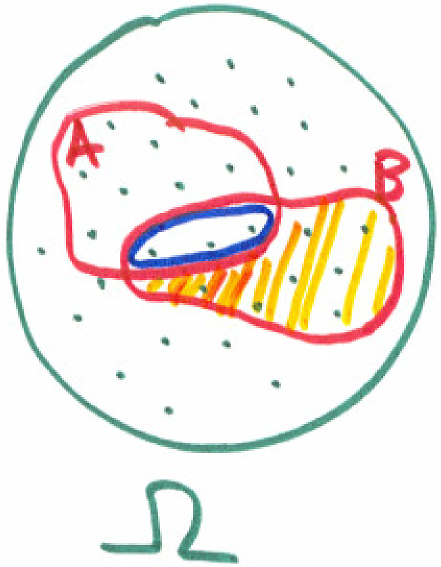
- We view the calendar ~~days~~ dates as $n = 365$ bins and people as balls, which are equally likely to have any date as his/her birthday.

- Plugging in $n = 365$ and r into our formula, we get:

r	$Pr(A) = Pr$(no common birthdays)
5	0.972..
10	0.880
15	0.742
20	0.581
25	0.422
30	0.284
35	0.178
40	0.103

45	0.055
50	0.027

- It turns out that when $r=23$, the probability that no two people have a common birthday is 0.484, so it is more likely than not that there will be a common birthday if at least 23 people are gathered.



$A \cap B$

$\bar{A} \cap B$

$$Pr(A \mid B) = \frac{Pr(A \cap B)}{Pr(B)}$$

$$\begin{aligned} Pr(A \cap B) + Pr(\bar{A} \cap B) &= Pr((A \cap B) \cup (\bar{A} \cap B)) \\ &= Pr(B) \end{aligned}$$

Extra Page