Calcolo delle Probabilità e Statistica 2021/2022

Probabilità condizionata e indipendenza

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

1 Probabilità condizionata

1.1 Definizione e proprietà

Consideriamo un evento A riguardante l'esito di un qualche esperimento aleatorio, e indichiamo con $\mathbb{P}(A)$ la sua probabilità. Se veniamo a conoscenza del fatto che un altro evento B si è verificato, come è sensato aggiornare il valore di $\mathbb{P}(A)$ per tenere conto di questa nuova informazione?
Introduciamo un simbolo per indicare la probabilità dell'evento A sapendo che l'evento B si è verificato:

$$\mathbb{P}(A|B).$$

Chiameremo $\mathbb{P}(A|B)$ la ***probabilità condizionata*** (o ***condizionale***) di A dato B. Quanto vale $\mathbb{P}(A|B)$? Prima di rispondere vediamo due esempi.

Esempio 1.1. *Lanciamo un dado (a sei facce).*

Qual è la probabilità che esca un numero maggiore o uguale a 3 *sapendo che è uscito un numero pari?*

Soluzione. In primo luogo, introduciamo uno spazio di probabilità $(\Omega, \mathbb{P})$ che descriva l'esperimento aleatorio: $\Omega = \{1, 2, 3, 4, 5, 6\}$ e $\mathbb{P}$ *probabilità uniforme*[1].
Nel testo dell'esercizio si fa riferimento ai due eventi seguenti:

$$\begin{aligned} A &= \text{“esce un numero maggiore o uguale a 3”}, \\ B &= \text{“esce un numero pari”}. \end{aligned}$$

Essi sono rappresentati dai seguenti sottoinsiemi di Ω:

$$\begin{aligned} A &= \{3, 4, 5, 6\}, \\ B &= \{2, 4, 6\}. \end{aligned}$$

La probabilità richiesta dall'esercizio è la probabilità condizionata di A dato B:

$$\mathbb{P}(A|B).$$

Poiché non abbiamo ancora detto come si calcola tale probabilità, possiamo solo provare a indovinare quanto dovrebbe valere: *dato che si è verificato B, cioè è uscito un numero pari, i “veri” casi possibili sono* 2, 4, 6*; dunque i “veri” casi favorevoli sono* 4 *e* 6*; supponendo che sia lecito utilizzare la formula* “veri” casi favorevoli/“veri” casi possibili, *otteniamo*

$$\mathbb{P}(A|B) = \frac{2}{3}.$$

[1] Quindi $\mathbb{P}(\{1\}) = \cdots = \mathbb{P}(\{6\}) = \frac{1}{6}$ e vale la *formula di Laplace:*

$$\mathbb{P}(A) = \frac{\text{n}^{\text{o}} \text{ di eventi elementari che compongono } A}{6} = \frac{\textit{casi favorevoli}}{\textit{casi possibili}}.$$

In effetti questa è la risposta corretta, come si potrà verificare utilizzando la *formula per la probabilità condizionata.* □

Vediamo ora un secondo esempio in cui la probabilità non è uniforme.

Esempio 1.2. *Lanciamo un dado* ***truccato*** *a quattro facce, per cui la probabilità che esca* 1 *è il doppio della probabilità che esca* 2*, che a sua volta è il doppio della probabilità che esca* 3*, che a sua volta è il doppio della probabilità che esca* 4*.*

Qual è la probabilità che esca un numero maggiore o uguale a 3 *sapendo che è uscito un numero pari?*

Soluzione. Abbiamo già studiato questo esperimento aleatorio e sappiamo che uno spazio di probabilità $(\Omega, \mathbb{P})$ che lo descrive è $\Omega = \{1, 2, 3, 4\}$ e $\mathbb{P}$ tale che

$$\mathbb{P}(\{1\}) \;=\; \frac{8}{15}, \qquad \mathbb{P}(\{2\}) \;=\; \frac{4}{15}, \qquad \mathbb{P}(\{3\}) \;=\; \frac{2}{15}, \qquad \mathbb{P}(\{4\}) \;=\; \frac{1}{15}.$$

I due eventi

$$\begin{aligned} A &= \text{“esce un numero maggiore o uguale a 3”}, \\ B &= \text{“esce un numero pari”}, \end{aligned}$$

sono rappresentati dai seguenti sottoinsiemi di Ω:

$$\begin{aligned} A &= \{3, 4\}, \\ B &= \{2, 4\}. \end{aligned}$$

Resta da trovare $\mathbb{P}(A|B)$. Come nell'esempio precedente, non avendo ancora a disposizione la formula di $\mathbb{P}(A|B)$, possiamo solo provare a indovinare quanto dovrebbe valere. Il ragionamento è lo stesso di prima: *dato che si è verificato* B*, i “veri” casi possibili sono* 2 *e* 4*; dunque c'è un solo “vero” caso favorevole che è* 4. Come seguirà dalla formula, $\mathbb{P}(A|B)$ è pari al *rapporto tra la probabilità dei “veri” casi favorevoli e la probabilità dei “veri” casi possibili:*

$$\mathbb{P}(A|B) \;=\; \frac{\mathbb{P}(\{4\})}{\mathbb{P}(\{2\}) + \mathbb{P}(\{4\})} \;=\; \frac{1}{5}.$$

□

Definizione 1.1. *Siano* A *e* B *due eventi con* $\mathbb{P}(B) > 0$*. La* ***probabilità condizionata*** *(o* ***condizionale****)* ***di*** A ***dato*** B *è*

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

Osservazione 1. *Possiamo esprimere a parole la formula che definisce* $\mathbb{P}(A|B)$ *dicendo che la probabilità condizionata di* A *dato da* B *è pari al rapporto tra la probabilità dei “veri” casi favorevoli e la probabilità dei “veri” casi possibili:*

$$\mathbb{P}(A|B) \;=\; \frac{\textit{probabilità dei “veri” casi favorevoli}}{\textit{probabilità dei “veri” casi possibili}}.$$

Nel caso in cui Ω sia finito e gli esiti siano equiprobabili, dunque $\mathbb{P}$ è uniforme, allora $\mathbb{P}(A|B)$ è pari al rapporto tra i "veri" casi favorevoli e i "veri" casi possibili:

$$\mathbb{P}(A|B) \;=\; \frac{n^o \text{ dei "veri" casi favorevoli}}{n^o \text{ dei "veri" casi possibili}}.$$

Osservazione 2. *In generale,* ***non***[2] *vale l'uguaglianza $\mathbb{P}(A|B) = \mathbb{P}(B|A)$.*

È interessante studiare due casi particolari.

- Se $B = \Omega$, allora dalla definizione segue che

  $$\mathbb{P}(A|\Omega) \;=\; \frac{\mathbb{P}(A \cap \Omega)}{\mathbb{P}(\Omega)} \;=\; \mathbb{P}(A).$$

  Questo è naturale, infatti se l'evento che sappiamo essersi verificato è Ω, non possediamo alcuna informazione aggiuntiva (i casi possibili non sono cambiati). Infatti Ω è l'*evento certo* e sappiamo già che si verificherà sicuramente.

- Se scegliamo $A = \Omega$ (in altri termini, ci chiediamo quale sia la probabilità di Ω sapendo che B si è verificato), allora $\mathbb{P}(\Omega|B)$ è uguale a 1, infatti

  $$\mathbb{P}(\Omega|B) \;=\; \frac{\mathbb{P}(\Omega \cap B)}{\mathbb{P}(B)} \underset{\substack{\uparrow \\ \Omega \cap B = B}}{=} \frac{\mathbb{P}(B)}{\mathbb{P}(B)} \;=\; 1.$$

  Più in generale, se $B \subset A$, si ottiene $\mathbb{P}(A|B) = 1$. I due casi limite sono $A = \Omega$ (che abbiamo appena visto) e $A = B$. Anche in quest'ultimo caso vale che $\mathbb{P}(B|B) = 1$, infatti

  $$\mathbb{P}(B|B) \;=\; \frac{\mathbb{P}(B \cap B)}{\mathbb{P}(B)} \;=\; \frac{\mathbb{P}(B)}{\mathbb{P}(B)} \;=\; 1.$$

La probabilità condizionata è anch'essa una probabilità (la "vera" probabilità se sappiamo che l'evento B si è verificato) definita per tutti i sottoinsiemi dello spazio campionario Ω. Ciò significa innanzitutto che essa è una funzione che ha come dominio l'insieme delle parti di Ω, ovvero $\mathscr{P}(\Omega)$, e come codominio l'intervallo $[0, 1]$. In simboli:

$$\mathbb{P}(\,\cdot\,|B)\colon \mathscr{P}(\Omega) \;\longrightarrow\; [0, 1].$$

Inoltre, essa possiede tutte le proprietà di una probabilità, come affermato nel seguente teorema.

[2] Vedremo in seguito che relazione esiste tra $\mathbb{P}(A|B)$ e $\mathbb{P}(B|A)$ (*formula di Bayes*).

Teorema 1.1. *Sia B un evento tale che $\mathbb{P}(B) > 0$. Valgono le seguenti proprietà.*

I) *Per ciascun sottoinsieme (o evento) A di Ω, la probabilità condizionata verifica*
$$0 \ \le \ \mathbb{P}(A|B) \ \le \ 1.$$

II) $\mathbb{P}(\Omega|B) = 1$.

III) *Vale la proprietà di **additività numerabile** (o σ-**additività**): sia $A_1, A_2, \dots, A_n, \dots$ una successione di sottoinsiemi di Ω* tra loro disgiunti *e sia*
$$A \ = \ \bigcup_{n=1}^{\infty} A_n.$$
Allora
$$\mathbb{P}(A|B) \ = \ \sum_{n=1}^{\infty} \mathbb{P}(A_n|B).$$

IV) $\mathbb{P}(\emptyset|B) = 0$.

V) ***Additività finita:*** *se A_1 e A_2 sono disgiunti allora*
$$\mathbb{P}(A_1 \cup A_2|B) \ = \ \mathbb{P}(A_1|B) + \mathbb{P}(A_2|B).$$
Più in generale, se $A_1, \dots, A_n$ sono tra loro disgiunti allora
$$\mathbb{P}(A_1 \cup \dots \cup A_n|B) \ = \ \sum_{i=1}^{n} \mathbb{P}(A_i|B).$$

VI) $\mathbb{P}(A^c|B) = 1 - \mathbb{P}(A|B)$.

VII) ***Monotonia:*** *se $A_1 \subset A_2$, allora $\mathbb{P}(A_1|B) \le \mathbb{P}(A_2|B)$.*

Dimostrazione. Notiamo innanzitutto che è sufficiente dimostrare le proprietà I-II-III, in quanto le altre proprietà si dimostrano a partire da I-II-III come è stato fatto nel caso della probabilità $\mathbb{P}$.

Per quanto riguarda I-II-III, riportiamo, a titolo di esempio, solo la dimostrazione della proprietà I.

I) Dato che $A \cap B$ è un sottoinsieme di B, cioè $A \cap B \subset B$, per la *monotonia della probabilità non condizionata* $\mathbb{P}$ si ha che
$$\mathbb{P}(A \cap B) \ \le \ \mathbb{P}(B).$$
Perciò
$$\mathbb{P}(A|B) \ = \ \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \ \le \ 1.$$
Inoltre $\mathbb{P}(A|B) \ge 0$. Quindi $0 \le \mathbb{P}(A|B) \le 1$.

$\square$

1.2 Utilizzo della probabilità condizionata

Quando si studiano esperimenti aleatori reali, spesso molte probabilità condizionate sono note. Analogamente, negli esercizi, la probabilità condizionata è spesso data dal testo dell'esercizio (anche se non esplicitamente, come vedremo), mentre sarà nostro compito determinare la probabilità dell'intersezione $\mathbb{P}(A \cap B)$, che nella formula di $\mathbb{P}(A|B)$ compare a numeratore:

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

In altri termini, spesso (anche se non sempre) utilizzeremo tale formula riscritta come segue:

$$\mathbb{P}(A \cap B) \;=\; \mathbb{P}(A|B)\,\mathbb{P}(B), \tag{1.1}$$

dove $\mathbb{P}(B)$ e $\mathbb{P}(A|B)$ saranno note, mentre $\mathbb{P}(A \cap B)$ sarà l'incognita. Data l'importanza della formula (1.1) (che si chiama *regola della catena* o *formula della probabilità composta*), è utile riportarla come teorema.

Teorema 1.2. *Siano A e B due eventi con $\mathbb{P}(B) > 0$. Vale la* ***regola della catena*** *(o* ***formula della probabilità composta****):*

$$\mathbb{P}(A \cap B) \;=\; \mathbb{P}(A|B)\,\mathbb{P}(B).$$

Più in generale, dati n eventi $A_1, A_2, \ldots, A_n$, con $\mathbb{P}(A_1 \cap \cdots \cap A_{n-1}) > 0$, vale la ***regola della catena****:*

$$\mathbb{P}(A_1 \cap A_2 \cap \cdots \cap A_n) \;=\; \mathbb{P}(A_n|A_1 \cap A_2 \cap \cdots \cap A_{n-1}) \cdots \mathbb{P}(A_2|A_1)\,\mathbb{P}(A_1).$$

Osservazione. *Nella regola della catena per n insiemi, la richiesta $\mathbb{P}(A_1 \cap \cdots \cap A_{n-1}) > 0$ garantisce che tutte le probabilità condizionate che compaiono siano definite. Infatti, dato che*

$$A_1 \cap \cdots \cap A_{n-1} \;\subset\; A_1 \cap \cdots \cap A_{n-2} \;\subset\; \cdots \;\subset\; A_1,$$

dalla monotonia *della probabilità segue che anche le probabilità $\mathbb{P}(A_1 \cap \cdots \cap A_{n-2}), \ldots, \mathbb{P}(A_1)$ sono strettamente maggiori di zero.*

Dimostrazione. La dimostrazione della regola della catena nel caso di due insiemi A e B è riportata appena prima dell'enunciato del teorema e segue immediatamente dalla definizione di $\mathbb{P}(A|B)$ (è infatti un modo di riscrivere la definizione di $\mathbb{P}(A|B)$).

La regola della catena per n insiemi segue anch'essa dalla definizione di $\mathbb{P}(A_n|A_1 \cap A_2 \cap \cdots \cap A_{n-1})$, ..., $\mathbb{P}(A_2|A_1)$, infatti

$$\begin{aligned} &\mathbb{P}(A_n|A_1 \cap A_2 \cap \cdots \cap A_{n-1})\,\mathbb{P}(A_{n-1}|A_1 \cap A_2 \cap \cdots \cap A_{n-2}) \cdots \mathbb{P}(A_2|A_1)\,\mathbb{P}(A_1) \\ &\qquad = \frac{\mathbb{P}(A_1 \cap A_2 \cap \cdots \cap A_n)}{\mathbb{P}(A_1 \cap A_2 \cap \cdots \cap A_{n-1})}\frac{\mathbb{P}(A_1 \cap A_2 \cap \cdots \cap A_{n-1})}{\mathbb{P}(A_1 \cap A_2 \cap \cdots \cap A_{n-2})} \cdots \frac{\mathbb{P}(A_1 \cap A_2)}{\mathbb{P}(A_1)}\mathbb{P}(A_1). \end{aligned}$$

Elidendo i termini uguali tra loro, si ottiene $\mathbb{P}(A_1 \cap A_2 \cap \cdots \cap A_n)$. □

La *regola della catena* è particolarmente utile quando si studiano esperimenti aleatori costituiti da più sotto-esperimenti aleatori, come accade nell'esercizio seguente.

Esercizio 1.1. *Un'urna contiene tre palline bianche, due palline nere e una pallina rossa. Si eseguono tre estrazioni senza reimmissione.*

Qual è la probabilità di estrarre nell'ordine una bianca, una rossa e una nera?

Soluzione. Si noti che l'esperimento aleatorio è costituito da tre sotto-esperimenti aleatori che corrispondono alle tre estrazioni dall'urna. L'evento di cui dobbiamo calcolare la probabilità è

$$A = \text{“si estraggono nell'ordine una bianca, una rossa e una nera”}.$$

Per risolvere questo esercizio non è necessario introdurre uno spazio campionario Ω, elencando dunque tutti gli esiti dell'esperimento aleatorio. Esiste infatti un'altra via, più semplice e veloce, per risolvere l'esercizio, la quale utilizza solo gli eventi e le proprietà della probabilità (in particolare, la *regola della catena*). Vediamone i dettagli nel risolvere questo esercizio. Invece di elencare gli esiti dell'esperimento aleatorio, elenchiamo gli *eventi* di cui conosciamo la probabilità (condizionata oppure non condizionata). Essi sono in generale eventi che si riferiscono ai singoli sotto-esperimenti aleatori. Nell'esercizio in questione, sono gli eventi che si riferiscono alle singole estrazioni:

$$\begin{aligned} B_i &= \text{“si estrae una pallina bianca alla } i\text{-esima estrazione”}, \\ N_i &= \text{“si estrae una pallina nera alla } i\text{-esima estrazione”}, \\ R_i &= \text{“si estrae una pallina rossa alla } i\text{-esima estrazione”}, \end{aligned}$$

con $i = 1, 2, 3$. Dato che il testo dell'esercizio non dice nulla a riguardo, si suppone che le palline all'interno dell'urna abbiano tutte la stessa probabilità di essere estratte. Dunque, nota la composizione dell'urna, vale l'equiprobabilità, ovvero vale la formula *casi favorevoli/casi possibili.* Ad esempio

$$\mathbb{P}(B_1) \underset{\substack{\uparrow \\ \text{è l'urna iniziale}}}{=} \frac{3}{6} = \frac{1}{2}$$

oppure

$$\mathbb{P}(B_2|B_1) \underset{\substack{\uparrow \\ \text{nell'urna ci sono} \\ \text{2 b, 2 n e 1 r}}}{=} \frac{2}{5}$$

oppure

$$\mathbb{P}(B_3|B_1 \cap B_2) \underset{\substack{\uparrow \\ \text{nell'urna ci sono} \\ \text{1 b, 2 n e 1 r}}}{=} \frac{1}{4}.$$

Queste tre probabilità sono date (anche se non esplicitamente) dal testo dell'esercizio, infatti seguono dalla sola ipotesi secondo cui ad ogni estrazione le palline hanno tutte la stessa probabilità di essere estratte[3]. In conclusione, le seguenti probabilità sono note (non sono riportate tutte quelle relative agli eventi B_3, N_3, R_3):

$$
\begin{aligned}
\mathbb{P}(B_1) &= \frac{3}{6} = \frac{1}{2},\\
\mathbb{P}(N_1) &= \frac{2}{6} = \frac{1}{3},\\
\mathbb{P}(R_1) &= \frac{1}{6},\\
\mathbb{P}(B_2|B_1) &= \frac{2}{5},\\
\mathbb{P}(B_2|N_1) &= \frac{3}{5},\\
\mathbb{P}(B_2|R_1) &= \frac{3}{5},\\
\mathbb{P}(N_2|B_1) &= \frac{2}{5},\\
\mathbb{P}(N_2|N_1) &= \frac{1}{5},\\
\mathbb{P}(N_2|R_1) &= \frac{2}{5},\\
\mathbb{P}(R_2|B_1) &= \frac{1}{5},\\
\mathbb{P}(R_2|N_1) &= \frac{1}{5},\\
\mathbb{P}(R_2|R_1) &= 0,\\
\mathbb{P}(B_3|B_1 \cap B_2) &= \frac{1}{4},\\
\mathbb{P}(B_3|B_1 \cap N_2) &= \frac{2}{4} = \frac{1}{2},\\
\mathbb{P}(B_3|B_1 \cap R_2) &= \frac{2}{4} = \frac{1}{2},\\
&\text{e così via} \ldots
\end{aligned}
$$

L'evento A di cui è richiesta la probabilità può essere espresso in termini degli eventi che abbiamo introdotto come segue:

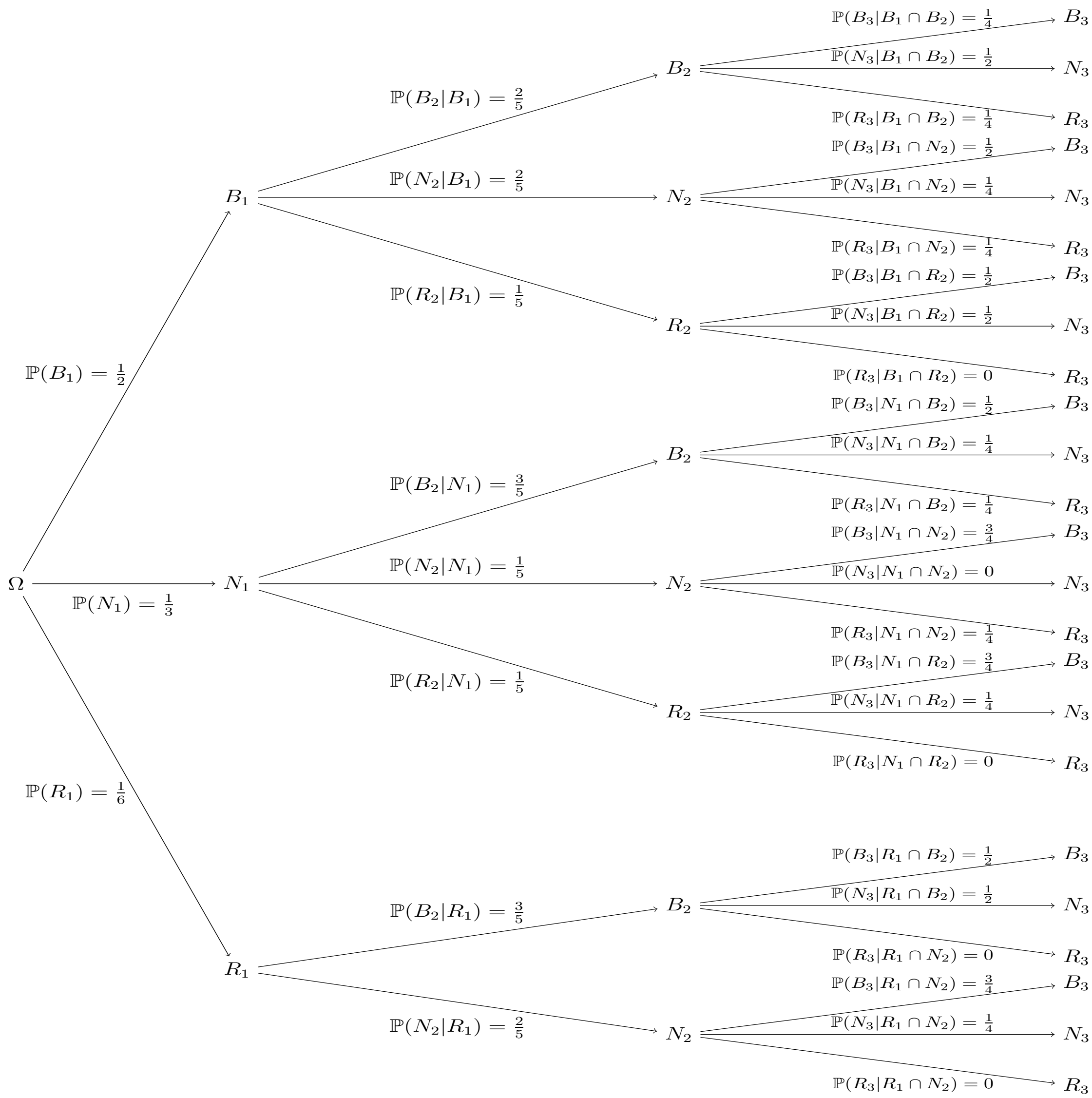
$$A = B_1 \cap R_2 \cap N_3.$$

Dunque, per la *regola della catena*,

$$\mathbb{P}(A) = \underset{\substack{\uparrow \\ \text{nell'urna ci sono} \\ \text{2 b, 2 n e 0 r}}}{\mathbb{P}(N_3|B_1 \cap R_2)} \quad \underset{\substack{\uparrow \\ \text{nell'urna ci sono} \\ \text{2 b, 2 n e 1 r}}}{\mathbb{P}(R_2|B_1)} \quad \underset{\substack{\uparrow \\ \text{è l'urna iniziale}}}{\mathbb{P}(B_1)} = \frac{2}{4} \cdot \frac{1}{5} \cdot \frac{3}{6} = \frac{1}{20}. \ \square$$

[3]Come abbiamo detto, l'ipotesi secondo cui *le palline hanno tutte la stessa probabilità di essere estratte* è assunta seppur tacitamente. Se infatti non fosse assunta, il testo dell'esercizio dovrebbe segnalarlo e dovrebbe spiegare come cambiano le probabilità (come accade nell'esercizio riguardante il dado truccato).

Nello studio di un esperimento aleatorio costituito da più sotto-esperimenti aleatori, in cui come abbiamo visto nel precedente esercizio sono note alcune probabilità condizionate e altre non condizionate, è utile servirsi del *diagramma ad albero*. Ad esempio, il diagramma ad albero relativo all'Esercizio 1.1 è il seguente:



Le caratteristiche di un diagramma ad albero sono le seguenti:

- ogni nodo corrisponde ad un evento[4] (il primo nodo, la *radice*, corrisponde sempre all'evento certo Ω);
- ogni ramo corrisponde ad una probabilità: nella prima ramificazione ci sono probabilità non condizionate; dalla seconda ramificazione in poi ci sono probabilità condizionate;

[4] Tipicamente, la prima ramificazione descrive il primo sotto-esperimento aleatorio, la seconda il secondo sotto-esperimento aleatorio e così via.

- i rami che escono da un medesimo nodo conducono ad eventi tra loro disgiunti la cui unione è Ω; per tale ragione, le probabilità dei rami che escono da un medesimo nodo sommano a uno;

- scelto un cammino che collega la radice ad un evento, ad esempio $\Omega \to R_1 \to B_2 \to N_3$, moltiplicando le probabilità lungo i rami del cammino, cioè

$$\mathbb{P}(N_3|R_1 \cap B_2)\,\mathbb{P}(B_2|R_1)\,\mathbb{P}(R_1),$$

  sappiamo dalla *regola della catena* che si ottiene la probabilità dell'intersezione degli eventi, infatti

$$\mathbb{P}(R_1 \cap B_2 \cap N_3) \;=\; \mathbb{P}(N_3|R_1 \cap B_2)\,\mathbb{P}(B_2|R_1)\,\mathbb{P}(R_1).$$

  La probabilità $\mathbb{P}(R_1 \cap B_2 \cap N_3)$ si chiama *probabilità del cammino* $\Omega \to R_1 \to B_2 \to N_3$.

La terza proprietà di un diagramma ad albero riportata qui sopra, ovvero che *i rami che escono da un medesimo nodo conducono ad eventi tra loro disgiunti la cui unione è* Ω, può essere abbreviata dicendo che *i rami che escono da un medesimo nodo conducono ad una **partizione*** di Ω. La definizione di *partizione* di Ω è la seguente.

Definizione 1.2. *Si dice che n eventi (o sottoinsiemi) $B_1, \dots, B_n$ di Ω sono una **partizione** (anche detta **schema di alternative**) di Ω se:*

1) *gli insiemi $B_1, \dots, B_n$ sono tra loro disgiunti;*

2) *l'unione degli insiemi $B_1, \dots, B_n$ è Ω:*

$$\Omega \;=\; \bigcup_{i=1}^{n} B_n.$$

Osservazione. *Se una partizione è costituita solamente da **due** insiemi B_1 e B_2, allora B_1 e B_2 sono necessariamente uno il **complementare** dell'altro.*

Esercizio 1.2. *Ci sono due urne: la prima contiene due palline rosse e una bianca; la seconda contiene tre palline rosse e due bianche. Si lancia una moneta: se esce testa si estrae una pallina dalla prima urna, se esce croce si estrae una pallina dalla seconda urna.*

Qual è la probabilità che l'esito del lancio della moneta sia testa e la pallina estratta sia bianca?

Soluzione. L'esperimento aleatorio è costituito da due sotto-esperimenti aleatori: il lancio della moneta seguito dall'estrazione dall'urna che è stata scelta. L'evento di cui dobbiamo calcolare la probabilità è

$$A \;=\; \text{“l'esito del lancio della moneta è testa e la pallina estratta è bianca”}.$$

Elenchiamo gli eventi riguardanti i due sotto-esperimenti aleatori:

$$
\begin{aligned}
T &= \text{“l'esito del lancio della moneta è testa”}, \\
C &= \text{“l'esito del lancio della moneta è croce”} \ = \ T^c, \\
B &= \text{“la pallina estratta è bianca”}, \\
R &= \text{“la pallina estratta è rossa”} \ = \ B^c.
\end{aligned}
$$

Si noti che per ogni sotto-esperimento aleatorio abbiamo considerato una *partizione*[5] o *schema di alternative*. Infatti, gli insiemi T e C sono una partizione di Ω, come anche gli insiemi B ed R. Le probabilità note sono le seguenti:

$$
\begin{aligned}
\mathbb{P}(T) &= \frac{1}{2}, \\
\mathbb{P}(C) &= \frac{1}{2}, \\
\mathbb{P}(B|T) &\underset{\substack{\uparrow \\ 1^{\text{a}}\ \text{urna}}}{=} \frac{1}{3}, \\
\mathbb{P}(R|T) &\underset{\substack{\uparrow \\ 1^{\text{a}}\ \text{urna}}}{=} \frac{2}{3}, \\
\mathbb{P}(B|C) &\underset{\substack{\uparrow \\ 2^{\text{a}}\ \text{urna}}}{=} \frac{3}{5}, \\
\mathbb{P}(R|C) &\underset{\substack{\uparrow \\ 2^{\text{a}}\ \text{urna}}}{=} \frac{2}{5},
\end{aligned}
$$

Il diagramma ad albero associato all'esperimento aleatorio è dunque il seguente:

$$
\begin{tikzcd}
 & & R \\
 & T \arrow[ur, "\mathbb{P}(R|T) = \frac{2}{3}"] \arrow[r, "\mathbb{P}(B|T) = \frac{1}{3}"'] & B \\
\Omega \arrow[ur, "\mathbb{P}(T) = \frac{1}{2}"] \arrow[dr, "\mathbb{P}(C) = \frac{1}{2}"'] & & R \\
 & C \arrow[ur, "\mathbb{P}(R|C) = \frac{3}{5}"] \arrow[r, "\mathbb{P}(B|C) = \frac{2}{5}"'] & B
\end{tikzcd}
$$

L'evento A di cui è richiesta la probabilità può essere espresso in termini degli eventi che abbiamo introdotto come segue:

$$A \ = \ T \cap B.$$

[5] Se non fosse così per qualche sotto-esperimento aleatorio, vorrebbe dire che gli eventi considerati non tengono conto di tutti i possibili risultati del sotto-esperimento in questione.

Quindi, per la *regola della catena*,

$$\mathbb{P}(A) \;=\; \mathbb{P}(B|T)\,\mathbb{P}(T) \;=\; \frac{1}{3}\cdot\frac{1}{2} \;=\; \frac{1}{6}.$$

$\square$

2 Eventi indipendenti

La probabilità condizionata $\mathbb{P}(A|B)$ rappresenta la probabilità dell'evento A sapendo che l'evento B si è verificato. Può succedere che l'informazione che l'evento B si è verificato non alteri la probabilità di A, cioè $\mathbb{P}(A|B) = \mathbb{P}(A)$? Quando questo accade, diremo che A e B sono eventi "indipendenti", nel senso che verificano quanto affermato nella definizione seguente (si noti che nella Definizione 2.1 non compare la probabilità condizionata; questo punto sarà chiarito dal Teorema 2.1).

Definizione 2.1. *Due eventi A e B si dicono* ***indipendenti*** *se*

$$\mathbb{P}(A\cap B) \;=\; \mathbb{P}(A)\,\mathbb{P}(B). \tag{2.1}$$

Notazione. *Se A e B sono eventi indipendenti scriviamo*

$$A \perp\!\!\!\perp B.$$

Si noti che la proprietà di essere indipendenti è *simmetrica*, ovvero se il fatto di sapere che un evento si è verificato "non influenza" la probabilità di un altro evento, vale anche il viceversa, come enunciato nel teorema seguente.

Teorema 2.1.

1) *Se $\mathbb{P}(B) > 0$ allora*

$$A \perp\!\!\!\perp B \qquad\Longleftrightarrow\qquad \mathbb{P}(A|B) \;=\; \mathbb{P}(A).$$

2) *Se $\mathbb{P}(A) > 0$ allora*

$$A \perp\!\!\!\perp B \qquad\Longleftrightarrow\qquad \mathbb{P}(B|A) \;=\; \mathbb{P}(B).$$

Osservazione. *In altri termini, se $\mathbb{P}(A) > 0$ e $\mathbb{P}(B) > 0$, le tre uguaglianze seguenti sono equivalenti:*

$$\mathbb{P}(A\cap B) \;=\; \mathbb{P}(A)\,\mathbb{P}(B), \qquad \mathbb{P}(A|B) \;=\; \mathbb{P}(A), \qquad \mathbb{P}(B|A) \;=\; \mathbb{P}(B).$$

Viene adottata come definizione di indipendenza la (2.1) *in quanto è simmetrica rispetto ad A e B, inoltre non necessita di assunzioni su $\mathbb{P}(A)$ o $\mathbb{P}(B)$.*

Dimostrazione (del Teorema 2.1). Dimostriamo solo l'affermazione 1), dato che la 2) si dimostra allo stesso modo.
La 1) è una conseguenza della seguente catene di equivalenze:

$$
\begin{aligned}
A \perp\!\!\!\perp B \quad \Longleftrightarrow \quad \mathbb{P}(A \cap B) \;=\; \mathbb{P}(A)\,\mathbb{P}(B) \quad &\Longleftrightarrow \quad \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \;=\; \frac{\mathbb{P}(A)\,\mathbb{P}(B)}{\mathbb{P}(B)} \\
&\Longleftrightarrow \quad \mathbb{P}(A|B) \;=\; \mathbb{P}(A).
\end{aligned}
$$

□

Osservazione. *La nozione di* indipendenza *non è da confondersi con quella di* insiemi disgiunti*. Due eventi A e B sono contemporaneamente disgiunti e indipendenti solo quando $\mathbb{P}(A) = 0$ oppure $\mathbb{P}(B) = 0$. Infatti*

$$
0 \;=\; \mathbb{P}(\emptyset) \underset{\substack{\uparrow \\ A \text{ e } B \text{ disgiunti}}}{=} \mathbb{P}(A \cap B) \underset{\substack{\uparrow \\ A \text{ e } B \text{ indip.}}}{=} \mathbb{P}(A)\,\mathbb{P}(B).
$$

L'uguaglianza $\mathbb{P}(A)\,\mathbb{P}(B) = 0$ è verificata solo quando $\mathbb{P}(A) = 0$ oppure $\mathbb{P}(B) = 0$.

Teorema 2.2. *Siano A e B due eventi indipendenti. Allora anche A^c e B, oppure A e B^c, oppure A^c e B^c sono coppie di eventi indipendenti.*

Dimostrazione. Dimostriamo ad esempio che A^c, B è una coppia di eventi indipendenti. Dobbiamo mostrare che vale l'uguaglianza seguente:

$$
\mathbb{P}(A^c \cap B) \;=\; \mathbb{P}(A^c)\,\mathbb{P}(B).
$$

Dato che

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

e gli eventi $A \cap B$ e $A^c \cap B$ sono disgiunti, per l'*additività finita* abbiamo che

$$
\mathbb{P}(B) \;=\; \mathbb{P}(A \cap B) + \mathbb{P}(A^c \cap B) \underset{\substack{\uparrow \\ A \text{ e } B \text{ indip.}}}{=} \mathbb{P}(A)\,\mathbb{P}(B) + \mathbb{P}(A^c \cap B).
$$

Quindi

$$
\mathbb{P}(A^c \cap B) \;=\; \mathbb{P}(B) - \mathbb{P}(A)\,\mathbb{P}(B) \;=\; (1 - \mathbb{P}(A))\,\mathbb{P}(B) \;=\; \mathbb{P}(A^c)\,\mathbb{P}(B).
$$

□

La seguente definizione generalizza la nozione di indipendenza ad una famiglia di tre o più eventi.

Definizione 2.2. *Tre eventi A, B, C si dicono **indipendenti** se valgono simultaneamente le quattro uguaglianze seguenti:*

$$\begin{aligned}\mathbb{P}(A \cap B) &= \mathbb{P}(A)\,\mathbb{P}(B), \\ \mathbb{P}(A \cap C) &= \mathbb{P}(A)\,\mathbb{P}(C), \\ \mathbb{P}(B \cap C) &= \mathbb{P}(B)\,\mathbb{P}(C), \\ \mathbb{P}(A \cap B \cap C) &= \mathbb{P}(A)\,\mathbb{P}(B)\,\mathbb{P}(C).\end{aligned}$$

*Più in generale, n eventi $A_1, A_2, \ldots, A_n$ si dicono **indipendenti** se valgono simultaneamente le uguaglianze seguenti:*

$$\mathbb{P}(A_{i_1} \cap \cdots \cap A_{i_k}) = \mathbb{P}(A_{i_1}) \cdots \mathbb{P}(A_{i_k}),$$

per ogni $k = 2, \ldots, n$ e per ogni scelta di indici $i_1, \ldots, i_k$, tutti distinti tra loro e compresi tra 1 e n.

Concludiamo questa sezione con alcuni esercizi.

Esercizio 2.1. *Si lancia un dado regolare a sei facce. Siano*

$$\begin{aligned}A &= \text{“esce un numero maggiore di 4”}, \\ B &= \text{“esce un numero pari”}.\end{aligned}$$

Quanto valgono $\mathbb{P}(A)$ e $\mathbb{P}(A|B)$?

Soluzione. Consideriamo $(\Omega, \mathbb{P})$ con $\Omega = \{1, 2, 3, 4, 5, 6\}$ e $\mathbb{P}$ probabilità uniforme. Allora

$$\begin{aligned}A &= \{5, 6\}, \\ B &= \{2, 4, 6\}.\end{aligned}$$

Quindi

$$\begin{aligned}\mathbb{P}(A) &= \frac{1}{3}, \\ \mathbb{P}(A|B) &= \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(\{6\})}{\mathbb{P}(B)} = \frac{1}{3}.\end{aligned}$$

Perciò $\mathbb{P}(A) = \mathbb{P}(A|B)$, quindi A e B sono indipendenti. □

Osservazione. *Detto in modo poco preciso, due eventi sono indipendenti se e solo se “si intersecano nelle giuste proporzioni”. Come abbiamo appena visto nell'Esercizio 2.1, questo può accadere anche per eventi riferiti allo stesso sotto-esperimento aleatorio.*
Un'altra situazione particolarmente importante in cui ci possono essere eventi indipendenti si verifica quando tali eventi riguardano sotto-esperimenti aleatori distinti che “non si influenzano tra loro” (come avviene nell'esempio che segue). Si noti che il testo dell'esercizio generalmente non dice in maniera esplicita che l'indipendenza vale; tuttavia, se

non valesse, dovrebbe spiegare in che modo i sotto-esperimenti aleatori si influenzano tra loro (come accade ad esempio negli Esercizi 1.1 e 1.2).

Esercizio 2.2. *Lanciamo*[a] *una moneta e un dado a quattro facce, entrambi non truccati. Determinare uno spazio di probabilità che descriva l'esperimento aleatorio.*

[a]Dato che i due sotto-esperimenti aleatori "non si influenzano tra loro", non ha alcuna importanza quale si effettua per primo (possono anche svolgersi contemporaneamente).

Soluzione. Uno spazio campionario naturale è il seguente:

$$\begin{aligned}\Omega &= \{\mathrm{t},\mathrm{c}\} \times \{1,2,3,4\} \\ &= \big\{(\mathrm{t},1),(\mathrm{t},2),(\mathrm{t},3),(\mathrm{t},4),(\mathrm{c},1),(\mathrm{c},2),(\mathrm{c},3),(\mathrm{c},4)\big\}.\end{aligned}$$

Resta da determinare $\mathbb{P}$, che significa assegnare la probabilità di tutti gli eventi elementari $\mathbb{P}(\{(\mathrm{t},1)\}), \mathbb{P}(\{(\mathrm{t},2)\}), \ldots, \mathbb{P}(\{(\mathrm{c},4)\})$. Intuitivamente è naturale aspettarsi che gli eventi elementari (che sono otto) siano *equiprobabili:*

$$\mathbb{P}(\{(\mathrm{t},1)\}) \;=\; \mathbb{P}(\{(\mathrm{t},2)\}) \;=\; \cdots \;=\; \mathbb{P}(\{(\mathrm{c},4)\}) \;=\; \frac{1}{8}. \tag{2.2}$$

Ciò significa che $\mathbb{P}$ è la *probabilità uniforme*, quindi vale la formula *casi favorevoli/casi possibili:*

$$\mathbb{P}(A) \;=\; \frac{\mathrm{n^o \ di\ eventi\ elementari\ che\ compongono\ } A}{8},$$

per ogni sottoinsieme A di Ω.
Dimostriamo dunque la validità di (2.2). Ciò che sappiamo è solamente che dado e moneta non sono truccati, e inoltre che i due sotto-esperimenti aleatori "non si influenzano tra loro" (sono indipendenti), infatti diversamente sarebbe descritto nel testo dell'esercizio il modo in cui si influenzano. Esprimiamo tutto questo in formule introducendo gli eventi che riguardano i singoli sotto-esperimenti aleatori:

$$\begin{aligned} T &= \text{“l'esito del lancio della moneta è testa”}, \\ C &= \text{“l'esito del lancio della moneta è croce”} \;=\; T^c, \\ E_i &= \text{“l'esito del lancio del dado è } i\text{”}, \qquad i = 1,2,3,4. \end{aligned}$$

Dato che dado e moneta non sono truccati, abbiamo che

$$\mathbb{P}(T) \;=\; \mathbb{P}(C) \;=\; \frac{1}{2}$$

e

$$\mathbb{P}(E_1) \;=\; \mathbb{P}(E_2) \;=\; \mathbb{P}(E_3) \;=\; \mathbb{P}(E_4) \;=\; \frac{1}{4}.$$

Dire che i due sotto-esprimenti aleatori "non si influenzano tra loro" significa dire, in termini matematici, che *gli eventi ad essi riferiti sono tra loro indipendenti.* Quindi vale che

$$\mathbb{P}(T \cap E_1) \;=\; \mathbb{P}(T)\,\mathbb{P}(E_1) \;=\; \frac{1}{8},$$

$$
\begin{aligned}
\mathbb{P}(T \cap E_2) &= \mathbb{P}(T)\,\mathbb{P}(E_2) = \frac{1}{8}, \\
\mathbb{P}(T \cap E_3) &= \mathbb{P}(T)\,\mathbb{P}(E_3) = \frac{1}{8}, \\
\mathbb{P}(T \cap E_4) &= \mathbb{P}(T)\,\mathbb{P}(E_4) = \frac{1}{8}, \\
\mathbb{P}(C \cap E_1) &= \mathbb{P}(C)\,\mathbb{P}(E_1) = \frac{1}{8}, \\
\mathbb{P}(C \cap E_2) &= \mathbb{P}(C)\,\mathbb{P}(E_2) = \frac{1}{8}, \\
\mathbb{P}(C \cap E_3) &= \mathbb{P}(C)\,\mathbb{P}(E_3) = \frac{1}{8}, \\
\mathbb{P}(C \cap E_4) &= \mathbb{P}(C)\,\mathbb{P}(E_4) = \frac{1}{8}.
\end{aligned}
$$

Poiché

$$T \cap E_i = \big\{(\text{testa}, i)\big\} \qquad \text{e} \qquad C \cap E_i = \big\{(\text{croce}, i)\big\}$$

per ogni $i = 1, 2, 3, 4$, abbiamo dimostrato che tutti gli eventi elementari hanno la stessa probabilità pari a $\frac{1}{8}$, ovvero che vale (2.2). $\square$

Osservazione. *Il risultato ottenuto nell'Esercizio 2.2 vale in generale: se un esperimento aleatorio si compone di più sotto-esperimenti aleatori tra loro "indipendenti", ognuno dei quali ha esiti equiprobabili, anche l'esperimento aleatorio nel suo complesso avrà esiti equiprobabili.*

Vediamo infine un esercizio in cui si utilizza anche quanto visto nella prima sezione.

Esercizio 2.3. *Nel gioco del lotto si estraggono* ***senza reimmissione*** *cinque numeri da un'urna che contiene* 90 *palline numerate da* 1 *a* 90.

1) *Determinare uno spazio di probabilità che descriva l'esperimento aleatorio.*

2) *Come cambia la risposta al punto precedente se le estrazioni avvengono* ***con reimmissione****?*

Soluzione.

1) Uno spazio campionario naturale è l'insieme di tutte le cinquine *ordinate* di numeri *distinti* da 1 a 90:

$$\Omega = \big\{(x_1, x_2, x_3, x_4, x_5)\colon x_i \text{ è un numero naturale da 1 a 90, con } x_i \neq x_j \text{ se } i \neq j\big\}.$$

Si noti che

$$|\Omega| = 90 \cdot 89 \cdot 88 \cdot 87 \cdot 86.$$

Resta da determinare $\mathbb{P}$. Intuitivamente, ci aspettiamo che $\mathbb{P}$ sia la *probabilità uniforme*, ovvero

$$\mathbb{P}(\{(x_1, x_2, x_3, x_4, x_5)\}) = \frac{1}{90 \cdot 89 \cdot 88 \cdot 87 \cdot 86}, \qquad \forall\, (x_1, x_2, x_3, x_4, x_5) \in \Omega. \tag{2.3}$$

Dimostriamo questo risultato. Introduciamo gli eventi che riguardano i singoli sotto-esperimenti aleatori, ovvero le singole estrazioni:

$$E_{i,n} \;=\; \text{“all'}i\text{-esima estrazione esce il numero } n\text{”}, \qquad i = 1, 2, 3, 4, 5,\; n = 1, \dots, 90.$$

Consideriamo ad esempio la cinquina $(17, 54, 2, 76, 45)$. Allora possiamo esprimere l'evento elementare $\{(17, 54, 12, 76, 45)\}$ in termini degli eventi $E_{i,n}$ come segue:

$$\big\{(17, 54, 2, 76, 45)\big\} \;=\; E_{1,17} \cap E_{2,54} \cap E_{3,12} \cap E_{4,76} \cap E_{5,45}.$$

Per determinare la probabilità di $\{(17, 54, 12, 76, 45)\}$ (e dimostrare la validità di (2.3)) possiamo usare la *regola della catena:*

$$\begin{aligned}\mathbb{P}(\{(17, 54, 2, 76, 45)\}) &= \mathbb{P}(E_{1,17} \cap E_{2,54} \cap E_{3,12} \cap E_{4,76} \cap E_{5,45}) \\ &= \mathbb{P}(E_{5,45}|E_{1,17} \cap E_{2,54} \cap E_{3,12} \cap E_{4,76})\, \mathbb{P}(E_{4,76}|E_{1,17} \cap E_{2,54} \cap E_{3,12}) \times \\ &\quad \times\, \mathbb{P}(E_{3,12}|E_{1,17} \cap E_{2,54})\, \mathbb{P}(E_{2,54}|E_{1,17})\, \mathbb{P}(E_{1,17}).\end{aligned}$$

Tutte queste probabilità sono note, poiché ad ogni estrazione conosciamo la composizione dell'urna. In particolare, si ha che

$$\begin{aligned}\mathbb{P}(E_{1,17}) &= \frac{1}{90}, \\ \mathbb{P}(E_{2,54}|E_{1,17}) &= \frac{1}{89}, \\ \mathbb{P}(E_{3,12}|E_{1,17} \cap E_{2,54}) &= \frac{1}{88}, \\ \mathbb{P}(E_{4,76}|E_{1,17} \cap E_{2,54} \cap E_{3,12}) &= \frac{1}{87}, \\ \mathbb{P}(E_{5,45}|E_{1,17} \cap E_{2,54} \cap E_{3,12} \cap E_{4,76}) &= \frac{1}{86}.\end{aligned}$$

Quindi

$$\mathbb{P}(\{(17, 54, 2, 76, 45)\}) \;=\; \frac{1}{90 \cdot 89 \cdot 88 \cdot 87 \cdot 86}.$$

È chiaro che questo ragionamento vale per qualunque cinquina, non solo per $(17, 54, 2, 76, 45)$. Possiamo dunque concludere che (2.3) è valida, quindi che $\mathbb{P}$ è la *probabilità uniforme.*

2) Se le estrazioni avvengono *con reimmissione*, uno spazio campionario naturale è l'insieme di tutte le cinquine *ordinate* di numeri da 1 a 90, *non necessariamente distinti:*

$$\Omega \;=\; \big\{(x_1, x_2, x_3, x_4, x_5)\colon x_i \text{ è un numero naturale da 1 a 90}\big\}.$$

Si noti che in tal caso Ω può anche essere scritto come segue:

$$\begin{aligned}\Omega &= \{1, 2, 3, 4, \dots, 88, 89, 90\}^5 \\ &= \underbrace{\{1, 2, 3, 4, \dots, 88, 89, 90\} \times \cdots \times \{1, 2, 3, 4, \dots, 88, 89, 90\}}_{\text{5 volte}}.\end{aligned}$$

Notiamo inoltre che

$$|\Omega| \;=\; 90^5.$$

Resta da determinare $\mathbb{P}$. Dato che in questo caso le estrazioni non si influenzano tra di loro, sono dunque indipendenti, inoltre ogni estrazione ha esiti equiprobabili, possiamo concludere senza fare conti (come già osservato alla fine dell'Esercizio 2.2) che la probabilità $\mathbb{P}$ è *uniforme*, ovvero

$$\mathbb{P}(\{(x_1, x_2, x_3, x_4, x_5)\}) \ = \ \frac{1}{90^5}, \qquad \forall\, (x_1, x_2, x_3, x_4, x_5) \in \Omega. \tag{2.4}$$

Tuttavia, per maggiore chiarezza, possiamo comunque procedere come prima e dimostrarlo. Utilizziamo le stesse notazioni introdotte al punto precedente, quindi

$$E_{i,n} \ = \ \text{“all'}i\text{-esima estrazione esce il numero } n\text{''}, \qquad i = 1, 2, 3, 4, 5,\ n = 1, \dots, 90.$$

Consideriamo ad esempio la cinquina $(52, 34, 65, 34, 52)$. Allora

$$\big\{(52, 34, 65, 34, 52)\big\} \ = \ E_{1,52} \cap E_{2,34} \cap E_{3,65} \cap E_{4,34} \cap E_{5,52}.$$

A differenza del punto precedente, nel caso *con reimmissione* gli eventi che si riferiscono ad estrazioni differenti sono tra loro *indipendenti*, quindi vale che

$$\begin{aligned}\mathbb{P}(\{(52, 34, 65, 34, 52)\}) \ &= \ \mathbb{P}(E_{1,52} \cap E_{2,34} \cap E_{3,65} \cap E_{4,34} \cap E_{5,52}) \\ &= \ \mathbb{P}(E_{1,52})\, \mathbb{P}(E_{2,34})\, \mathbb{P}(E_{3,65})\, \mathbb{P}(E_{4,34})\, \mathbb{P}(E_{5,52}) \ = \ \frac{1}{90^5}.\end{aligned}$$

Abbiamo dunque dimostrato la validità di (2.4). □

3 Formula delle probabilità totali

Abbiamo già introdotto il concetto di *partizione* (o *schema di alternative*), si veda la Definizione 1.2. Possiamo dunque enunciare la *formula delle probabilità totali*.

Teorema 3.1 (Formula delle probabilità totali). *Sia $B_1, \dots, B_n$ una partizione di Ω. Allora per ogni evento A vale la formula:*

$$\underbrace{\mathbb{P}(A)}_{\text{prob. totale di } A} \ = \ \sum_{i=1}^n \underbrace{\mathbb{P}(A \cap B_i)}_{\text{prob. parziale di } A}.$$

Se inoltre si ha che $\mathbb{P}(B_i) > 0$, per ogni $i = 1, \dots, n$, allora vale che

$$\mathbb{P}(A) \ = \ \sum_{i=1}^n \mathbb{P}(A|B_i)\, \mathbb{P}(B_i).$$

Dimostrazione. Si noti che

$$A \ = \ A \cap \Omega \ = \ A \cap \left(\bigcup_{i=1}^n B_i\right) \ = \ \bigcup_{i=1}^n (A \cap B_i),$$

dove l'ultima uguaglianza segue dalla *proprietà distributiva dell'intersezione rispetto all'unione*. Inoltre, dato che gli eventi $B_1, \ldots, B_n$ sono disgiunti (ovvero, non hanno elementi in comune), segue che anche gli eventi $A \cap B_1, \ldots, A \cap B_n$ sono disgiunti. Quindi, per la proprietà di *additività finita* della probabilità,

$$\mathbb{P}(A) \ = \ \sum_{i=1}^{n} \mathbb{P}(A \cap B_i).$$

Infine, supponiamo che $\mathbb{P}(B_i) > 0$, per ogni $i = 1, \ldots, n$. Allora, dalla *regola della catena*,

$$\mathbb{P}(A \cap B_i) \ = \ \mathbb{P}(A|B_i)\,\mathbb{P}(B_i), \qquad \forall\, i = 1, \ldots, n.$$

Quindi

$$\mathbb{P}(A) \ = \ \sum_{i=1}^{n} \mathbb{P}(A \cap B_i) \ \overset{\textit{regola catena}}{=} \ \sum_{i=1}^{n} \mathbb{P}(A|B_i)\,\mathbb{P}(B_i).$$

□

OSSERVAZIONE. *Grazie alla formula delle probabilità totali, possiamo dire qualcosa di più sul diagramma ad albero di un esperimento aleatorio. Infatti, segue dalla formula delle probabilità totali la seguente importante proprietà di un diagramma ad albero:*

- *la probabilità di un qualunque evento che compare nel diagramma ad albero (in altri termini, un evento che corrisponde ad un nodo dell'albero) è la somma delle probabilità di tutti i cammini che dalla radice Ω conducono ad esso.*

Il modo migliore per rendersi conto della validità di questa proprietà è con un esercizio, come quello che segue.

Esercizio 3.1. *Un'urna contiene* 10 *palline di cui* 6 *bianche e* 4 *rosse. Si estraggono due palline* ***senza reimmissione***. *Calcolare la probabilità dell'evento*

$$B_2 \ = \ \text{“}\textit{la seconda estratta è bianca}\text{”}.$$

Soluzione. Introduciamo gli eventi riguardanti le due estrazioni:

$$\begin{aligned} B_1 \ &= \ \text{“la prima estratta è bianca”}, \\ R_1 \ &= \ \text{“la prima estratta è rossa”} \ = \ B_1^c, \\ B_2 \ &= \ \text{“la seconda estratta è bianca”}, \\ R_2 \ &= \ \text{“la seconda estratta è rossa”} \ = \ B_2^c. \end{aligned}$$

Le probabilità note sono le seguenti (in realtà, per trovare la probabilità di B_2, come richiesto dall'esercizio, non serve riportare tutte queste probabilità):

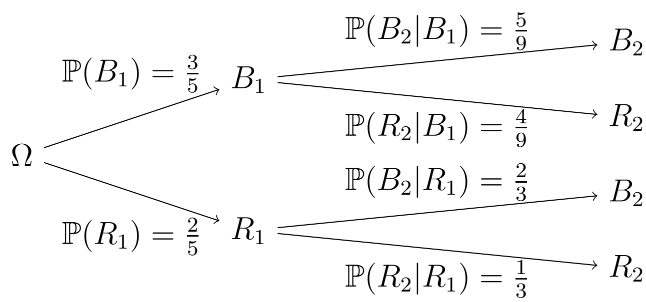
$$\begin{aligned} \mathbb{P}(B_1) \ &= \ \frac{6}{10} \ = \ \frac{3}{5}, \\ \mathbb{P}(R_1) \ &= \ 1 - \mathbb{P}(B_1) \ = \ \frac{2}{5}, \end{aligned}$$

$$
\begin{aligned}
\mathbb{P}(B_2|B_1) &= \frac{5}{9}, \\
\mathbb{P}(R_2|B_1) &= 1 - \mathbb{P}(B_2|B_1) \;=\; \frac{4}{9}, \\
\mathbb{P}(B_2|R_1) &= \frac{6}{9} \;=\; \frac{2}{3}, \\
\mathbb{P}(R_2|R_1) &= 1 - \mathbb{P}(B_2|R_1) \;=\; \frac{3}{9} \;=\; \frac{1}{3}.
\end{aligned}
$$

Allora, dalla *formula delle probabilità totali* otteniamo

$$
\mathbb{P}(B_2) \;=\; \mathbb{P}(B_2|B_1)\,\mathbb{P}(B_1) + \mathbb{P}(B_2|R_1)\,\mathbb{P}(R_1) \;=\; \frac{5}{9}\cdot\frac{3}{5} + \frac{2}{3}\cdot\frac{2}{5} \;=\; \frac{3}{5}.
$$

A tale risultato si può arrivare utilizzando il *diagramma ad albero:*



Infatti, sappiamo che *la probabilità dell'evento B_2 è la somma delle probabilità di tutti i cammini che dalla radice Ω conducono a B_2 stesso.* Ci sono solo due cammini: $\Omega \to B_1 \to B_2$ e $\Omega \to R_1 \to B_2$. Quindi

$$
\mathbb{P}(B_2) \;=\; \mathbb{P}(\text{cammino } \Omega \to B_1 \to B_2) + \mathbb{P}(\text{cammino } \Omega \to R_1 \to B_2).
$$

Ricordando che la probabilità di un cammino è il prodotto delle probabilità dei suoi rami, otteniamo

$$
\mathbb{P}(B_2) \;=\; \mathbb{P}(B_2|B_1)\,\mathbb{P}(B_1) + \mathbb{P}(B_2|R_1)\,\mathbb{P}(R_1),
$$

che corrisponde alla formula delle probabilità totali. □

4 Formula di Bayes

Le formule che seguono dalla definizione di probabilità condizionata sono tre: la *regola della catena*, la *formula delle probabilità totali* e la *formula di Bayes*, che ora presentiamo. Sono molto importanti in quanto permettono di risolvere problemi di Calcolo delle probabilità utilizzando solo gli eventi, senza dover introdurre esplicitamente lo spazio campionario. Sono inoltre strettamente collegate al *diagramma ad albero*.

Veniamo dunque alla *formula di Bayes*, che stabilisce la relazione tra le probabilità condizionate $\mathbb{P}(A|B)$ e $\mathbb{P}(B|A)$.

Teorema 4.1 (Formula di Bayes). *Siano A e B due eventi tali che $\mathbb{P}(A) > 0$ e $\mathbb{P}(B) > 0$, allora vale la formula*

$$\mathbb{P}(B|A) \;=\; \frac{\mathbb{P}(A|B)\,\mathbb{P}(B)}{\mathbb{P}(A)}.$$

Dimostrazione. Per definizione di $\mathbb{P}(B|A)$ si ha che

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

Utilizzando la *regola della catena*, possiamo riscrivere il numeratore come segue

$$\mathbb{P}(A\cap B) \;=\; \mathbb{P}(A|B)\,\mathbb{P}(B).$$

Quindi

$$\mathbb{P}(B|A) \underset{\substack{\uparrow \\ \text{defn. di } \mathbb{P}(B|A)}}{=} \frac{\mathbb{P}(A\cap B)}{\mathbb{P}(A)} \underset{\substack{\uparrow \\ \text{reg. catena}}}{=} \frac{\mathbb{P}(A|B)\mathbb{P}(B)}{\mathbb{P}(A)},$$

che conclude la dimostrazione. □

Osservazione (utilizzo della formula di Bayes). *Come già detto, in un problema di Calcolo delle probabilità molte probabilità condizionate sono note. Questo accade ad esempio quando in $\mathbb{P}(A|B)$ l'evento B riguarda il primo (in ordine di tempo) sotto-esperimento aleatorio, mentre l'evento A riguarda il secondo sotto-esperimento aleatorio. Se ora consideriamo la probabilità condizionata $\mathbb{P}(B|A)$, in essa gli eventi B ed A non sono disposti nell'ordine temporale naturale[6]. Per questo motivo la probabilità $\mathbb{P}(B|A)$ non è in generale nota (a meno che i due sotto-esperimenti aleatori non siano indipendenti, nel qual caso $\mathbb{P}(B|A) = \mathbb{P}(B)$). È in situazioni del genere che è utile utilizzare la formula di Bayes per calcolare $\mathbb{P}(B|A)$.*

Esercizio 4.1. *Ci sono due urne: la prima urna contiene una pallina bianca e due palline rosse, mentre la seconda contiene due palline bianche e cinque palline rosse. Si lancia una moneta: se esce testa si estrae una pallina dalla prima urna, se esce croce si estrae una pallina dalla seconda.*

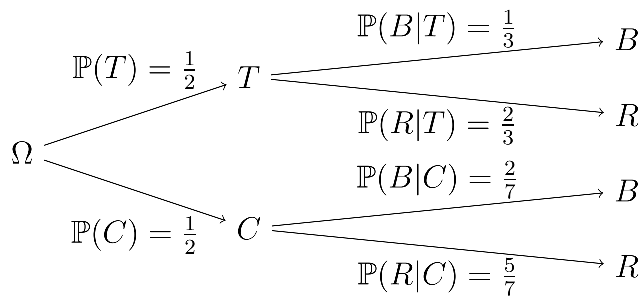
Sapendo che è stata estratta una pallina bianca, calcolare la probabilità che l'esito del lancio della moneta sia stato testa.

Soluzione. Introduciamo gli eventi riguardanti i due sotto-esperimenti aleatori:

$$\begin{aligned}
T &= \text{“l'esito del lancio della moneta è testa”} \;=\; \text{“si sceglie la prima urna”},\\
C &= \text{“l'esito del lancio della moneta è croce”} \;=\; \text{“si sceglie la seconda urna”} \;=\; T^c,\\
B &= \text{“si estrae una pallina bianca”},\\
R &= \text{“si estrae una pallina rossa”} \;=\; B^c.
\end{aligned}$$

Il diagramma ad albero dell'esperimento aleatorio è il seguente:

[6] Con riferimento al diagramma ad albero, il nodo A è un *“figlio”* del nodo B.



La probabilità richiesta è la seguente probabilità condizionata

$$\mathbb{P}(T|B),$$

che si calcola con la *formula di Bayes:*

$$\mathbb{P}(T|B) \;=\; \frac{\mathbb{P}(B|T)\,\mathbb{P}(T)}{\mathbb{P}(B)}.$$

Le due probabilità a numeratore sono note, mentre (come accade spesso quando si usa la formula di Bayes) il denominatore va calcolato con la *formula delle probabilità totali:*

$$\mathbb{P}(B) \;=\; \mathbb{P}(B|T)\,\mathbb{P}(T) + \mathbb{P}(B|C)\,\mathbb{P}(C).$$

Quindi

$$\mathbb{P}(T|B) \;=\; \frac{\mathbb{P}(B|T)\,\mathbb{P}(T)}{\mathbb{P}(B|T)\,\mathbb{P}(T) + \mathbb{P}(B|C)\,\mathbb{P}(C)} \;=\; \frac{\frac{1}{3}\cdot\frac{1}{2}}{\frac{1}{3}\cdot\frac{1}{2}+\frac{2}{7}\cdot\frac{1}{2}} \;=\; \frac{7}{13}.$$

□

5 *Esempi e paradossi*

Esempio 5.1. *In un'urna ci sono due palline che possono essere rosse (R) o bianche (B). La composizione esatta non è nota, quindi le composizioni possibili sono:*

$$RR, \quad RB, \quad BB.$$

Supponiamo che, in base alle informazioni a disposizione, sia ragionevole assegnare uguale probabilità pari a $1/3$ *alle tre composizioni (ipotesi) possibili, che denotiamo* H_0, H_1 *e* H_2.

1) *Se si estrae una pallina dall'urna, qual è la probabilità che sia bianca?*

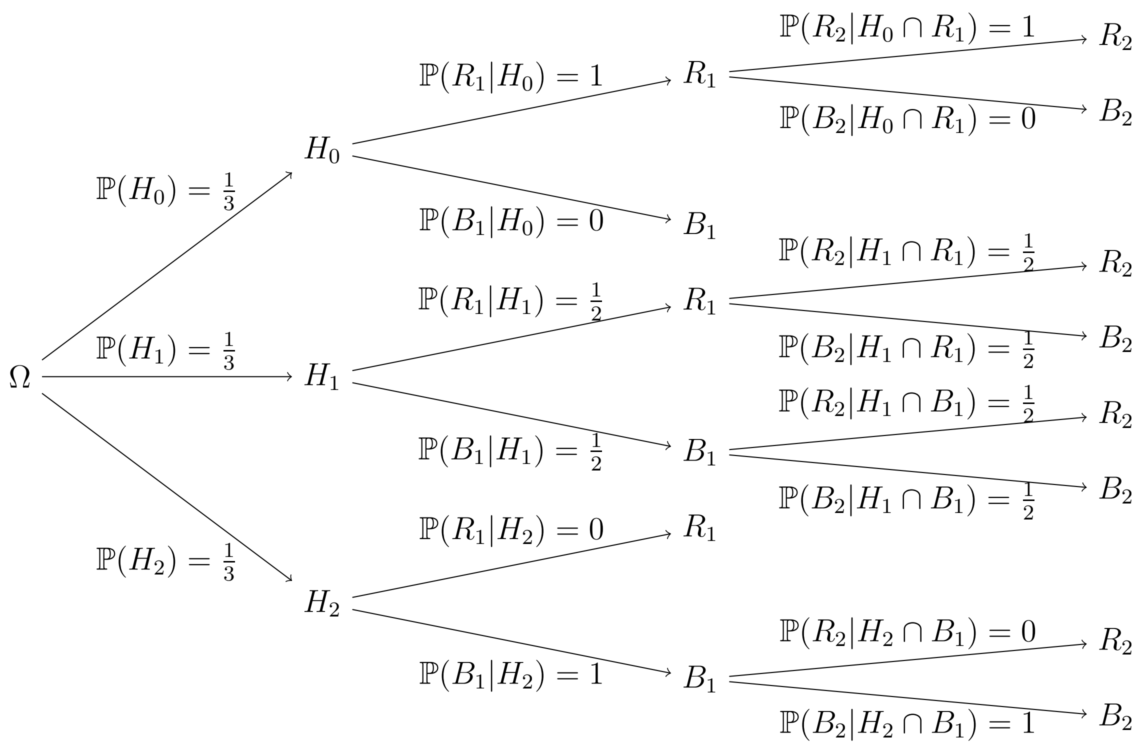
2) *Si effettuano tre estrazioni con reimmissione: sapendo che le prime due palline estratte sono bianche, qual è la probabilità che anche la terza pallina estratta sia bianca?*

Soluzione. Introduciamo i seguenti eventi:

$H_0 \;=\;$ "nell'urna ci sono due palline rosse",

$$
\begin{aligned}
H_1 &= \text{“nell’urna ci sono una pallina rossa e una pallina bianca”}, \\
H_2 &= \text{“nell’urna ci sono due palline bianche”}, \\
R_i &= \text{“all’}i\text{-esima estrazione esce una pallina rossa”}, && i = 1, 2, 3, \\
B_i &= \text{“all’}i\text{-esima estrazione esce una pallina bianca”} \;=\; R_i^c, && i = 1, 2, 3.
\end{aligned}
$$

Il diagramma ad albero dell’esperimento aleatorio fino alla *seconda* estrazione è il seguente:



1) La probabilità richiesta è $\mathbb{P}(B_1)$. Dalla *formula delle probabilità totali* (o, equivalentemente, dal diagramma ad albero), si ottiene

$$\mathbb{P}(B_1) \;=\; \mathbb{P}(B_1|H_0)\,\mathbb{P}(H_0) + \mathbb{P}(B_1|H_1)\,\mathbb{P}(H_1) + \mathbb{P}(B_1|H_2)\,\mathbb{P}(H_2) \;=\; \frac{1}{2}.$$

2) La probabilità richiesta è $\mathbb{P}(B_3|B_1 \cap B_2)$. Dalla definizione di probabilità condizionata, si ha che

$$\mathbb{P}(B_3|B_1 \cap B_2) \;=\; \frac{\mathbb{P}(B_1 \cap B_2 \cap B_3)}{\mathbb{P}(B_1 \cap B_2)}$$

Calcoliamo denominatore e numeratore con la *formula delle probabilità totali:*

$$\mathbb{P}(B_1 \cap B_2) \;=\; \mathbb{P}(B_1 \cap B_2 \cap H_0) + \mathbb{P}(B_1 \cap B_2 \cap H_1) + \mathbb{P}(B_1 \cap B_2 \cap H_2)$$

e

$$
\begin{aligned}
\mathbb{P}(B_1 \cap B_2 \cap B_3) \;=\; &\mathbb{P}(B_1 \cap B_2 \cap B_3 \cap H_0) + \mathbb{P}(B_1 \cap B_2 \cap B_3 \cap H_1) \\
&+ \mathbb{P}(B_1 \cap B_2 \cap B_3 \cap H_2).
\end{aligned}
$$

Le probabilità a destra del segno di uguaglianza si calcolano con la *regola della catena*. Si ottiene

$$\begin{aligned}\mathbb{P}(B_1 \cap B_2) &= \mathbb{P}(B_1 \cap B_2 \cap H_0) + \mathbb{P}(B_1 \cap B_2 \cap H_1) + \mathbb{P}(B_1 \cap B_2 \cap H_2) \\ &= 0 + \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{3} + \frac{1}{3} = \frac{5}{12}\end{aligned}$$

e

$$\begin{aligned}\mathbb{P}(B_1 \cap B_2 \cap B_3) &= \mathbb{P}(B_1 \cap B_2 \cap B_3 \cap H_0) + \mathbb{P}(B_1 \cap B_2 \cap B_3 \cap H_1) \\ &\quad + \mathbb{P}(B_1 \cap B_2 \cap B_3 \cap H_2) \\ &= 0 + \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{3} + \frac{1}{3} = \frac{3}{8}.\end{aligned}$$

Quindi

$$\mathbb{P}(B_3 | B_1 \cap B_2) = \frac{\mathbb{P}(B_1 \cap B_2 \cap B_3)}{\mathbb{P}(B_1 \cap B_2)} = \frac{\frac{3}{8}}{\frac{5}{12}} = \frac{9}{10}.$$

□

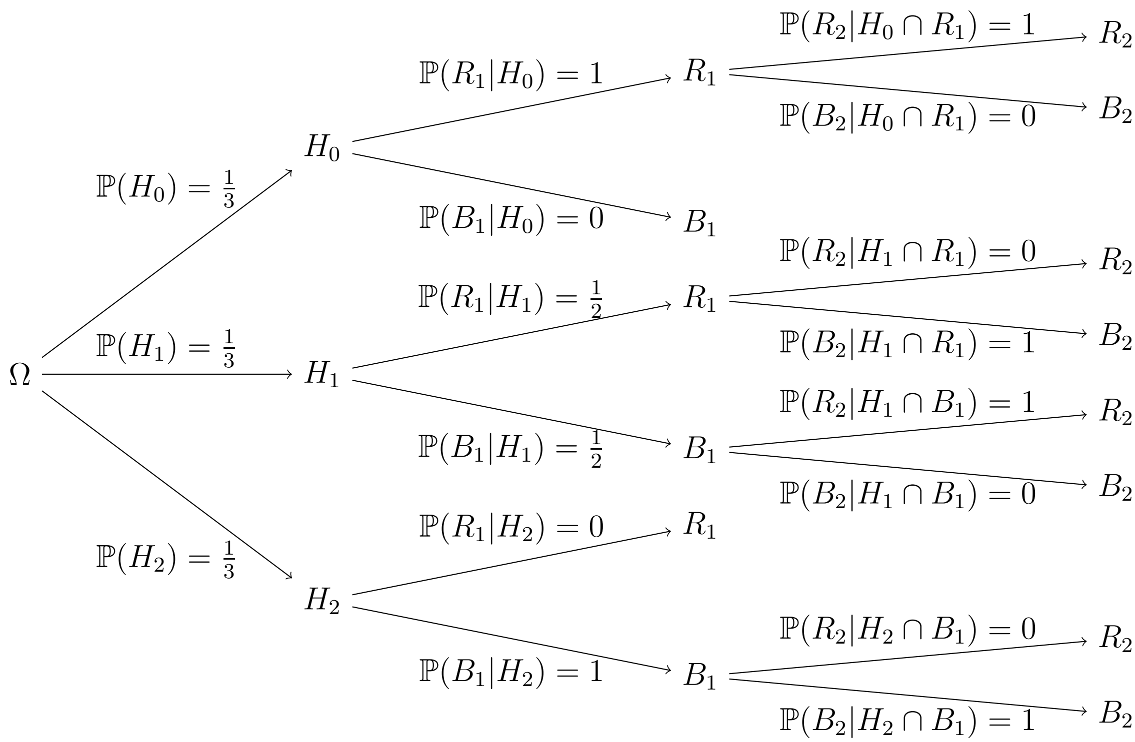
Esempio 5.2 (Paradosso delle tre carte). *Giochiamo con tre carte. Una è bianca su entrambi i lati, una è rossa su entrambi i lati e una è bianca da un lato e rossa dall'altro. Ogni carta è nascosta in una scatoletta nera. Il giocatore sceglie una delle tre scatolette, estrae la carta e la posa sul tavolo in modo che sia visibile un solo lato.*

Sapendo che il lato superiore della carta è bianco, qual è la probabilità che l'altro lato sia rosso?

Soluzione. Possiamo utilizzare gli stessi eventi introdotti nell'Esercizio 5.1 per descrivere questo esperimento aleatorio:

$$\begin{aligned}H_0 &= \text{“la carta scelta dal giocatore ha entrambi i lati rossi”}, \\ H_1 &= \text{“la carta scelta dal giocatore ha un lato rosso e l'altro bianco”}, \\ H_2 &= \text{“la carta scelta dal giocatore ha entrambi i lati bianchi”}, \\ R_1 &= \text{“il lato superiore è rosso”}, \\ B_1 &= \text{“il lato superiore è bianco”} = R_1^c, \\ R_2 &= \text{“il lato inferiore è rosso”}, \\ B_2 &= \text{“il lato inferiore è bianco”} = R_2^c.\end{aligned}$$

Il diagramma ad albero dell'esperimento aleatorio è il seguente (si noti che nella terza ramificazione le probabilità sui rami sono diverse rispetto all'albero dell'Esercizio 5.1; sarebbero state le stesse se nell'Esercizio 5.1 l'estrazione fosse stata *senza reimmissione*):



La probabilità richiesta è $\mathbb{P}(R_2|B_1)$. Si procede dunque come al punto 2 dell'Esercizio 5.1. Si ottiene quindi

$$
\begin{aligned}
\mathbb{P}(R_2|B_1) &= \frac{\mathbb{P}(B_1 \cap R_2)}{\mathbb{P}(B_1)} \\
&= \frac{\mathbb{P}(B_1 \cap R_2 \cap H_0) + \mathbb{P}(B_1 \cap R_2 \cap H_1) + \mathbb{P}(B_1 \cap R_2 \cap H_2)}{\mathbb{P}(B_1)} \\
&= \frac{0 + 1 \cdot \frac{1}{2} \cdot \frac{1}{3} + 0}{\frac{1}{2}} = \frac{1}{3}.
\end{aligned}
$$

$\square$

Esempio 5.3 (Dilemma di Monty-Hall). *Sei a un gioco a premi, e devi scegliere fra tre porte. Dietro a una porta c'è un'automobile, mentre dietro alle altre troverai solo delle capre. Tu scegli, diciamo, la porta n^o 1, e il presentatore, che sa dov'è l'automobile, ne apre un'altra, dietro a cui c'è una capra. A questo punto, ti dà la possibilità di scegliere tra il restare fedele alla porta n^o 1 o il passare all'altra.*

Che cosa ti conviene fare?

(a) *Restare fedele alla porta n^o 1.*

(b) *Passare all'altra porta.*

Come cambia la risposta se il presentatore non sa dov'è l'automobile?

Soluzione. L'automobile potrebbe essere dietro alla porta n^o 1, n^o 2 oppure n^o 3. Ci sono dunque tre eventi possibili:

$$\begin{aligned} H_1 &= \text{“l'automobile si trova dietro alla porta n}^o\text{ 1”}, \\ H_2 &= \text{“l'automobile si trova dietro alla porta n}^o\text{ 2”}, \\ H_3 &= \text{“l'automobile si trova dietro alla porta n}^o\text{ 3”}. \end{aligned}$$

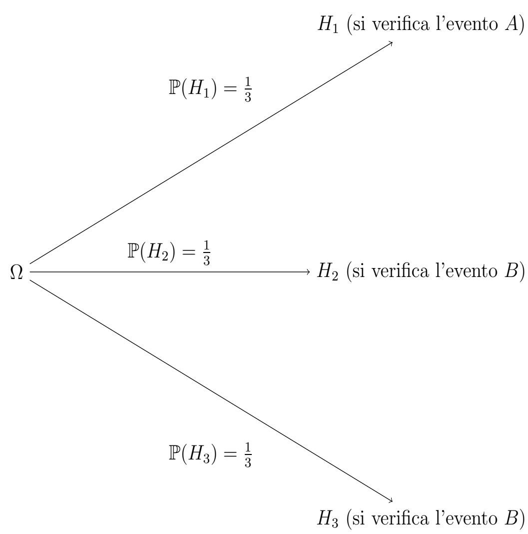
Dato che il testo dell'esercizio non dice nulla a riguardo, non c'è alcuna ragione per assegnare probabilità differenti agli eventi H_1, H_2, H_3, perciò supponiamo che essi siano *equiprobabili*. Quindi

$$\mathbb{P}(H_1) \; = \; \mathbb{P}(H_2) \; = \; \mathbb{P}(H_3) \; = \; \frac{1}{3}.$$

Introduciamo inoltre i due eventi seguenti:

$$\begin{aligned} A &= \text{“vinci restando fedele alla porta n}^o\text{ 1”}, \\ B &= \text{“vinci passando all'altra porta”} \; = \; A^c. \end{aligned}$$

Il diagramma che descrive l'esperimento aleatorio è il seguente:



Quindi

$$A \; = \; H_1, \qquad\qquad B \; = \; H_2 \cup H_3.$$

La strategia migliore è quella che ha probabilità maggiore. Dobbiamo dunque calcolare $\mathbb{P}(A)$ e $\mathbb{P}(B)$. Si ha che

$$\mathbb{P}(A) \; = \; \mathbb{P}(H_1) \; = \; \frac{1}{3}$$

e

$$\mathbb{P}(B) \; = \; \mathbb{P}(H_2) + \mathbb{P}(H_3) \; = \; \frac{2}{3}.$$

In conclusione $\mathbb{P}(A) < \mathbb{P}(B)$, quindi *conviene passare all'altra porta.* $\square$

Esempio 5.4 (Paradosso dei tre prigionieri). *Tre prigionieri A, B, C sono condannati all'impiccagione. Il sovrano decide di graziare uno dei tre scelto a caso, ma il nome del fortunato verrà comunicato soltanto alla vigilia dell'esecuzione. Il prigioniero A si avvicina al secondino, che conosce il nome del graziato, e gli dice: "Per favore, comunicami un nome, tra B e C, che verrà sicuramente impiccato. È noto che almeno uno di loro due sarà impiccato, pertanto non mi fornisci alcuna informazione dicendomelo". Il secondino ci pensa, trova l'argomento sensato e risponde: "B verrà impiccato". A questo punto A esclama: "Evviva! Visto che B verrà impiccato, restiamo in gioco solo io e C, pertanto ho il 50% di probabilità di essere graziato, mentre in precedenza ne avevo solo 1/3."*

Questo argomento è corretto?

Soluzione. Il prigioniero graziato potrebbe essere A, B o C. Ci sono dunque tre eventi possibili:

$$\begin{aligned} H_A &= \text{“il prigioniero graziato è } A\text{”}, \\ H_B &= \text{“il prigioniero graziato è } B\text{”}, \\ H_C &= \text{“il prigioniero graziato è } C\text{”}. \end{aligned}$$

Dato che il sovrano sceglie a caso quale prigioniero graziare, si ha che

$$\mathbb{P}(H_A) \;=\; \mathbb{P}(H_B) \;=\; \mathbb{P}(H_C) \;=\; \frac{1}{3}.$$

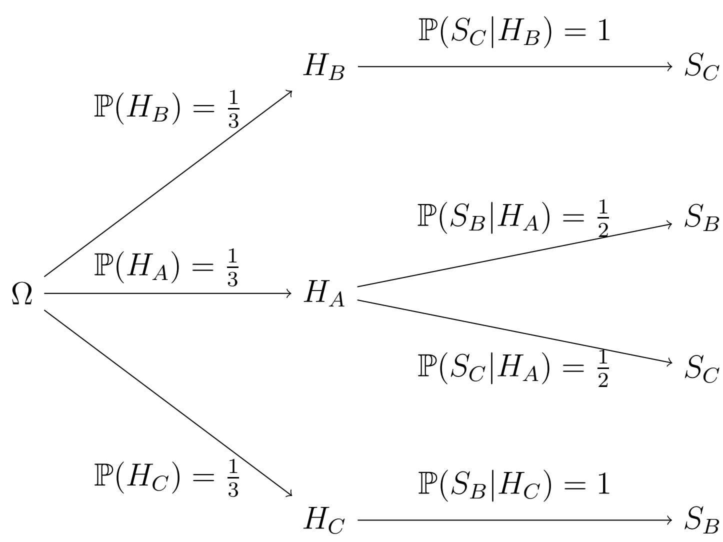
Introduciamo ora gli eventi riguardanti il secondino:

$$\begin{aligned} S_B &= \text{“il secondino comunica che } B \text{ verrà impiccato”}, \\ S_C &= \text{“il secondino comunica che } C \text{ verrà impiccato”} \;=\; S_B^c. \end{aligned}$$

Poiché il secondo trova sensato l'argomento del prigioniero A, è ragionevole assegnare le seguenti probabilità condizionate:

$$\mathbb{P}(S_B|H_A) \;=\; \frac{1}{2}, \qquad \mathbb{P}(S_B|H_B) \;=\; 1, \qquad \mathbb{P}(S_B|H_C) \;=\; 0.$$

Il diagramma ad albero dell'esperimento aleatorio in esame è allora il seguente:



Pertanto, per la formula di Bayes,

$$\mathbb{P}(H_A|S_B) \;=\; \frac{\mathbb{P}(S_B|H_A)\,\mathbb{P}(H_A)}{\mathbb{P}(S_B)} \;=\; \frac{\frac{1}{2}\cdot\frac{1}{3}}{\frac{1}{2}\cdot\frac{1}{3}+\frac{1}{3}} \;=\; \frac{1}{3}.$$

L'argomento di A è *sbagliato*. $\square$