18.10.2021

YZV 231E

Probability Theory & Stats

Gözde Ünal

Recap: Conditional Probability: $P(A|B) = \dfrac{P(A \cap B)}{P(B)}$

ex: Consider a random experiment: measuring heights & weights of students in a class

$A \cap C \to P(A \cap C) = 0.22$

H \ W	W_1 (45-50)	W_2 (50-55)	W_3	W_4	W_5	$P(H_i)$ → marginal probability
H_1	0.08	0.04	.	.		$P(H_1)$ 0.14
H_2		0.12				0.26
H_3		0.06				0.26
H_4		0.02				0.22
H_5		0				0.12
these sums $P(W_i)$	.	0.24	.	.	.	

$P(C) = 0.66$

$P(B) = 0.6$

$P(A \cap B) = 0.08$

$P(A)$; Event $A = \{$student has weight $w = w_2\}$

$B = \{$student has $H \geq H_3\}$

$H = H_3, H = H_4, H = H_5$

$P(A|B) = ?$

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{0.08}{0.6} \approx 0.15$$

Compare to $P(A)$ 0.24

In this case

$P(A|B) < P(A)$

Now consider another event

→

$A = \{\omega = \omega_2\}$

$C = \{H \leq H_3\}$

$P(A|C) = \frac{P(A \cap C)}{P(C)} = \frac{0.22}{0.66}$

$P(A|C) = 0.33$

Now: $P(A|C) > P(A)$;

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

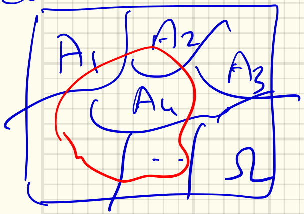
Recall: $P(\bigcap_i A_i) =$

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

Multiplication (Prob. Chain) Rule)

$P(\bigcap_{i=1}^{n} A_i) = P(A_1).P(A_2|A_1).P(A_3|A_1 \cap A_2) \ldots P(A_n | \bigcap_{i=1}^{n-1} A_i)$

Recall:



$\bigcup_i A_i = \Omega \rightarrow$ then $\{A_i\}_{i=1}^{n}$ are a partition of Ω.

$A_i \cap A_j = \emptyset, \forall i,j, i \neq j.$

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

Total Prob. Law.

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

A & B independent

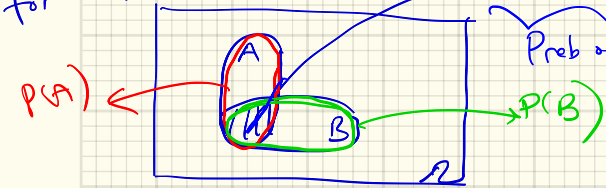
$$\to P(A|B) = P(A) \Rightarrow \frac{P(A\cap B)}{P(B)}, \quad \boxed{P(B) \neq 0} \leftarrow \text{denominator}$$

$+\varepsilon$

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

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

for independence, check; whether $\underbrace{P(A\cap B)}_{\text{Prob of the shaded area}} = P(A)P(B)$.

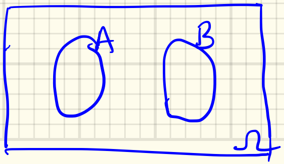


Q: Are Disjoint events independent? No!

Say A & B are disjoint $\to A\cap B = \emptyset$

$$P(A\cap B) = 0 \neq P(A).P(B)$$

A & B are NOT INDEPENDENT.



Don't confuse disjointness w/ independence.

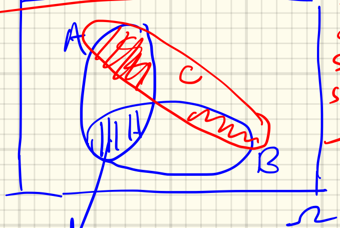
→ Conditional Independence: (same rules apply) A & B conditionally indep. & C occured.

A & B are conditionally indep.

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

∴ independence applies in a conditional universe.

Assume A & B are independent; (Now we are given that C occured.



in the original sample space Ω.

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

Now A & B are independent ? No!

In the new sample space w/ conditioning by C.

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

→ they are not independent

Conditioning may affect independence.

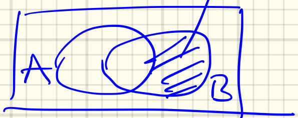
Having independence in the original space does not imply independence in the conditional sample space.

- These statements are equivalent.

1) A & B are independent
2) A^c & B^c " "
3) A & B^c " "
4) A^c & B " "

Show 4):

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

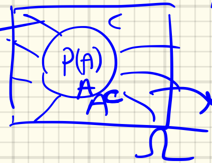


A & B are independent:
$$P(A \cap B) = P(A).P(B)$$

$$\Rightarrow P(A^c \cap B) = P(B)(1 - P(A))$$

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

$\therefore$ A^c & B are independent.



$P(\Omega) = 1$

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

Exercise: Show the equiv. of these 4 statements

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

Events: A_1: H on first toss

A_2: H on 2^{nd} toss

A_3: same outcome on both tosses

Sample Space Ω (S) $= \{HH, HT, TH, TT\}$; $|\Omega| = 4$

A_3

each outcome is equally like. $P(\cdot) = \frac{1}{4}$

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

$$P(A_1 \cap A_2) \underset{=\{HH\}}{=} \frac{1}{4} = P(A_1).P(A_2)$$

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

$$P(A_2 \cap A_3) \underset{\{HH\}}{=} \frac{1}{4}$$

Q: Are A_i pairwise independent?

Yes ✓ all 3 are pairwise indep.

Q. What about mutually independent?

$$P(A_1 \cap A_2 \cap A_3) \overset{?}{=} P(A_1).P(A_2).P(A_3)$$

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

∴ Not mutually indep. event.

$\Rightarrow$ Pairwise indep. does not imply mutual independence of multiple events.

Multiple events independent $\quad P(\bigcap_i A_i) = P(A_1) . P(A_2) \ldots P(A_n)$

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

Multiple Random Experiments :

$\rightarrow$ A coin toss $\rightarrow$ one random experiment : Sample Space $S_1 = \{H, T\}$ $\quad A_1 = H$

$\rightarrow$ Two coin tosses $\rightarrow$ two $\boxed{\text{independent}}$ experiments. $\quad S_2 = \{H, T\}$ $\quad A_2 = H$

$\rightarrow$ $S = S_1 \times S_2 = \{HH, HT, TH, TT\}$

1st toss $\quad$ 2nd toss

An event in S ; $A = (A_1, A_2)$ $\rightarrow P(A) = P(A_1) . P(A_2) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$

A Bernouilli trial (experiment) $\equiv$ One coin toss

$\equiv$ A binary outcome experiment

$\Rightarrow$ Bernouilli Sequence $\equiv$ Experiments composed of subexperiments, which are INDEPENDENT.

"Success" , "Failure"

$P(A) = p$ $\qquad$ $A^c \Rightarrow 1-p = P(A^c)$

$p \in [0,1]$

$\Rightarrow$ Multiple coin tosses $\equiv$ Bernouilli sequence.

M times: (M independent coin tosses)

say 7 coin tosses $\Leftrightarrow$ Bernouilli seq. w/ prob of success $= p$.

$P(H) = p$.

$$P(HTHHTTT) = ?$$

$$p \cdot (1-p) \cdot p \cdot p \, (1-p)^3 = p^3 (1-p)^4$$

$P(\text{sequence}^{\text{Bernouilli}}) = \underbrace{p^{\#\text{Heads}} (1-p)^{\#\text{Tails}}}$.

$P(k \text{ Heads in } n \text{ tosses}) = \binom{n}{k} \underbrace{p^k \cdot (1-p)^{n-k}}$ **Binomial Law.**

HHT . . .

T	H	H	T	T	H	T

P(3 Heads in 7 tosses)

Recall: Counting w/ replacement
w/o replacement. (49 in Turkish lottery)

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

Winning # is randomly selecting 6 #'s from 44

ordered.

w/ replacement $N = 44$ balls

$44 \cdot 44 \cdot 44 \cdot 44 \cdot \ldots = N^{r=6} = 44^6 \rightarrow \frac{1}{44^6} \approx 1$ in a 7 Billion.

w/o replacement.

$\boxed{44 \cdot 43 \cdot 42 \cdots} = N_r = (44)_6 = \frac{N!}{(N-r)!} = \frac{44!}{38!} \approx 5$ Billion.

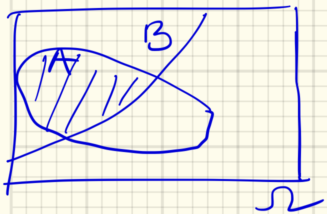
Unordered w/o replacement. $r!$ ways of ordering of r elements from total N elements. (account for)

$$\Rightarrow \frac{N!}{(N-r)!\cdot r!} = \binom{N}{r} \quad \text{combination.}$$

Ex: Coin tossing experiment: 10 tosses of independent coin tosses $\longrightarrow$ Bernouilli experiment.

Event B = 3 out of 10 tosses were "Heads" $\rightarrow$.

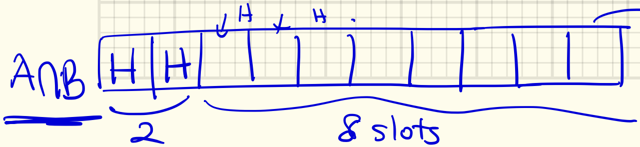
Event A = The 1st two tosses were "Heads"



Given B occured $\rightarrow$ $P(A|B) = \dfrac{P(A\cap B)}{P(B)} = \dfrac{\binom{8}{1}}{\binom{10}{3}}$

$|B| = \binom{10}{3} \rightarrow p^3\cdot(1-p)^7$.

prob. of each outcome in B.



3rd Head is placed in those 8 slots in $\binom{8}{1}$ ways $= \underline{8}$.

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

$$P(k \text{ successes in } M \text{ Bernoulli trials})$$
$$\equiv P(k \text{ successes in } M \text{ binary attempts (outcomes) in any order})$$
$$= \binom{M}{k} p^k (1-p)^{M-k}$$

Geometric Prob Law: Another Bernoulli trial

F = Fail (Tails)

$$\text{Events} = \underbrace{F \cdot F \cdot F \cdot \ldots \cdot F}_{k-1 \text{ Fails}} \cdot \underset{\uparrow\, k^{th} \text{ is a success}}{S}) = ?$$

$$P(\text{success at the } k^{th} \text{ attempt in a Bernoulli experiment})$$
$$= \underbrace{(1-p)^{k-1}} \cdot p$$

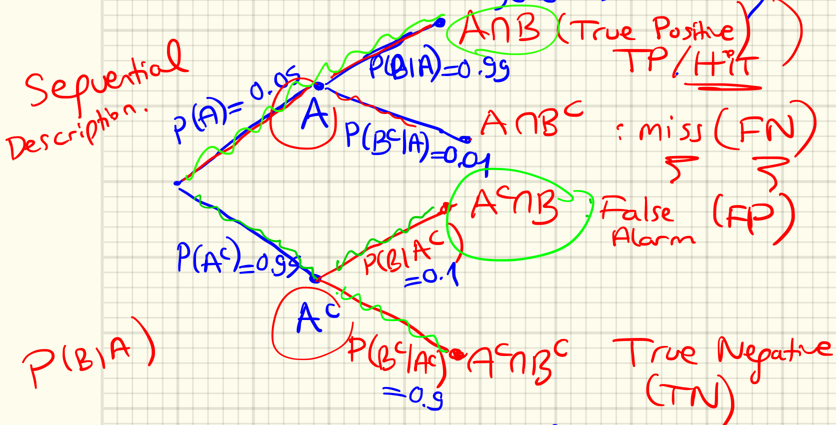
(Ex. 4.8) [SKay] Fax machine dials a phone # that is busy 80% of the time. $\rightarrow p = 0.2$, $(1-p) = 0.8$

$$P(\text{success at the } \underbrace{9}_{k}{}^{th} \text{ trial}) = \underbrace{(0.8)^8 \cdot 0.2}_{(1-p)^{k-1} \cdot p}$$

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

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

" $B = \{$ radar fires = generates an alarm $\} \rightarrow B^c = \{$ radar does not gen. an alarm $\}$



Let's consider event

$A \cap B$ $\{$ aircraft present & radar fires $\}$

$$P(A \cap B) = P(B|A)\,P(A)$$
$$= 0.05 \times 0.99 = 0.0495$$

$$P(B) = P(B|A)\,P(A) + P(B|A^c)\,P(A^c)$$

$$P(B) = P(TP) + P(FP) = 0.0495 + 0.99 \times 0.1 = 0.1445$$

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

$$P(A|B) = ?\ \frac{P(A \cap B)}{P(B)} = \frac{0.0495}{0.1445} = 0.34$$

This is inference!

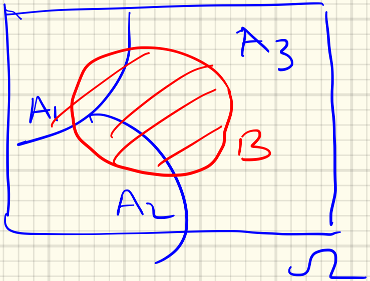
Bayes Rule (Theorem):

— "Prior" probabilities : $P(A_i)$: our initial belief about how likely each event A_i are to occur.

— We know $P(B|A_i)$ $\forall i$, "conditional" probabilities

Now, we're told that event B occured.

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



$$\boxed{\underbrace{P(A_i|B)}_{\text{posterior prob.}} = \frac{P(B|A_i) \cdot \overbrace{P(A_i)}^{\text{prior}}}{\underbrace{\sum_j P(B|A_j)P(A_j)}_{P(B)}}}$$

Bayes Thm (Rule).

We reversed the order of conditioning!

Cause A_i (aircraft present) $\xrightarrow{P(B|A_i)}$ B (effect) → Radar fires.

A_i $\xleftarrow{\text{inference}}$ B : B occured. we are given the effect.

how likely was scenario A_i given B.

We Reversed it in Bayes rule:

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

R B

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

Hypothesis $\to B = \{$urn is fair$\} = (\#\text{Reds} = \#\text{Blacks})$

$$P(B|A) = ? \qquad P(A|B) = \frac{P(A\cap B)}{P(B)} \qquad P(B) = 0.9 \text{ our prior belief.}$$

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

$$P(A|B) = \binom{10}{10}\left(\frac{1}{2}\right)^{10}\cdot\left(\frac{1}{2}\right)^{0} = \left(\frac{1}{2}\right)^{10}$$

$$P(B|A) = \frac{P(A|B)\,P(B)}{P(A) = \left[P(A|B)P(B) + P(A|B^c).P(B^c)\right]}$$

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

$(B^c = \{$unfair urn$\} =$ all Red balls for simplicity.)

$$P(A) = \left(\frac{1}{2}\right)^{10}(0.9) + 1.\underbrace{(1-0.9)}_{0.1}$$

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

posterior prob. (after 10 Red balls are drawn) that the urn is fair is only 0.0087.

→ reject the assumption of a fair urn!
(hypothesis)

A Quantity "Odds Ratio": Odds against the hypothesis.
(a fair urn)

$$\text{odds} = \frac{P(B^c|A)}{P(B|A)} = \frac{1-P(B|A)}{P(B|A)} = \frac{1-0.0087}{0.0087} = 113.$$

∴ having an unfair urn is 113 times more likely (probable) than fair urn.