Probability Review

Probability theory used to model real-world experiments in which outcomes can be uncertain

Sample space of experiment is the finest grain mutually ~~exclusive~~ exclusive and collective exhaustive set of all possible outcomes

e.g. flip a coin $\Omega = \{h, t\}$

flips a coin 3 times, observe sequence of heads/tails: $\{hhh, hht, \ldots, ttt\}$

Event – a subset of sample space, set of outcomes

e.g. flip a coin 3 times

Event $E = \{$ at least 2 heads $\} = \{hhh, hht, hth, thh\}$

Event space – a collectively exhaustive, mutually exclusive set of events.

Probability Space $(\Omega, \mathcal{F}, P)$

Ω: sample space

$\mathcal{F}$: set of events

P: probability measure $\quad P: \mathcal{F} \to [0,1]$

1) $P(E) \geq 0$ for any $A \in \mathcal{F}$

2) $P(\Omega) = 1$

3) For any countable collection $E_1, E_2, \ldots$ ($\in \mathcal{F}$) of ME events,
$$P\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} P(E_i)$$

- From axioms, can show $P(\phi) = 0$.
- For finite ME collection $E_1, \ldots, E_m$, $E = \bigcup_{i=1}^{m} E_i \in \mathcal{F}$
$$P(E) = \sum_{i=1}^{m} P(E_i)$$
- $P(\cancel{E}E^c) = 1 - P(E)$, $E \in \mathcal{F}$.

e.g. ~~ex~~ flip coin ~~of~~ coin 3 times

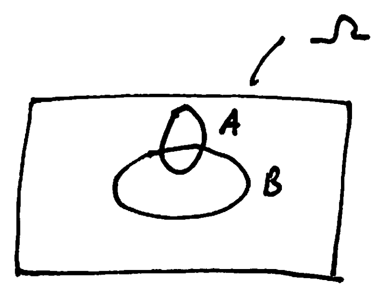
$\Omega = \{hhh, hht, \ldots, ttt\}$

Prob. of each outcome $= 1/8$

Let $E = \{$exactly 2 heads occur$\}$

$= \{hht, hth, thh\}$

$$P(E) = \frac{1}{8} + \frac{1}{8} + \frac{1}{8} = \frac{3}{8}.$$



<u>Conditional Probability</u>

- $A, B \in \mathcal{F}$ events

$P(A)$ = a priori prob of A

$P(A|B)$ = cond. prob. of A given B

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

defined only when $P(B) > 0$

Treat conditional probs. as prob. law defined on new universe B.

e.g. toss fair coin 3 times

$A = \{$ more heads than tails come up $\}$

$B = \{$ 1st toss is head $\} = \{hhh, hht, hth, htt\}$

$P(B) = \frac{4}{8} = \frac{1}{2}$

$A \cap B = \{hhh, hht, hth\}$

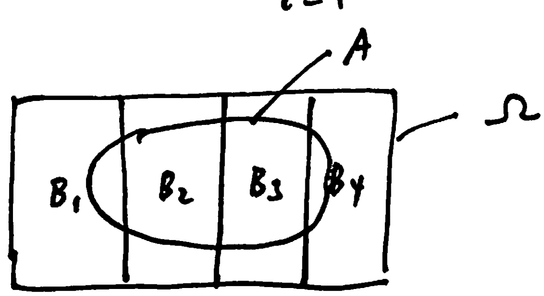
$P(A \cap B) = \frac{3}{8}$

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{3/8}{4/8} = \frac{3}{4}$$

<u>Law of total probability</u>

For event space $\{B_1, B_2, \ldots, B_m\}$, $P(B_i) > 0 \ \forall i$

$$P(A) = \sum_{i=1}^{m} P(A \mid B_i) P(B_i) \quad \text{for any } A \in \mathcal{F}.$$



<u>Bayes Theorem</u>

Have info. about $P(A \mid B)$, need to find $P(B \mid A)$

$$P(B \mid A) = \frac{P(A \mid B) P(B)}{P(A)}$$

let $\{B_1, \ldots, B_m\}$ be event space with $P(B_i) > 0 \ \forall i$

$$P(B_i | A) = \frac{P(A|B_i)P(B_i)}{P(A)} = \frac{P(A|B_i)\, P(B_i)}{\sum_{i=1}^{m} P(A|B_i)\, P(B_i)}$$

used for inference ; $B_1, \ldots, B_m$ = "causes"
A = "effect" :

$P(B_i)$ = a priori prob. of event B_i
$P(B_i|A)$ = a posteriori prob. of event B_i given A

independence : events governed by distinct non-interacting physical processes
Events A and B are statistically indep. iff
$$P(A \cap B) = P(A)\, P(B)$$
when $P(A) > 0$, $P(B) > 0$, equivalent to
$$P(A|B) = P(A), \quad P(B|A) = P(B).$$

conditional independence

Next best thing to independence
A, B conditionally independent given C if
$$P(A \cap B | C) = P(A|C)\, P(B|C)$$

– Now $P(A \cap B|C) = P(B|C)\, P(A|B \cap C)$
So conditional indep $\iff$ $P(A|C) = P(A|B \cap C)$
"Markov condition"

independence $\overset{\Rightarrow}{\not\Leftarrow}$ conditional indep.