

Capítol 6

Probabilitat: teoria i exemples

Contents

6.1	Bla, bla, bla sobre l'atzar	7
6.2	Espai mostral	8
6.3	Successos	9
6.3.1	Operacions amb successos	9
6.4	Probabilitat	12
6.4.1	Propietats de la probabilitat	13
6.5	Probabilitat condicionada	15
6.5.1	Definició i càlcul	15
6.5.2	Diagrames d'arbre i probabilitat total	17
6.5.3	Teorema de Bayes i falsos positius	19
6.6	Independència de successos	20

6.1 Bla, bla, bla sobre l'atzar

Què és l'atzar? L'atzar s'associa a allà que no podem predir amb exactitud. Però, existeix l'atzar? O és que no en sabem prou com per predir-ho tot? Aquesta és un vella discussió lligada al determinisme, que no s'ha resolt fins fa relativament poc.

Quan parlem d'atzar, sovint ens preguntem si l'atzar pur existeix a la natura o si, simplement, les coses ens semblen aleatòries perquè ens falta informació. A principis del segle XX, aquest va ser l'origen d'un dels debats científics més importants de la història. Albert Einstein es resistia profundament a creure en un univers governat per l'atzar i va pronunciar la seva cèlebre frase: “*Déu no juga als daus amb l'univers*”. Ell defensava que el món havia de ser estrictament determinista i que, si en la mecànica quàntica hi havia fenòmens aparentment aleatoris, era perquè hi havia unes “variables ocultes” que encara no havíem descobert.

Per contra, físics com Niels Bohr i Erwin Schrödinger defensaven que la natura a escala subatòmica és intrínsecament probabilística. L'any 1964, el físic John Stewart Bell va proposar un test matemàtic i experimental (les anomenades **desigualtats de Bell**) per resoldre el debat.

Durant dècades es van fer experiments que donaven la raó a la quàntica (i que van culminar amb el Premi Nobel de Física de 2022). No obstant això, sempre quedava un petit problema teòric: la *llibertat d'elecció*. Per garantir que l'experiment fos vàlid, la decisió de *com* mesurar les partícules s'havia de prendre de forma absolutament aleatòria i independent just abans de la mesura. Si es feien servir màquines (generadors de nombres aleatoris) per prendre aquesta decisió, sempre quedava el dubte filosòfic: i si la pròpia màquina estava governada per les mateixes “variables ocultes” que volíem descartar?

Per tancar aquesta escaleta d'una vegada per totes, l'any 2016 es va dur a terme **The BIG Bell Test** (El Gran Test de Bell), un experiment global coordinat des de Catalunya per l'Institut de Ciències Fotòniques (ICFO), a Castelldefels.

La idea era brillant: utilitzar el lliure albir humà com a font suprema d'impredictibilitat. El 30 de novembre de 2016, més de 100.000 voluntaris de tot el món (els *Bellsters*) van participar en un videojoc en línia on se'ls demanava que tequlegessin seqüències de zeros i uns de la manera més impredecible possible. Aquests milions de decisions humanes aleatòries s'enviaven en temps real a dotze laboratoris d'arreu del món (Brisbane, Xangai, Viena, Munic, Niça, Boulder, Buenos Aires, etc.). Allà, els físics feien servir aquests zeros i uns per configurar instantàniament els seus làsers i mesurar partícules quàntiques (fotons, àtoms superconductors).

Els resultats van ser concloents¹: les desigualtats de Bell es van violar sistemàticament. L'experiment va descartar definitivament l'existència de les variables ocultes d'Einstein. Avui podem afirmar amb rigor científic que l'univers quàntic és veritablement probabilístic i que l'atzar absolut, efectivament, existeix.

6.2 Espai mostral

Definició: Experiment aleatori

No es pot saber el seu resultat amb seguretat abans de fer-lo. El resultat pot canviar si repetim l'experiment.

Alguns exemples per entendre si una situació és aleatòria o no:

- Temps que tarda una moneda en caure des de 20 metres: No és aleatori, és determinista, tot i que cal considerar un marge d'error.
- Guanyador d'una cursa hípica de 7 cavalls o el resultat d'un partit de futbol: Són aleatoris, però no depenen exclusivament de l'atzar.
- Resultat de llençar una moneda (cara/creu) o el nombre guanyador de la grossa de Nadal: Són completament aleatoris i depenen només de l'atzar.

Definició: Espai mostral

Conjunt de tots els resultats possibles d'un experiment aleatori. Es representa amb la lletra E .

Exemple 1: El dau

Experiment aleatori: llançar un dau cúbic i mirar la cara que ha sortit.

Espai mostral: $E = \{1, 2, 3, 4, 5, 6\}$.

Exemple 2: Tres iogurts

Experiment aleatori: una persona prova el sabor de tres iogurts, A, B i C, i els ordena segons les seves preferències.

Espai mostral: $E = \{ABC, ACB, BAC, BCA, CAB, CBA\}$.

Exemple 3: Dues monedes

Experiment aleatori: llançar dues monedes a l'aire i veure si surt cara (C) o creu ($+$).

Espai mostral: $E = \{(CC), (C+), (+C), (++)\}$.

¹The BIG Bell Test Collaboration. *Challenging local realism with human choices*. Nature 557, 212–216 (2018).
<https://doi.org/10.1038/s41586-018-0128-1>

Exemple 4: Moneda i dau

Experiment aleatori: Llancem una moneda a l'aire. Si el resultat és cara, es torna a llançar la moneda; si el resultat és creu, llavors, es llança un dau. Apuntem el resultat dels dos llançaments.
Espai mostral: $E = \{CC, C+, +1, +2, +3, +4, +5, +6\}$.

6.3 Successos

Definició: Succés o esdeveniment

succ És qualsevol subconjunt de l'espai mostral d'un experiment aleatori. Dit d'una altra forma: és qualsevol resultat possible que puguem enunciar.

Definició: Tipus de successos

Distingim diversos tipus de successos:

- **Succés elemental o simple:** És un únic element de l'espai mostral.
- **Succés compost:** Engloba més d'un element de l'espai mostral.
- **Succés segur:** Tindrà lloc sempre. És igual a l'espai mostral E .
- **Succés impossible:** No tindrà lloc mai. És igual a l'espai buit \emptyset .

Exemple 5: Dues monedes

Considerem l'experiment de llançar dues vegades una moneda. L'espai mostral és $E = \{(CC), (C+), (+C), (++)\}$.

- $A =$ "obtenir dues creus" $= \{++\}$ és un succés elemental.
- $B =$ "obtenir, com a mínim, una cara" $= \{CC, C+, +C\}$ és un succés compost.

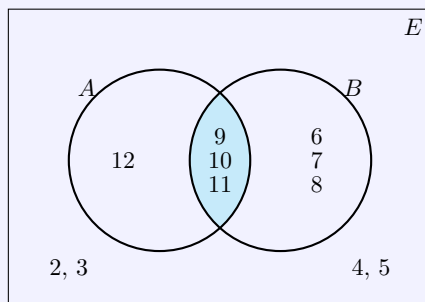
6.3.1 Operacions amb successos

A vegades ens interessa fer operacions amb dos o més successos, per obtenir successos nous i més complexos. Il·lustrem-ho amb el següent exemple: considerem l'experiment aleatori de llançar un dau dues vegades i sumar els punts obtinguts.

Definim els següents successos: $A =$ "treure més de 8" $= \{9, 10, 11, 12\}$ i $B =$ "treure entre 6 i 11" $= \{6, 7, 8, 9, 10, 11\}$.

Definició: Intersecció de successos ($A \cap B$)

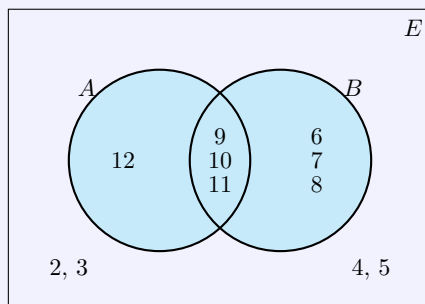
Es llegeix “A i B”. És el succés que conté tots els elements que pertanyen tant a A com a B.
 $A \cap B =$ “treure 9, 10 o 11” = {9, 10, 11}.



La intersecció de dos successos es tradueix com a **i**, i són els successos compleixen dues condicions alhora.

Definició: Unió de successos ($A \cup B$)

Es llegeix “A o B”. És el succés que conté tots els elements que es troben en A, en B o en ambdós.
 $A \cup B =$ “treure més de 6” = {6, 7, 8, 9, 10, 11, 12}.



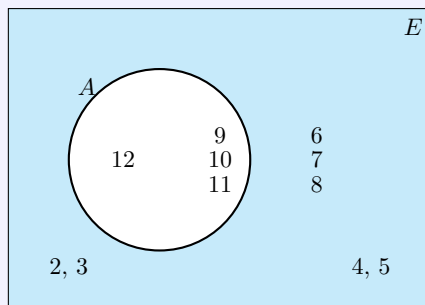
La unió de successos es tradueix com a **o** inclusiva: successos que compleixen una cosa, o una altra o les dues alhora. Consisteix en agafar els successos compleixen una condició i ajuntar-los amb els compleixen l'altra (els repetits només cal que hi siguin un cop).

La unió de successos apareix associada a conceptes com “alguna” o “com a mínim una”.

Definició: Succés complementari o contrari (\bar{A})

Es llegeix “no A”. És el format per tots els elements de l’espai mostral que no pertanyen a A. Recordem que $A \cup \bar{A} = E$.

\bar{A} = “treure 8 o menys” = {2, 3, 4, 5, 6, 7, 8}.



Fixem-nos que el complementari no es de fet una operació entre successos ja que només intervé un succés. A l’hora de calcular el complementari o contrari d’un succés és molt important recordar quin és el nostre espai mostral, E , cosa que es representa incloent un “marc” que delimita quines són els successos que estan fora de A .

Lleis de Morgan

És important recalcar que el succés contrari de la unió és la intersecció de les negacions:

$$\overline{A \cup B} = \bar{A} \cap \bar{B}$$

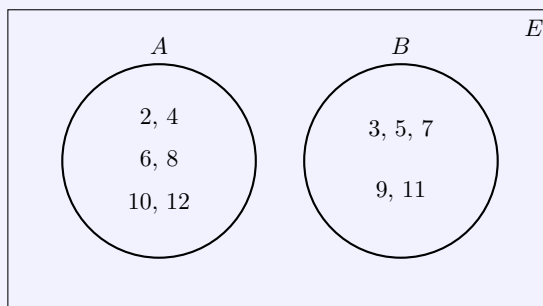
És dir, el contrari de “algun” o “com a mínim un” és “cap” o “ni” (ni a A ni a B). Això es coneix com llei de Morgan. D’aquí es dedueix també la segona llei de Morgan:

$$\overline{A \cap B} = \bar{A} \cup \bar{B}$$

Definició: Successos incompatibles

Diem que dos successos són incompatibles si no tenen cap element en comú i, per tant, és impossible que tinguin lloc a la vegada. És a dir, la seva intersecció és buida: $A \cap C = \emptyset$.

Per exemple, si A = “treure parell” = {2, 4, 6, 8, 10, 12} i B = “treure senar” = {3, 5, 7, 9, 11} aleshores A i B no poden passar alhora:



Exemple 6: Spotify

Triem una persona a l'atzar al carrer, i li demanem si fa servir Spotify i Instagram regularment. Sigui I el succés “utilitza Instagram”, i S el succés “utilitza Spotify”. Expressa mitjançant operacions amb successos els casos següents:

- Utilitza alguna de les dues aplicacions: $S \cup I$
- Utilitza Spotify, però no Instagram: $S \cap \bar{I}$
- No utilitza cap de les dues aplicacions: $\bar{S} \cap \bar{I} = \overline{S \cup I}$
- Només utilitza una de les dues aplicacions: $(S \cap \bar{I}) \cup (\bar{S} \cap I)$

6.4 Probabilitat**Definició: probabilitat**

És una mesura de quant d'esperable és un determinat succés en un experiment aleatori. Per a un succés S , la seva probabilitat es troba compresa entre 0 i 1, és a dir: $0 \leq p(S) \leq 1$.

- Succés impossible $\implies p = 0$.
- Succés segur $\implies p = 1$.

Troben tres maneres principals d'assignar una probabilitat:

Mètode de la freqüència relativa (Llei dels grans nombres)

Quan podem repetir un experiment tantes vegades com es vulgui en condicions idèntiques, la proporció de vegades (o freqüència relativa) que té lloc un determinat resultat és una estimació de la probabilitat d'aquest resultat. Matemàticament: $P(S) = \lim_{N \rightarrow \infty} f_r(S)$.

Per exemple, imaginem que volem calcular la probabilitat que, en llançar un dau, surti un 5. Podem llançar el dau moltes vegades i calcular-ne la freqüència relativa:

Nre. de llançaments	n	$f_i(5)$
30	4	0,13...
50	8	0,166...
100	15	0,15
300	51	0,17
500	83	0,166
1000	166	0,166

La llei dels grans nombres ens diu que la freqüència relativa s'estabilitza a mida que augmentem el nombre d'experiments, i que el valor que obtenim s'apropa cada vegada més a la probabilitat que aquell succés succeeixi. En aquest cas observem que el dau no sembla trucat i que la probabilitat de sortir un 5 és $\frac{1}{6} = 0.16666\dots$

Mètode subjectiu

Ens podem trobar en què el calcul de la probabilitat d'un succés sigui extramadament complicat i que, a més, no sigui possible ni tan sols realitzar l'experiment. Per exemple: ha aparegut un fisura al sostre d'una casa. Quina és la probabilitat que caigui? En aquest cas, enginyers i arquitectes estimen un valor que, tot i que es pot acompanyar amb càlculs, és subjectiu.

Llei de Laplace

Aquest és el mètode clàssic. Si podem calcular el nombre de resultats possibles i el nombre de resultats favorables a un succés S (i tots els resultats són equiprobables!), llavors podem calcular la probabilitat per la llei de Laplace:

$$P(S) = \frac{\text{nre. de casos favorables al succés } S}{\text{nre. de casos possibles}}$$

6.4.1 Propietats de la probabilitat

A partir de la definició, es dedueixen algunes propietats clau:

Fórmula: Propietats de la probabilitat

1. Probabilitat d'un succés impossible: $P(\emptyset) = 0$.
2. Probabilitat de tot l'espai mostral: $P(E) = 1$.
3. Probabilitat del succés contrari: $P(\bar{A}) = 1 - P(A)$.
4. Probabilitat del succés unió: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

1. $P(\emptyset) = 0$. El succés impossible, per definició, no pot passar. Per exemple, si llancem un dau de 6 cares numerades de l'1 al 6, és impossible que surti un 7. La probabilitat de succés impossible ha de ser nul·la perquè, altrament, podria passar.

Observació: friqui

Que un succés tingui probabilitat nul·la no el fa impossible! Per exemple, imaginem que tinguéssim una barra de mesurar amb precisió infinita. Aleshores, si triem una persona a l'atzar, la probabilitat que aquesta medeixi EXACTAMENT 1.75m és 0, però en canvi no és impossible.

2. $P(E) = 1$. Dins l'espai mostral hi tenim tot el que es considera possible. Per tant, la probabilitat que passi algun succés dels considerats possibles és 1, ja que és segur.

Observació: friqui

Alerta! A l'inrevés no funciona! Que un succés tingui probabilitat 1 no implica que sigui segur. Per exemple, la probabilitat de triar un nombre irracional d'entre tots els nombres reals entre 0 i 1 és 1, però no és segur perquè podria ser irracional.

3. $P(\bar{A}) = 1 - P(A)$. Aquesta propietat és molt evident. Si la probabilitat de tot l'espai mostral (succés segur) és 1, aleshores la probabilitat que no passi una cosa és el complementari a 1. D'aquí es després

que la probabilitat que passi una cosa o la contrària és 1:

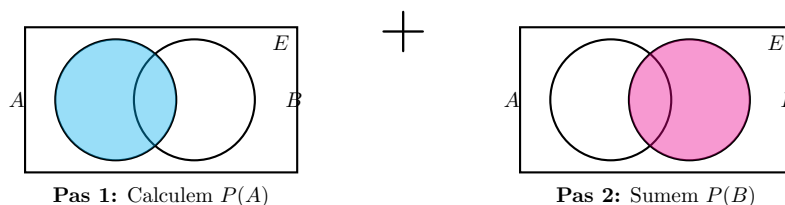
$$P(A \cup \bar{A}) = 1$$

Observació: important!

Aquesta propietat, de fet, ens pot ser de molta ajuda. Sovint és més senzill calcular la probabilitat que no passi una cosa que que passi. Per exemple, en general, la probabilitat de successos com “algun” o “com a mínim un” solen comportar molts casos i pot ser difícil de comptar. En canvi, com el contrari d’“algun” és “cap”, serà més fàcil calcular la probabilitat de “cap” i després fer el complementari a 1 que no pas calcular directament la probabilitat que passi almenys un.

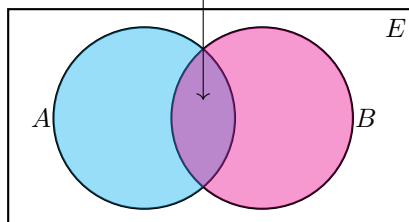
4. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$. Aquesta propietat és contraintuïtiva, però té una demostració ben natural.

Imaginem que tenim dos successos, A i B , i els volem “ajuntar” (fer la unió). Si en sumem les probabilitats obtenim el següent:

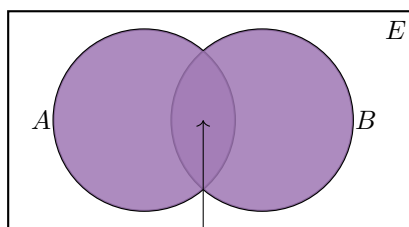


↓ Resultat provisional: $P(A) + P(B)$

ATENCIÓ! Zona de doble comptabilitat
 Hem sumat la intersecció **dues vegades**:
 una amb el cian (A) i una altra amb el magenta (B).



↓ Correcció final: Restem $P(A \cap B)$



Resultat correcte: $P(A \cup B)$
En restar una vegada la intersecció,
ara cada zona s'ha comptat exactament **una vegada**.

Observació: Important!

En general no hi ha una fórmula per calcular la “i”. És a dir, si coneixem $P(A)$ i $P(B)$, en general, no podem calcular $P(A \cup B)$. Més endavant veurem que, si A i B són independents, aleshores aquesta probabilitat serà el producte de les dues.

Exemple 7: Pluja i neu

’ha estimat que la probabilitat de pluja (P) per demà és del 70%, i la de neu (N) del 40%. Així mateix, la probabilitat que plogui i nevi és del 20%. Calcula aquestes probabilitats:

- a) Que nevi o plogui: $P(P \cup N) = P(P) + P(N) - P(P \cap N) = 0.7 + 0.4 - 0.2 = 0.9$
- b) Que no nevi ni plogui. $P(\overline{P \cup N}) = P(\overline{P \cap N}) = 1 - 0.9 = 0.1$
- c) Que o bé nevi o bé plogui (però no les dues coses). $P(P \cup N \cup \overline{P \cap N}) = P(\overline{P \cap N}) = 1 - P(P \cap N) = 1 - 0.2 = 0.8$
- d) Que nevi però no plogui. $P(N \cap \overline{P}) = P(N) - P(P \cap N) = 0.4 - 0.2 = 0.2$

6.5 Probabilitat condicionada

6.5.1 Definició i càlcul

A vegades, tenim informació prèvia sobre un experiment que pot alterar la probabilitat d’un succés. Per exemple, la probabilitat de tenir una certa malaltia és una cosa, però la probabilitat de tenir aquesta mateixa malaltia sabent que hem donat positiu al test és una altra. La segona és més probable, oi?

Doncs a això s’anomena **probabilitat condicionada**: probabilitat que passi A si sabem segur que ja ha passat B , i s’escriu $P(A|B)$:

Definició: Probabilitat condicionada

$P(A|B)$: és la probabilitat que passi A sabent que ha passat B .

Per veure com es calcula, considerem el següent exemple.

Exemple 8: Situació laboral de graduats

Les dades sobre la situació laboral d’una mostra de 800 graduats, transcorregut un any des de la finalització dels seus estudis, apareixen reflectides en aquesta taula:

	Treballant (T)	A l'atur (A)	Total
Homes (H)	520	60	580
Dones (D)	180	40	220
Total	700	100	800

Si escollim una persona a l'atzar, podem calcular les probabilitats globals fàcilment:

- Probabilitat que sigui home: $P(H) = \frac{580}{800} = 0,725$
- Probabilitat que estigui treballant: $P(T) = \frac{700}{800} = 0,875$
- $P(T|D)$? És a dir, i si ens faciliten la informació que el graduat escollit a l'atzar **és una dona**? Quina és la probabilitat que estigui treballant? Com que sabem que és dona, el nostre "total" ja no són els 800 graduats, sinó només les 220 dones. D'aquestes, sabem que n'hi ha 180 que treballen. Per tant:

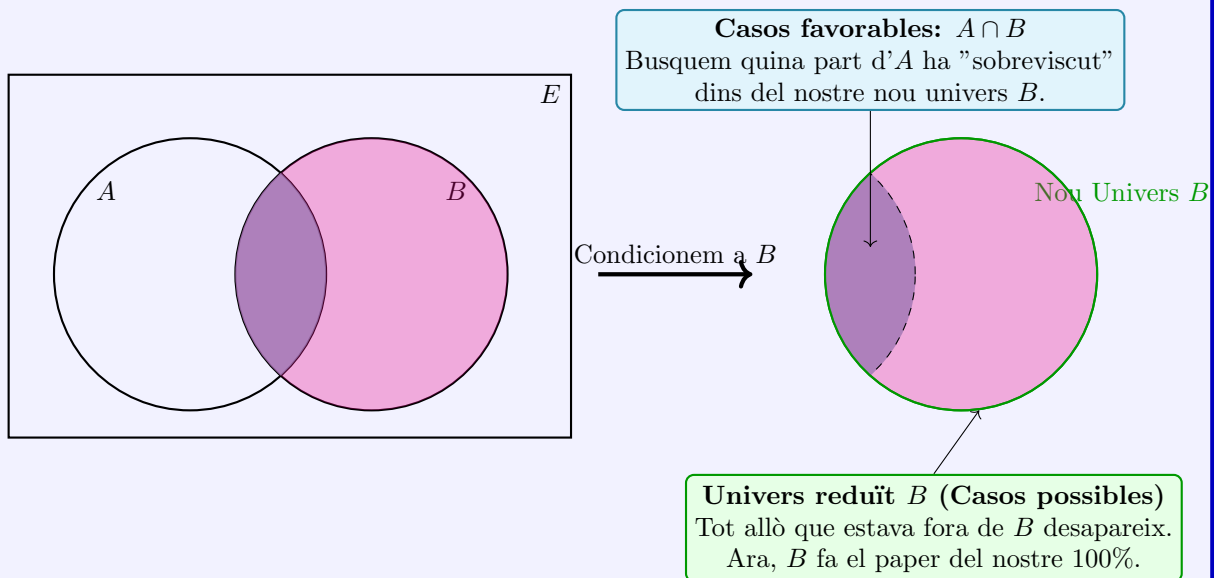
$$P(T|D) = \frac{180}{220} \approx 0,818$$

- $P(D \cap T)$? Alerta! Això es llegeix diferent: escollim una persona a l'atzar (pot ser home, dona, treballar o estar a l'atur). Quina és la probabilitat que la persona escullida sigui dona **i** treballi? Noteu que això és diferent a la de l'aprtat anterior, perquè ara la person escullida podria ser dona o no ser-ho. Per tant, la probabilitat es calcula sobre un total de 800 persones. Aquest probabilitat serà:

$$P(D \cap T) = \frac{180}{800}$$

Càlcul de la probabilitat condicionada

A partir de l'exemple anterior, podem generalitzar. Si sabem que ha passat B , el nostre espai mostral es redueix i els casos possibles "passen a ser B ". Els casos favorables "per a A seran aquells elements d' A que també compleixen B , és a dir, $A \cap B$:"



Per tant:

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

sempre que $P(B) > 0$.

Fórmula: Probabilitat condicionada

Tal com acabem de veure, la probabilitat condicionada es calcula segona la fórmula:

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

Ara bé, sovint, la dada que tindrem serà precisament la probabilitat condicionada, $P(A|B)$, i el que ens demnaran serà la probabilitat que passi A i passi B que, com hem vist abans, no és el mateix. En aquest cas, farem servir la mateixa fórmula però amb la intersecció aïllada:

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

Com veurem a continuació, aquesta fórmula ens diu que, quan dins d'un arbre volem seguir un cert camí, aleshores les probabilitats cal multiplicar-les.

Exemple: Graduats

graduats2 Recuperem l'exemple dels graduats per fer servir aquestes fórmules.

Per exemple, calculem $P(T|D)$ d'una altra manera. Segons la fórmula obtenim:

$$P(T|D) = \frac{P(T \cup D)}{P(D)} = \frac{\frac{180}{800}}{\frac{220}{800}} = \frac{180}{220}$$

que és el que havíem obtingut directament dividint les dones que treballen entre el total de dones. També podem calcular $P(T \cup D)$ d'una altra manera. Segons la fórmula:

$$P(T \cup D) = P(D) \cdot P(T|D)$$

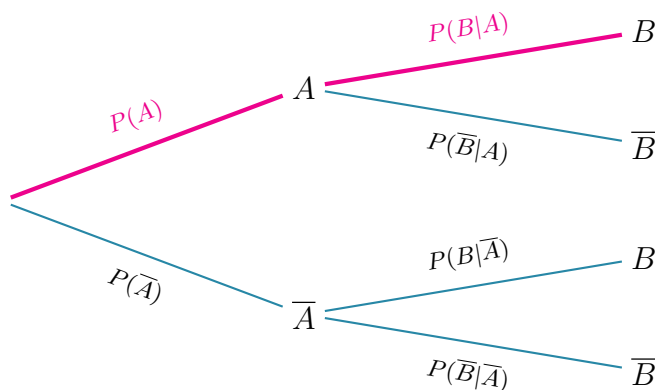
En aquesta versió de la fórmula el que coneixem són $P(D) = \frac{220}{800}$ i $P(T|D) = \frac{180}{220}$. Si ho apliquem obtenim:

$$P(T \cup D) = \frac{220}{800} \cdot \frac{180}{220} = \frac{180}{800}$$

que és el que havíem obtingut dividint el nombre de dones que treballen entre el total de gent enquesta.

6.5.2 Diagrames d'arbre i probabilitat total

Els diagrames d'arbre són molt útils per representar experiments compostos o seqüencials. Cal tenir molt clar que el segon nivell de l'arbre **no** representa probabilitats totals d'un succés, sinó probabilitats condicionades al fet que hagi passat el succés de la branca anterior. És a dir, les branques que surten d'un succés són de fet probabilitat condicionada a que aquell succés hagi passat. El següent diagrama representa aquesta idea:



Fixem-nos que cal llegir bé el diagrama. La primera B que trobem al segon nivell (marcada en magenta) no és el succés “que passi B ” sinó el succés “que passi B havent passat A ”. Així, per tal d’arribar a B , el recorregut total pel primer camí (marcat en magenta) es llegeix “que passi A i (després) B ”. Això se simbolitza amb $A \cap B$ i, segons el càlcul de la probabilitat condicionada, obtenim que cal multiplicar les probabilitat que apareixen a les branques:

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

Teorema de les probabilitats totals

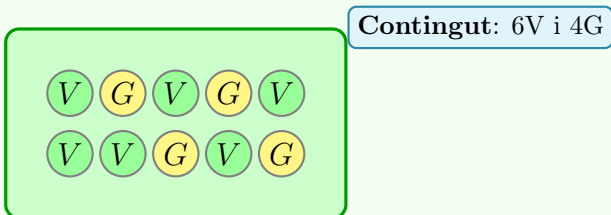
Si observem l’arbre i volem saber la probabilitat total d’un succés del segon nivell (per exemple, B), hem de sumar totes les maneres (camins) d’arribar-hi. Com que els camins de l’arbre són completament excloents (la intersecció entre branques inicials és buida, no pots anar per A i \bar{A} a la vegada), no cal restar cap intersecció:

$$P(B) = P(B \cap A) + P(B \cap \bar{A}) = P(B|A) \cdot P(A) + P(B|\bar{A}) \cdot P(\bar{A})$$

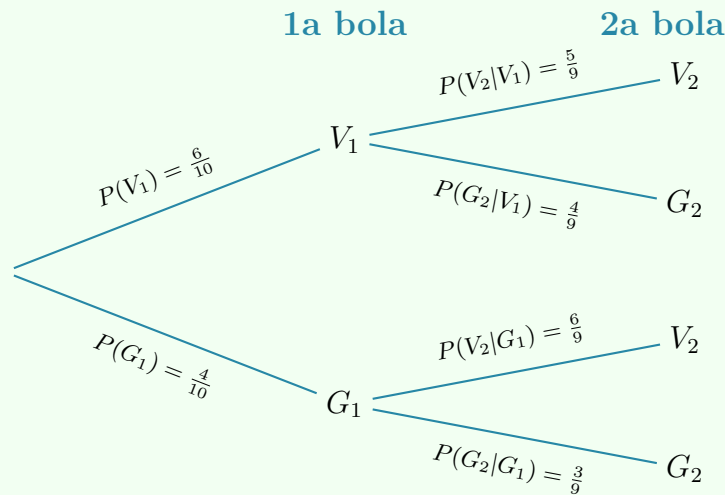
Això es coneix com el **Teorema de les probabilitats totals**.

Exemple 9: Extreure dues boles amb reposició d’una caixa

Considerem l’experiment d’extreure dues boles, una rere l’altra i **amb reposició**, d’una caixa que conté dues boles verdes (V) i una bola groga (G).



Per calcular les probabilitats dels diferents resultats compostos (per exemple, de treure una bola de cada color, de treure dues boles del mateix color, etc.), utilitzarem un diagrama d’arbre:



Fixem-nos que hem afegit un subíndex per indicar si es tracta de la primera extracció o de la segona. A partir de l'arbre, calculem les probabilitats dels diferents resultats compostos si escollim una parella a l'atzar:

- Probabilitat d'obtenir Verda i després Groga: $P(V_1 \cap G_2) = P(V_1) \cdot P(G_2|V_1) = \frac{6}{10} \cdot \frac{4}{9} = \frac{24}{90} = \frac{4}{15}$
- Probabilitat d'obtenir Groga i després Verda: $P(G_1 \cap V_2) = P(G_1) \cdot P(V_2|G_1) = \frac{4}{10} \cdot \frac{6}{9} = \frac{24}{90} = \frac{4}{15}$
- Probabilitat d'obtenir dues Grogues: $P(G_1 \cap G_2) = P(G_1) \cdot P(G_2|G_1) = \frac{4}{10} \cdot \frac{3}{9} = \frac{12}{90} = \frac{2}{15}$
- Probabilitat d'obtenir dues Verdes: $P(V_1 \cap V_2) = P(V_1) \cdot P(V_2|V_1) = \frac{6}{10} \cdot \frac{5}{9} = \frac{30}{90} = \frac{1}{3}$
- Probabilitat que la segona bola sigui Verda (Probabilitat Total): Sumem tots els camins que acaben en Verda.

$$P(V_2) = P(V_1 \cap V_2) + P(G_1 \cap V_2) = \frac{30}{90} + \frac{24}{90} = \frac{54}{90} = \frac{3}{5}$$

- Probabilitat que la segona bola sigui Groga (Probabilitat Total): Sumem tots els camins que acaben en Groga.

$$P(G_2) = P(V_1 \cap G_2) + P(G_1 \cap G_2) = \frac{24}{90} + \frac{12}{90} = \frac{36}{90} = \frac{2}{5}$$

Si sumem les probabilitats de totes les branques, comprovem que obtenim la probabilitat total de l'espai mostral E , que és 1:

$$\frac{24}{90} + \frac{24}{90} + \frac{12}{90} + \frac{30}{90} = \frac{90}{90} = 1$$

6.5.3 Teorema de Bayes i falsos positius

A vegades coneixem $P(A|B)$ però el que realment ens interessa és saber $P(B|A)$. El Teorema de Bayes ens permet "donar la volta" a la condició. Per exemple, en un estudi clínic, si sabem quants dels que han estat medicats (sense placebo) s'han curat, podem estimar quants dels que s'han curat han pres el medicament? Si sabem la probabilitat que algú que ha donat positiu en un test tingui la malaltia, podem saber quina és la probabilitat de donar positiu tenint la malaltia?

Tot això ens ho respon el Teorema de Bayes.

Teorema de Bayes

A partir de la definició de probabilitat condicionada, sabem que:

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

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

Com que la intersecció és commutativa ($A \cap B = B \cap A$), podem igualar les dues expressions:

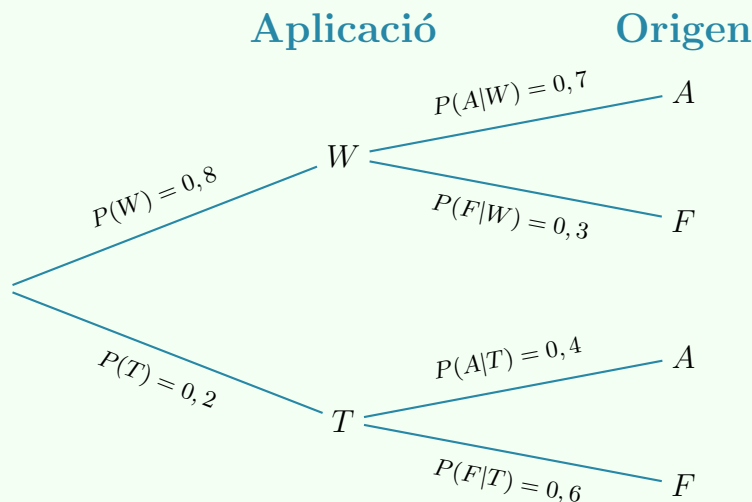
$$P(B|A) \cdot P(A) = P(A|B) \cdot P(B) \implies P(B|A) = \frac{P(A|B) \cdot P(B)}{P(A)}$$

Aquest resultat és el **Teorema de Bayes**. El denominador, $P(A)$, normalment es calcula utilitzant el teorema de les probabilitats totals.

Exemple 10: Exemple d'aplicació del Teorema de Bayes

L'Eva utilitza durant els dies laborables dos serveis de missatgeria: WhatsApp (W) i Telegram (T). El primer el fa servir en un 80% de les ocasions, i el segon, en el 20% dels casos. D'altra banda, el 70% dels missatges de WhatsApp procedeixen d'amics (A), i el 30%, de feina (F); per altra banda, a Telegram, el 40% són d'amics i el 60% de feina.

A partir d'aquestes dades, podem construir el següent diagrama d'arbre:



A continuació, donem resposta a les dues preguntes plantejades:

a) Calcula la probabilitat que un dia laborable l'Eva rebi un missatge de feina.

Per trobar la probabilitat total de rebre un missatge de feina (F), hem de sumar la probabilitat d'arribar a F per la branca de WhatsApp i la probabilitat d'arribar a F per la branca de Telegram (Teorema de les probabilitats totals):

$$P(F) = P(W \cap F) + P(T \cap F) = P(W) \cdot P(F|W) + P(T) \cdot P(F|T)$$

$$P(F) = 0,8 \cdot 0,3 + 0,2 \cdot 0,6 = 0,24 + 0,12 = 0,36$$

Per tant, la probabilitat total que el missatge sigui de feina és del 36%.

b) Si un dia rep un missatge de feina, calcula la probabilitat que sigui a través de WhatsApp.

Aquesta pregunta ens demana la probabilitat condicionada inversa, $P(W|F)$. Apliquem el Teorema de Bayes, aprofitant que al denominador hi va justament la probabilitat total $P(F)$ que acabem de calcular a l'apartat anterior:

$$P(W|F) = \frac{P(W) \cdot P(F|W)}{P(F)} = \frac{0,8 \cdot 0,3}{0,36} = \frac{0,24}{0,36} = \frac{24}{36} = \frac{2}{3} \approx 0,667$$

Si el missatge que acaba de rebre és de feina, hi ha un 66,7% de probabilitats que hagi arribat per WhatsApp.

Exemple 11: Proves mèdiques: Falsos positius i falsos negatius

Suposem una malaltia que afecta només a l'1% de la població ($P(M) = 0,01$). Tenim un test per detectar-la. Si estàs malalt, el test dona positiu el 99% de les vegades (això és un *veritable positiu*: $P(+|M) = 0,99$). Si estàs sa (\bar{M}), el test dona negatiu el 95% de les vegades (*veritable negatiu*: $P(-|\bar{M}) = 0,95$), però un 5% de les vegades dona positiu erròniament. Aquest 5% és el que anomenem **fals positiu** ($P(+|\bar{M}) = 0,05$).

Si et fas el test i **et dona positiu**, quina és la probabilitat que realment tinguis la malaltia? Molta gent intuiria que és un 99%, però apliquem Bayes:

$$P(M|+) = \frac{P(+|M) \cdot P(M)}{P(+)}$$

Primer, calculem la probabilitat total de donar positiu (denominador):

$$P(+) = P(+|M) \cdot P(M) + P(+|\bar{M}) \cdot P(\bar{M}) = (0,99 \cdot 0,01) + (0,05 \cdot 0,99) = 0,0099 + 0,0495 = 0,0594$$

Ara ja podem calcular la probabilitat de tenir la malaltia sabent que el test ha estat positiu:

$$P(M|+) = \frac{0,0099}{0,0594} \approx 0,166 \quad \text{és a dir, un 16,6\%}$$

Sorprenent! Tot i donar positiu, només tens un 16,6% de probabilitat d'estar malalt. Això passa perquè, com que la malaltia és molt rara, el volum absolut de gent sana que rep un fals positiu és molt més gran que els malalts reals detectats. Això il·lustra molt bé que **no és el mateix** la probabilitat de donar positiu sabent que s'està malalt ($P(+|M)$) que la probabilitat d'estar malalt sabent que s'ha donat positiu ($P(M|+)$).

6.6 Independència de successos

Quan parlem de probabilitat condicionada, ens podem preguntar si l'ocurrència del succés B afecta sempre la probabilitat del succés A . La resposta és no; quan l'ocurrència de l'un no aporta cap informació sobre l'altre, parlem de successos independents.

Per exemple, ens podem preguntar per la probabilitat que avui plougui. La probabilitat que avui plougui i la probabilitat que avui plougui sabent que hi ha núvols grisos no són la mateixa. En el segon cas hi ha informació que és rellevant perquè perquè plougui ens calen núvols. Per tant:

$$P(\text{"pluja"}) \neq P(\text{"pluja"} | \text{"hi ha núvols grisos"})$$

En canvi, la probabilitat que avui plogui i la probabilitat que avui plogui sabent que ahir vam anar a dormir tard sí que és la mateixa, perquè el fet que plogui o no no depèn de l'hora a la que haguem anat a dormir. Per tant podem dir que

$$P(\text{"pluja"}) = P(\text{"pluja"} | \text{"ahir vam anar a dormir tard"})$$

Així, de manera més formal, podem definir el següent

Definició: Successos independents

successos Dos successos A i B són **independents** si, i només si, el fet que passi un no altera la probabilitat que passi l'altre. Matemàticament s'expressa com:

$$P(A|B) = P(A)$$

Fórmula: Intersecció de successos independents

Si agafem la fórmula de la probabilitat condicionada $P(A|B) = \frac{P(A \cap B)}{P(B)}$ i aïllem la intersecció, tenim:

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

Com que A i B són independents, podem substituir $P(A|B)$ directament per $P(A)$, obtenint la següent propietat fonamental:

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

Observació: Important!

La fórmula de la probabilitat de la unió ($P(A \cup B)$) és sempre vàlida, independentment de si A i B són independents o no. En canvi, la fórmula de la intersecció ($P(A \cap B)$) només és vàlida si A i B són independents. En cas que sigui dependents cal utilitzar la fórmula de la probabilitat condicionada.

Exemple 12: Quatre sisos seguits

l típic exemple de successos independents és el fet de repetir un experiment diverses vegades sense que el resultat d'un afecti el següent. Per exemple el llançament d'un dau. Si llancem un dau 4 vegades, quina és la probabilitat que surtin 4 sisos seguits? Com el resultat d'una tirada no depèn del resultat anterior, obtenim que:

$$P(\text{"4 sisos seguits"}) = P(\text{"un sis"}) \cdot P(\text{"un sis"}) \cdot P(\text{"un sis"}) \cdot P(\text{"un sis"}) = \left(\frac{1}{6}\right)^4$$

Un altre exemple clàssic és de l'extracció de boles amb reposició:

Exemple 13: Extracció de boles amb reposició (Successos independents)

ecuperem l'urna de l'exemple anterior amb 6 boles verdes (V) i 4 de grogues (G). Aquesta vegada, però, extreurem dues boles **amb reposició**. Això vol dir que, un cop extreta i anotada la primera bola, la **tornem a introduir** a l'urna abans de treure la segona.

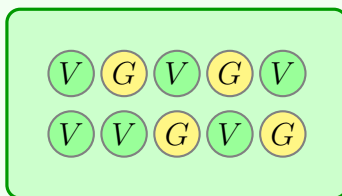
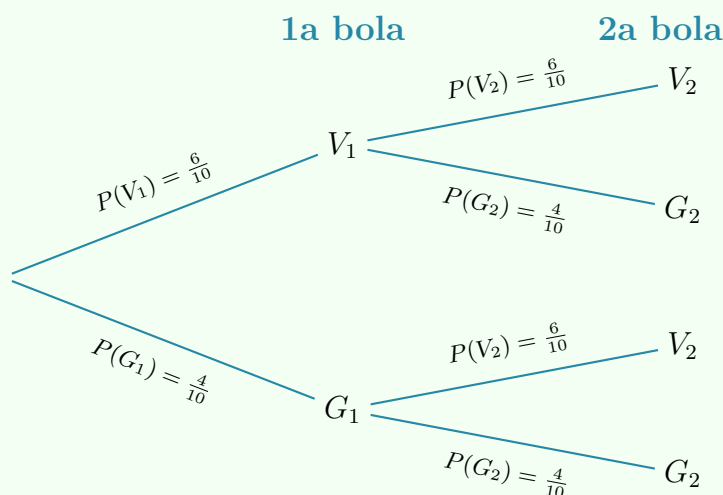


Figura 6.1: Urna amb 6 boles verdes (V) i 4 de grogues (G).

Com que l'urna torna a estar exactament igual per a la segona extracció (hi torna a haver 10 boles en total), el resultat de la primera no altera les probabilitats de la segona. Els successos són **independents**: $P(V_2|V_1) = P(V_2)$ i $P(G_2|G_1) = P(G_2)$. Observem com l'arbre reflecteix això mantenint les mateixes fraccions al segon nivell:



Si volem calcular, per exemple, la probabilitat de treure dues boles verdes, només ens cal multiplicar les probabilitats (ja que són independents):

$$P(V_1 \cap V_2) = P(V_1) \cdot P(V_2) = \frac{6}{10} \cdot \frac{6}{10} = \frac{36}{100} = \frac{9}{25} \quad (36\%)$$

Això contrasta amb el $\frac{1}{3}$ (33,3%) que ens donava en l'exemple *sense* reposició.

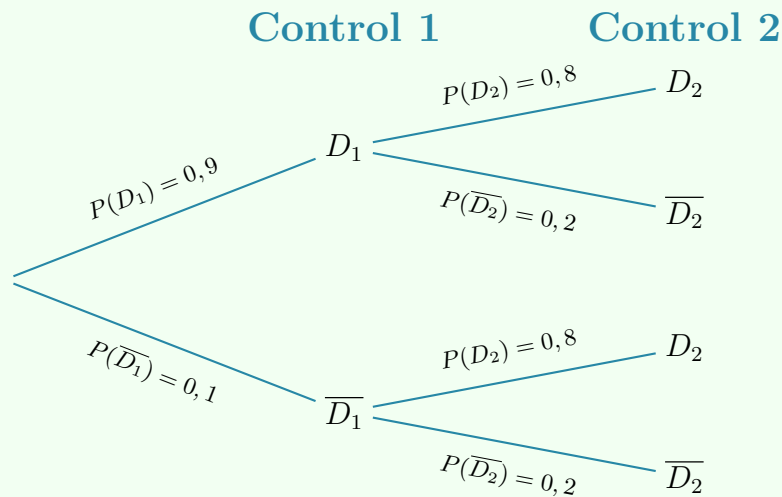
També és un exemple de probabilitat de successos independents controls de qualitat que actuen de manera independent:

Exemple 14: Controls de qualitat en sèrie

na peça defectuosa de fàbrica ha de passar per dos controls de qualitat consecutius. Els controls funcionen de manera **independent** l'un de l'altre.

- La probabilitat que el primer control detecti el defecte és del 90% ($P(C_1) = 0,9$).
- La probabilitat que el segon control detecti el defecte és del 80% ($P(C_2) = 0,8$).

Com que són successos independents, podem construir l'arbre de probabilitats on el segon nivell no depèn del primer. L'arbre tindrà la branca de "Detectat" (D) i "No Detectat / Passa" (\bar{D}):



Responem a les següents preguntes aplicant la regla del producte per a successos independents:

- a) **Quina és la probabilitat que la peça sigui detectada per tots dos controls?**

$$P(D_1 \cap D_2) = P(D_1) \cdot P(D_2) = 0.9 \cdot 0.8 = 0.72 \quad (72\%)$$

- b) **Quina és la probabilitat que la peça sigui detectada per algun dels dos controls?**

$$P(D_1 \cup D_2) = P(D_1) + P(D_2) - P(D_1 \cap D_2) = 0.9 + 0.8 - 0.72 = 0.98 \quad 98\%$$

- c) **Quina és la probabilitat que la peça esquivi els dos controls i surti al mercat defectuosa?** Perquè això passi, ha de fallar el primer control ($\overline{D_1}$) i ha de fallar el segon ($\overline{D_2}$).

$$P(\overline{D_1} \cap \overline{D_2}) = P(\overline{D_1}) \cdot P(\overline{D_2}) = 0.1 \cdot 0.2 = 0.02 \quad (2\%)$$

o bé també podem fer:

$$P(\overline{D_1} \cap \overline{D_2}) = P(\overline{D_1 \cup D_2}) = 1 - P(D_1 \cup D_2) = 1 - 0.98 = 0.02$$