ECSE 2500
Lec 7
Feb 2

Topic: Conditional Probability (Con't)
Notion of independent events

☐ Example on Spam email Filtering

We get tons of emails especially Russian spam emails. So I will train my spam filter by putting Russian Spam to my junk mailbox. The spam filter compute the conditional probability

$$P(\text{Russian} \mid \text{spam})$$

It also knows $P(\text{spam})$ and $P(\text{Russian})$ by counting my mail/junk boxes. Then once we have a good estimate of the above three probabilites. we are interested in

$$P(\text{Spam} \mid \text{Russian}) = \frac{P(\text{Russian} \mid \text{spam})\, P(\text{Spam})}{P(\text{Russian})}$$

Remark: If the above probability is high ($\geq 80\%$), apply a rule to automatically delete emails.

$\square$ Independent Events

In high level, Events A and B are independent if we know that Event A occurs has nothing to do with the probability of Event B.

In previous example,

Rolling two 6-sided Dice.

$$P(\text{Die}1 = 2, \text{Die}2 = 3) = P(\text{Die}1 = 2) \cdot P(\text{Die}2 = 3)$$

$$= \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}$$

This matches the definition of independent events:

formally, A and B are independent when

$$\boxed{P(A \cap B) = P(A) \cdot P(B).}$$

Based on this definition, we can also have (if $P(B) \neq 0$)

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

That is, the conditional probability of A given B is the same as $P(A)$ – meaning that knowing Event B doesn't have any effect on How likely A is.

$\boxed{\text{Note}}$ Independence is not same as mutually exculsive

i.e., $A \cap B = \phi$, which means that A and B cannot happen at the same time.

To explain this point, consider the following example:

Ex. Roll a fair die. Define two events.

$$A = \{4\}, \quad B = \{2, 4, 6\}.$$

$$P(A) = \frac{1}{6}, \quad P(B) = \frac{3}{6} = \frac{1}{2}$$

$$P(A \cap B) = P(\{4\}) = \frac{1}{6}$$

$$\neq P(A) \cdot P(B) = \frac{1}{12}$$

$\rightarrow$ A and B not independent

$$P(A|B) = \frac{1}{3} \neq P(A)$$

$$P(B|A) = 1 \neq P(B)$$

$\boxed{\text{Property}}$ In fact, if $A \subset B$ and

$$0 < P(A) \leq P(B) < 1$$

then Events A and B will be never independent.

$$P(B|A) = 1 \neq P(B) < 1$$

$\boxed{\text{Note}}$ Events can be independent while still seeming "related."

Example: Roll a die. Define two events

$$A = \{3, 4\}, \quad B = \{2, 4, 6\}.$$

$$P(A) = \frac{1}{3}, \quad P(B) = \frac{1}{2}$$

$$P(A \cap B) = P(\{4\}) = \frac{1}{6}$$

$$= P(A) \cdot P(B) = \frac{1}{2} \times \frac{1}{3}$$

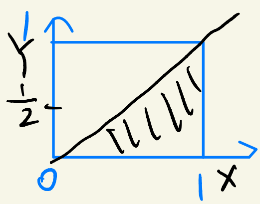
✓ Yes, they are independent.

Remark: So knowing A happened doesn't change probability of Event B, even thought they have outcomes in common. This is the place that seems counter-intuitive. The only way to be sure is to check the definition of independence

$$P(A \cap B) = P(A)\,P(B).$$

☐ Similar ideas with continuous sample spaces.

Pick X and Y uniformly at random from $[0, 1]$.



$A = \{X > Y\}$, $P(A) = \frac{1}{2}$

$B = \{Y > \frac{1}{2}\}$, $P(B) = \frac{1}{2}$

$C = \{X > \frac{1}{2}\}$, $P(C) = \frac{1}{2}$

$$P(A \cap B) = P\left(X > Y,\ Y > \frac{1}{2}\right)$$

$$= \frac{1}{8} \neq P(A) \cdot P(B)$$

$$P(A \cap C) = \frac{3}{8} \neq P(A) \cdot P(C)$$

$$P(B \cap C) = \frac{1}{4} = P(B) \cdot P(C)$$

Remark: B and C are independent, while A is not independent with others.

In general, for a set of events $\{A_1, A_2, \cdots, A_k\}$. To be independent, we will have

pairwise $\quad P(A_i A_j) = P(A_i) P(A_j), \ \forall i \neq j$

Triples $\quad P(A_i A_j A_k) = P(A_i) P(A_j) P(A_k),$

$\forall i \neq j \neq k$

$\vdots$

k events $\quad P\left(\bigcap_{i=1}^{k} A_i\right) = \prod_{i=1}^{k} P(A_i)$

Example: Roll 4-sided die. Define three events

$A = \{1,2\}$, $B = \{1,3\}$, $C = \{1,4\}$

$P(A) = P(B) = P(C) = \frac{1}{2}$.

$$P(A \cap B) = P(\{1\}) = \frac{1}{4}$$

$$P(B \cap C) = P(\{1\}) = \frac{1}{4}$$

$$P(A \cap C) = P(\{1\}) = \frac{1}{4}$$

$$P(A \cap B \cap C) = P(\{1\}) = \frac{1}{4} \neq P(A)P(B)P(C)$$

They are not independent!

- ☐ Many real systems can be viewed as sequence of independent experiments.

i.e., Experiments $E_1, E_2, \cdots, E_n$

Outcomes $(s_1, s_2, \cdots, s_n)$

Sample space $S_1 \times S_2 \times \cdots \times S_n$

Special case: Bernoulli Trial

Each sample space S_i is binary

The example of flipping the coins multiple times is an example of Bernoulli Trial.

$$S_i = \{ \text{Head}, \text{Tail} \}$$

$$= \{ \text{Success}, \text{Fail} \}$$

Assume $\underset{H}{P(\text{success})} = P$, $\underset{T}{P(\text{Failure})} = 1-P$.

Example: Flip a coin 5 times.

What is $P(\underbrace{HTHTH}_{\text{sequence of outcomes}})$

Since each coin flip is independent,

$$P(HTHTH) = P(H)P(T)P(H)P(T)P(H)$$

$$= P^3 (1-P)^2$$

Generalize this, we can calculate

$$P(\text{any sequence}) = P^{\#\text{ of heads}} (1-P)^{\#\text{ of tails}}$$

What is

$$P(3 \text{ out of } 5 \text{ heads}) = C^5_3 \, P^3 (1-P)^2$$

Generalize

$$\Longrightarrow \boxed{P(k \text{ out } n \text{ flips heads}) = C^n_k \, P^k (1-P)^{n-k},}$$

$k = 1, 2, \cdots n$

This is called the Bernoulli Probability Law.