Week 2

14/1 & 16/1

Probability : Branch of mathematics that provides the language necessary for situations where uncertainty is involved.

Basic set theory:

Set : Collection of distinct objects known as its "elements".

Finite set : Finite no. of elements.

No. of elements of a set A : $\#A$

Set operations:

$A \cap B$ = the set of elements that belong to both A & B. (at the same time)

$A \cup B$ = the set of elements that belong to either A or B (or both)

$\emptyset$ = empty set (set with no element)

$\#\emptyset = 0$

---

① ("omega") $\Omega \neq \emptyset$ Sample space (set of all possible outcomes.

Eg. rolling a die

$\Omega = \{1, 2, 3, 4, 5, 6\}$ ← the elements of S are called "outcomes".

$\mathcal{F} \subseteq 2^{\Omega}$ → set of events

② Possible events are the subsets of S

event $\neq$ outcome !

subset of S   element of S

The Probability of an event

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

Axioms of Probability :

(i) for any $A \subseteq \Omega$, $P(A) \geq 0$

(ii) $P(\Omega) = 1$

(iii) $A_i \cap A_j = \emptyset$ (A_i & A_j are mutually exclusive)

then $P(A_i \cup A_j) = P(A_i) + P(A_j)$

Remarks :

(i) The above axioms must always hold.

(ii) When outcomes are "equally likely", then → "same chance" of occurring

we define $P(A) = \dfrac{\# A}{\# S}$

(iii) All the properties of the probability funcⁿ can be derived from the Axioms.

Propⁿ : (i) $P(A^c) = 1 - P(A)$ for any event

(ii) $P(\emptyset) = 0$

(iii) if $A \subseteq B$, then $P(A) \leq P(B)$.

(iv) for any event A, $0 \leq P(A) \leq 1$.

(v) for generic event A & B,

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

Classical Perspective of Probability

"equally likely outcomes"

Ex. ① Experiment : Flipping 2 coins

Sample space

Ω = {ordered pairs of the 2 outcomes}

{ (H,H), (H,T), (T,H), (T,T)}

E_1 = { both flips come up as head}

= { (H,H)}

E_2 = { the 2 flips come up different }

= { (H,T), (T,H)}

$P(E_1) = 1/4$; $P(E_2) = 1/2$

② Experiment : Repeatedly rolling a die until we first get 6

Ω = { sequences of numbers betⁿ 1 & 5 & then a 6}

$|\Omega| = \infty$

E_1 = { roll 4 first, get 6 on the 3rd roll }

{ (4, m, 6) , m = 1, 2, 3, 4, 5}

Frequentist Interpretation of Probability :

Defines probability via a thought experiment.

Ex. Roughly speaking, suppose that we have a die which is weighted, but we dont know how it is weighted.

We could get a rough idea of the probability of each outcome by tossing the die a large no. of times & using the proportion of times that the die gives that outcome to estimate the probability of that outcome.

each 'repetition' is called a "TRIAL".

In the FREQUENTIST INTERPRETATION, P(E) : Proportion of times, in repeated & independent trials of an experiment, in which

. E occurs (by "independent trials we mean that they dont influence each other).

Precisely, let n = # of times the process is performed, (the trial no.)

and

m = # of times the outcome A happens out of the first n trials.

Remark (1) Notice that m & n stand for different things in this def.^{n}.

If $\frac{m}{n}$ approaches some value for large n, in the 'frequentist perspective' to probability,

$$P(A) = \lim_{n \to \infty} \frac{m}{n}$$

Remark 2: In this approach, the probability of an event A, i.e, the no. $P(A) \to$ can be objectively measured with a sequence of repeated & independent trials of an experiment.

Remark 3: F generalizes C (i.e if we have a fair dice) get same answer

" 4: F works when outcomes are not equally likely, such as the weighted die.

" 5: F is in principle just a 'thought' experiment in some cases.

The experiment can never in practice be carried out more than once. *

e.g. the probability that Kohli's avg. will go up — *

" " of going to war with country X. — *

Subjective perspective : Individual's measure of belief that an event will occur.

With this view, it makes sense to talk about events.

Limitations : subjective (individual centric)
" (must obey certain 'coherence consistency conditions in order to be workable)

Axiomatic Perspective

Unifying perspective. Codifies these coherence conditions & so can be used with any of the above perspective.

Thursday (18/1)

Independent events : A, B are independent if
$P(A \cap B) = P(A)P(B)$ (knowing whether one occurs provides no useful information about the other)

Remark: Different from A & B being mutually exclusive / disjoint. (if A occurs then B cannot occur)

Ex. Experiment : Drawing 1 card at random from a deck of 52 cards.

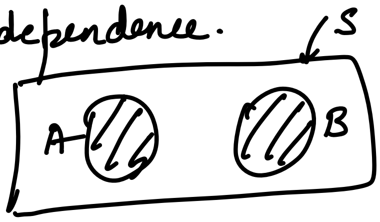
A : Card is an ace
B : " " a heart
} are they independent?

- $P(A) = 4/52 = 1/13$
- $P(B) = 13/52 = 1/4$
- $P(A \cap B) = P($ "the card is the ace of heart" $) = 1/52$

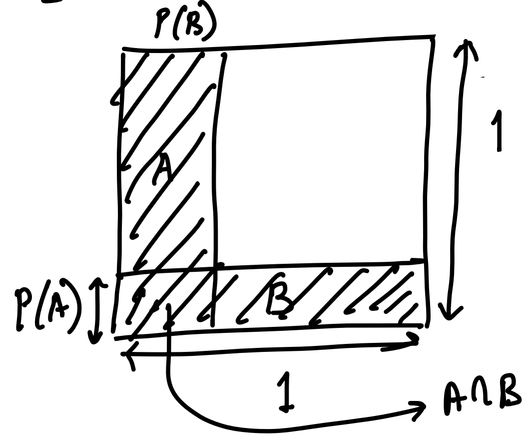
$\frac{1}{52} = \frac{1}{13} \cdot \frac{1}{4}$, A & B are independent.

Venn diagram representation of Independence.

Mutually exclusive / disjoint ←



2 events A & B are independent



Interpret area of a set as the probability of that set.

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

Note: Independence of the sets A & B is depicted through the sets A & B being orthogonal to each other.

Conditional Probability:

Defn: The conditional probability of A given B is

$P(A|B) = \frac{P(A \cap B)}{P(B)}$ (assuming that $P(B) > 0$)

↓

probability that A happens given that B has happened.

E.g. Tossing a coin 3 times:

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

A: head atleast twice $= \{HHH, HHT, HTH, THH\}$

B = the 1st toss lands head $= \{HHH, HHT, HTH, HTT\}$

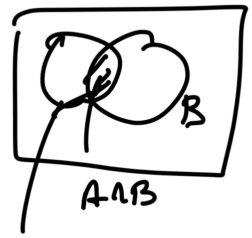
"Win" if A happens

$P(A) = 4/8 = 1/2$

$P(A|B) = 3/4$

Geometric interpretation of

$$P(A|B) = \frac{P(A\cap B)}{P(B)} \longrightarrow \text{Proportion of } B \text{ which is occupied by } A.$$



Rem 1: If A & B are independent, then

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

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

& vice-versa is also true

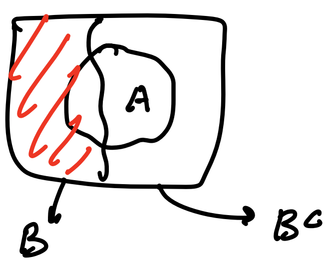
General way to solve a problem. There are different strategies.

(i) try simple & extreme cases (useful in many problems)

(Feynman's strategy
Write down the problem
↓
Think hard
↓
Write the sol^n)

(ii) Useful in statistics → break the problem into simpler problems (more problems but each easier)

Let A & B be generic events in S:



Find $P(A)$?
Break it into pieces using B.

$$A = (A\cap B) \cup (A\cap B^c)$$

mutually exclusive

$$(A\cap B) \cap (A\cap B^c) = \emptyset$$

Therefore, by Axiom 3:

$$P(A) = P(A\cap B) + P(A\cap B^c) \longrightarrow (*)$$

But by conditional probability

$$P(A\cap B) = P(A|B)\cdot P(B) \quad \&$$

$$P(A\cap B^c) = P(A|B^c)\cdot P(B^c)$$

hence (*) becomes

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

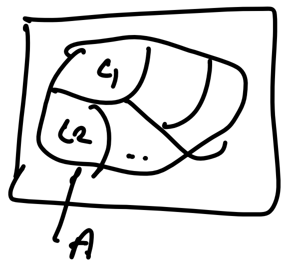
Total probability thm.

<u>Generalized total probability</u> thm.

Def.^{n}: The event $C_1, C_2, \ldots, C_n$ are called a 'Partition of A' if

(i) $C_1 \cup C_2 \cup \ldots \cup C_n = A$

(ii) for $i \neq j$, $C_i \cap C_j = \emptyset$



Prop.^{n}: If $C_1, C_2 \ldots, C_n$ is a partition of S.
Then:

$$P(A) = P(C_1) + \cdots + P(C_n) \quad \text{—— } (**)$$

If $B_1, B_2 \ldots B_n$ is a partition of S.
let A be a generic event in S

then $A \cap B_1, \ldots, A \cap B_n$ is a partition of A.

(since $(A \cap B_1) \cup \cdots \cup (A \cap B_n) = A$
for $i \neq j$, $(A \cap B_i) \cap (A \cap B_j) = \emptyset$)

Therefore for $C_i = A \cap B_i$,
(**) yields

$$P(A) = P(A \cap B_1) + \cdots + P(A \cap B_n)$$

$$\Rightarrow \boxed{P(A) = P(A|B_1)P(B_1) + \cdots + P(A|B_n)P(B_n)}$$

Generalized total probability thm.

Ex. We have 2 urns

urn #1 : 9 white balls & 1 red ball

urn #2 : 1 " " & 4 " "

Experiment: (i) Toss a fair coin

- if heads, choose urn #1
- if tails, choose urn #2

(ii) We then pick a ball at random from the chosen urn.

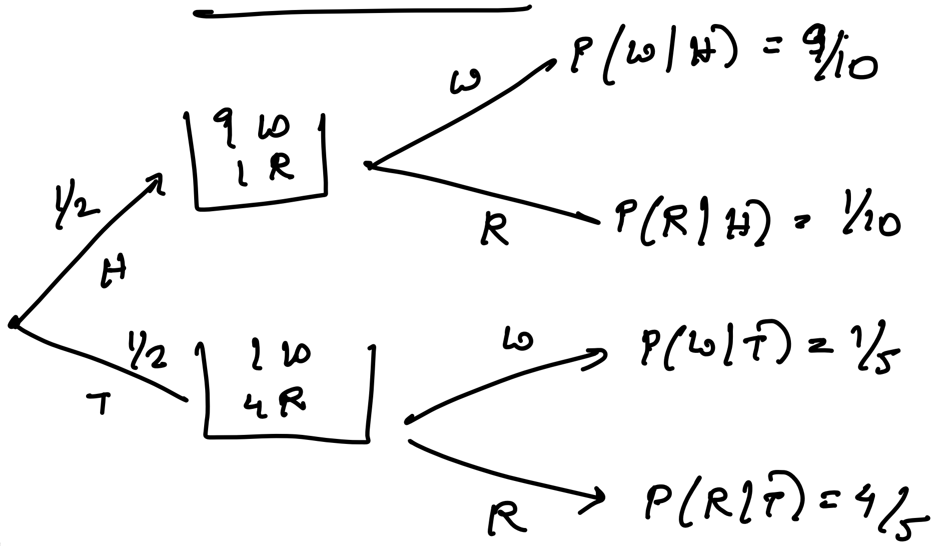
Denote W = "pick a white ball"

H = "coin shows head"

R = "pick a red ball"

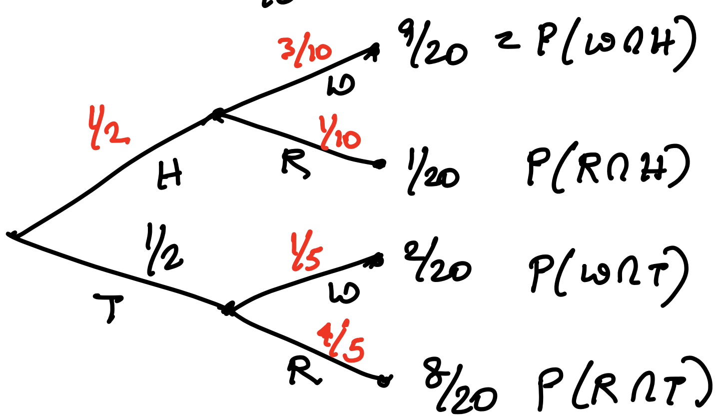
T = coin shows tail

Use a <u>TREE DIAGRAM</u>:



So:

$$P(W) = P(W|H)\,P(H) + P(W|H^c)\,P(H^c)$$
$$= P(W|H)\,P(H) + P(W|T)\,P(T)$$
$$= \frac{9}{10}\cdot\frac{1}{2} + \frac{1}{5}\cdot\frac{1}{2} = \frac{9}{20} + \frac{1}{10} = \frac{11}{20}$$



Q. Now assume that at the end, we have picked a white ball.

What are the chances that we tossed heads in the 1st place?

What is $P(H|W)$?

Use Bayes' Rule:

If A & B are events in S, then:

$$\boxed{P(A|B) = \frac{P(B|A)P(A)}{P(B)}}$$

denominator can be computed using total probability thm.

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

Using Bayes' Rule

$$P(H|W) = \frac{P(W|H)P(H)}{P(W)} = 9/11 > P(H) = 1/2$$

Using generalized total probability thm., one can get

generalized Bayes' Rule

If $B_1, B_2 \ldots B_n$ is a partition of S, then

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{P(A)} \quad \text{By Bayes' rule}$$

$$\Rightarrow \boxed{P(B_i|A) = \frac{P(A|B_i)P(B_i)}{P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + \cdots + P(A|B_n)P(B_n)}}$$

(using total Probability thm.)

This will help in problems with tree diagrams with multiple branches.

E.g. 3 urns

#1 5W, 1R

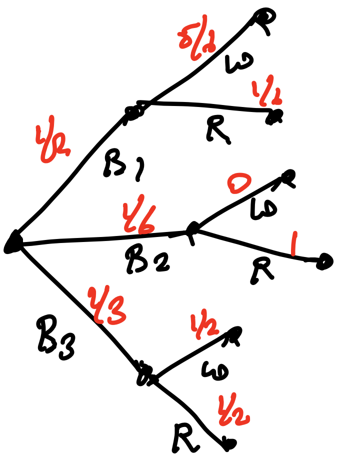
#2 0W, 7R

#3 3W, 3R

Game: Roll a fair die.

- if $B_1 = \{1,2,3\}$ choose urn 1
- if $B_2 = \{4\}$ " " 2
- if $B_3 = \{5,6\}$ " " 3

then pick a random marble.



By generalized total probability thm

$$P(R) = P(R|B_1)P(B_1) + P(R|B_2)P(B_2) + P(R|B_3)P(B_3)$$

$$= \frac{5}{12}$$

By Bayes rule

$$P(B_2|R) = 2/5$$

E.g. 2 Suppose that we are testing for a disease

Patient gets tested for a disease which afflicts 1% of population

Suppose the patient tests positive

Suppose the test (as advertised) is 95% accurate

↳ what does this mean?

if the patient has the disease, then 95% time, the test will correctly report +ve.

same as if the patient does not have the disease, then 95% time, the test will correctly report -ve.

Notation:

D : Patient has disease

T : " tests +ve

$P(T|D) = 0.95 = P(T^c|D^c)$

What does the patient care? $P(D|T)$

$$P(D|T) = \frac{P(T|D)P(D)}{P(T)} \leftarrow \text{dont know}$$

$$= \frac{P(T|D)P(D)}{P(T|D)P(D) + \underbrace{P(T|D^c)}_{0.05}P(D^c)} \approx 0.16$$

Even though the test is 95% accurate in this sense, only 16% chance that patient has the disease.