03.10.2022

YZV 231E

Probability Theory & Stats

Week 3

Gü.

Recap: Conditional Probability of A given B: prob that A happens when we know that event B has already happened.

Ex: Throw a die (fair) : $\Omega = \{1, 2, 3, 4, 5, 6\}$

$A = \{\text{getting a } 3\}$ $\longrightarrow$ P: discrete uniform prob law

$B = \{\text{getting an odd } \#\}$

$P(A|B) = ?$

$$= \frac{P(A \cap B)}{P(B)} = \frac{1/6}{1/2} = \frac{1}{3}$$

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

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

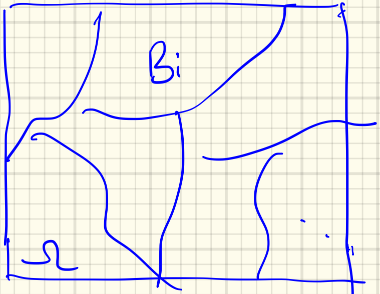
$P(A \cap B) = \frac{1}{6}$

Total Probability Law:

Let B_i be a partition of $\Omega = \bigcup_{i=1}^{N} B_i$, $\quad B_i \cap B_j = \emptyset \quad i \neq j \quad \forall i,j.$

(also valid for an ∞-seq of sets/events)

sample space



$$P(A) = \sum_{i=1}^{N} P(A \cap B_i) = \sum_{i=1}^{N} P(A|B_i)P(B_i)$$

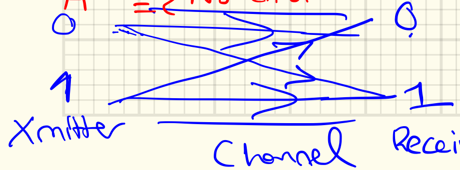
Ex: Prob of error in a digital communication system.

$B_1 = \{0 \text{ xmitted}\}$

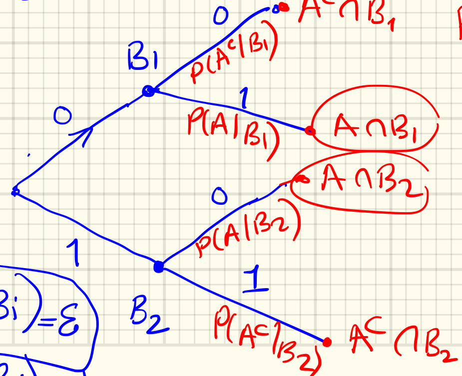
$B_2 = \{1 \text{ xmitted}\}$

$A = \{\text{error at the receiver}\}$

$A^c = \{\text{No error}\}$



$\boxed{P(A|B_i) = \varepsilon}$ $\quad P(B_i) = \frac{1}{2}$



$P(A) = P(A \cap B_1) + P(A \cap B_2)$

$P(A) = P(A|B_1) \cdot P(B_1) + P(A|B_2)P(B_2)$

$P(A) = \varepsilon \cdot \frac{1}{2} + \varepsilon \cdot \frac{1}{2}$

$P(A) = \varepsilon$

Recall: $P\left(\bigcap_i A_i\right) = ?$; $P(A\cap B) = P(A|B)\,P(B)$

$$P(\underbrace{A\cap B}\cap C) = P(C|A\cap B)\,P(\underbrace{A\cap B}) = P(C|A\cap B)\,P(A|B)\,P(B)$$

Multiplication Rule (prob. Chain rule)

$$P\left(\bigcap_{i=1}^{n} A_i\right) = P(A_1)\,P(A_2|A_1)\,P(A_3|(A_2\cap A_1)\ldots\ldots P\left(A_n \Big| \underbrace{\bigcap_{i=1}^{n-1} A_i}\right)$$

Compare to Total prob. law:

$$\bigcup_{i=1}^{n} A_i = \Omega \qquad A_i\cap A_j = \emptyset,\ i\neq j$$

$$P(B) = \sum_{i=1}^{n} \underbrace{P(B|A_i)\,P(A_i)}_{P(B\cap A_i)}$$

Independence: A & B are independent events when occurence of A (or B) does not influence occurence of the other

A & B independent: $\Rightarrow P(A \mid B) = P(A) = \dfrac{P(A\cap B)}{P(B)}$ ← $P(B) \neq 0$

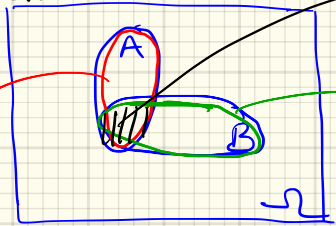
Def: A & B are independent iff (if and only if)

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

check A & B indep

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

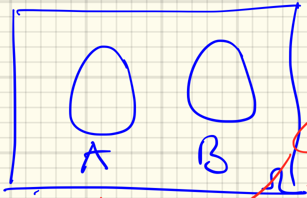
for independence check whether $\underbrace{P(A\cap B)}_{\text{prob of that shaded area}} \overset{?}{=} P(A)\cdot P(B)$



Q. Are disjoint events independent? No!

Don't confuse disjointness w/ independence ★

Let A & B be disjoint (& both are non-zero prob. events.)



$$A\cap B = \emptyset$$

$$P(A\cap B) = P(\emptyset) = 0$$

$\Rightarrow$ A & B are NOT independent

$$\underset{0}{P(A\cap B)} \underset{\neq}{\neq} \underset{>0}{P(A)} \cdot \underset{>0}{P(B)}$$

Conditional Independence: (same rules apply)

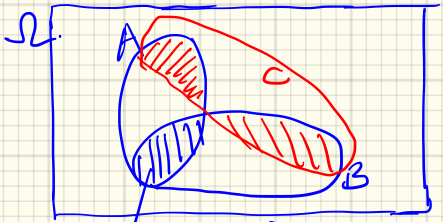
A & B are (conditionally) independent & C occured

$$P(A \cap B \mid C) = P(A \mid C) \cdot P(B \mid C)$$

∴ independence applies in a conditional universe.

— Assume A & B are independent:

Now, we are given that C occured.



In the new sample space w/ conditioning by C

Are $A \mid C$ · $B \mid C$ independent?

$A \mid C$ & $B \mid C$ are disjoint

∴ → they are not independent

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

⟹ Conditioning may affect independence!

* Having independence in the original space does not imply independence in the conditional sample space.

* These statements are equivalent:

(1) A & B are independent

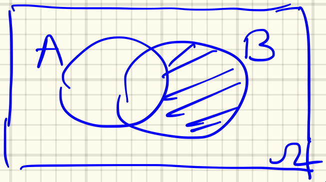
(2) A^c & B^c are independent

(3) A & B^c " "

(4) A^c & B " "

Show (4), assume (1)

$$P(A^c \cap B) = P(B) - P(A \cap B)$$



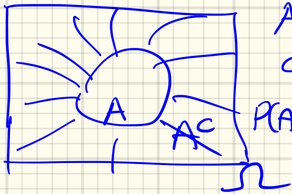
A & B are independent:

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

$$\rightarrow P(A^c \cap B) = P(B) \underbrace{(1 - P(A))}_{P(A^c)}$$

$$\rightarrow P(A^c \cap B) = P(B) . P(A^c)$$

$\Longrightarrow$ A^c & B are independent ✓



A & A^c are disjoint

$P(A^c) = 1 - P(A)$

$A \cup A^c = \Omega$

Exercise: Show the equivalence of these 4 statements,

e.g. start w/ 2

Ex: Tossing a fair coin two times w/ $p = 0.5$ $\nearrow H$, $\searrow T$

Events: A_1 : H on the 1st toss

A_2 : H on the 2nd toss

A_3 : same outcome on both tosses

Sample Space $\Omega = \{HH, HT, TH, TT\}$, $|\Omega| = 4$

$\underbrace{HH, TT}_{A_3}$ $\to P(\cdot) = 1/4$

each outcome is equally likely

$$P(A_1) = \frac{1}{2} = P(A_2) = P(A_3)$$

Q: Are A_i pairwise independent? Yes, all 3 are pairwise indep.

$$\underbrace{P(A_1 \cap A_2)}_{1/4\ \{HH\}} \overset{?}{=} P(A_1) \cdot P(A_2) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$$

$$P(A_1 \cap A_3) = P(A_1) \cdot P(A_3) = \frac{1}{4}$$

$$P(A_2 \cap A_3) = P(A_2) \cdot P(A_3) = 1/4$$

Q. What about mutually independent?

$$\underbrace{P(A_1 \cap A_2 \cap A_3)}_{\frac{1}{4}} \overset{?}{=} P(A_1) P(A_2) P(A_3)$$

$$\frac{1}{4} \neq \frac{1}{8}$$

$\therefore$ Not mutually independent

$\Longrightarrow$ Pairwise independence does not imply mutual independence of multiple events

Multiple Events are independent $\quad P\left(\bigcap_i A_i\right) = P(A_1) \cdot P(A_2) \cdots P(A_n)$

Def: A collection of events A_i are called mutually independent iff every sub-collection consists of independent events.

Multiple Random Experiments:

- A coin toss $\longrightarrow$ one random experiment : Sample Space $\Omega_1 = \{\underbrace{H}_{\text{success}}, \underbrace{T}_{\text{failure}}\}$
- Two coin tosses $\rightarrow$ two $\boxed{\text{independent}}$ experiments $\quad \Omega_2 = \{H, T\}$

$$\Omega = \Omega_1 \times \Omega_2 = \{HH, HT, TH, TT\}$$

1st toss A_1 2nd toss A_2

An event in Ω : $A = (A_1, A_2) \rightarrow P(A) = \underbrace{P(A_1)}_{=\frac{1}{2}} \underbrace{P(A_2)}_{\frac{1}{2}}$

$$= \frac{1}{2} \cdot \frac{1}{2}$$

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

One coin toss $\equiv$ A binary Outcom experiment

$\leftrightarrow$ A <u>Bernouilli</u> experiment.

Bernouilli Sepuence $\equiv$ Experiments composed of sub-experiments which are <u>INDEPENDENT</u>

Binary outcome : "<u>Success</u>" $\Leftrightarrow$ Failure

$$P(A) = p \quad ; \quad P(A^c) = 1-p = 1-P(A)$$

$$p \in [0,1]$$

$\rightarrow$ Multiple coin tosses $\equiv$ Bernouilli sepuence

M times ; M <u>independent</u> coin tosses.

say 7 coin tosses $\rightarrow$ Bernouilli sep. w/ prob. of success $= p$

$M=7$

$P(H) = p$

$$P(\underset{p}{H}\,\underset{(1-p)}{T}\,\underset{p}{H}\,\underset{p}{H}\,\underbrace{T\,T\,T}_{(1-p)^3}) = p^3 \cdot (1-p)^4$$

$$P(T\,T\,T\,H\,T\,H\,H) = p^3 (1-p)^4$$

all 7 toss (Bernouilli) sep. w/ 3 heads have equal probability

→ $P(\underbrace{\text{Bernoulli sequence}}) = p^{\#\text{Heads}} \cdot (1-p)^{\#\text{Tails}}$

$$P(k \text{ heads in } n \text{ tosses}) = \sum_{k\text{-head sequences}} p(\text{sequence}) = \underbrace{\#\, k\text{-head sequences}}_{\binom{n}{k}} \; p^k (1-p)^{n-k}$$

T	H	H	T	H	T	T	T

$P(3 \text{ heads in } 7 \text{ tosses})$

$$P(k \text{ heads in } n \text{ tosses}) = \binom{n}{k} p^k (1-p)^{n-k}$$

Binomial Law for $k = 0, 1, \dots, n$

Q. $$\sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} = 1.$$

Recall: Counting w/ replacement
w/o replacement

Ex: Lottery (NY state 44 balls) → pick 6 numbers for your ticket.

Turkish lottery 49 balls.

Winning # is randomly selecting 6 #s from 44

w/ replacement: | | | | | | | 44 $\quad 44^6 = N^{r=6} \rightarrow \frac{1}{44^6} \approx$ 1 in a 7 billion

w/o replacement |44|43|42| | |39| $\quad N_r = (44)_6 = \frac{N!}{(N-r)!} = \frac{44!}{38!} \sim 5$ billion.

<u>unordered</u> w/o replacement : account for $r!$ ways of ordering r elements out of N elements:

$$\frac{N!}{(N-r)!} \cdot \frac{1}{r!} = \binom{N}{r}$$: N choose r.

<u>Ex:</u> Coin tossing experiment : 10 tosses of independent coin tosses

→ <u>Bernouilli</u> experiment

Event B = 3 out of 10 tosses were "Heads"
Event A = The 1st two tosses were "Heads"

Given B occured

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{\binom{8}{1}}{\binom{10}{3}}$$

remaining \#H to be placed

$A \cap B$: | H | H | | | | | | | | |

$\underbrace{\text{H H}}_{\text{2 slots fixed}}$ $\underbrace{\quad\quad\quad\quad}_{\text{8 slots}}$ → $\binom{8}{1}$ ways $= 8$

$$|B| = \binom{10}{3}$$

$$p^3 \cdot (1-p)^7$$ prob of each outcome

Binomial Prob Law: $p(\text{success}) = p$

$P(k \text{ successes in } M \text{ Bernoulli trials})$
experiments

$= P(k \text{ successes in } M \text{ (binary outcome) attempts in any order})$

$$= \binom{M}{k} p^k (1-p)^{M-k}$$

Geometric Prob Law:

Another Bernoulli sep. experiment.

F: Fail (Tails), S: Success (Heads)

$p(S) = p$

$p(F) = (1-p)$

Events: $\underbrace{F \cdot F \cdot F \cdots F}_{k-1 \text{ Fails}} \cdot S$

k^{th} is a sucess

$P(\text{success at the } k^{th} \text{ attempt in a Bernoulli experiment}) = ?$

$$= (1-p)^{k-1} \cdot p$$

Ex : (Ex 4.8 (Skay (Book))). Fax machine dials a phone # that is busy; 80% of the time. $\rightarrow$ $p = 0.2$, $(1-p) = 0.8$

$$P(\text{success at the } \underbrace{9^{th}}_{k^{th}} \text{ trial}) = (0.8)^8 (0.2)$$

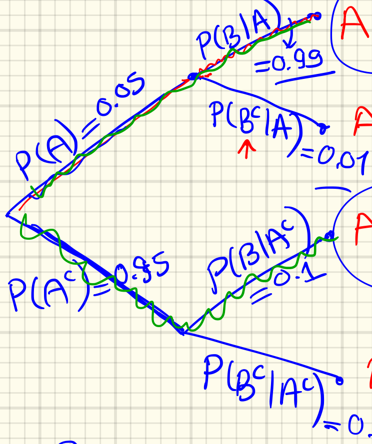
$$P(\text{success at the } k^{th} \text{ trial}) = (1-p)^{k-1} \cdot p \, .$$

Ex : [ex 1.9 Bertsekas] Radar Detection : Detecting an Aircraft

Event $A = \{\text{aircraft present}\}$ $\qquad A^c = \{\text{aircraft not present}\}$

Event $B = \{\text{radar fires} = \text{generates an alarm}\}$ $\qquad B^c = \{\text{radar does not fire}\}$

Sequential Description



$A \cap B$ (True Positive) TP

$A \cap B^c$ (Miss = False Negative) FN

$A^c \cap B$ (False Alarm = False Positive) FP

$A^c \cap B^c$ (True Negative) TN

Let's consider an event

$A \cap B$: aircraft present & radar fires

$$P(A \cap B) = P(B|A)P(A) = (0.99)\cdot(0.05)$$

$$\boxed{P(A \cap B) = 0.0495}$$

$$\rightarrow P(B) = ? \quad \underbrace{P(B|A)P(A)}_{0.0495} + P(B|A^c)P(A^c) = P(TP) + P(FP)$$

$$= 0.0495 + (0.95)\cdot(0.1) = \boxed{0.1445}$$

Q. Given that your radar fired, how likely is it that there is an airplane out there?

$$P(A|B) = ? \quad \frac{P(A \cap B)}{P(B)} = \frac{0.0495}{0.1445} = 0.34 \longrightarrow \text{This is inference!}$$

$\boxed{\text{Bayes Theorem}}$ allows us to switch $P(A|B)$ vs $P(B|A)$

- "Prior" probabilities : $P(A_i)$ our initial belief about how likely each event A_i are to occur.
- We know $P(B|A_i)$ $\forall i$: likelihood probabilities. "conditional" probabilities

Now, we're told that event B occured

$$\underbrace{P(A_i | B)}_{\text{posterior probability}} = \frac{P(A_i \cap B)}{P(B)}$$

$$\rightarrow \boxed{\underbrace{P(A_i|B)}_{\text{posterior prob.}} = \frac{\overbrace{P(B|A_i)}^{\text{cond. likelihood}}\overbrace{P(A_i)}^{\text{prior}}}{\underbrace{\sum_j P(B|A_j)P(A_j)}_{P(B)\text{ : from Total Prob. Thm.}}}}$$ Bayes Thm (Rule)

⚹ → We reversed the order of conditioning! ∴ B occured. we are given the effect.

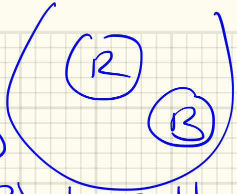
$\boxed{\text{Cause}}$ A_i (aircraft present / not) ⟶ B $\boxed{\text{effect}}$ (radar fires)

Inference

how likely was A_i

↑ w/ Bayes rule we reversed Cause → Effect : Effect → Cause.

Ex: We have an urn w/ Red & Black balls. [Skay]



$A = \{$ observe 10 Red balls in a row w/ replacement $\}$

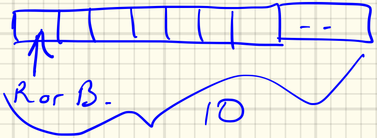
Hypothesis: $B = \{$ urn is fair $\}$ $=$ # Red Balls $=$ # Black Balls

$B^c = \{$ urn is not fair $\}$ $\Rightarrow$ $\boxed{\text{all Red Balls}}$.

$$P(B|A) = ? = \frac{P(A\cap B)}{P(A)} \qquad P(A|B) = \frac{P(A\cap B)}{P(B)}$$

given $P(B) = 0.9$ prior belief

$$\underset{\text{fair urn}}{P(A|B)} = \binom{10}{10}\underbrace{\left(\frac{1}{2}\right)^{10}}_{\text{urn is fair}}\cdot\left(\frac{1}{2}\right)^{0} = \left(\frac{1}{2}\right)^{10}$$



$$P(A) = P(A|B)\overset{1}{P(B)} + P(A|B^c)P(B^c)$$

$P(A|B^c) = 1$

$$P(A) = \left(\frac{1}{2}\right)^{10}\cdot 0.9 + 1\cdot(0.1)$$

$$P(B|A) = \frac{P(A|B)\,P(B)}{P(A|B)P(B) + P(A|B^c)\cdot P(B^c)} = \frac{\left(\frac{1}{2}\right)^{10}\cdot(0.9)}{\left(\frac{1}{2}\right)^{10}(0.9) + (0.1)} \simeq 0.0087$$

posterior prob. that the urn is fair is only ~1%.

$\Rightarrow$ reject the hypothesis of a fair urn!

Quantity: "Odds ratio" : Odds against the hypothesis
e.g. "fair urn"

$$\text{Odds} = \frac{P(B^c|A)}{P(B|A)} = \frac{1-P(B|A)}{P(B|A)} = \frac{1-0.0087}{0.0087} \sim 113,$$

$\therefore$ having an unfair urn is 113 times more likely (probable) than a fair urn.