5. Pensare in termini elementari... legge della moltiplicazione

Nei capitoli precedenti abbiamo discusso e risolto vari problemi, interpretando spesso la situazione mediante lo schema classico, e valutando dunque la probabilità come rapporto tra il numero dei casi "favorevoli" e quello dei "casi possibili". Gli strumenti probabilistici introdotti si sono rivelati pertanto efficaci; ma non bastano per interpretare situazioni più articolate. Ad esempio, non sono sufficienti per rispondere a varie questioni motivanti che abbiamo posto a fondamento del nostro Percorso nel capitolo 1, come le domande relative ai test clinici oppure ai misconcetti.
È dunque per arricchire il Percorso che introdurremo in questo capitolo la legge della moltiplicazione e fisseremo l'attenzione sulla probabilità dell'evento complementare. Questi sono gli aspetti sui quali abbiamo scelto di concentrare l'attenzione, evitando di proporre troppi termini o troppi risultati: in particolare presenteremo nella forma di domanda aperta, ossia come questione sulla quale discutere, la formula che viene spesso indicata come teorema della probabilità totale. Piuttosto preferiamo investire del tempo per guidare gli studenti a (ri)scoprire i due risultati prima evidenziati, in contesti opportunamente semplificati, e a saggiarne poi la potenza in vari ambiti, dalla giurisprudenza alla genetica.
L'idea sottesa agli strumenti matematici che ci apprestiamo ad introdurre è di interpretare gli eventi in termini di eventi più elementari, dei quali sia più semplice valutare la probabilità. Proprio allo sviluppo di questa abilità interpretativa daremo rilevanza, perché in essa e nell'abilità di esprimere gli eventi mediante registri diversi, prima che nei risultati, risiede la valenza didattica delle attività che proporremo nel capitolo.

5.1 Unione

Se alla roulette (europea[1]) punto su un numero pari o nero, qual è la probabilità che io vinca?

Coerentemente con l'approccio metodologico che anima il Percorso, in un primo momento è opportuno lasciare spazio agli interventi degli studenti.

Si può rispondere semplicemente contando i casi favorevoli; si ottiene così che la probabilità richiesta è

$$\mathbf{\frac{26}{37}}$$

Ma si può investigare ulteriormente la situazione, proponendo agli studenti una questione del tipo:

"Possiamo ricorrere all'uguaglianza seguente? È vera? Perché?"

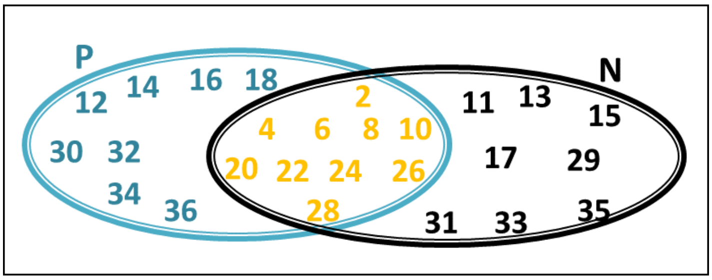
$$p(\text{"pari"} \cup \text{"}\mathbf{nero}\text{"}) = p(\text{"pari"}) + p(\text{"}\mathbf{nero}\text{"}) \qquad (5.1)$$

L'esame che ne segue risulta più coinvolgente e didatticamente significativo se è arricchito dall'effettuazione di prove materiali. E dovrebbe portare a concludere che
l'uguaglianza è falsa poiché vale:

$$\frac{26}{37} \neq \mathbf{\frac{18}{38}} + \frac{18}{37}$$

[1] Ovvero la roulette costituita da 37 caselle, delle quali una è indicata come 0 (zero). Altre tipologie di roulette, come ad esempio quelle americane, hanno 38 caselle: oltre alla casella 0 hanno la casella 00.

Per comprendere le ragioni profonde sottese a tale conclusione, è utile ricorrere a una rappresentazione grafica, quale la seguente, dell'insieme P dei numeri pari della roulette e dell'insieme N costituito dai numeri neri del gioco:



Essa permette di esprimere # (P U N) nella forma[2]:

$$\#\,(\mathbf{P} \cup \mathbf{N}) = \#\,\mathbf{P} + \#\,\mathbf{N} - \#\,(\mathbf{P} \cap \mathbf{N}) \qquad (5.2)$$

Si conclude quindi che l'uguaglianza (5.1) è errata perché il suo secondo membro conta due volte gli elementi di $(\mathbf{P} \cap \mathbf{N})$.

*Quesiti di tale tipo, che non avrebbero una grande importanza di per sé visto che si possono risolvere semplicemente contando con attenzione il numero dei casi favorevoli, possono fornire l'occasione per sviluppare abilità trasversali. Essi, infatti, stimolano l'uso di più forme di rappresentazione e l'abilità di **passare** da un registro all'altro coinvolgendo:*

- *il linguaggio degli **insiemi** (simboli e termini, operazioni),*
- *la schematizzazione grafica mediante **diagrammi** di Venn,*
- *il linguaggio **logico** … i connettivi "**o**", "**e**", "**non**" (significato logico e uso nel linguaggio naturale).*

*Anche in questo contesto è importante **scegliere** oculatamente quali formule e quali termini proporre agli studenti. A tale proposito, si consideri l'assioma della probabilità:*

$$\text{se } A \cap B = \emptyset, \text{ vale } p(A \cup B) = p(A) + p(B) \qquad (5.3)$$

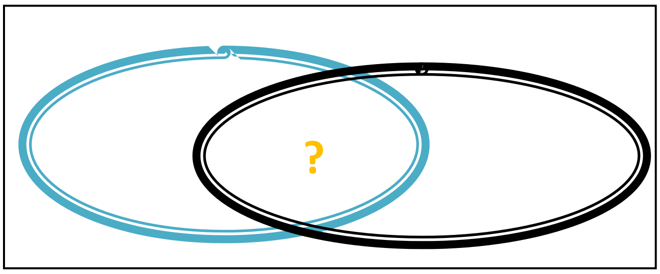
Il docente dovrebbe essere consapevole della sua rilevanza culturale, ma anche del fatto che non è opportuno, almeno in questa fase, presentarlo come una formula di cui disporre immediatamente o, peggio, da imparare a memoria.
Analogamente, per risolvere la maggior parte degli esercizi proposti sui libri di testo, non è necessario ricorrere alla formula (5.2): è sufficiente contare i casi possibili con attenzione. In quest'ottica è utile spostare l'attenzione da mere questioni di calcolo a questioni come la seguente

In una scuola la probabilità che uno studente, scelto a caso, sappia pattinare è del *31%, q*uella che uno studente sappia arrampicare è del *24%*.
Qual è la probabilità che uno studente della scuola sappia pattinare *e* arrampicare?

Anche in questa situazione è opportuno lasciare spazio agli interventi degli studenti e discutere con loro: alcuni forniranno dei valori di probabilità, ma altri potrebbero osservare che è necessario assumere ipotesi aggiuntive e/o dire che non vi sono gli elementi sufficienti per rispondere...

[2] Come indicato nel capitolo 4, dato un insieme E, indichiamo con #E, la sua cardinalità.

In ogni caso, può essere utile far riferimento alla ormai collaudata schematizzazione grafica



*Si comprende così che le informazioni fornite dal problema **non** sono sufficienti per rispondere!*
Vale la pena allora esplorare più a fondo la questione:

Fornisci un esempio di informazione aggiuntiva, mediante la quale si possa determinare la probabilità richiesta.

Osservazione
Un approccio diverso a questi problemi, ma con gli stessi obiettivi, è proposto nell'approfondimento relativo alla sperimentazione in classe.

5.2 Complementare

Lanciamo tre dadi "onesti" che hanno le facce numerate da 1 a 6.
Qual è la probabilità che il punteggio (somma dei tre numeri usciti) sia *almeno "5"?*

*Gli studenti si dovrebbero presto accorgere che vi sono **molti** casi da esaminare e raccogliere così il suggerimento di considerare l'evento complementare ossia l'evento "il punteggio sia almeno 5". Possono così apprezzare il suggerimento del docente di seguire una strategia risolutiva alternativa di esprimere la probabilità richiesta nella forma:*

$$p(somma\ almeno\ 5) = 1 - p(somma\ minore\ di\ 5)$$

Più precisamente, ciò comporta percorrere i passi seguenti:

- considerare l'evento **complementare** (o contrario)
  $$\text{“}il\ punteggio\ \text{è}\ minore\ di\ \mathbf{5}\text{”}\text{, ossia “}il\ punteggio\ \text{è}\ \mathbf{3}\ oppure\ \text{è}\ \mathbf{4}\text{”}$$
- contare il numero dei casi favorevoli a quest'ultimo evento, ossia $1 + 3 = 4$ (infatti il punteggio 3 si può ottenere in un solo modo, mentre il punteggio 4 in tre modi)
- dedurre che il numero dei **casi favorevoli** all'evento "*almeno 5*" è:
  $\mathbf{216} - 4 = \mathbf{212}$, essendo $\mathbf{216} = 6 \cdot 6 \cdot 6$ il numero di punteggi possibili.

Dunque, in definitiva

$$p(\text{“}somma\ almeno\ 5\text{”}) = \frac{\mathbf{212}}{216} = \frac{53}{54}$$

Osservazione

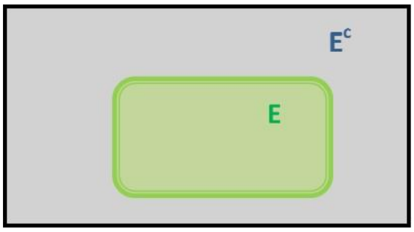
$$p(\text{“}somma\ almeno\ 5\text{”}) = p(\text{"}\geq 5\text{"}) = \frac{212}{216}, \qquad mentre \qquad p(\text{"}< 5\text{"}) = \frac{4}{216}$$

quindi

$$p(\text{"}\geq 5\text{"}) + p(\text{"}< 5\text{"}) = \frac{212}{216} + \frac{4}{216} = \mathbf{1}$$

In generale, tale relazione vale per ogni evento *E, ossia*

$$p(E) = 1 - p(E^c) \qquad (5.4)$$



Infatti

$$p(E) + p(E^c) = 1$$

dato che

$$p(E \cup E^c) = 1$$

Alcuni esercizi

Alcuni esempi significativi sulla probabilità di eventi non elementari (modellizzazione anche mediante *circuiti elettrici*), eventualmente per il lavoro autonomo, si trovano in fondo al capitolo all'appendice A1.

5.3 Legge della moltiplicazione

Al solito, preferiamo motivare l'introduzione della legge della moltiplicazione mediante la discussione di problemi, quali la valutazione della probabilità di errore di un test diagnostico. Comunque per costruire il risultato lavoreremo in un contesto più elementare, quello delle estrazioni da un'urna, che può modellizzare quasi tutte le situazioni in cui interviene la probabilità nel primo biennio della scuola secondaria di secondo grado.

Compleanni
Ti trovi ad una festa a cui partecipano 23 persone.
Scommetteresti che vi sono almeno due tra esse che compiono gli anni nello stesso giorno (anche se sono nate in anni diversi)?

Test clinici
Il test "Elisa" relativo all'HIV ha una sensibilità del 99,9% e una specificità del 99,9%.
Se la malattia ha una prevalenza dello 0,3%, qual è la probabilità che il test dia indicazioni errate su un individuo scelto a caso nella popolazione?

L'esame diretto di tali questioni è delicato. Pertanto puntiamo a sviluppare gli strumenti matematici per risolverli, nel contesto più elementare dell'urna.

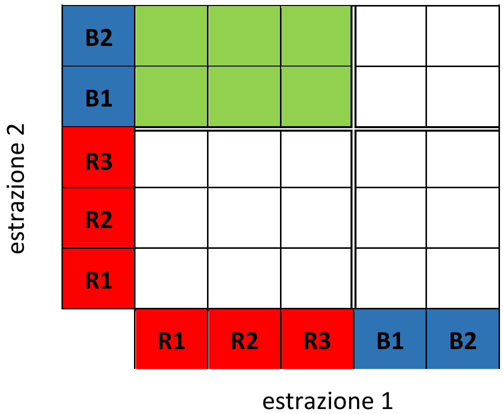
a) Un problema modello: l'urna con reinserimento

In un'urna vi sono 5 palline, diverse solo per il colore:

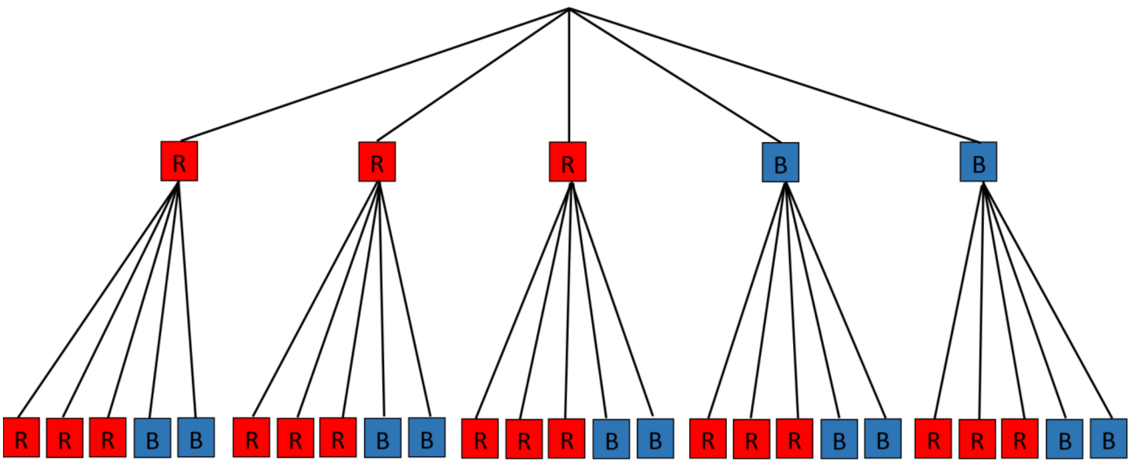
3 sono rosse e **2** blu.

Si estrae in modo casuale una pallina alla volta e la si *reinserisce* nell'urna prima dell'estrazione successiva.
Qual è la probabilità che la prima pallina estratta sia **rossa** e la seconda sia **blu**?

Gli studenti esplorano il problema: effettuano ***prove*** *dell'estrazione, poi magari* ***elencano*** *i casi possibili:*
$\boldsymbol{R}1\boldsymbol{R}1, \boldsymbol{R}1\boldsymbol{R}2, \dots\ \boldsymbol{R}1\boldsymbol{B}1, \dots$
Eventualmente possono schematizzare la situazione mediante una **tabella:**



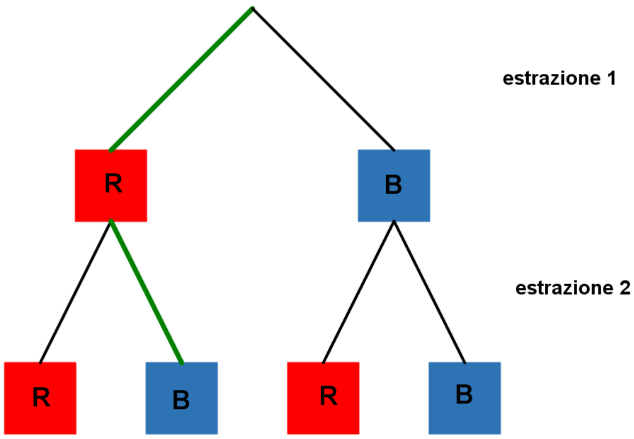
A questo punto, semplicemente contando il numero di "celle favorevoli" si ottiene la probabilità richiesta dal quesito in esame, ossia la probabilità dell'evento che indichiamo[3] con *"**R** e **B**"*: $\frac{\mathbf{6}}{\mathbf{25}}$
Ma, ci chiediamo, c'è un approccio più efficace per determinare il valore di probabilità richiesto? E che sia valido anche nel caso di tre o più estrazioni?
Il primo passo che si può compiere è ricorrere ad una diversa schematizzazione, nella forma di **grafo ad albero:**



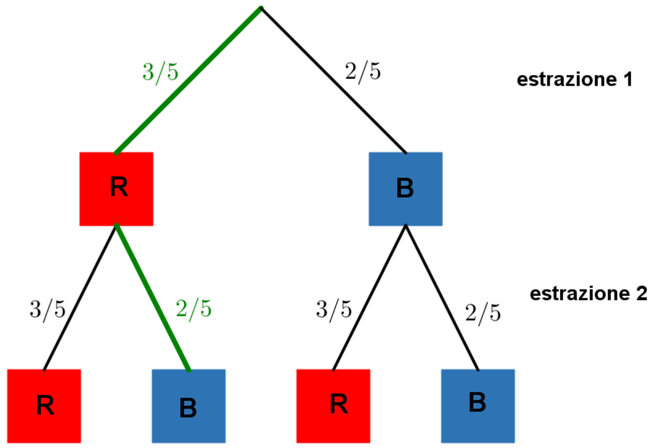
[3] Osserviamo che nel linguaggio naturale con la frase "viene estratta la pallina rossa e la pallina blu" non si precisa l'ordine nel quale sono estratte le due palline, ossia si contempla sia il caso in cui la prima estratta è rossa e la seconda è blu sia il caso in cui la prima estratta è blu e la seconda rossa. Invece il quesito in esame richiede una condizione più forte, ossia impone di tener conto dell'ordine di estrazione: prima la pallina rossa e poi la pallina blu.
Pertanto con la scrittura "R e B" conveniamo di tener conto anche dell'ordine di estrazione, ossia intendiamo "la prima pallina estratta è rossa, la seconda è blu".

Ora, per quanto tale rappresentazione sia espressiva, essa diventa poco leggibile non appena aumenta il numero di palline. Pertanto è più conveniente considerare un grafo "contratto", tipo quello rappresentato nella seguente figura[4].

Il cammino "**favorevole**" è evidenziato in verde.



Accanto a ogni ramo si indica la probabilità dell'evento che esso rappresenta.



Il passo successivo consiste nell'esplorare la situazione lavorando sul grafo: si invitano gli studenti a cercare relazioni tra $p(\boldsymbol{R} \text{ e } \boldsymbol{B}) = \frac{6}{25}$ e le probabilità sul grafo, per arrivare ad accorgersi che vale

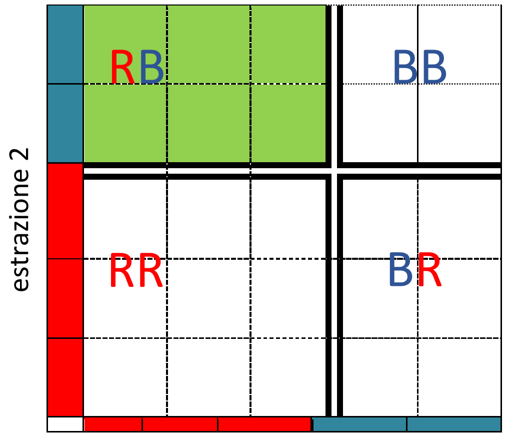
$$p(\boldsymbol{R} \text{ e } \boldsymbol{B}) = \frac{\mathbf{3}}{\mathbf{5}} \cdot \frac{\mathbf{2}}{\mathbf{5}} \qquad \text{ossia } p(\boldsymbol{R} \text{ e } \boldsymbol{B}) = p(\boldsymbol{R}) \cdot p(\boldsymbol{B}) \qquad (5.5)$$

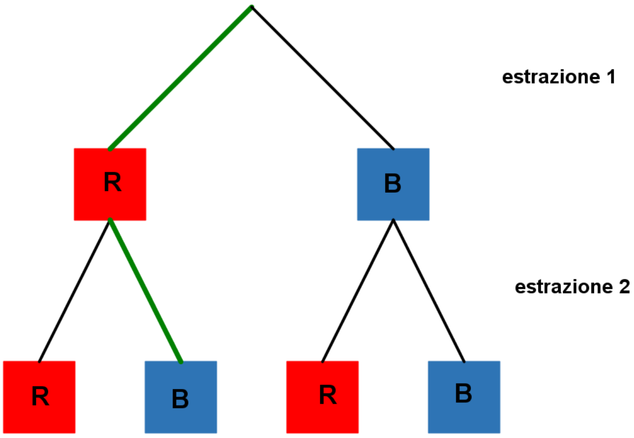
Gli studenti possono così provare la soddisfazione di scoprire che, nella situazione in esame, la probabilità dell'evento R e B è uguale alla probabilità degli eventi componenti R, B. In realtà, tale risultato vale più in generale, ma ciò che interessa in questa fase è che gli studenti sappiano leggere l'uguaglianza $p(\boldsymbol{R} \text{ e } \boldsymbol{B}) = p(\boldsymbol{R}) \cdot p(\boldsymbol{B})$ direttamente sul grafo ad albero.

[4] Didatticamente è opportuno che il passaggio alla rappresentazione contratta avvenga solo nel momento in cui gli studenti ne avvertono realmente l'esigenza. Ossia prima dovrebbero cimentarsi nella realizzazione di grafi ad albero con molti rami, esperirne le difficoltà e dunque riflettere sulla necessità (dello studente, non del docente) di una schematizzazione più compatta Naturalmente ciò non significa che alcuni studenti non propongano di costruire direttamente l'albero contratto; in questo caso è compito del docente premurarsi che essi abbiano compreso a fondo i vantaggi che derivano da tale rappresentazione.

Osservazione

Infine è utile mettere a confronto i due modelli introdotti. Analogamente al grafo, si può pensare di "contrarre" anche la tabella a doppia entrata rappresentando solo gli esiti Rosso, Blu; in questa rappresentazione le celle tratteggiate mostrano anche una misura della probabilità dei 4 casi RR, RB, BR, BB.





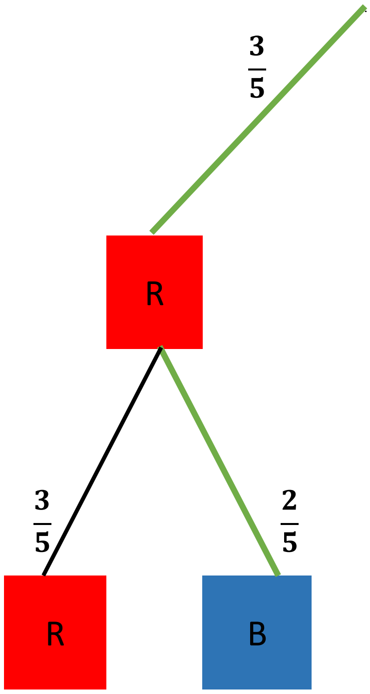
In particolare si può notare che ad ogni **cammino** sull'albero corrisponde una **cella** della tabella contratta; ad esempio al cammino evidenziato in verde corrisponde la cella verde RB.
A sua volta, ogni cammino rappresenta un evento intersezione; ad esempio quello evidenziato in verde corrisponde all'evento $\boldsymbol{R}$ *e* $\boldsymbol{B}$.

Una giustificazione mediante un'analogia

Per giustificare il fatto che vale il prodotto delle probabilità sui rami, si può ricorrere a un'analogia con un condotto per l'acqua. Precisamente:

si può immaginare che il grafo rappresenti un condotto per l'**acqua** e i valori di probabilità rappresentino la portata[5] in un'opportuna unità di misura

- se il tubo verde in alto porta a litri, allora nel tubo verde sotto scorrono i $\frac{2}{5}$ di $\boldsymbol{a}$ litri, ossia $\frac{2}{5}\cdot \boldsymbol{a}$ litri
- se il tubo in alto porta $\frac{3}{5}$ di litro, allora …



Analogamente nel problema in esame si ha che:

il ramo in alto si percorre con probabilità $\frac{3}{5}$, quindi
quello verde in basso si percorre con probabilità globale $\frac{2}{5}\cdot\frac{3}{5}$

Attenzione: è solo un'analogia. Non una dimostrazione!
Inoltre tale modello richiama le percentuali iterate.

[5] Per portata del tubo intendiamo la quantità massima d'acqua che può fluire attraverso una sua sezione nell'unità di tempo. Supponiamo inoltre che ogni tubo venga sfruttato al massimo della sua portata.

Una risoluzione più formale

Per concludere l'esame del problema mediante lo schema del grafo ad albero, osserviamo che la legge della moltiplicazione si può leggere direttamente sull'albero "non contratto", precedentemente rappresentato, anche in modo formale.
Infatti, fissiamo l'esito della prima estrazione, ad esempio la prima rossa a sinistra. La probabilità che esso si realizzi è $\frac{\mathbf{1}}{\mathbf{5}}$ *, quindi la probabilità dell'insieme dei 5 esiti della estrazione successiva deve essere* $\frac{\mathbf{1}}{\mathbf{5}}$. Ma *gli esiti sono equiprobabili, dunque la probabilità di "ciascun cammino" è* $\frac{\frac{\mathbf{1}}{\mathbf{5}}}{\mathbf{5}} = \frac{\mathbf{1}}{\mathbf{5}}\frac{\mathbf{1}}{\mathbf{5}} = \frac{\mathbf{1}}{\mathbf{25}}$ *. Si può concludere quindi che* $p(\boldsymbol{R} \text{ e } \boldsymbol{B}) = 6 \cdot \frac{1}{25}$ *lo stesso risultato ottenuto seguendo l'approccio discusso precedentemente.*

b) Un ulteriore problema modello: l'urna senza reinserimento

In un'urna vi sono 5 palline, diverse solo per il colore:

3 sono rosse e **2** blu.

Si estrae in modo casuale una pallina alla volta e NON la si *reinserisce* nell'urna prima dell'estrazione successiva.
Qual è la probabilità che la prima pallina estratta sia **rossa** e la seconda sia **blu**?

Come nel problema dell'urna con reinserimento, dovremmo creare le condizioni affinché gli studenti esplorino a fondo il problema ed effettuino ***prove*** *dell'esperimento.*
E, ancora una volta, è istruttivo schematizzare il problema mediante una **tabella:**

estrazione 2

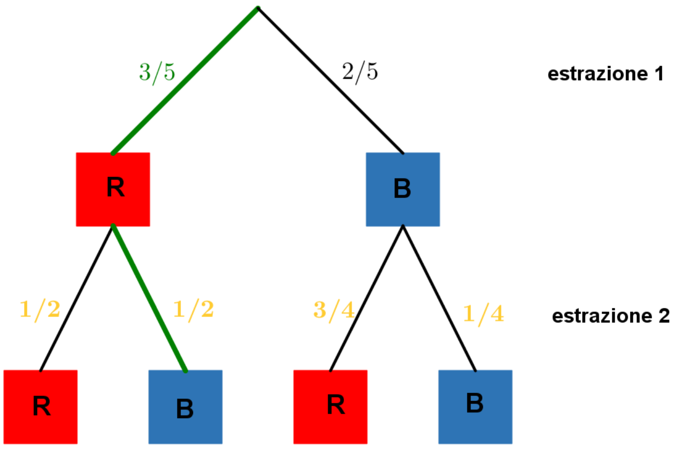
B2					X
B1				X	
R3			X		
R2		X			
R1	X				
	R1	**R2**	**R3**	**B1**	**B2**

estrazione 1

Attenzione però: dato che ora la pallina non viene reinserita nell'urna, alcune celle (non rappresentano più dei casi possibili! Si tratta delle celle segnate con la lettera X in figura.

Dunque, anche in questa situazione, pur di prestare un po' di attenzione, si può dedurre dalla tabella che la probabilità dell'evento *"**R** e **B**"* : $\frac{\mathbf{6}}{\mathbf{25-5}} = \frac{3}{10}$

Cambiando angolazione, possiamo considerare anche la rappresentazione mediante grafo ad albero[6].



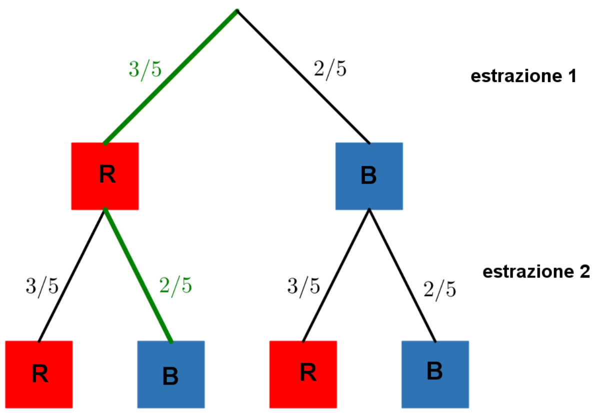
Attenzione: rispetto al problema dell'urna con reimmissione cambiano le probabilità della seconda estrazione. Infatti, senza reinserimento, dopo la prima estrazione, cambia la composizione dell'urna!

Ad esempio dopo aver estratto un pallina rossa rimangono nell'urna quattro palline delle quali solo due rosse[7]. Quindi la probabilità di estrarre una pallina blu (alla seconda estrazione) è $\frac{2}{4} = \frac{1}{2}$.
A questo punto non dovrebbe essere difficile osservare che vale

$$p(\boldsymbol{R} \text{ e } \boldsymbol{B}) = \frac{\mathbf{3}}{\mathbf{5}} \cdot \frac{\mathbf{1}}{\mathbf{2}}$$

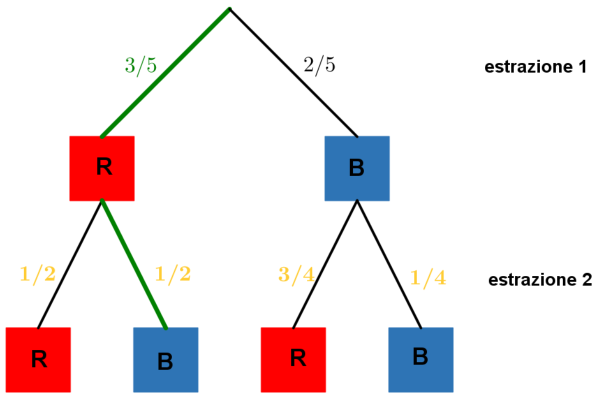
Ma se a tale conclusione giungono gli studenti in un'attività esplorativa, allora tale scoperta è … "per sempre" e contribuisce a rafforzare il senso di autoefficacia.
In definitiva, come accade per l'estrazione con reimmissione, la probabilità richiesta è il **prodotto** delle probabilità "elementari", pur di tener presente che la composizione dell'urna è cambiata dopo la prima estrazione. Questo è anche quanto evidenzia il confronto tra le due situazioni:

con reimmissione



$$p(\boldsymbol{R} \text{ e } \boldsymbol{B}) = \frac{\mathbf{3}}{\mathbf{5}} \cdot \frac{\mathbf{2}}{\mathbf{5}}$$

senza reimmissione



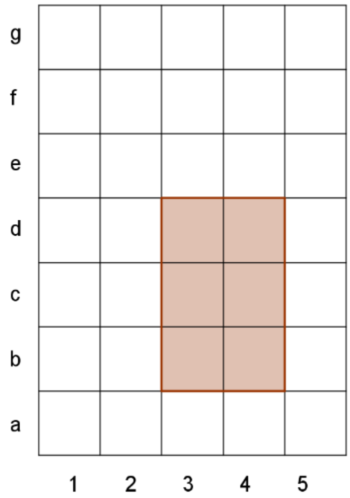
$$p(\boldsymbol{R} \text{ e } \boldsymbol{B}) = \frac{\mathbf{3}}{\mathbf{5}} \cdot \frac{\mathbf{1}}{\mathbf{2}}$$

[6] Per semplicità abbiamo considerato direttamente il grafo contratto; naturalmente può aver senso, a seconda delle esigenze della classe, invitare gli studenti a schematizzare prima l'intero albero.
[7] Tale situazione è descritta in modo espressivo mediante il grafo ad albero "*non contratto*", il quale è costituito dagli stessi 5 rami di quello con reinserimento relativi alla prima estrazione. Ma, in questo caso, per ciascuno di essi, vi sono 4 rami per la seconda estrazione.

c) Un'altra giustificazione della Legge*

Un altro approccio[8] *per giustificare la legge della moltiplicazione può essere quello di ricorrere al contesto della battaglia navale.*

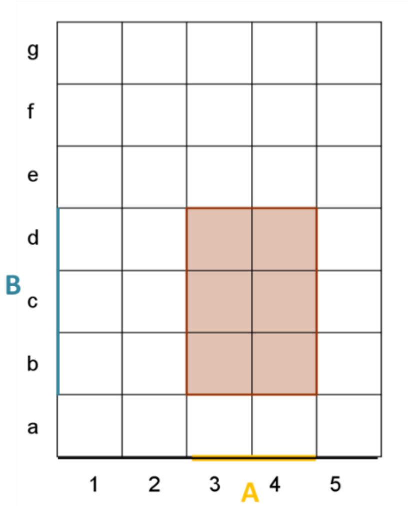
Giochiamo a **battaglia navale**[9]. Qual è la probabilità di colpire la portaerei in figura con un solo colpo?



Contando esplicitamente il numero di casi favorevoli e quello dei casi possibili si ha che la probabilità è

$$\frac{\mathbf{6}}{\mathbf{35}}$$

Un altro modo per giungere al risultato è seguire un approccio per componenti:



- sparare un colpo si può pensare come:
  1) indicare un *numero*
  2) indicare una *lettera*

- "colpire la portaerei" è l'evento ***A e B***, dove A, B sono gli insiemi rappresentati in figura ossia ***A*** è l'insieme dei numeri favorevoli e ***B*** l'insieme delle lettere favorevoli

- numero casi *favorevoli* $= \mathbf{2 \cdot 3}$
  numero casi *possibili* $= \mathbf{5 \cdot 7}$

Pertanto:

$$\mathrm{p}(A\ e\ B) = \frac{\mathbf{2 \cdot 3}}{\mathbf{5 \cdot 7}} = \frac{\mathbf{2}}{\mathbf{5}} \cdot \frac{\mathbf{3}}{\mathbf{7}} = \mathrm{p}(A) \cdot \mathrm{p}(B)$$

[8] Ispirato ad un'attività proposta da Giovanni Prodi, in "Matematica come scoperta", ed. D'Anna.
[9] La questione si può esprimere più formalmente in termini di *prodotto cartesiano.*

d) Legge della moltiplicazione – Facciamo il punto

Mediante attività analoghe a quelle discusse nelle sezioni precedenti, gli studenti hanno modo di prendere confidenza con la legge della moltiplicazione e con i concetti ad essa sottesi, in particolare la dipendenza-indipendenza di eventi. Su queste basi dovrebbero poi essere in grado di apprezzare la precisazione di tali aspetti matematici, almeno nella forma che proponiamo di seguito: si tratta di un primo livello di formalizzazione che, eventualmente, potrà essere ripreso e affinato nelle classi successive al biennio. In sostanza, ancora una volta, vorremmo che la definizione costituisca il punto di arrivo del Percorso, non quello di partenza. Anzi, spingendoci oltre, ci piacerebbe che, prima delle parole, lo studente disponga di efficaci rappresentazioni degli oggetti matematici e, a partire da esse, trovi volta per volta le parole per descriverli in modo compiuto.

Due eventi si dicono **indipendenti** se la conoscenza del fatto che uno di essi si è verificato non modifica la probabilità di verificarsi dell'altro. Altrimenti si dicono **dipendenti.**
Un modello di riferimento per tali concetti è l'urna da noi considerata negli esempi precedenti[10]. Precisamente:

l' **urna** **con** reimmissione è il modello di riferimento per eventi indipendenti

l' **urna** **senza** reimmissione è il modello di riferimento per eventi dipendenti

Legge della moltiplicazione:
dati due eventi A, B, la probabilità dell'evento A e B è uguale al prodotto della probabilità dell'evento A per la probabilità di B valutata assumendo **l'ipotesi che A si sia verificato**.

Tale formulazione vale sia per eventi indipendenti che dipendenti.

5.4 Alcune osservazioni per il docente

Indipendenza tra eventi

a) Prima la definizione o il concetto?

Una definizione equivalente di indipendenza tra due eventi A, B si può esprimere mediante la condizione

$$p(A \cap B) = p(A) \cdot p(B) \tag{5.6}$$

Però in questa fase del Percorso non ci sembra opportuno insistere su questa formulazione in quanto potrebbe venir interpretata semplicemente come una formula da imparare a memoria e da ripetere a parole senza magari preoccuparsi di comprenderne il significato. Piuttosto, come già osservato, è molto più importante che lo studente disponga del ***modello*** *di riferimento dell'urna (con o senza reinserimento) e sviluppi l'abilità trasversale di passare da una forma di rappresentazione all'altra, in una sorta di traduzione nel linguaggio naturale.*

[10] Ci sono casi particolari, come ad esempio l'urna contenente solo palline nere, in cui anche l'estrazione senza reimmissione dà luogo a eventi probabilisticamente indipendenti.

b) È sempre intuitiva?

D'altro canto, se si opta per un approccio meno formale, si deve accettare il fatto di perdere l'univocità e la precisione che viene assicurata dalla formula (5.6). Infatti, non sempre è chiaro a livello intuitivo se la probabilità di un evento sia o meno modificata dal realizzarsi di un altro, come richiesto dalla definizione di eventi indipendenti al paragrafo 5.3 sezione d).
Ad esempio, si consideri l'esperimento del lancio di un dado a 6 facce e i due eventi:

A = "*esce un numero pari*"
B = "*esce il numero 1 o il numero 2*"

Intuitivamente i due eventi A, B vi sembrano indipendenti? Difficile stabilirlo a una prima lettura… alcuni diranno che sono dipendenti...
Per dirimere la questione c'è bisogno di fermarsi a valutare con attenzione le probabilità che intervengono:

$$p(A) = \frac{1}{2}$$

$$p(A\ nell'ipotesi\ che\ si\ sia\ verificato\ B) = p(\text{"2 o 4 o 6" sapendo che è uscito "1 o 2"}) = \cdots = \frac{1}{2}$$

quindi gli eventi A, B sono indipendenti! Ma non è stato immediato concluderlo.

Invece l'applicazione della definizione (5.6) permette di giungere in modo più diretto alla conclusione, anche se in modo meccanico e poco intuitivo:

$$p(A) \cdot p(B) = \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6} = p(A \cap B)$$

c) Influenza tra eventi[11]?

Si deve prestare attenzione a non confondere la relazione di dipendenza tra eventi con quella di causa - effetto. Un esempio significativo a tale proposito è costituito da un'analisi statistica effettuata in Inghilterra, dopo la seconda guerra mondiale. Su un campione di case, si è rilevata la presenza di:

- *un nuovo nato,*
- *un nuovo nido di cicogna.*

Si sono considerati gli eventi:
A = "c'è almeno un nuovo nido di ***cicogna*** *sul tetto di una casa fissata"*
B = "c'è almeno un nuovo ***nato*** *in una casa fissata"*

Dall'elaborazione dei dati è emerso che:

$$p(B\ nell'ipotesi\ che\ si\ sia\ verificato\ A) > p(B)$$

ossia gli eventi A,B sono dipendenti.

In altre parole dove c'è un ***nido di cicogna*** *è maggiore la probabilità di una* ***nascita.***
Questo significa che l'evento *A* ***influenza*** *l'evento B?*

No! *In ultima analisi, la ragione per cui vale la disuguaglianza in esame risiede nel fatto che A, B hanno una causa comune: la fine della guerra, in particolare la fine dei bombardamenti.*

[11] L'esempio seguente è tratto dal testo "La matematica dell'incertezza" di Carla Rossi, ed. Zanichelli.

La notazione $p(A|B)$

In una prima fase è preferibile non utilizzare notazioni specifiche per la probabilità che dipende da altre, tipo quella che solitamente compare sui libri di testo della scuola secondaria: p(A|B) per indicare la probabilità dell'evento A nell'ipotesi che B si sia verificato. La sua introduzione si può tranquillamente posticipare al momento in cui sarà avvertita dagli studenti come una necessità per descrivere la situazione e i procedimenti in modo univoco e conciso. Così sarà percepita come un vantaggio e non come una complicazione inutile imposta dall'insegnante.
In ogni caso, riteniamo più espressiva la notazione $p_B(A)$. Ma di questo ci occuperemo più avanti.

*Piuttosto in questa prima fase è più importante curare aspetti sostanziali, come **l'ambiguità** nel linguaggio. Ad esempio dalla confusione "probabile = possibile" c'è il rischio che segua l'identificazione "**non** probabile = **non** possibile".*

Come applicare la legge della moltiplicazione?

Alcuni libri di testo riportano due diverse formulazioni della legge della moltiplicazione: una per eventi dipendenti, un'altra per eventi dipendenti. Noi però riteniamo didatticamente più efficace un'unica formulazione, come quella precedentemente proposta, nella quale non serve chiedersi a priori se A, B siano dipendenti o indipendenti[12]. Con l'unica accortezza di prestare attenzione a valutare la "nuova" probabilità dell'evento B, nell'ipotesi che A si sia verificato, anche se in certi casi essa potrebbe non cambiare.

La formula

$$p(B) \cdot p(A|B) = p(A \cap B) \qquad (5.7)$$

Su vari libri di testo viene introdotta tale formula ed è considerata unicamente per calcolare $p(A \cap B)$, *ossia come un modo formale per esprimere la legge della moltiplicazione.*
Però il docente dovrebbe tener presente che essa ha un ruolo fondamentale:
- *è la **definizione** di probabilità condizionata*
- *da essa **deriva***
  - *la definizione di dipendenza ed indipendenza di eventi*
  - *la legge della moltiplicazione*
  - *il significato di probabilità condizionata nei tre approcci*

Si tornerà su questa formula quando si affronterà il tema della probabilità condizionata.

5.5 Consolidamento

Alcuni esempi significativi sulla legge della moltiplicazione si trovano in fondo al capitolo nell'appendice A2.

[12] Aggiungiamo che in questa fase sembra prematuro affrontare esercizi in cui si chiede di stabilire se due eventi siano indipendenti o dipendenti. È opportuno che gli studenti prendano prima confidenza con tali concetti in un contesto più generale.

5.6 Torniamo ai problemi iniziali

A questo punto gli studenti dovrebbero disporre degli strumenti matematici per rispondere alle questioni proposte all'inizio del percorso nonché per esaminarle criticamente. La risoluzione è accompagnata da commenti, video, esplorazioni con il software... per favorire una comprensione più profonda delle situazioni esaminate e permettere così ai ragazzi di prendere decisioni consapevoli.

a) Regolarità

Lanciamo 10 volte una moneta "onesta".
Su quale tra le due sequenze di esiti scommettete?

$$T\ T\ T\ T\ T\ T\ T\ T\ T\ T \qquad T\ C\ T\ C\ C\ T\ C\ T\ T\ C$$

Ricorrendo alla legge della moltiplicazione si ottiene che le probabilità dei due eventi sono uguali:

$$p(T\ T\ T\ T\ T\ T\ T\ T\ T\ T) = p(T)\cdot p(T)\cdot \ldots \cdot p(T) = \left(\frac{\mathbf{1}}{\mathbf{2}}\right)^{\mathbf{10}}$$

$$p(T\ C\ T\ C\ C\ T\ C\ T\ T\ C) = p(T)\cdot p(C)\cdot \ldots \cdot p(C) = \left(\frac{\mathbf{1}}{\mathbf{2}}\right)^{\mathbf{10}}$$

Ma, al di là della valutazione numerica, le probabilità dei due eventi sono uguali essenzialmente perché *i lanci sono **indipendenti.***

Più in generale si può osservare che ogni sequenza di 10 lanci della moneta ha probabilità $\left(\frac{1}{2}\right)^{\mathbf{10}}$

b) Compensazione (riformulata)

Lanciamo 10 volte una moneta "onesta".
L'esito dei primi 9 lanci è

$$T\ T\ T\ T\ T\ T\ T\ T\ T$$

È vero che al decimo lancio è più probabile ottenere C?

Al decimo lancio, come al primo, vale $p(T) = p(C) = \frac{1}{2}$ in quanto i lanci sono indipendenti[13] ovvero "la moneta **non ha memoria**".

Osservazione
Si è risposto alla questione iniziale, ma è interessante investigare insieme agli studenti l'origine del misconcetto.

- L'evento "i primi nove lanci hanno tutti esito testa" è **poco probabile**, infatti:

$$p(T\ T\ T\ T\ T\ T\ T\ T\ T) = \left(\frac{1}{2}\right)^9 < \frac{1}{500}$$

  Questa è anche la probabilità di una qualsiasi sequenza fissata costituita da 9 simboli ciascuno scelto tra T, C.

[13]Ossia, più formalmente, gli eventi "esito del lancio i-esimo", "esito del lancio i+1-esimo" sono indipendenti.

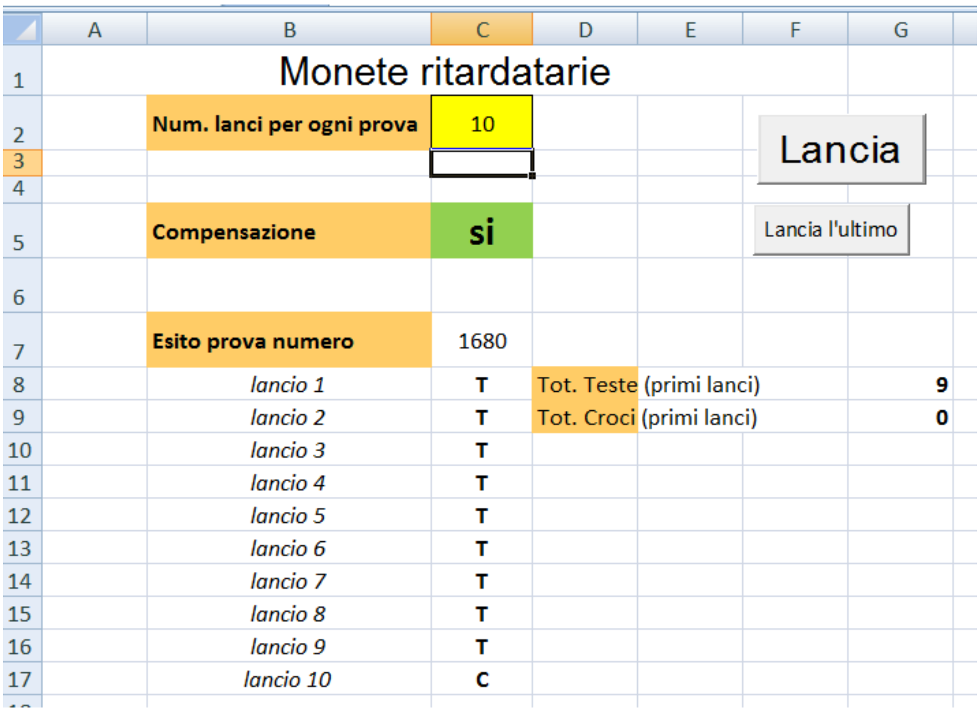
Però, secondo il testo del problema, questo evento ormai **è accaduto**, è un evento certo, è storia. Solo ciascuno dei due esiti del decimo lancio (ossia "esce T", "esce C" è un evento aleatorio, su cui ha senso scommettere.

Quali sono dunque i fraintendimenti sottesi?

- Il primo, come appena visto, risiede nel considerare globalmente i 10 esiti e non il solo decimo esito.
- Il secondo è nell'interpretare in modo distorto la **Legge dei grandi numeri**. Più precisamente, tale legge si applica a sequenze di esiti e non al singolo esito: dunque non si può applicare al problema in esame. In ogni caso proviamo a considerare comunque l'insieme dei 10 lanci: la legge essa afferma, sostanzialmente, che su "molti" lanci è "grande" la probabilità che la frequenza relativa delle croci sia $\frac{1}{2}$; ciò è coerente con il fatto che, sui 10 lanci, è maggiore la probabilità di ottenere una C rispetto a quella di non ottenerne nessuna (vi sono infatti 10 sequenze che contengono una sola C, mentre solo una sequenza è costituita interamente da T); ma la legge non precisa a quale lancio è più probabile esca la C, mentre nel problema in esame interessa l'uscita di C ad un lancio ben preciso: il decimo.

*Tutto questo in teoria, ma nella **pratica** cosa succede? Ecco una simulazione mediante il foglio elettronico per investigare la questione.*

L'idea dell'attività e una traccia di lavoro si trovano in fondo al capitolo nell'appendice (A3 Attività). Entrambe fanno riferimento al file Excel MoneteRitardatarieConsecutive.xlsm, predisposto per studenti.



Gli studenti possono ora rispondere in modo autonomo al quesito iniziale "Marta e i bambini".

c) Numeri ritardatari

Quanto appena discusso a proposito dei lanci di moneta si può applicare anche alle estrazioni del gioco del Lotto. Pertanto, in questo contesto potrebbe aver senso far lavorare gli studenti in modo più autonomo e piuttosto soffermarsi a discutere collettivamente le conclusioni individuali.

Il “53” non è uscito per 182 estrazioni consecutive sulla ruota di Venezia.
Qual è la probabilità che esca su tale ruota alla 183-esima estrazione?

La probabilità che esca il “53” ad **una** data estrazione su tale ruota è

$$p = \frac{5}{90} = \frac{1}{18}$$

La probabilità di uscita alla **183-esima** estrazione è ancora p: le estrazioni sono indipendenti (per il meccanismo fisico di estrazione).

Osservazioni

- **Ritardo del 53**

Qual è la probabilità che il “53” non esca per 182 estrazioni consecutive?

La probabilità che, ad una (qualunque) estrazione fissata, non esca il “53” è 1-p. Pertanto, per la legge della moltiplicazione la probabilità che il “53” non esca per tutte le 182 estrazioni consecutive è:

$$(1-p)^{182} \approx 0{,}000030$$

L’evento è dunque “poco probabile”, ma ormai è storia.
Comunque, per la cronaca, … il “53” è uscito su Venezia il 9 febbraio 2005, alla 183 – esima estrazione…

- **Numeri spia**

Ecco un'attività che gli studenti dovrebbero trovare stimolante...
Sul sito della Lottomatica (https://www.lottomaticaitalia.it/it/prodotti/lotto/statistiche/numero-spia) vengono elencati dei numeri speciali e sono presentati con queste parole:
“La tradizione vuole che l’estrazione di certi numeri “preannunci” l’uscita di altri. La tabella indica, per ogni numero estratto (o “**numero spia**”), 5 numeri che avrebbero la maggior probabilità di uscire.”

Venezia	
27	42 58
	65
	86
	90

Ad esempio:

L’8 gennaio 2018 è uscito il 27 sulla ruota di Venezia. Allora è vero che ciò aumenta la probabilità dell’uscita del 42 all’estrazione successiva?

d) Test clinici

Prima di affrontare i problemi motivanti proposti all'inizio del percorso relativi ai test clinici, è importante che gli studenti prendano confidenza con la situazione affrontando dei problemi più semplici, come i seguenti.

- Test di gravidanza:

si veda in fondo al capitolo il testo e la risoluzione all'appendice A4

- Un problema relativo alla diagnosi di una malattia:

Una popolazione di 10.000 individui è stata sottoposta a un test per diagnosticare una certa malattia.
Sono risultate positive al test 1.726 persone e si assume che il test sia risultato positivo per il 99,0% dei malati.
Inoltre si assume che il 2,0% della popolazione avesse la malattia.
Qual è la probabilità che il test abbia fornito **indicazioni errate** su un individuo scelto a caso in tale popolazione?

Il video di questo problema e della risoluzione realizzato dal Laboratorio DicoMat Lab dell'Università degli studi di Trento si trova all'indirizzo http://youtu.be/N_sdkLtECps.

Ecco allora uno dei problemi posti ad inizio percorso relativamente a questo contesto.

Qual è la probabilità che il test "Elisa" fornisca indicazioni errate?

Il quesito, così formulato, non ammette risposta univoca: era stato volutamente presentato in forma contratta per fare emergere le convinzioni di senso comune. Pertanto, prima di procedere è opportuno precisare alcune ipotesi sul test.

Una popolazione è sottoposta al test "Elisa" per la diagnosi dell'HIV.
La probabilità che il test sia positivo sull'individuo che ha il virus è del 99,9% (sensibilità del test).
Assumiamo che la probabilità di essere negativo al test per l'individuo "sano" sia del 99,8% (specificità).
Inoltre assumiamo che lo 0,3% della popolazione abbia la malattia (prevalenza).
Qual è la probabilità che il test fornisca **indicazioni errate** su un individuo scelto a caso in tale popolazione?

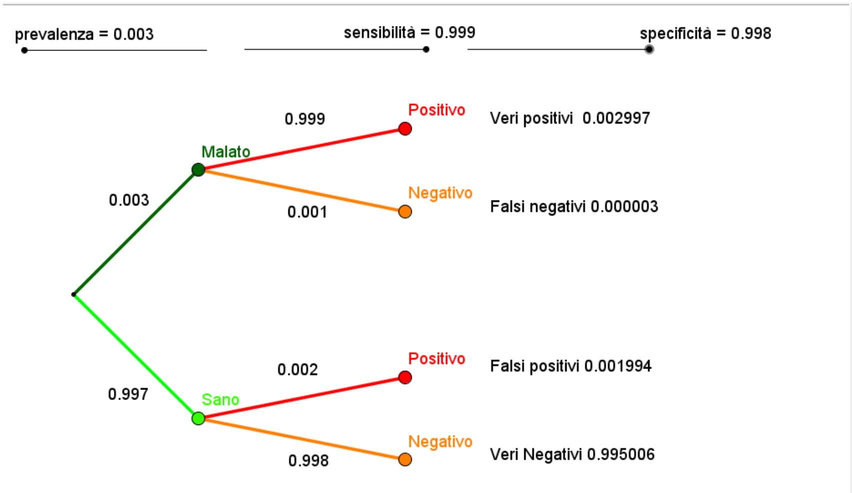
Questo problema si risolve analogamente ai due problemi precedenti. Così, modellizzando la situazione mediante un grafo ad albero, si arriva a concludere che:

$$p(\text{“}esito\ errato\text{”}) = 0{,}997 \cdot 0{,}002 + 0{,}003 \cdot 0{,}001 \approx \mathbf{0{,}002}$$

Più in generale, ha senso chiedersi come varia la probabilità richiesta al variare dei valori in ipotesi. Immergiamoci dunque in un'attività esplorativa, con il supporto del file Geogebra Test_Malattia.ggb[14]. È importante però che la risposta ottenuta empiricamente sia accompagnata da una giustificazione algebrica.

[14] Il materiale è stato ideato dalla collega Francesca Arrigoni.

Ad esempio, si può iniziare indicando con a la prevalenza della malattia ed esprimere la probabilità richiesta, cioè dell'evento "esito errato", in funzione di a; si ottiene p $=0{,}002 - a \cdot 0{,}001$; dunque si conclude che al crescere della prevalenza a, la probabilità p di ottenere un esito errato diminuisce e volendo si può precisare che diminuisce linearmente.



	A	B	C	D
1		Positivi	Negativi	
2	Malati	0.002997	0.000003	0.003
3	Sani	0.001994	0.995006	0.997
4		0.004991	0.995009	1
5				
6	Errati	0.001997		
7				
8				
9				
10				
11				
12				
13				
14				
15				
16				
17				

5.7 Contesti significativi

Gli strumenti matematici introdotti nei paragrafi precedenti, in particolare la legge della moltiplicazione, permettono di esaminare criticamente le situazioni interessanti che proponiamo di seguito. D'altra parte, tali attività consentono di chiarire ulteriormente il significato e la portata dei nuovi oggetti matematici. Queste letture possono anche essere assegnate per l'approfondimento individuale ad alcuni studenti che le potranno poi discutere con l'intera classe.

a) Il daltonismo

Noto il patrimonio genetico dei genitori, gli eventi "avere un figlio daltonico" e "avere un figlio maschio" sono indipendenti?

Nell'appendice A5 alla fine del capitolo si trova una lettura di approfondimento della questione.

*Al di là degli aspetti contenutistici, l'attività è molto interessante perché permette di sviluppare alcune **abilità**, quali:*

- ***interpretare*** *un testo scientifico-matematico*
- ***modellizzare*** *con schemi con frecce, diagramma di Punnet, grafo ad albero…*
- ***effettuare collegamenti*** *con le altre discipline, come raccomandato nelle Indicazioni nazionali*
- ***giustificare*** *e* ***argomentare.***

b) Un caso giudiziario diventato un classico

Consideriamo la seguente sintesi di un caso giudiziario divenuto, ormai, un classico per lo studio della probabilità.

- 18 giugno 1964. Los Angeles. Juanita Brooks viene derubata.
- I testimoni individuano sei caratteristiche dei due responsabili del furto:
  - *uomo di colore con la barba* $\frac{1}{10}$
  - *uomo con i baffi* $\frac{1}{4}$
  - *donna bianca con capelli biondi* $\frac{1}{3}$
  - *donna con la coda di cavallo* $\frac{1}{10}$
  - *coppia mista in un'automobile* $\frac{1}{1.000}$
  - *automobile gialla* $\frac{1}{10}$
- Viene arrestata la coppia Malcom e Janet Collins che presentale le precedenti sei caratteristiche.
- L'accusa stima la probabilità che una persona a Los Angeles possieda una di tali caratteristiche. L'insieme di tali valori è riportato nella tabella precedente.
- Qual è la probabilità p che una coppia qualunque possieda le 6 caratteristiche?
  La stima che fornisce il consulente della difesa è

$$p = \frac{1}{10} \cdot \frac{1}{4} \cdot \frac{1}{3} \cdot \frac{1}{10} \cdot \frac{1}{1.000} \cdot \frac{1}{10} = \frac{1}{12.000.000}$$

- 1964. Essenzialmente sulla base di tale stime, la giuria dichiara colpevole la coppia arrestata.
- 1968. La corte suprema dello Stato della California annulla la sentenza, in seguito a un'accorta riflessione sugli errori commessi nella valutazione di probabilità fornita dall'accusa.

Quali errori sono stati commessi nel primo processo? Più di uno, ma quello che riteniamo didatticamente più significativo è legato all'uso distorto che si è fatto della legge della moltiplicazione: infatti essa è stata applicata nella forma p(A e B) = p(A) · p(B), ma tale uguaglianza vale **solo se gli eventi** A, B sono indipendenti. Invece le sei caratteristiche considerate nel processo ad esempio A = "uomo di colore con la barba"[15] …) **non sono indipendenti!**

Per comprendere meglio dove risiede l'errore consideriamo l'esempio seguente:

In un istituto scolastico 1 studente su 30 pratica lo scialpinismo, 1 su 10 l'arrampicata.
La probabilità che un suo studente scelto a caso pratichi entrambi gli sport è $\frac{1}{30} \cdot \frac{1}{10}$?

[15] Più precisamente A ="esce un uomo di colore con la barba, in un'estrazione casuale tra tutti gli abitanti maschili di Los Angeles"

- Proviamo a rispondere applicando la legge della moltiplicazione: la probabilità che uno studente (scelto a caso) pratichi lo scialpinismo e l'arrampicata è uguale al prodotto della probabilità che uno studente pratichi lo scialpinismo, cioè $\frac{1}{30}$, e della probabilità che uno studente **che pratica lo scialpinismo**, pratichi l'arrampicata.
- Dunque, per valutare la probabilità dell'evento richiesto, occorre conoscere la percentuale di arrampicatori **tra gli scialpinisti** dell'Istituto .
- Però tale dato non è fornito nel testo! Infatti viene indicato solamente che $\frac{1}{10}$ degli studenti **dell'istituto** pratica l'arrampicata, ma non è detto che tra gli scialpinisti, gli arrampicatori siano ancora $\frac{1}{10}$; anzi ragionevolmente ce ne possiamo aspettare di più. In altre parole praticare lo scialpinismo ed arrampicare **non** sono eventi **indipendenti!** Così non possiamo concludere che la probabilità richiesta nel quesito sia $\frac{1}{30} \cdot \frac{1}{10}$. Potremmo determinare la probabilità richiesta solo se sapessimo che tra gli scialpinisti, gli arrampicatori sono ad esempio $\frac{1}{4}$. In tal caso la legge della moltiplicazione permette di affermare che la probabilità richiesta è

$$\frac{1}{30} \cdot \mathbf{\frac{1}{4}} = \mathbf{\frac{1}{120}} \neq \frac{1}{300}$$

Quindi serve attenzione nell'applicare la legge della moltiplicazione!

5.8 Un punto d'arrivo: il problema dei compleanni[16]

Ti trovi ad una festa a cui partecipano 23 persone.
Qual è la probabilità che almeno due tra esse compiano gli anni in uno stesso giorno (anche se sono nate in anni diversi)?

La situazione dei compleanni può essere esaminata anche da una prospettiva matematicamente più interessante, ma che può rivelarsi troppo complessa per la classe:

Ad una festa scommetti che almeno due partecipanti compiano gli anni in uno stesso giorno. Affinché la tua probabilità di vittoria sia maggiore del 50%, i partecipanti devono essere più della metà del numero di giorni dell'anno, cioè più di 182?

Per quanto osservato, preferiamo però affrontare solo in un secondo momento questa diversa formulazione del problema.
Prima di affrontare la questione dal punto di vista teorico è importante che gli studenti la esplorino per prendere confidenza con la situazione, per provare a stimare il risultato, e per far emergere conoscenze pregresse o legate al senso comune. A tale scopo possono risultare utili le attività che descriviamo sinteticamente di seguito.

a) Attività esplorative

1) Verificare se vi sono compleanni nello stesso giorno all'interno di una stessa classe dell'Istituto.
2) Simulare i «compleanni», facendo scrivere a ogni studente un naturale "a caso" tra 1 e 365; poi confrontare i numeri scritti.

[16] Il problema è stato formulato da Richard von Mises.

3) Esaminare i compleanni dei titolari e dell'arbitro (22+1) di alcune partite di calcio della squadra del cuore.

Osservazione

Ogni squadra ammessa alla fase finale del mondiale di calcio 2014 ha dovuto convocare 23 giocatori.
Per **15** squadre sulle **32 partecipanti a tale fase vi erano** almeno due giocatori che compiono gli anni nello stesso giorno[17].

A conclusione di queste attività, ci sia aspetta che gli studenti formulino delle congetture sul risultato del problema dei compleanni.

b) Risoluzione del problema dei compleanni

Conviene esaminare prima il caso, più semplice, in cui alla festa partecipano solo 3 persone. Se non emerge dalla discussione, il docente può suggerire di calcolare la probabilità dell'evento complementare, più semplice da esprimere in termini di eventi elementari.

c) Un problema semplificato

Consideriamo prima la situazione in cui alla festa partecipano **solo 3 persone**.
Vogliamo cioè determinare la probabilità dell'evento

E = "Almeno due delle tre persone date compiono gli anni lo stesso giorno"

- Tale evento si può realizzare in più modi: quando le persone sono nate tutte nello stesso giorno o quando esattamente due di esse sono nate nello stesso giorno.
Pertanto il calcolo *diretto* della probabilità di *E* deve essere articolato in più casi e può rivelarsi lungo[18].
Proviamo allora a considerare l'evento **complementare** di E

E^c = *"**Le 3 persone compiono gli anni in giorni diversi**"*

Una volta determinata la probabilità di E^c, resta individuata anche quella di *E:* infatti la somma delle due probabilità deve essere 1.

- Vediamo innanzitutto di esprimere l'evento E^c in termini di eventi "elementari", dei quali si possa calcolare più agevolmente la probabilità.
Identificando, per semplicità di linguaggio, ciascuna persona con un numero[19] da 1 a 3, decidiamo di considerare gli eventi:

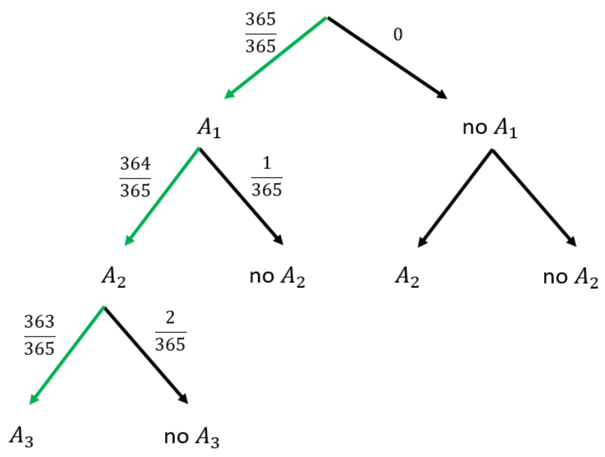
A_1 = *"La prima persona compie gli anni in un **giorno qualsiasi**"*
A_2 = *"La seconda compie gli anni in un giorno diverso dalla prima"*
A_3 = *"La terza compie gli anni in un giorno **diverso dalle prime due**"*

L'evento E^c si può così esprimere in termini di tali eventi:

$$E^c = A_1 \cap A_2 \cap A_3$$

[17] Dati desunti da Wikipedia, consultata in data 01/06/2014.
[18] A più forte ragione se alla festa partecipano più di 3 persone.
[19] In alternativa possiamo identificarle con il nome proprio. Ad esempio Luca, Paolo, Giovanni.

- Possiamo rappresentare la situazione mediante un diagramma ad albero, in modo analogo a quanto visto nei problemi precedenti.



Dunque per la legge della moltiplicazione

$$p(E^c) = \frac{365}{365}\cdot\frac{364}{365}\cdot\frac{363}{365} \approx 0{,}9918$$

- In definitiva il valore di probabilità richiesto è

$$p(E) = 1 - p(E^c) \approx 1\text{-}0{,}9918 = \mathbf{0{,}0082}$$

d) Il problema

Il problema iniziale si può affrontare in modo analogo.
Siamo interessati a determinare la probabilità dell'evento

E = "Almeno due delle 23 persone date compiono gli anni lo stesso giorno"

- Consideriamo l'evento **complementare** di E

  E^c = "Le 23 persone compiono gli anni in giorni diversi"

- Esprimiamo l'evento E^c in termini degli eventi "elementari seguenti, identificando, per semplicità di linguaggio, ciascuna persona con un numero da 1 a 23

  *A_1 = "La prima persona compie gli anni in un **giorno qualsiasi**"*
  A_2 = "La seconda compie gli anni in un giorno diverso dalla prima"
  *A_3 = "La terza compie gli anni in un giorno **diverso dalle precedenti**[20] nella numerazione"*
  …
  *A_{23} = "La 23-esima compie gli anni in un giorno **diverso dalle precedenti** nella numerazione"*

  L'evento E^c si può così esprimere in termini di tali eventi

  $$\mathbf{E^c = A_1 \cap A_2 \cap A_3 \cap_{\dots} \cap A_{23}}$$

- Dunque per la legge della moltiplicazione

$$p(E^c) = \frac{365}{365}\cdot\frac{364}{365}\cdot\frac{363}{365}\cdot \ldots \cdot \frac{343}{365} \approx 0{,}4927$$

Per effettuare il calcolo conviene utilizzare il foglio elettronico.

[20] Cioè dalla prima e dalla seconda.

- In definitiva la probabilità in esame vale

$$p(E) = 1 - p(E^c) \approx 1\text{-}0{,}4927 = \mathbf{0{,}5073}$$

Il risultato ottenuto è… sorprendente:
ad una festa alla quale partecipano 23 persone, è più probabile che almeno due dei presenti compiano gli anni lo stesso giorno, piuttosto che tutti abbiano compleanni diversi.

e) Alcune precisazioni

- **Ipotesi**
  Per semplicità supponiamo che tutti gli anni siano di 365 giorni, che le date di nascita siano indipendenti e che siano equidistribuite su un periodo di 365 giorni[21].

- **Non solo linguaggio naturale**
  Il problema in esame ci sembra adatto per motivare l'uso del linguaggio simbolico, di termini specifici della matematica e di alcune regole di ragionamento. Infatti, in questa situazione didattica, il ricorso a tale approccio è motivato dall'**esigenza** di descrivere e giustificare il procedimento con **precisione, concisione** ed espressività.
  Attenzione però a tenere conto del livello della classe: se gli studenti non sono ancora in grado di affrontare con sicurezza quesiti analoghi, allora il ricorso alla formalizzazione può venir percepito come una inutile complicazione imposta dall'esterno.

- **Quale formalizzazione per gli eventi?**
  Nella risoluzione abbiamo scelto di caratterizzare gli eventi in termini di insiemi. In particolare abbiamo fatto ricorso all'operazione di intersezione e al complementare.
  Una scelta altrettanto valida poteva essere quella di ricorrere alle **proposizioni logiche,** facendo invece riferimento alla congiunzione e all'evento contrario.

f) Alcuni numeri…

Ad una festa scommetti che almeno due partecipanti compiano gli anni in uno stesso giorno. Affinché la tua probabilità di vittoria sia maggiore del 50%, i partecipanti devono essere più della metà del numero di giorni dell'anno, cioè più di 182?

Si può rispondere per tentativi, fissando volta per volta il numero di partecipanti e calcolando, mediante il foglio elettronico, la probabilità che almeno due tra loro compiano gli anni nello stesso giorno.
Se però gli studenti hanno già affrontato il problema posto ad inizio del paragrafo, relativo cioè alla festa che prevede 23 partecipanti, la questione perde di significato.
In questo caso può essere interessante determinare (ancora per tentativi) il numero minimo di partecipanti affinché la probabilità di compiere gli anni nello stesso giorno sia maggiore di un valore dato: ad esempio del 90% oppure del 99%. Anche in questi casi la risposta non è per nulla scontata…

[21] L'ipotesi di equidistribuzione delle nascite nell'arco dell'anno è un'ipotesi semplificativa. Ad esempio nel 2012 in Italia, secondo i dati Istat, il numero di nati al giorno varia tra gli 895 del 25 dicembre e i 1949 del 4 ottobre.

Più in generale ecco alcuni valori della probabilità p che almeno due partecipanti compiano gli anni lo stesso giorno, in funzione del numero n di persone presenti alla festa:

n. persone	p (≥ 2 compleanni = giorno)
10	*0,12*
20	*0,41*
23	***0,51***
30	*0,71*
40	*0,90*
50	*0,97*
56	***0,99***

g) Per comprendere

Perché la probabilità del problema iniziale è "**grande**"?

Ti trovi ad una festa a cui partecipano 23 persone. Qual è la probabilità che almeno una tra esse compia gli anni **nel tuo** stesso giorno (oltre a te)?

La probabilità richiesta è:

$$p = 1 - \left(\frac{364}{365}\right)^{22} \approx 0{,}0586$$

Qual è il motivo per cui le due probabilità relative al problema dei compleanni ("in **uno** stesso giorno") e a quello ora proposto ("nel **tuo** stesso giorno") sono diverse?

L'idea è che:

- nel problema iniziale i casi favorevoli non sono 23; intervengono le **coppie di persone**: $23 \cdot \frac{22}{2}$
- nel problema ora proposto le coppie sono 22

Un approfondimento interessante relativo alle coincidenze è quello proposto da G. Anichini all'indirizzo https://passcienzeunitn2014.files.wordpress.com/2014/04/coincidenze.pdf.

APPENDICE

A1 Eventi non elementari

1. Problema amici - lavoro[22]

Tizio e Caio sono amici: Tizio è libero solo il lunedì, mercoledì, sabato e domenica. Caio è libero solo il martedì, mercoledì, venerdì e domenica.

- **Quando possono incontrarsi?**
- **Come rappresenteresti l'insieme dei giorni possibili?**

Indichiamo con U l'insieme dei giorni della settimana:

$$\mathrm{U} = \{\text{lunedì}, \text{martedì}, \text{mercoledì}, \text{giovedì}, \text{venerdì}, \text{sabato}, \text{domenica}\};$$

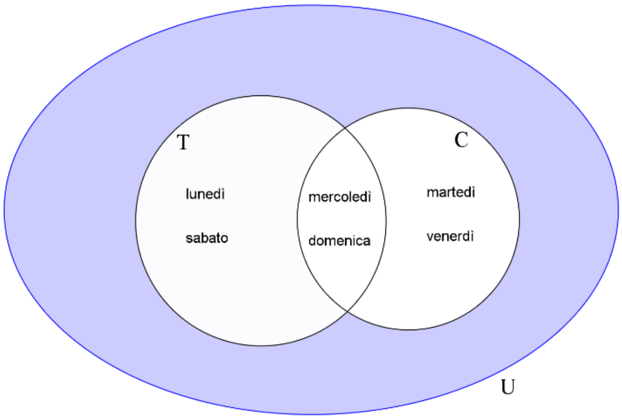
con C il sottoinsieme dei giorni della settimana in cui è libero Caio

$$\mathrm{C} = \{\text{lunedì}, \text{mercoledì}, \text{sabato}, \text{domenica}\};$$

con T il sottoinsieme dei giorni della settimana in cui è libero Tizio

$$\mathrm{T} = \{\text{martedì}, \text{mercoledì}, \text{venerdì}, \text{domenica}\}.$$

Possiamo così schematizzare la situazione mediante i diagrammi di Venn, utilizzando le notazioni ora introdotte.



Tizio e Caio possono trovarsi nei giorni della settimana in cui sono *entrambi* liberi. Tali giorni sono rappresentati nello schema grafico dal sottoinsieme comune all'insieme C e all'insieme T.
E tale sottoinsieme si può scrivere nella forma

$$C \cap T = \{\text{mercoledì}, \text{domenica}\}$$

<u>Osservazione</u>
Al di là delle richieste del quesito, è interessante **rappresentare graficamente** e descrivere mediante il linguaggio degli insiemi:

- l'insieme dei giorni in cui Tizio è libero, ma Caio è occupato
- l'insieme dei giorni in cui Tizio e Caio sono entrambi occupati

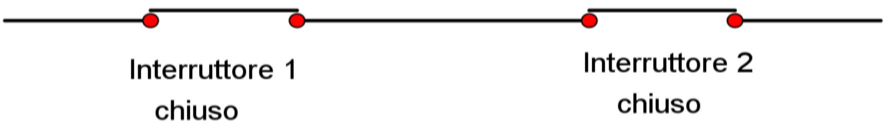
Altrettanto significativa è la questione, per alcuni aspetti inversa, di descrivere l'insieme

$$(T \cap C) \cup T^c$$

nel **linguaggio naturale** e di **rappresentarlo** graficamente.

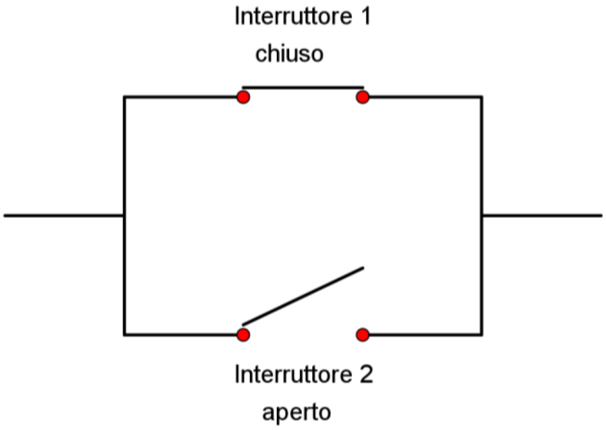
[22] Problema tratto da Giovanni Prodi, "Matematica come scoperta", ed D'Anna, vol. 1, seconda edizione, pag. 27 n. 1.

2. **Problema degli interruttori**[23]
   a) si **collegano due interruttori così (cioè "in serie").**



   In quali casi passa la corrente? Fare il grafo.

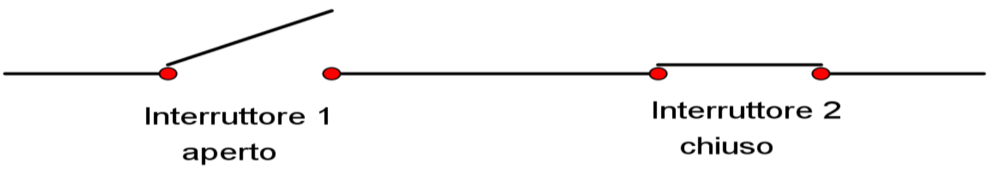
   b) **si collegano due interruttori così (cioè "in parallelo").**



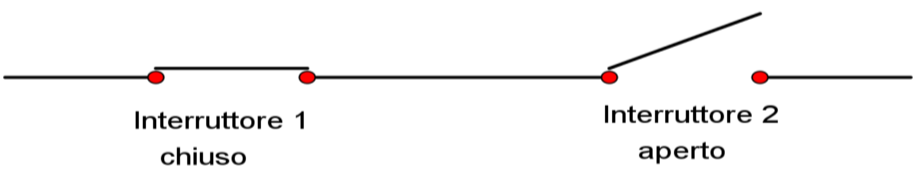
   In quali casi passa la corrente? Fare il grafo. Supponendo che vi sia probabilità 1/2 per ciascun interruttore di essere chiuso (o aperto senza alcuna dipendenza fra i due interruttori) quale è la probabilità che passi corrente?

Nel **collegamento in serie** gli interruttori sono disposti uno di seguito all'altro, con un solo estremo in comune.
Esaminiamo i casi che si possono presentare, cominciando da quello in cui entrambi gli interruttori sono chiusi. Indichiamo tale situazione con la notazione C_1C_2.
In questo caso, come suggerisce la figura, la corrente può passare.

Se invece il primo interruttore è aperto (indichiamo tale situazione con la notazione A_1C_2) la corrente non passa



Simmetricamente, anche nel caso C_1A_2 in cui è aperto solo il secondo interruttore, non può passare la corrente.



Dovrebbe essere chiaro che la corrente non passa nemmeno nel caso A_1A_2 in cui sono aperti entrambi gli interruttori.

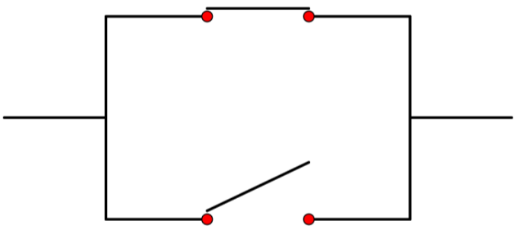
[23] Problema tratto da Giovanni Prodi, "Matematica come scoperta", ed D'Anna, vol. 1, seconda edizione, pag. 37 n. 2.

In definitiva nel collegamento in **serie** la corrente può passare solo se *entrambi gli interruttori sono chiusi,* ossia se

l'interruttore 1 è chiuso e l'interruttore 2 è chiuso

Invece il **collegamento in parallelo** è strutturalmente diverso, in quanto i due interruttori hanno entrambi gli estremi in comune.

Si comprende che la corrente passa in tre situazioni: nel caso C_1C_2 in cui i due interruttori sono chiusi, ma anche nei due casi in cui è chiuso solo un interruttore, cioè C_1A_2 o A_1C_2. Quest'ultima situazione è rappresentata nella figura seguente



La corrente non passa nel caso A_1A_2 in cui sono aperti tutti due gli interruttori.

Così nel collegamento **in parallelo** la corrente passa se *almeno* uno degli interruttori è chiuso, ossia se

l'interruttore 1 è chiuso o l'interruttore 2 è chiuso

Osservazione

L'esercizio proposto è particolarmente istruttivo perché permette di cogliere a fondo il significato dei connettivi *e* ed *o* del linguaggio naturale. Infatti è espressiva **l'analogia fra i connettivi e ed o** e i collegamenti, rispettivamente in serie e in parallelo. Notiamo che il connettivo *o* usato in senso disgiuntivo corrisponde all'operatore logico *XOR*.
Possiamo ora calcolare la probabilità che passi corrente nel **circuito in serie.**
Le possibili configurazioni degli interruttori sono *4*:

$$A_1A_2,\ A_1C_1,\ C_1A_1,\ C_1C_2$$

ed esse sono equiprobabili.
Per quanto osservato nel punto a) *l'unico caso* in cui passa corrente è C_1C_2.
Pertanto la probabilità richiesta è

$$p(passa\ corrente) = \frac{\mathbf{1}}{\mathbf{4}}$$

Cosa accade invece nel **collegamento in parallelo**?
Le possibili disposizioni degli interruttori sono ancora *4*, come nel collegamento in serie.
Però si ha passaggio di corrente quando *almeno* uno dei due interruttori è chiuso. E ciò avviene in 3 casi:

$$A_1C_1,\ C_1A_1,\ C_1C_2$$

Pertanto la probabilità richiesta è

$$p(passa\ corrente) = \frac{\mathbf{3}}{\mathbf{4}}$$

Osservazione

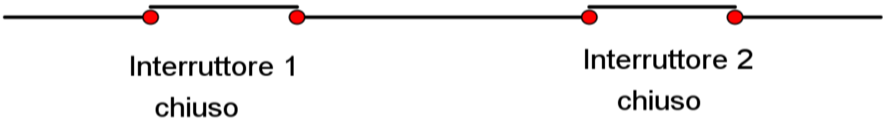
In alternativa, si può considerare la probabilità dell'**evento complementare** "non passa corrente", cioè A_1A_2
Seguendo questo approccio si ottiene:

$$p(passa\ corrente) = 1 - p(A_1A_2) = 1 - \frac{1}{4} = \frac{3}{4}$$

A2 Legge della moltiplicazione

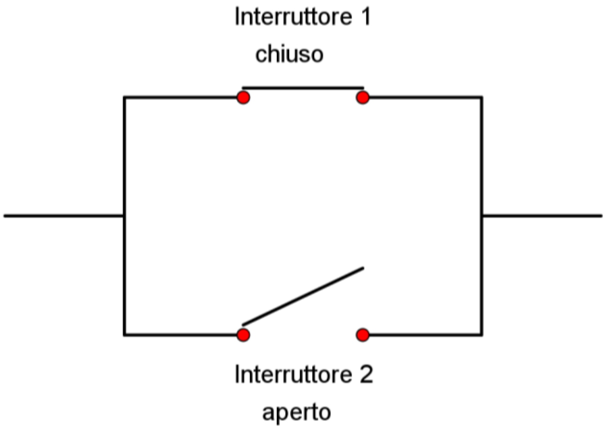
1. Problema degli interruttori[24]

a) si collegano due interruttori così (cioè "in serie").



In quali casi passa la corrente? Fare il grafo.

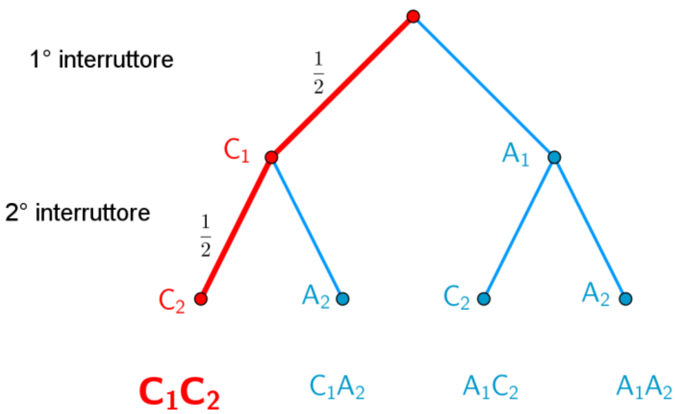
b) si collegano due interruttori così (cioè "in parallelo").



In quali casi passa la corrente? Fare il grafo. Supponendo che vi sia probabilità 1/2 per ciascun interruttore di essere chiuso (o aperto senza alcuna dipendenza fra i due interruttori) quale è la probabilità che passi corrente?

Abbiamo già mostrato come si può risolvere il problema nella Sezione A1. Ora esaminiamo come si possa affrontare la questione anche mediante opportuni grafi.

Ogni interruttore è chiuso o aperto *a caso*. Questo comporta che la probabilità che esso sia chiuso è ½ e quella che sia aperto è ancora ½.
La situazione relativa al **collegamento in serie** si può rappresentare efficacemente con un grafo ad albero, come il seguente



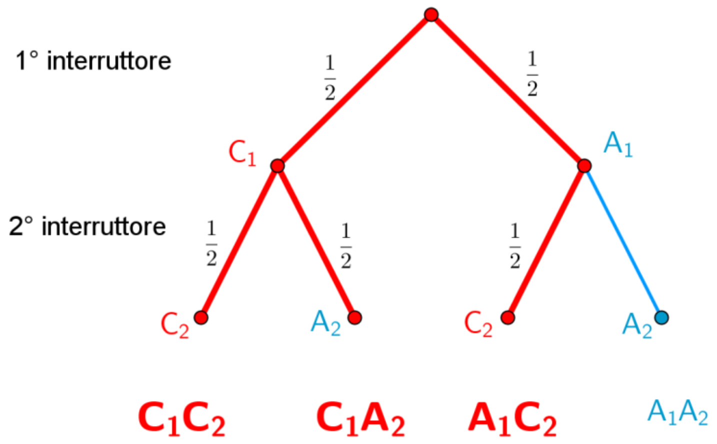
Sul grafo è evidenziato in rosso l'*unico* cammino per il quale si ha passaggio di corrente.
Dato che i 4 cammini sono equiprobabili[25], la probabilità richiesta è $\frac{1}{4}$.

[24] Problema tratto da Giovanni Prodi, "Matematica come scoperta", ed D'Anna, vol. 1, seconda edizione, pag. 37 n. 2.

Possiamo anche seguire un approccio di validità più generale, riportando accanto a ciascuno dei rami del cammino "favorevole" i valori di probabilità dei corrispondenti eventi costituenti C_1 , C_2.
Disponiamo così di tutti gli elementi per *leggere* direttamente *dal grafo* la probabilità dell'evento *C_1 e C_2:* essa è il prodotto delle probabilità relative a ciascun ramo del cammino in considerazione; ossia ancora

$$p(passa\ corrente) = \frac{1}{2} \cdot \frac{1}{2} = \mathbf{\frac{1}{4}}$$

Invece il **collegamento in parallelo** si può così rappresentare



I tre cammini sul grafo per i quali si ha passaggio di corrente sono quelli evidenziati in rosso.
Pertanto la probabilità richiesta sarà $\mathbf{\frac{3}{4}}$.

Analogamente a quanto visto per il collegamento in serie, la probabilità che passi corrente si può anche calcolare come somma delle probabilità dei 3 cammini "favorevoli"

$$p(passa\ corrente) = p(C_1C_2) + p(C_1A_2) + p(A_1C_2) = \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} = \mathbf{\frac{3}{4}}$$

<u>Osservazioni</u>

- Per calcolare la probabilità nel collegamento in parallelo si può, in alternativa, considerare la probabilità dell'**evento complementare** "non passa corrente", ossia l'evento A_1A_2.
  Si ha così:
  $$p(passa\ corrente) = 1 - p(A_1A_2) = 1 - \frac{1}{4} = \frac{3}{4}$$

- Abbiamo denotato con C_1C_2 l'evento: "sono chiusi gli interruttori 1 e 2".
  Potevamo invece indicare, in modo più preciso, con C_1 l'evento "*è chiuso l'interruttore 1*", con C_2 l'evento "*è chiuso l'interruttore 2*" e **denotare** l'evento in esame con C_1 **e** C_2.
  O, ancora più precisamente, considerare gli **insiemi** C_1, C_2 che corrispondono rispettivamente agli eventi C_1 , C_2. E denotare l'evento in esame con $C_1 \cap C_2$.
  La scelta tra le due notazioni dipende dalla situazione didattica e dalla sensibilità del docente. Ma, qualsiasi sia la notazione utilizzata, lo studente deve essere consapevole del fatto che essa si riferisce all'evento: "è chiuso l'interruttore 1 **ed** è chiuso l'interruttore 2" e disponga del significato del connettivo "e", anche operativamente.
  Naturalmente analoghe considerazioni valgono a proposito delle altre notazioni: A_1C_2 …

[25] Attenzione: l'ipotesi di equiprobabilità non si può estendere a qualsiasi situazione. Ad esempio, vedremo che non si può applicare nel problema delle lampadine fulminate che proponiamo di seguito.

- Abbiamo **letto** direttamente **sul grafo ad albero** il procedimento di calcolo delle probabilità richieste, ad esempio di $p(C_1C_2)$. Lo abbiamo cioè utilizzato, non solo per rappresentare la situazione ma anche come **modello** di calcolo.
  Però, anche in tale contesto, lo studente deve essere consapevole che in sostanza non si è fatto altro che applicare la ben nota legge della moltiplicazione
$$p(C_1C_2) = p(C_1) \cdot p(C_2)$$

2. Problema dei cacciatori[26]

Due cacciatori bravi (che non sbagliano mai il bersaglio!) vanno a caccia assieme. Passano due fagiani e ciascuno spara ad un fagiano, a caso. Quale è la probabilità che un fagiano si salvi? (Lo sparo di ciascun cacciatore è un esperimento con due risultati...)

Ovviamente un fagiano potrà salvarsi solo se l'altro fagiano sarà colpito con due colpi...

Anche in questo caso si può schematizzare la situazione con una tabella a doppia entrata.

Cacciatore 1

Cacciatore 2

	Fagiano 1	**Fagiano 2**
Fagiano 1		
Fagiano 2		

Cerchiamo di comprendere il significato di tale rappresentazione. Consideriamo, ad esempio, la cella evidenziata in grigio, che si trova sulla prima riga e seconda colonna: essa rappresenta l'evento "il cacciatore 1 ha sparato al fagiano 2 e il cacciatore 2 ha sparato al fagiano 1".
A questo punto, servendoci di tale schematizzazione, possiamo classificare i possibili esiti:

Cacciatore 1

Cacciatore 2

	Fagiano 1	**Fagiano 2**
Fagiano 1	*Colpito solo fagiano 1*	*Colpiti entrambi*
Fagiano 2	*Colpiti entrambi*	*Colpito solo fagiano 2*

Pertanto la probabilità *p* che uno qualsiasi dei due fagiani si salvi è
$$p = \frac{2}{4} = \mathbf{\frac{1}{2}}$$

<u>**Osservazione**</u>
Senza ricorrere alla tabella, si può pensare di esprimere l'evento
A: "*I due cacciatori colpiscono lo stesso fagiano*",
in termini degli eventi
C: "*Il cacciatore 1 colpisce un fagiano (uno qualsiasi tra i due)*"
D: "*Il cacciatore 2 colpisce lo stesso fagiano che colpisce il cacciatore 1*".

[26] Problema tratto da Giovanni Prodi, "Matematica come scoperta", ed D'Anna, vol. 1, seconda edizione, pag. 37 n. 5.

Precisamente, utilizzando queste notazioni, si può dire che

$$A = C \; e \; D.$$

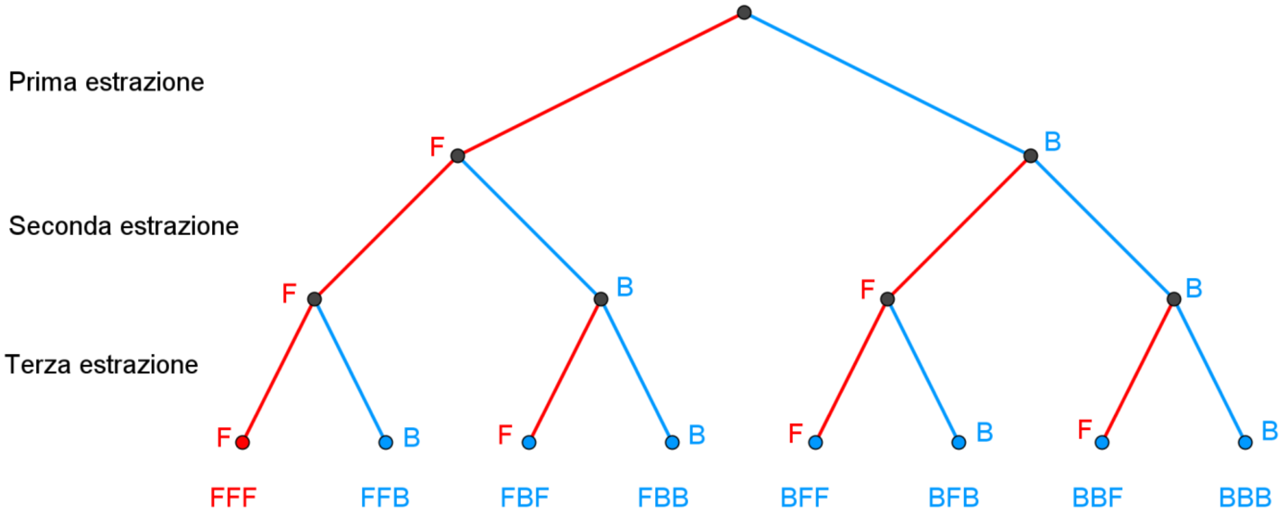
E si conclude ancora che

$$p(A) = p(C \; e \; D) = p(C) \cdot p(D) = 1 \cdot \frac{1}{2} = \frac{1}{2}$$

3. Problema delle lampadine[27]

Una scatola contiene 3 lampadine elettriche fulminate e sette buone. Se ne prendono 3 a caso. Quale è la probabilità che una lampadina almeno sia buona? Disegnare il grafo e calcolare le probabilità dei vari risultati.

Iniziamo schematizzando la situazione mediante un diagramma ad albero a 3 "livelli", ciascuno dei quali rappresenta un'estrazione. Denotiamo con *F* l'evento "*la lampadina estratta è fulminata*" e con *B* l'evento "*la lampadina estratta è buona*".

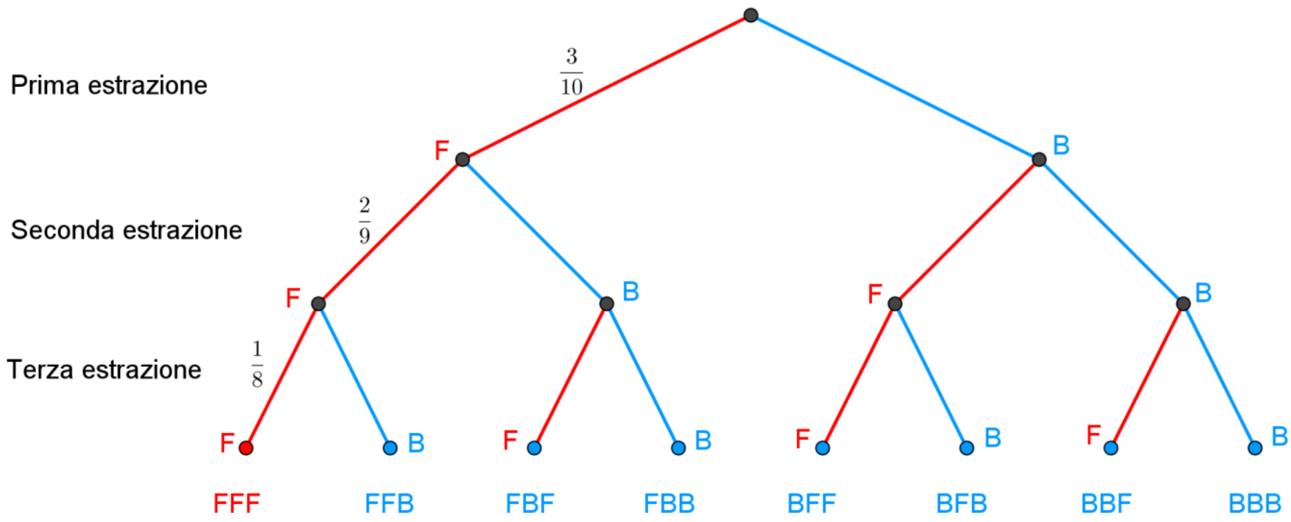


Come chiariremo nelle osservazioni, per risolvere il problema conviene considerare l'evento *complementare* dell'evento "*viene estratta almeno una lampadina buona*". Si tratta dell'evento: "*tutte le lampadine estratte sono fulminate*", che possiamo anche denotare schematicamente con *FFF,* analogamente a quanto visto nell'esercizio dei circuiti. Sul grafo è rappresentato dal cammino più a sinistra.

Per calcolare la probabilità di tale evento, riportiamo accanto a ciascuno dei rami che formano il cammino *FFF*, le probabilità delle relative estrazioni. Allo scopo dobbiamo tener presente che **dopo ogni estrazione la composizione** della scatola **risulta modificata**, dato che le lampadine *non* vengono reinserite al suo interno[28]. Di conseguenza anche il numero dei "casi favorevoli", come del resto quello dei "casi possibili", varia in seguito ad ogni estrazione.

[27] Problema tratto da Giovanni Prodi, "Matematica come scoperta", ed D'Anna, vol. 1, seconda edizione, pag. 44 n. 3. Calcoleremo la probabilità dell'evento "Almeno una lampadina è buona". Le probabilità degli altri esiti si possono ricavare in modo analogo.

[28] In altre parole, gli esiti delle varie estrazioni sono eventi dipendenti.



Così analogamente agli esercizi precedenti[29], possiamo *leggere dal grafo* che la probabilità dell'evento *FFF* è

$$p(FFF) = \frac{3}{10} \cdot \frac{2}{9} \cdot \frac{1}{8} = \frac{1}{120}$$

e quindi la probabilità richiesta è

$$p(almeno\ una\ buona) = 1 - p(FFF) = 1 - \frac{1}{120} = \frac{119}{120} = \mathbf{0,992}$$

Osservazioni

- Per rispondere al quesito si può anche optare per un *approccio diretto*. Cioè si può pensare di calcolare direttamente la probabilità dell'evento indicato nel testo, cioè l'evento: "*viene estratta almeno una lampadina buona*". Ma, esso si realizza in ben 7 differenti modalità:
  FFB, FBF, FBB, BFF, BFB, BBF, BBB
  Per questo il procedimento si rivela più complesso di quello che abbiamo precedentemente proposto e che prevede il ricorso all'evento complementare.

- Le 8 possibili tipologie di estrazione (le 7 ora elencate oltre al caso *FFF*) **non sono equiprobabili**.
  Senza calcolare esplicitamente le probabilità dei vari casi, ci si può rendere conto di ciò riflettendo sul fatto che estrarre una lampadina fulminata è meno probabile che estrarne una buona, visto che quelle difettose sono di meno. E quindi, ad esempio, l'evento *FFB* avrà probabilità minore dell'evento BBB.

- Per il **significato delle notazioni** del tipo *FFF* e per i risultati sottesi all'uso del grafo ad **albero come modello di calcolo**, rimandiamo alle osservazioni scritte a commento di altri esercizi (in particolare, quello relativo ai *circuiti elettrici*).

- Come osservato relativamente ad altri esercizi, l'uso dei grafi ad albero è didatticamente significativo: può aiutare a comprendere più a fondo la situazione, ma *non è necessario* per rispondere al quesito.
  Ad esempio, lo studente può iniziare a investigare sul problema producendo un **elenco** dei vari casi possibili (ma, ricordiamo, non equiprobabili!): *FFF, FFB*, … e da qui rendersi conto che è più conveniente passare all'evento contrario *FFF*.

[29] A tale proposito si veda, in particolare, l'esercizio dei circuiti elettrici, collegamento in serie.

4. Problema del quadrigetto[30]

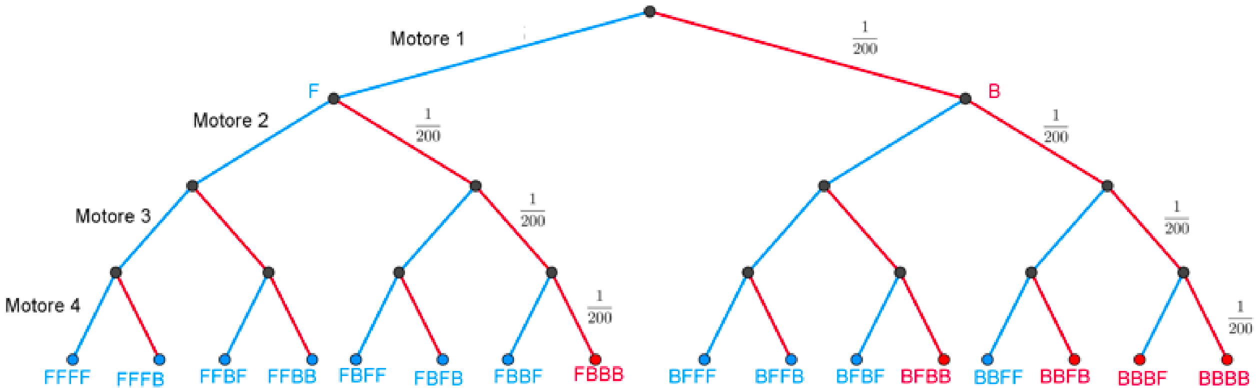
Un quadrigetto compie la traversata dell'Atlantico. La probabilità che un motore si blocchi si può valutare 1/200; perché l'aereo arrivi a destinazione basta che due motori funzionino regolarmente. Quale è la probabilità che l'aereo non arrivi a destinazione?
(Indicare l'espressione esatta della probabilità e poi darne una valutazione approssimata con un po' di buon senso).

Si assuma che il funzionamento di ogni motore sia indipendente da quello degli altri.

Esaminiamo cosa significa l'ipotesi "*perché l'aereo arrivi a destinazione basta che due motori funzionino*". Vuol dire che per *non* arrivare a destinazione devono bloccarsi **almeno** 3 motori. Ossia devono bloccarsi 3 dei quattro motori oppure tutti e 4 i motori. In particolare, allora, la probabilità richiesta **non** è

$P(BBB) = \left(\frac{1}{200}\right)^3 = 0{,}005^3.$

Pertanto è necessario esaminare più alternative. E questo si può fare in modo espressivo con l'ausilio di un grafo ad albero[31], sul quale denotiamo con *F* l'evento "*il motore è funzionante*" e con *B* l'evento "il *motore è bloccato*".



I cammini sul grafo che ci interessano sono di due tipi e sono scritti in rosso alla base dell'albero. Uno è costituito dai cammini che rappresentano la situazione in cui esattamente *3 motori* non funzionano:

FBBB, BFBB, BBFB, BBBF[32]

Invece l'altro tipo di cammino corrisponde alla situazione in cui tutti *4 i motori* non funzionano, cioè

BBBB

Non ci resta allora che calcolare le probabilità di ciascuno di questi 5 eventi.
Iniziamo da *BBBB*. Procedendo in modo analogo ai quesiti precedenti, otteniamo subito che

$$p(BBBB) = \left(\frac{1}{200}\right)^4$$

Consideriamo poi un cammino relativo al blocco di 3 motori, ad esempio *FBBB*.

---

[30] Problema tratto da Giovanni Prodi, "Matematica come scoperta", ed. D'Anna, vol. 1, seconda edizione, pag. 54 n. 1.
[31] Il grafo ad albero è uno strumento molto espressivo, ma non è necessario ricorrere ad esso. Lo studente "esperto" può, ad esempio, produrre un **elenco** dei casi che si possono presentare oppure visualizzare anche solo mentalmente una parte di tale grafo.
[32] Per il significato di tale notazione rimandiamo agli esercizi precedenti, in particolare a quello relativo ai circuiti.

Dato che

$$p(F) = 1 - \frac{1}{200}$$

vale

$$p(FBBB) = \left(1 - \frac{1}{200}\right) \cdot \frac{1}{200^3} = \frac{199}{200^4}$$

In modo analogo potremmo calcolare esplicitamente le probabilità dei restanti 3 eventi in cui si bloccano esattamente 3 motori. Ma non è necessario: infatti ciascuno di tali valori è uguale a $p(FBBB)$, visto che anch'esso deve essere il prodotto del fattore $p(F)$ e di 3 fattori $p(B)$, pur considerati in ordine diverso.
Concludiamo allora che la probabilità che l'aereo *non* arrivi a destinazione è

$$P(non\ arriva) = 4 \cdot p(FBBB) + p(BBBB) = 4 \cdot \frac{199}{200^4} + \frac{1}{200^4} = \frac{797}{200^4} =$$

$$= 4{,}98125 \cdot 10^{-7} \approx \mathbf{5 \cdot 10^{-7}}$$

Osservazioni

- Si può affrontare il problema in esame anche ricorrendo all'*evento complementare*, "*l'aereo arriva a destinazione*". Tale approccio comporta però una maggior complessità nel calcolo.
  Infatti, l'aereo arriva a destinazione se funzionano 2 oppure 3 oppure 4 motori e i modi in cui ciò si può verificare sono più numerosi rispetto a quelli considerati nella risoluzione diretta, proposta precedentemente.[33] In ogni caso, si ha:
  $P(arriva\ ) = 6 \cdot P(FFBB) + 4 \cdot P(FFFB) + P(FFFF) = 6 \cdot 2{,}48^{-5} + 4 \cdot 4{,}93 \cdot 10^{-3} + 9{,}8 \cdot 10^{-1}$
  $= 9{,}999995 \cdot 10^{-1}$
  La probabilità che l'aereo arrivi è *quasi* 1.

- Per il **significato delle notazioni** del tipo *FFF* e per i risultati sottesi all'uso del grafo ad **albero come modello di calcolo**, rimandiamo alle osservazioni scritte a commento di alcuni altri esercizi (in particolare, quello relativo ai circuiti elettrici).

[33] In particolare, la sola situazione con due motori bloccati e due funzionanti si può verificare in 6 modi: FBFB, FBBF, FFBB, BFFB,BFBF e BBFF.

A3 Attività. Compensazione nel lancio di una moneta

L'idea

Stabiliamo il numero totale di lanci per ogni prova, ad esempio 10 ed effettuiamo più prove, ciascuna costituita da 10 lanci.
Siamo interessati alle prove i cui **primi 9 lanci** hanno tutti esito ***"Testa"*** (*). Se al **decimo lancio** di una prova di tale tipo si ottiene "*Croce*" diciamo che nella prova c'è stata **compensazione**.

L'implementazione

Per velocizzare l'esecuzione delle prove ci serviamo del file Excel MoneteRitardatarieConsecutive.xlsm sul quale abbiamo implementato una procedura VBA.

- Fissiamo il numero di lanci nella cella C2.
- Con il bottone ***Lancia*** iniziamo le prove: il software simula i lanci e ferma il procedimento quando **i primi 9 lanci** di una prova sono nelle condizioni (*).
  Il numero di prove effettuato prima di ottenere tale prova è indicato nella cella C7.
- A questo punto, con il bottone ***Lancia l'ultimo**,* effettuiamo **l'ultimo lancio** della prova in esame[34].
  Nella cella C5 si **segnala se nella prova c'è stata compensazione.**

Possiamo così iniziare una nuova serie di prove, utilizzando il bottone *Lancia*.

L'uso didattico

Il docente illustra l'attività ed indica come possono utilizzare il file *MoneteRitardatarieConsecutive.xlsm.*
Gli studenti possono lavorare *direttamente* sul file predisposto dall'insegnante.
Eseguono varie prove e **registrano** il numero di volte in cui vi è stata o non vi è stata compensazione. Dopo aver *provato* a rispondere alle questioni che seguono, discutono le conclusioni con i compagni e con il docente.

Traccia di lavoro

Utilizza il file *MoneteRitardatarieConsecutive.xlsm*. Effettua più prove costituite da 10 lanci ciascuna e registra il numero di volte in cui si ha "compensazione".[35]

- Ritieni che nelle **tue** prove globalmente ci sia "**compensazione"**? Giustifica.
- Se eseguissi un **numero maggiore di prove,** pensi che cambierebbero le tue conclusioni?
- Cosa permettono di affermare sull'esperimento i **risultati teorici** che conosci**?**

[34] Quella caratterizzata dalla condizione (*).
[35] Nel senso illustrato nella sezione *L'idea* di questa attività.

A4 Test gravidanza

Test clinici[36]

Alcuni test clinici, come per esempio quello di gravidanza, danno una risposta che può essere positiva o negativa: nell'esempio, test positivo significa che la donna è incinta, test negativo significa che la donna non è incinta.

Qualche volta il test dà risultati sbagliati; riportiamo una statistica di un laboratorio di analisi, relativamente a 3000 test di gravidanza effettuati:

- **fra i 1.000 test risultati negativi, 80 erano sbagliati (cioè la donna era incinta, mentre il test stabiliva che non lo era);**
- **fra i 2.000 test risultati positivi solo 5 erano sbagliati (cioè la donna non era incinta, mentre il test stabiliva che lo era).**

A partire da questi dati risolvere i seguenti quesiti:

a. calcolare la probabilità q che un test positivo indichi la gravidanza (cioè determinare *il valore predittivo del test positivo),*
b. calcolare la probabilità r che un test negativo escluda la gravidanza (cioè determinare *il valore predittivo del test negativo),*
c. calcolare la probabilità p che il test dia indicazioni esatte (cioè determinare *l'efficienza del test).*

Un primo approccio risolutivo

Per rispondere al quesito si può seguire una via puramente numerica.
Riteniamo però più significativo dal punto di vista didattico seguire anche altri due approcci che prevedono il ricorso ad opportune rappresentazioni grafiche. Li illustreremo esaminando la domanda c.

a) Il testo afferma che, dei 1000 test positivi (“casi possibili”), solo 5 test sono errati. Quindi sono 1995 gli esiti che indicano correttamente la gravidanza (“casi favorevoli”). La probabilità q richiesta è pertanto:

$$q = \frac{1995}{2000} \cong 0{,}998$$

b) Ragionando in modo analogo, si determina facilmente la probabilità r che un test negativo escluda la gravidanza:

$$r = \frac{920}{1000} = 0{,}920$$

[36] Tratto da Castelnuovo - Gori Giorgi – Valenti, Matematica oggi 2, ed. La Nuova Italia, 1992, pag. 821 n. 214.

Modellizzazione mediante tabella a doppia entrata

Possiamo **suddividere** la popolazione in esame in quattro sottoinsiemi, a seconda dell'esito del test e della presenza o meno di gravidanza[37] e rappresentare in modo espressivo la situazione mediante una tabella come la seguente:

	negativi	**positivi**
no gravidanze	*Diagnosi corretta*	***Falso Positivo**: test positivo **e** no gravidanza*
gravidanze	***Falso Negativo**: test negativo **e** gravidanza*	*Diagnosi corretta*

Riportiamo sulla tabella i dati forniti nel testo e a partire da essi ricaviamo presto gli altri che servono per rispondere al quesito[38] (a destra).

	negativi	***positivi***	
no gravidanze		5	
gravidanze	80		
Totale	**1000**	**2000**	**3000**

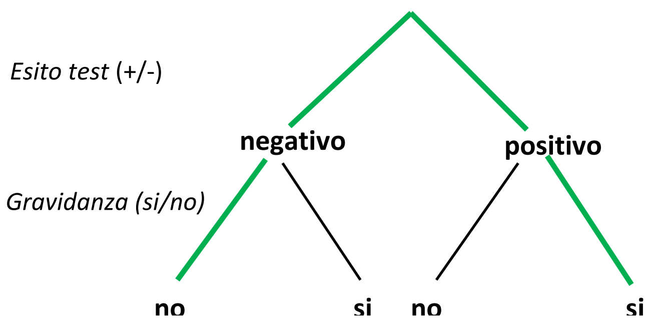
	negativi	***positivi***	
no gravidanze	920	5	
gravidanze	80	1995	
	1000	**2000**	**3000**

Il test fornisce indicazioni esatte per le donne che appartengono ai due sottoinsiemi evidenziati in verde. Pertanto i "**casi favorevoli**" sono $1995 + 920 = 2915$ su una popolazione di 3000 unità, i "**casi possibili**".
In definitiva la probabilità p che il test dia indicazioni esatte è

$$p = \frac{\mathbf{2915}}{\mathbf{3000}} \approx 0{,}972.$$

Modellizzazione mediante grafo ad albero

Per comprendere più in profondità la questione può essere utile schematizzare le situazioni che si possono presentare anche mediante un grafo ad albero:

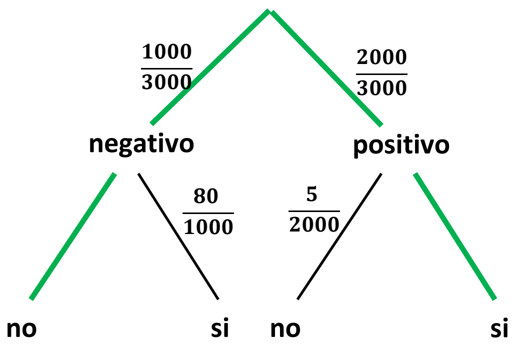


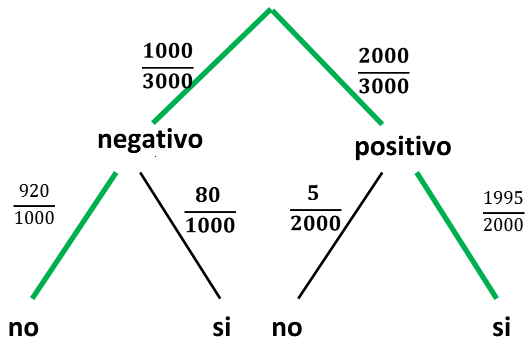
[37] Più precisamente si è ottenuta una *partizione* della popolazione in esame.
[38] Il testo afferma che "su 1000 test negativi, 80 erano sbagliati". Questo implica che i restanti 920 test negativi erano esatti, cioè indicavano effettivamente assenza di gravidanza.

Sono evidenziati in verde i due cammini che rappresentano i casi in cui il test fornisce un esito esatto:

- *test negativo e "no" gravidanza*
- *test positivo e gravidanza*

Per calcolare le probabilità di questi due eventi, cominciamo con l'indicare accanto ad ogni ramo il valore di probabilità dell'evento "elementare" corrispondente.
Sul grafo posto a sinistra riportiamo le probabilità che si desumono *direttamente* dai dati forniti. Deduciamo[39] poi le due probabilità che mancano e le aggiungiamo sul grafo a destra.





La probabilità dell'evento "*test negativo* **e** *no gravidanza*" si può ottenere in modo espressivo **moltiplicando** i valori di probabilità indicati lungo il cammino corrispondente:

$$p(\text{"}test\ negativo\ e\ no\ gravidanza\text{"}) = \frac{1000}{3000} \cdot \frac{920}{1000} = \frac{920}{3000}$$

Analogamente

$$p(\text{"}test\ positivo\ e\ gravidanza\text{"}) = \frac{2000}{3000} \cdot \frac{1995}{2000} = \frac{1995}{3000}$$

Il test dà un'*indicazione esatta* quando accade uno **oppure** l'altro dei due eventi rappresentati in verde. Pertanto la sua probabilità è la **somma** delle probabilità di questi due eventi:

$$p(\text{"}test\ esatto\text{"}) = \frac{920}{3000} + \frac{1995}{3000} = \frac{\mathbf{2915}}{\mathbf{3000}} \approx 0{,}972.$$

<u>Osservazioni</u>

- Ciascuna delle quattro celle della tabella (che rappresentano i quattro sottoinsiemi in cui è suddivisa la popolazione) **corrisponde** ad uno dei quattro possibili cammini sul grafo ad albero. Ad esempio alla cella in alto a sinistra nella tabella corrisponde il cammino più a sinistra sul grafo ad albero ("negativo e no gravidanza").
- Per i risultati sottesi all'uso del grafo ad **albero come modello di calcolo** rimandiamo alle osservazioni scritte a commento di altri esercizi (in particolare quello relativo ai *circuiti elettrici*).
- La somma delle probabilità dei rami relativi ad uno stesso nodo vale 1.
  E ancora 1 vale la somma delle probabilità di tutti i quattro cammini sull'albero:

$$\frac{1995}{3000} + \frac{5}{3000} + \frac{80}{3000} + \frac{920}{3000} = 1$$

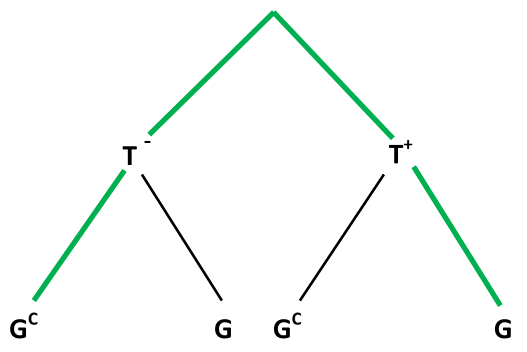
  Questi due fatti possono essere utilizzati per controllare i risultati.

[39] Il testo afferma che "su 1.000 test negativi, 80 erano sbagliati". Questo implica che i restanti 920 test negativi erano esatti, cioè indicavano assenza di gravidanza. Pertanto la probabilità che non vi sia gravidanza, sapendo che il test ha avuto esito negativo, è $\frac{920}{1.000}$. In modo analogo si calcola la probabilità di gravidanza sapendo che l'esito del test è positivo.

- Possiamo esprimere il procedimento in modo coinciso, utilizzando le seguenti **notazioni** per indicare gli eventi che intervengono:

  T^+ = "*la donna* è *risultata positiva al test*", T^- = "*la donna è risultata negativa al test*"
  G = "*la donna ha una gravidanza*", G^c = "*la donna non ha una gravidanza*"

  Ecco allora come diventa il grafo ad albero:



Ed ecco come si può esprimere in **forma compatta** l'evento "*test esatto*":

$$(T^- \text{ e } G^c) \text{ o } (T^+ \text{ e } G)$$

Gli errori dei test diagnostici

Il peso che può avere un errore nell'esito del test dipende dalla tipologia del test. Esaminiamo alcuni esempi al riguardo.

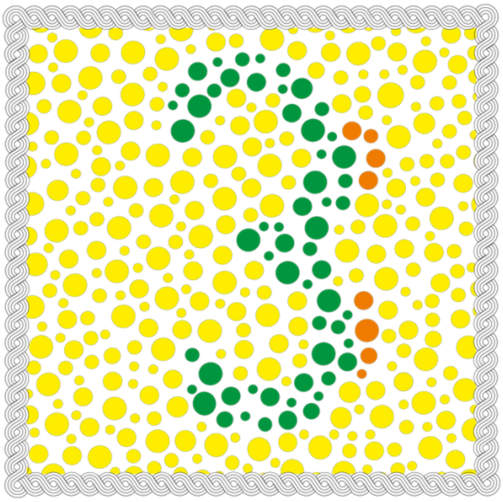
- Per un test di gravidanza (effettuato magari con un kit di farmacia) un **Falso Positivo**, può creare un po' d'ansia, ma può venir corretto da una successiva visita medica. Invece un **Falso Negativo** è potenzialmente più grave dato che la madre, in tal caso, non avrebbe motivo di prendere le precauzioni che normalmente si seguono in caso di gravidanza.

- Per un test relativo all'HIV un **Falso Positivo** può risultare molto grave, al punto che si sono verificati casi di suicidio di persone in realtà sane. Un **Falso Negativo** può avere conseguenze ancora più gravi dato che si rassicura il paziente sul suo stato di salute mentre in realtà risulta malato.

- È importante ridurre il numero dei **Falso Positivo** anche negli esperimenti in cui si testa l'efficacia di un farmaco. Altrimenti l'analisi potrebbe sancire erroneamente la bontà di un medicinale che in realtà non ha alcun valore terapeutico.

- Il giudizio di un tribunale è soggetto a valutazioni probabilistiche: è difficile avere la certezza assoluta della colpevolezza o dell'innocenza di un persona. Ma, se da una parte è importante evitare la condanna di un innocente (**Falso Positivo**), dall'altra è altrettanto importante non assolvere un colpevole (**Falso Negativo**).
  Ora, se si promuovono misure troppo garantiste per evitare la condanna di innocenti, si rischia di ottenere anche l'assoluzione di qualche colpevole. Il sistema giudiziario italiano ha introdotto un terzo grado di giudizio proprio per tutelare meglio gli innocenti; l'effetto è però anche quello di allungare vistosamente la durata del processo, con il rischio di rilasciare possibili colpevoli per decorrenza dei termini.

A5 Probabilità e genetica: il daltonismo[40]



Per capire che cos'è il daltonismo, si può osservare la figura: la maggior parte delle persone riesce a distinguere, dal fondo, un numero di colore diverso, perché nell'occhio sono ugualmente sviluppate le terminazioni nervose sensibili ai vari colori. Invece, un daltonico non riesce a distinguere il numero dal fondo, perché nel suo occhio sono inattive alcune terminazioni nervose.
Studi abbastanza recenti hanno portato a concludere che la caratteristica "attività delle terminazioni nervose sensibili ad un dato colore" **è ereditaria e si trova sul cromosoma X**. Si hanno perciò situazioni differenti per i maschi e per le femmine.
Infatti un **maschio** ha la coppia di cromosomi XY ed è **daltonico** se sull'unico cromosoma X manca la caratteristica relativa alla percezione dei colori; questa situazione cromosomica si può indicare con $\mathbf{X^*Y}$.
Invece, per una femmina che ha la coppia di cromosomi XX, è sufficiente avere la caratteristica su un solo cromosoma X per percepire normalmente i colori.

Quindi una **donna è daltonica** solo se ha i cromosomi $\mathbf{X^*X^*}$. Nella situazione X^*X non è daltonica, ma portatrice sana del daltonismo.

Si ha dunque che:
- il daltonismo modifica la percezione dei colori;
- il daltonismo è un'anomalia ereditaria, trasmessa attraverso il cromosoma X.

Le possibili situazioni genetiche sono:

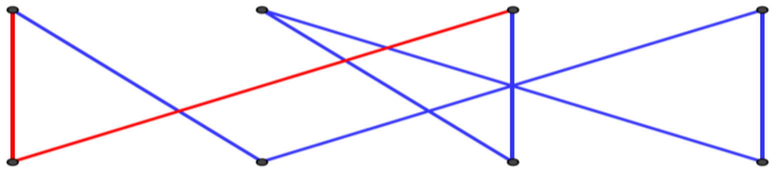
XY	maschio sano;
$\mathbf{X^*Y}$	maschio daltonico;
XX	femmina sana;
$\mathbf{X^*X}$	femmina sana, ma portatrice del daltonismo;
$\mathbf{X^*X^*}$	femmina daltonica.

[40] Dal libro di testo di Castelnuovo - Gori Giorgi – Valenti, Matematica oggi 2, ed. La Nuova Italia, 1992, pag. 472-473-474. La figura di riferimento è stata modificata.

Come si trasmette il daltonismo

Lo schema seguente aiuta a capire quali situazioni si possono presentare per i figli di una portatrice sana del daltonismo e di uomo non daltonico.

	Madre portatrice sana del daltonismo		Padre non daltonico	
Cellule	**X^*X**		**XY**	
Gameti	X^*	X	X	Y



Figli possibili	**X^*X**	**X^*Y**	**XX**	**XY**

Nello schema si è messo in evidenza il caso del figlio daltonico: *solo un figlio maschio può essere daltonico.*

Eventi dipendenti nella trasmissione del daltonismo

La precedente coppia di genitori potrebbe valutare la probabilità di avere un **figlio sano** in due situazioni diverse:

a. **prima di sapere il sesso del figlio**; in tal caso si hanno 4 casi possibili, fra cui 3 favorevoli, e perciò si trova:

$$s = \frac{3}{4} = 0{,}75$$

b. **dopo aver saputo che il figlio è maschio**; così sono rimasti due soli casi possibili, di cui uno favorevole, e perciò la probabilità r di avere un figlio sano, *subordinata* al fatto che il figlio sia maschio è:

$$r = \frac{1}{2} = 0{,}5$$

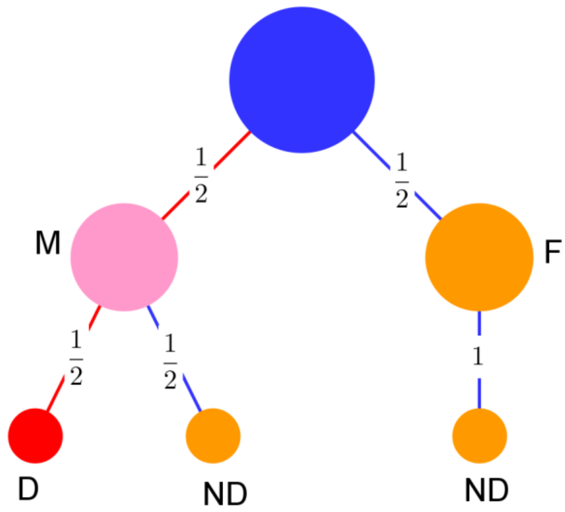
Si ritrova così che i due eventi «**avere un figlio maschio» e «avere un figlio non daltonico**» sono ***dipendenti***, perché sapere che il figlio è maschio altera la probabilità che il figlio non sia daltonico.

Probabilità composta per esaminare la trasmissione del daltonismo

Per valutare la **probabilità** p che un figlio sia maschio *e* non daltonico ci si può basare sul **diagramma ad albero** della figura che segue. Si trova:

$$p = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$$

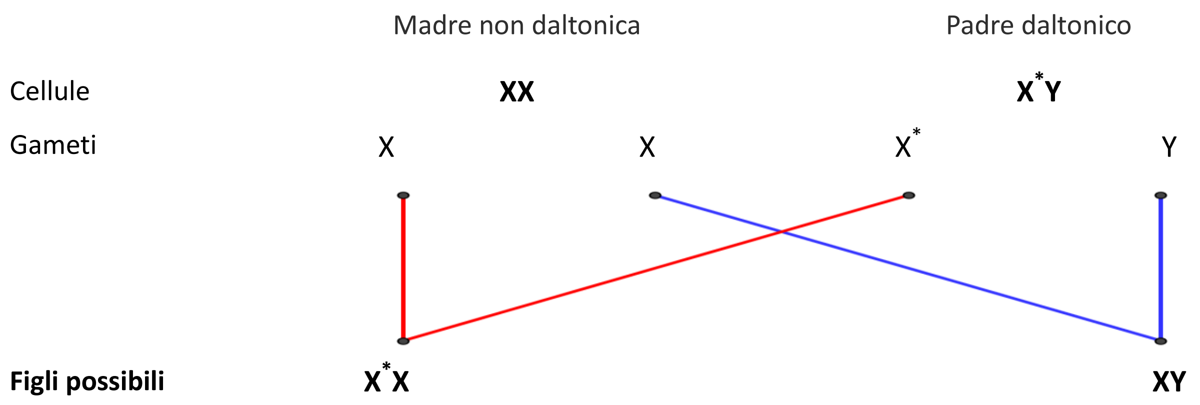
perché $\frac{1}{2}$ è la probabilità che il figlio sia maschio e $\frac{1}{2}$ è la probabilità che il figlio maschio sia anche non daltonico.



Il daltonismo si trasmette «per via di donna»

Altre situazioni interessanti da esaminare sono quelle che si presentano per i figli di un daltonico con una donna sana, anche geneticamente.

Ecco lo schema che mostra le varie situazioni possibili.



In questo caso non si trovano più situazioni di incertezza:

- i figli maschi sono tutti **XY,** cioè certamente non daltonici;
- le figlie femmine sono tutte **X*X,** cioè portatrici sane del daltonismo.

Dunque, un uomo daltonico che si unisce con una donna non daltonica non trasmette la sua anomalia ai figli; ma il daltonismo può ricomparire nei nipoti maschi dell'uomo daltonico, trasmesso da una sua figlia. Per questo si dice che il daltonismo si trasmette «per via di donna».

Altre anomalie che si trasmettono «per via di donna»

Gli stessi ragionamenti seguiti per la trasmissione del daltonismo valgono anche nel caso di altre malattie trasmesse attraverso il cromosoma X. Ecco altri due esempi:

- *l'emofilia,* una malattia per cui è notevolmente ritardata la coagulazione del sangue;
- *il favismo,* che determina devastanti distruzioni dei globuli rossi del sangue, come reazione alle piante di fava (da cui il nome).

L'emofilia è stata, fino alla fine del secolo scorso, una malattia molto diffusa presso alcune famiglie reali (fra cui i Borboni di Spagna e i Romanov di Russia), in cui, per motivi dinastici, erano frequenti i matrimoni fra consanguinei.

Il favismo è ancora molto comune in Sardegna. Favismo e microcitemia si trovano su cromosomi differenti; tuttavia sono piuttosto frequenti, soprattutto in Sardegna, i maschi affetti sia dalla microcitemia che dal favismo.