6. Probabilità di eventi che dipendono da altri

La probabilità che dipende da altre[1], ovvero la probabilità condizionata, è già comparsa in varie situazioni esaminate nei capitoli precedenti del Percorso: dalle estrazioni senza reinserimento nell'urna fino ai test clinici.
Di comparsa appunto si è trattato, visto che il concetto è rimasto sullo sfondo e non è stato precisato tramite una definizione nonché una notazione specifica. Ciò è frutto di una scelta meditata e coerente con l'approccio seguito nelle altre fasi del Percorso: passare alla formalizzazione solo dopo che gli studenti hanno preso confidenza con i nuovi oggetti matematici, attraverso attività che permettono di scorgerne significato e portata. Anzi, proprio tali situazioni didattiche dovrebbero far nascere l'esigenza di comunicare in modo coinciso ed univoco, motivando così il ricorso alla formalizzazione.
In sintesi, riteniamo che l'attività didattica nel contesto[2] della probabilità condizionata andrebbe articolata in tre fasi: esplorativa, formale, applicativa. E ciò dovrebbe avvenire per varie ragioni, che in parte abbiamo già discusso: innanzitutto affinché i nuovi oggetti matematici si possano radicare con continuità sui saperi pregressi dello studente, in secondo luogo perché i ragazzi possano attribuire un senso a tali oggetti e non li percepiscano come inutili complicazioni imposte dall'esterno, ed infine affinché i contenuti e i metodi così sviluppati rimangano effettivamente disponibili anche a lungo termine.

In quest'ottica, si partirà dall'esame di alcune situazioni allo scopo di investigare gli aspetti didatticamente più significativi della probabilità che dipende da altre. Si proseguirà precisando formalmente il significato di probabilità condizionata, ad un livello che riteniamo adeguato per lo studente di scuola secondaria.
Infine, lo strumento matematico così introdotto verrà impiegato per descrivere e risolvere efficacemente alcuni interessanti problemi ambientati in contesti reali: test clinici, casi giudiziari…

---

[1] Più precisamente dovremmo parlare di "probabilità di eventi che dipendono da altri", ma la formula linguistica utilizzata ci sembra più espressiva.
[2] Come in diversi altri contesti didattici, in particolare in quelli esaminati nel nostro percorso.

6.1 Situazioni motivanti

Dunque iniziamo ad esaminare come il valore di probabilità che si assegna a un dato evento possa venir modificato quando si disponga di <u>ulteriori informazioni</u> su di esso. Precisamente ci proponiamo di schematizzare tali situazioni mediante opportune rappresentazioni grafiche e intendiamo motivare l'introduzione di notazioni e termini specifici che contribuiscano a sintetizzarle in modo espressivo.

a) Sopravvivenza e tavole

L'Istat pubblica periodicamente le tavole di mortalità della popolazione, elaborando in modo opportuno i dati anagrafici rilevati. Esse costituiscono un riferimento fondamentale anche per le assicurazioni sulla vita.
I dati in tabella sono estratti dalle tavole pubblicate per l'anno 2013[3]: si considera una popolazione iniziale fittizia di 100.000 nati; su ogni riga è riportato il numero di individui della popolazione che sono in vita all'età specificata nella colonna a sinistra, suddivisi per sesso.

età	*numero viventi maschi*	*numero viventi femmine*
0	100.000	100.000
40	98.120	98.976
70	82.269	89.970

Qual è la probabilità per **una quarantenne,** scelta a caso nella popolazione, di vivere almeno fino a 70 anni?

Per rispondere alla questione assumiamo innanzitutto che la popolazione sia stazionaria[4] per quanto riguarda la durata della vita. Con questa ipotesi, possiamo interpretare la tabella dicendo che tra le 98.976 quarantenni della popolazione, 89.970 arriveranno a compiere 70 anni. Di conseguenza, la nostra valutazione della probabilità richiesta è

$$p(\text{"}70\ \boldsymbol{da}\ \mathbf{40}\text{"}) = \frac{\mathbf{89.970}}{\mathbf{98.976}} = 0{,}909\ \ldots$$

<u>Osservazione</u>
Con ciò abbiamo risposto, ma il nostro interesse per la questione si spinge oltre.
Consideriamo infatti l'evento "*una neonata vive almeno fino ai 70 anni*", e indichiamolo sinteticamente con "*70*"; coerentemente con quanto appena discusso, diciamo che la sua probabilità è

$$p(\text{“}70\text{”}) = \frac{89.970}{100.000} = 0{,}899\ \ldots$$

[3] Tali dati non sono stati effettivamente rilevati su una popolazione seguita costantemente nell'arco di molti anni. Comunque per ottenerli ci si basa su rilevazioni anagrafiche (reali) relative ad un anno recente e le si elabora opportunamente.
[4] Ossia il numero di nati è costante ogni anno e, per ogni classe di età, è costante il numero di individui che vi appartengono (questa seconda ipotesi si può anche esprimere in modo equivalente dicendo che, per ogni classe di età, il numero di ingressi è uguale al numero di uscite).

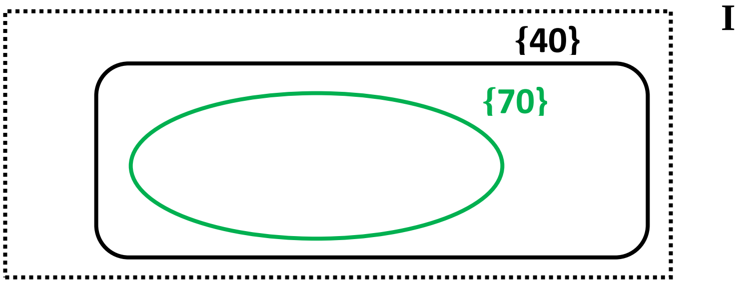
Abbiamo così ottenuto

$$p(\text{"70 } da \text{ 40"}) > p(\text{"70"})$$

Perché i due valori di probabilità $p(\text{"70 da 40"})$ e $p(\text{"70"})$ sono diversi?

Le ragioni profonde di tale differenza risiedono nel fatto che nel valutare la probabilità dell'evento "70 da 40" si sfrutta un'informazione in più: l'individuo è arrivato a compiere i 40 anni. Come vedremo, in situazioni di questo tipo si parla di probabilità condizionata.

Più in dettaglio, possiamo schematizzare la situazione nel modo seguente



dove I è l'insieme della popolazione iniziale (femmine), **{40}** l'insieme degli individui della popolazione I che arrivano a compiere 40 anni, e {70} l'insieme degli individui della popolazione I che arrivano a compiere 70 anni.

- L'insieme dei casi **favorevoli** è **{70}** , sia per l'evento "70 da 40" che per l'evento "70"
- l'insieme dei casi **possibili** invece è diverso nelle due situazioni: per "70 da 40" "è" **{40} ,** mentre per "70" è l'insieme I
- perciò, anche le probabilità dei due eventi in esame dovranno essere diverse.

In particolare, visto che tra i due insiemi di casi possibili vale la relazione $\{40\} \subset I$, si conclude che $p(\text{"70 da 40"}) > p(\text{"70"})$.

b) Fumatori

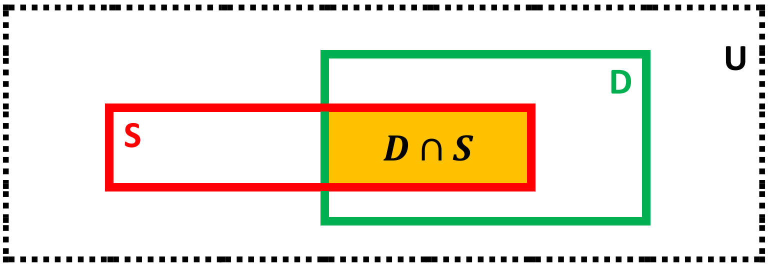
Sulla popolazione U dei 51,4 milioni italiani che hanno almeno 15 anni, ben 11,3 milioni sono fumatori, di questi 5,1 milioni sono donne[5].
Assumendo che le donne costituiscano il 52,5% della popolazione U, qual è la probabilità che **la donna** (estratta a caso tra le donne di U) sia fumatrice?

Per prima cosa osserviamo che abbiamo a che fare essenzialmente con tre insiemi di persone:

- l'insieme degli individui della popolazione, che denotiamo con **U**
- l'insieme dei fumatori della popolazione, che indichiamo con **S**
- l'insieme delle donne della popolazione, che chiamiamo **D**.

[5] Dati da indagine Doxa del 2014 per conto dell'ISS, Istituto Superiore Sanità http://www.iss.it/binary/pres/cont/schedeFUMO.pdf

La prima frase del testo permette di affermare che l'insieme **S** e l'insieme **D** hanno intersezione non vuota. Inoltre, l'informazione contenuta nelle seconda frase, consente di escludere che l'insieme **S** sia contenuto in **D**.
Pertanto possiamo schematizzare la situazione mediante i diagrammi di Venn, nel modo seguente:



A questo punto non resta che valutare la probabilità richiesta, ossia la probabilità di
"essere fumatore ***per la donna*** *della popolazione".*

Dunque, qual è l'insieme dei **"casi possibili"** relativamente a tale richiesta? Non può essere l'intero insieme U, dato che all'insieme iniziale si è aggiunta l'ipotesi "per la donna della popolazione", ma il più ristretto insieme **D.** Analogamente l'insieme dei **"casi favorevoli"** è **D ∩ S.**
Di conseguenza la probabilità richiesta è:

$$\frac{\mathbf{5,1 \cdot 10^6}}{\mathbf{27,0 \cdot 10^6}} \approx 0,\mathbf{189}\cdots$$

Osservazione
Proviamo ora a togliere l'ipotesi aggiuntiva, ossia prendiamo in esame l'evento "essere fumatore per un individuo scelto sull'intera popolazione". Con questa assunzione, l'insieme dei casi possibili diventa l'intero insieme iniziale U, l'insieme dei casi favorevoli l'insieme S e il valore di probabilità è

$$p(S) = \frac{11{,}3 \cdot 10^6}{51{,}4 \cdot 10^6} \approx 0{,}220$$

Come ci si aspetta, è un valore diverso da quello ottenuto considerando l'ipotesi aggiuntiva "per la donna". In altre parole, l'informazione aggiuntiva modifica il valore della probabilità dell'evento "essere fumatore", facendolo passare dal 22% al 18,9%.
Il divario risulta ancora più marcato se si confronta la probabilità di essere fumatore per una donna con quella di essere fumatore per un uomo: quest'ultima si attesta infatti sul 25,4%, come si può dedurre dai dati forniti nel testo del quesito. L'esame dei dati relativi al 2013 evidenzia una differenza ancora maggiore: 15,3% per le donne contro il 26,3% per gli uomini.

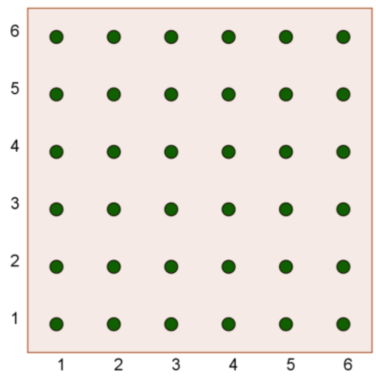
Possiamo riassumere quanto visto finora nella sezione, dicendo che la probabilità richiesta in questo quesito è una probabilità che dipende dal fatto che l'evento D si è realizzato e pertanto si indica spesso come "probabilità condizionata". Il suo valore è diverso da quello della probabilità dell'evento "essere fumatore", che usualmente si denota con $p(S)$. Dunque per la probabilità condizionata dovremmo utilizzare un simbolo diverso: come discuteremo in un contesto più generale in una delle prossime sezioni, utilizzeremo l'espressiva notazione $p_D(S)$ per indicare che la sua valutazione è condizionata al verificarsi dell'evento D.

c) Lancio di due dadi

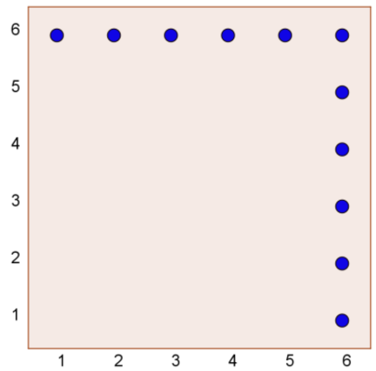
Calcola la probabilità che in un lancio di due dadi, uno bianco e l'altro giallo, escano due "6"

a) senza informazioni aggiuntive;
b) sapendo che nel lancio è uscito **almeno** un "6";
c) sapendo che l'esito del **dado giallo** è "6".

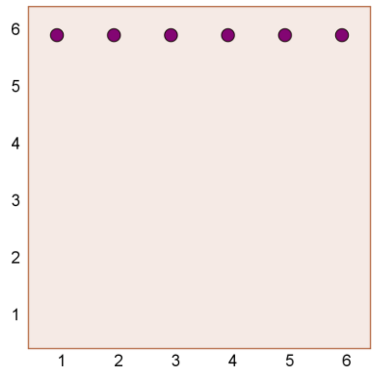
Le tre situazioni a), b), c) si possono modellizzare in modo espressivo mediante tabelle a doppia entrata, segnando con un puntino ogni cella che corrisponde ad uno dei casi possibili. Ecco un esempio di cosa si ottiene seguendo tale proposito:



a)



b)



c)

Tali modelli si possono visualizzare in modo ancora più efficace mediante opportune rappresentazioni dinamiche, come quelle offerte dal software Geogebra[6], per esempio usando il file LancioDadi_Condizionata.ggb.
Sulla base di tali rappresentazioni si dovrebbero determinare presto i valori di probabilità richiesti: essi sono

$$a)\ \frac{1}{36} \qquad b)\ \frac{1}{11} \qquad c)\ \frac{1}{16}$$

Osservazione
I tre schemi grafici mettono chiaramente in evidenza il fatto, a nostro avviso, didatticamente più rilevante: le nuove informazioni sugli esiti dei lanci modificano l'insieme dei "casi possibili".

[6] Il file proposto è stato ideato dalla collega Francesca Arrigoni.

6.2 Le attenzioni *

Nelle situazioni esaminate nel paragrafo precedente, il valore di probabilità di un evento A viene modificato dall'informazione che si è verificato un dato evento B. Già nel capitolo 5 abbiamo affrontato situazioni di questo tipo ed in particolare abbiamo osservato che ciò si può esprimere dicendo che i due eventi A, B sono dipendenti.
Prima di procedere oltre e descrivere in modo formale la probabilità condizionata, crediamo sia utile investigare ulteriormente sul concetto di dipendenza tra eventi, precisandone alcuni aspetti spesso miscompresi dagli studenti. Tali considerazioni, che riportiamo di seguito, sono già state rivolte agli insegnanti nel capitolo 5; in questa fase del Percorso possono essere apprezzate anche dagli studenti delle classi meno deboli.

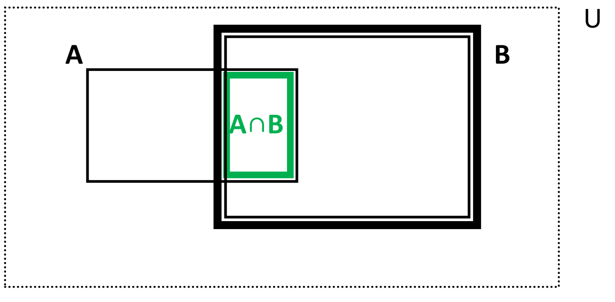
Per valutare se due eventi sono o meno indipendenti, in molte situazioni didattiche è sufficiente ricorrere all'intuizione. In alcuni casi però essa può essere fuorviante:
- la dipendenza tra due eventi non significa sempre che uno dei due influenza l'altro (a tale proposito si ricordi quanto osservato nel capitolo 5 relativamente alla statistica sulle case inglesi dopo la seconda guerra mondiale);
- l'indipendenza tra due eventi non è sempre intuitiva[7].

6.3 Facciamo il punto

A questo punto del Percorso, gli studenti dovrebbero ormai aver preso confidenza con la probabilità che dipende da altre; pertanto si trovano nella condizione di apprezzare una precisazione formale dei contenuti ad essa relativi.
La proponiamo di seguito ad un livello di rigore che riteniamo adeguato alla sensibilità e alla maturità culturale dello studente di scuola secondaria. Riserveremo invece al docente l'analisi critica dei concetti così introdotti e la discussione sulla loro collocazione nell'ambito della teoria assiomatica. Ma di questo ci occuperemo nella sezione successiva.

Dati due eventi A e B tali che $p(B) \neq 0$, diciamo **probabilità condizionata** di A dato B, la probabilità che si verifichi l'evento A, valutata sapendo che **l'evento B si è verificato**[8].
La indichiamo con la notazione $p_B(A)$.[9]

La situazione si può schematizzare nel modo seguente, dove U è un insieme in cui si considerano contenuti A, B.



[7] Si veda l'esempio del lancio di un dado, discusso nella sezione a) del capitolo 5.
[8] In altri termini, disponiamo dell'informazione che l'evento B è accaduto e decidiamo di tenerne conto nella valutazione della probabilità dell'evento A.
[9] È in uso anche la notazione $p(A\,|B)$, ma questa ci sembra meno espressiva e meno chiara per uno studente di scuola secondaria. Infatti l'ordine di scrittura, prima A e dopo B, può venir erroneamente interpretato come un ordine di condizionamento, ossia che l'evento A condizioni l'evento B.
Inoltre la scrittura $p(A\,|B)$ potrebbe indurre lo studente a pensare che A | B sia un evento; ma, come avremo modo di chiarire nella prossima sezione, tale scrittura non indica alcun evento.

Precisamente, in base alla nostra interpretazione di probabilità condizionata, si deve considerare che si è verificato l'evento B. Perciò:

l'insieme dei nuovi **"casi possibili" = B**
l'insieme dei **"casi favorevoli"[10] = A ∩ B**

Inoltre vale

$$\boldsymbol{p_B(A) = \frac{p(A \cap B)}{p(B)}} \qquad (*)$$

dove le probabilità *p* sono valutate rispetto all'insieme U.

Una giustificazione dell'uguaglianza (*)

- Per iniziare, ricordiamo che l'insieme dei nuovi "casi possibili" è B e l'insieme dei "casi favorevoli" ad A è $A \cap B$**.** Pertanto

$$p_B(A) = \frac{misura(\boldsymbol{A \cap B})}{misura(\boldsymbol{B})}$$

  Tali misure[11] sono effettuate entrambe rispetto all'insieme U.

- Ora, nell'interpretazione geometrica della probabilità, la **probabilità di un insieme** E **è una misura** *di E* rispetto all'insieme in cui lo si considera contenuto.
  Quindi possiamo sostituire alle misure, le probabilità degli insiemi. Ossia

  $misura\ (\boldsymbol{A \cap B}) = \boldsymbol{p(A \cap B)}$
  $misura\ (\boldsymbol{B}) = \boldsymbol{p(B)}$

  E con ciò abbiamo giustificato l'uguaglianza[12] (*).

Osservazione

Se nella formula (*) si esplicita il termine $\mathrm{p(A \cap B)}$, si ottiene la relazione che esprime la legge della moltiplicazione, già ampiamente discussa nel capitolo 5. In effetti la legge della moltiplicazione, che avevamo giustificato intuitivamente mediante l'analogia dell'acqua, è formalmente una conseguenza diretta della definizione di probabilità condizionata.

Per meglio comprendere il ruolo e il significato della formula (*), vale la pena approfondirne alcuni aspetti mediante attività che si possono collocare in più momenti del Percorso di probabilità (si veda l'appendice A1 in fondo al capitolo).

[10] Cioè quelli in cui può realizzarsi l'evento A.
[11] Nelle righe seguenti suggeriamo per mezzo di esempi cosa s'intende per *misura* di un insieme.
[12] A rigore, tale giustificazione vale quando si considera una distribuzione di probabilità uniforme, quale ad esempio si ha nel caso finito secondo l'approccio classico alla probabilità.

6.4 La definizione di probabilità condizionata *

Nella sezione precedente abbiamo guardato alla probabilità condizionata dal punto di vista dello studente di scuola secondaria. E il docente invece, quale punto di vista dovrebbe adottare? Ecco allora alcune considerazioni da tenere presenti per operare scelte didattiche consapevoli e interpretare criticamente il libro di testo[13].

- La formula $\mathbf{p}(\mathbf{B}) \cdot \mathbf{p}_{\mathbf{B}}(\mathbf{A}) = \mathbf{p}(\mathbf{A} \cap \mathbf{B})$(**) è in realtà la **definizione** di probabilità condizionata nell'ambito della teoria assiomatica e, in quanto definizione nel quadro di una teoria assiomatica, ha un ruolo strettamente formale. Ossia, a rigore, si opera su di essa in un solo modo: applicando le regole sintattiche fornite dagli assiomi, senza ricorrere esplicitamente a considerazioni di carattere semantico[14]. Ciò non significa che la definizione non si possa poi interpretare in uno o più modi; piuttosto si deve aver chiaro che ogni affermazione su di essa deriva unicamente da inferenze logiche e non da considerazioni basate solo sull'intuizione.

- Se si assume la relazione (**) come definizione di probabilità condizionata e si valuta la probabilità secondo l'approccio classico (o frequentista o soggettivo) si può **dimostrare** che $\mathrm{p}_{\mathrm{B}}(\mathrm{A})$ è la *probabilità dell'evento A sapendo che si è verificato l'evento B.* In altre parole, nell'ambito dell'approccio classico, la formula (*) e la definizione da noi proposta di probabilità condizionata sono equivalenti.

- Pertanto la nostra definizione fornisce il **significato** di probabilità condizionata nell'approccio classico. Normalmente la presentazione formale della probabilità condizionata prevede i due passi: 1) definizione come formula e 2) precisazione del significato. Invece, per ragioni didattiche riteniamo più efficace invertire tale ordine espositivo: prima prospettare il significato e poi la formula.

6.5 Ancora test clinici

Nei paragrafi precedenti del capitolo si è precisato formalmente il significato di probabilità che dipende da altre e si sono introdotti nomi, notazioni e schemi per rappresentarla in modo espressivo. Tali strumenti consentono di esaminare efficacemente situazioni articolate, anche più complesse rispetto a quelle discusse nel capitolo 5: infatti essi permettono di comunicare in modo univoco, coinciso ed espressivo, superando così i limiti imposti dal linguaggio naturale.
Per iniziare, mettiamo alla prova i nuovi strumenti nell'analisi di una delle domande stimolo poste all'inizio del capitolo 1 del Percorso.

> Risulto positivo al test per l'HIV. Sapendo che il test "Elisa" ha una sensibilità[15] del 99,9%, vuol dire che al 99,9% ho contratto l'HIV?

[13] Un ottimo riferimento al riguardo è il testo di Villani e altri "Non solo calcoli", ed. Springer.
[14] In un contesto analogo, nell'ambito della geometria euclidea, Hilbert sosteneva provocatoriamente "si deve poter dire ogni volta, in luogo di "punti, rette, piani": "tavoli", "sedie", "boccali di birra"." Con ciò intendeva che non interessa quale interpretazione intuitiva si vuole dare ai concetti geometrici fondamentali; l'unico aspetto matematicamente rilevante è la relazione formale che viene introdotta su di essi dagli assiomi.
[15] Probabilità che un individuo malato risulti positivo al test.

Il quesito non ammette risposta univoca sulla base delle informazioni fornite, pertanto prima di procedere è opportuno precisare alcune ipotesi sul test. La formulazione iniziale era volutamente contratta per far emergere le convinzioni di senso comune.

Il test “Elisa”, relativo all’HIV, può fornire indicazioni errate.
Assumiamo che vi sia una probabilità del 99,9% che il test dia esito positivo nei soggetti che effettivamente hanno contratto l’HIV (*sensibilità* del test) ed una probabilità del 99,8% che il test risulti negativo sui soggetti che non hanno l’HIV (*specificità* del test).
Consideriamo ora una data popolazione. Assumiamo che lo 0,3% della quantità di individui della popolazione abbia l’HIV (*prevalenza* della malattia).
Il test, applicato ad un individuo scelto a caso in tale popolazione, ha dato **esito positivo**. Qual è la probabilità che tale individuo sia in realtà **sano**, cioè non abbia l’HIV?

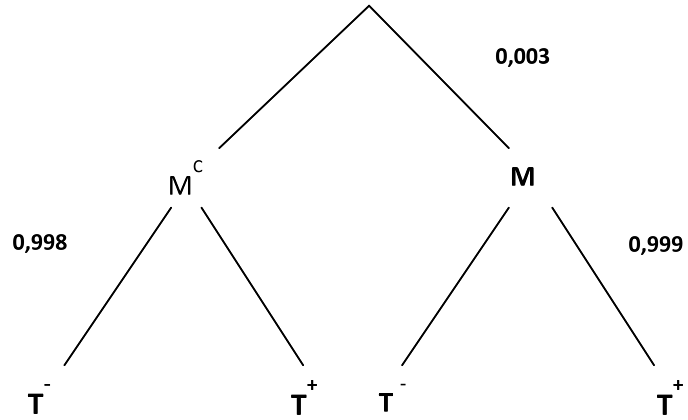
Come si vede, il testo risulta piuttosto articolato. Anche per questo, prima di discutere in classe una risoluzione del problema, è opportuno invitare gli studenti ad esaminarlo attentamente e a formulare delle congetture al riguardo, magari in forma scritta. L’esame delle produzioni dei ragazzi è cruciale per una azione didattica efficace, in quanto il docente ha così modo di fondare i nuovi saperi su quelli effettivamente disponibili per lo studente.
Tale attività rischia però di rivelarsi prematura se non è preceduta dall’analisi di alcuni problemi relativi ai test clinici, analoghi a quelli proposti nel capitolo 5 e che possiamo indicare come “problemi diretti”. Infatti, nell’affrontare tali questioni, gli studenti prendono confidenza con l’interpretazione del testo, con la schematizzazione mediante grafi opportuni, nonché con l’uso di notazioni specifiche. In sostanza, compiono un passo intermedio verso la risoluzione di un problema piuttosto articolato come quello in esame.

a) Modellizzazione

Cominciamo con il rappresentare i casi che si possono presentare per l’individuo della popolazione. E per farlo ci serviamo di un grafo ad albero. Per indicare le varie opzioni, denotiamo con

- M l’evento “*l’individuo è ammalato*”,
- M^C l’evento “*l’individuo non è ammalato*”,
- T^+ l’evento “*il test ha dato esito positivo*”,
- T^- l’evento “*il test ha dato esito negativo*”.

Proseguiamo riportando, accanto ai relativi rami del grafo ad albero, i valori di probabilità degli eventi forniti dal testo[16].



Osservazione

La raffigurazione sembra davvero molto espressiva. Prima di procedere però, è opportuno accertarsi che sia chiaro il preciso significato di ogni suo aspetto.

- I valori di probabilità che compaiono al secondo livello dell'albero sono probabilità condizionate. Ad esempio, il valore 0,998 più a sinistra, è $\boldsymbol{p_{M^c}(T^-)}$.
- Ogni cammino sull'albero rappresenta un evento intersezione. Ad esempio, quello raffigurato più a sinistra, identifica l'evento $\mathrm{M^c} \cap \mathrm{T^-}$.

Per comprendere meglio il testo del problema, è utile rappresentare la situazione anche mediante un altro modello: la **tabella**. L'insieme dei casi che si possono presentare[17] per l'individuo della popolazione (casi iniziali) può essere suddiviso nei quattro sottoinsiemi rappresentati dai quattro rettangoli più piccoli in figura. Ad esempio, il rettangolo che occupa la seconda riga e la prima colonna rappresenta l'insieme $\mathrm{M} \cap \mathrm{T^-}$, ossia l'evento *"l'individuo ha l'HIV e il suo test ha esito negativo"*.

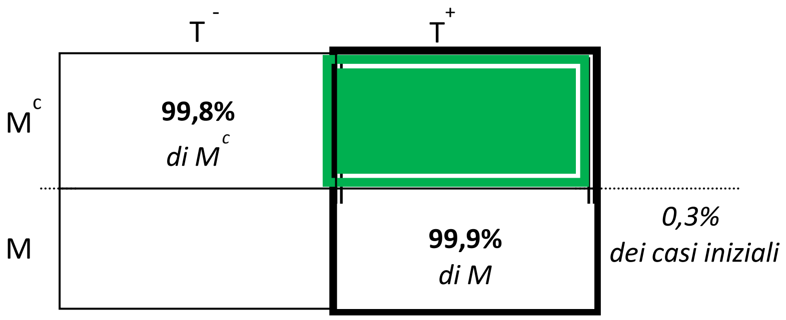
	T^-	T^+	
M^c	**99,8%** *di M^c*		
M		**99,9%** *di M*	*0,3% dei casi iniziali*

[16] Assumiamo che per la popolazione in esame valgano i due valori di probabilità indicati nel testo.
Inoltre osserviamo che, se la popolazione in esame è "sufficientemente numerosa", si possono esprimere tali ipotesi in termini di frequenze relative. Si può cioè assumere che risulti positivo al test circa il 99,9% degli individui della popolazione che hanno l'HIV. Analogamente possiamo assumere che risulti negativo al test circa il 99,8% degli individui della popolazione che non hanno l'HIV.
Volendo, in un primo momento, si può decidere di operare con le frequenze relative anziché con i valori di probabilità. Ciò è significativo dal punto di vista didattico perché permette agli studenti di prendere confidenza con la situazione; in particolare consente di approssimare la numerosità dei vari sottoinsiemi in cui si può partizionare la popolazione, secondo i due criteri: presenza (o meno) della malattia, esito del test. Quest'ultima schematizzazione dovrebbe risultare di più immediata comprensione per lo studente rispetto a quella che si ottiene considerando direttamente i valori di probabilità.
[17] Senza tener conto, per ora, del fatto che l'individuo è positivo al test.

Osserviamo che ciascuno dei quattro rettangoli corrisponde ad uno dei quattro possibili cammini nel grafo ad albero precedente.
È richiesta la probabilità che

"l'individuo non sia malato sapendo che il test ha avuto esito positivo"

Pertanto, l'insieme dei casi possibili è l'insieme T^+, il cui bordo è evidenziato in nero in figura. Mentre l'insieme dei casi favorevoli è l'insieme colorato in verde, ossia $\mathrm{M^c} \cap \mathrm{T^+}$.



Dunque la probabilità richiesta si denota con

$$\boldsymbol{p_{T^+}(M^C)}$$

<u>Osservazione</u>

A questo punto, il testo del problema dovrebbe risultare sufficientemente chiaro. Vogliamo però soffermarci ad investigare un ultimo aspetto ad esso legato, prima di passare alla risoluzione effettiva.

Consideriamo per un attimo una richiesta che può sembrare analoga, almeno dal punto di vista linguistico:

"sapendo che l'individuo non è malato, determinare la probabilità che il test abbia esito positivo"

Però ci si dovrebbe presto convincere che:
- l'insieme dei casi possibili è rappresentato dalla prima riga della tabella in figura, ossia da M^c,
- la probabilità richiesta si denota con $p_{M^c}(T^+)$.

Di conseguenza, concludiamo che la domanda è diversa da quella posta dal problema in esame.
In altri termini[18] $\boldsymbol{p_{T^+}(M^C)} \neq p_{M^c}(T^+)$

b) Risoluzione

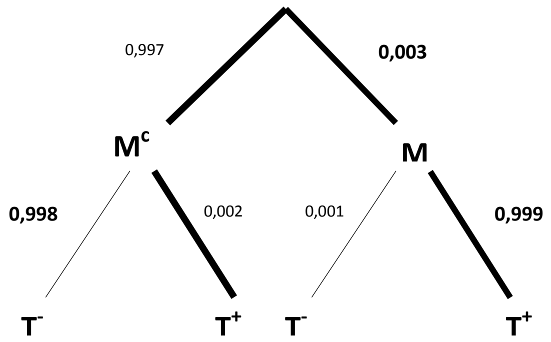
Si è illustrato in dettaglio come modellizzare la situazione e come passare dal registro linguistico del testo a quello grafico e simbolico. Potendo contare su tale attività preparatoria, gli studenti dovrebbero essere in grado di organizzare la risoluzione del problema, con consapevolezza e con una certa autonomia pur guidati dal docente. Dunque, mediante una lezione partecipata, dovrebbero ideare e precisare un procedimento analogo al seguente.

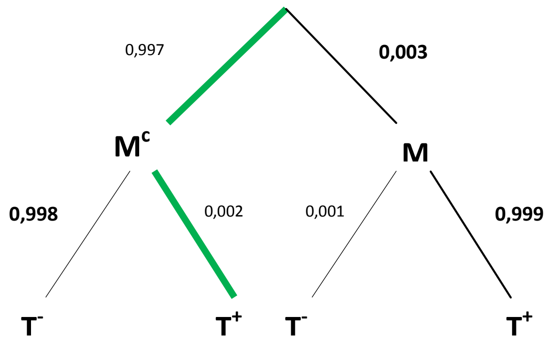
[18] L'importanza della questione non è da sottovalutare se, come sostiene Mlodinov in "La passeggiata dell'ubriaco", è un errore comune confondere $p_A(B)$ con $p_B(A)$. Quando poi l'errore è commesso da un medico, si comprende come le conseguenze dell'abbaglio vadano ben oltre i meri aspetti matematici della questione. A tale proposito, si pensi ai pazienti a cui è stata comunicata una probabilità del 90% di avere un cancro, per il solo fatto di essere risultati positivi al test diagnostico, mentre calcoli accurati mostrano come la probabilità di aver contratto la malattia si attesti "solo" al 9%. Torneremo ad esaminare nel dettaglio tale situazione nella prossima sezione.

È richiesta la probabilità dell'evento "l'individuo non è malato" sapendo che il test ha avuto esito positivo, ossia

$$p_{T^+}(M^C)$$

Evidenziamo **in grassetto** sull'albero i due cammini che rappresentano i **nuovi**[19] "**casi possibili**": "*individuo non malato e test positivo*", "*individuo malato e test positivo*". Otteniamo la figura riportata di seguito sulla sinistra.
Segniamo invece **in verde** il cammino che rappresenta i "**casi favorevoli**", "*individuo non malato e test positivo*", ottenendo la figura che segue, a destra. Aggiungiamo poi le probabilità non fornite nel testo, considerando che la somma delle probabilità relative ai rami uscenti da uno stesso nodo deve essere uguale ad 1.





Ora disponiamo di tutti gli elementi per calcolare la probabilità richiesta.
Leggiamo il processo di calcolo direttamente sull'albero ora costruito. Vale[20]

$$p_{T^+}(M^c) \; = \frac{p(\text{"}\textit{cammini favorevoli}\text{"})}{p(\text{"}\textit{cammini possibili}\text{"})}$$

Per calcolare la probabilità di ogni cammino moltiplichiamo le probabilità degli eventi rappresentati da ciascuno dei rami che lo costituiscono[21]:

$$\frac{p(\text{"}\textit{cammini favorevoli}\text{"})}{p(\text{"}\textit{cammini possibili}\text{"})} = \frac{\mathbf{0,997 \cdot 0,002}}{\mathbf{0,997 \cdot 0,002 + 0,003 \cdot 0,999}} \approx 0{,}40$$

<u>**Osservazione**</u>
Volendo, possiamo ripercorrere l'intero procedimento, esprimendo i singoli passi mediante il linguaggio degli insiemi:

$$p_{T^+}(M^c) = \frac{p(\boldsymbol{M^c} \cap \boldsymbol{T^+})}{p(\boldsymbol{T^+})} = \frac{p(\boldsymbol{M^c} \cap \boldsymbol{T^+})}{p(\boldsymbol{M^c} \cap \boldsymbol{T^+}) + p(\boldsymbol{M} \cap \boldsymbol{T^+})} = \frac{p(\mathbf{M^c}) \cdot p_{M^c}(\boldsymbol{T^+})}{p(\mathbf{M^c}) \cdot p_{M^c}(\mathbf{T^+}) + p(\mathbf{M}) \cdot p_M(\boldsymbol{T^+})}$$

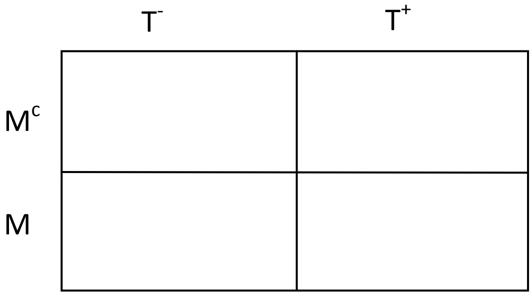
L'idea però è che lo studente non impari a memoria una qualche formula risolutiva, ma la ricostruisca di volta in volta, a partire dal grafo ad albero.

[19] Ossia una volta acquisita l'informazione che l'individuo è positivo al test.
[20] Per una giustificazione di tale uguaglianza, che riteniamo fondamentale, rimandiamo alla sezione 6.3.
[21] Si tratta in sostanza dell'interpretazione della legge della moltiplicazione mediante il modello dell'albero.

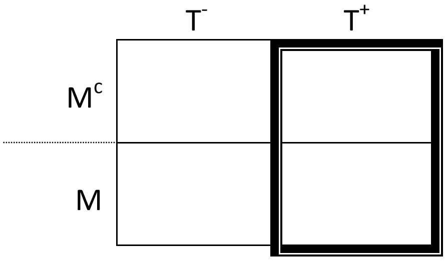
Risoluzione mediante tabella

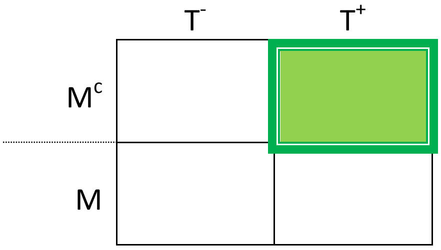
Possiamo, in alternativa, rappresentare la situazione mediante una **tabella**. L'insieme dei casi che si possono presentare[22] per l'individuo della popolazione (casi iniziali) può essere suddiviso nei quattro sottoinsiemi rappresentati dai quattro rettangoli più piccoli in figura. Ad esempio, il rettangolo che occupa la seconda riga e la prima colonna rappresenta l'insieme $\mathrm{M} \cap \mathrm{T}^{-}$, ossia l'evento *"l'individuo ha l'HIV e il suo test ha esito negativo"*.

Osserviamo che ciascuno dei quattro rettangoli corrisponde ad uno dei quattro possibili cammini nel grafo ad albero precedente.



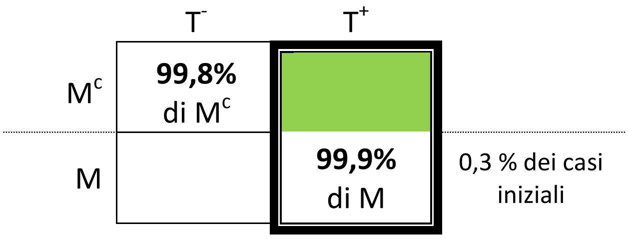
Sappiamo che è accaduto l'evento **T⁺**. Quindi l'insieme dei **nuovi** "**casi possibili**" è l'insieme **T⁺,** rappresentato con il bordo nero evidenziato nella figura seguente a sinistra.
Tenendo conto di ciò, si chiede di determinare la probabilità che si sia verificato l'evento M^c. Allora l'insieme dei "**casi favorevoli**", ossia $\boldsymbol{M^c} \cap \boldsymbol{T^+}$, è rappresentato dall'insieme colorato in verde.

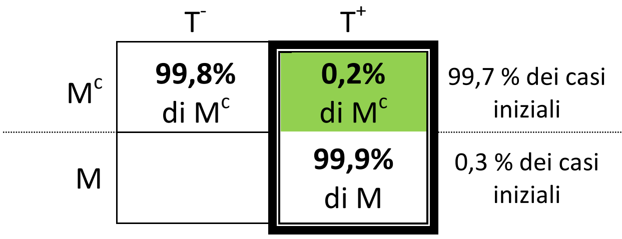




Prima di procedere con il calcolo, riportiamo su tale schema i valori di probabilità forniti in ipotesi, **prestando attenzione all'insieme** rispetto al quale essi sono espressi. In particolare l'ipotesi sulla prevalenza della malattia ci dice che l'insieme dei due rettangoli che costituisce la riga inferiore della tabella rappresenta lo 0,3% dei casi iniziali.

Nella figura posta a destra riportiamo anche il dato che manca per rispondere al quesito: la probabilità che *"l'individuo non malato risulti positivo al test[23] "*.





[22] Senza tener conto, per ora, del fatto che l'individuo è positivo al test.
[23] La somma delle probabilità relative ai due rettangoli della riga superiore deve essere 1, se esse sono riferite all'insieme M^c.

Ora disponiamo di tutti gli elementi per il calcolo della probabilità richiesta. **Riferendoci a tale rappresentazione**, abbiamo allora

$$p_{T^+}(M^c) = \frac{p(\boldsymbol{M^c} \cap \boldsymbol{T^+})}{p(\boldsymbol{T^+})}$$

Calcoliamo separatamente il valore del numeratore e quello del denominatore.

Vale[24]

$$p(\boldsymbol{M^c} \cap \boldsymbol{T^+}) = \frac{\mathbf{0,2}}{\mathbf{100}} \cdot \frac{\mathbf{99,7}}{\mathbf{100}} = \mathbf{0,997 \cdot 0,002}$$

Mentre[25]

$$p(\boldsymbol{T^+}) = \frac{\mathbf{0,2}}{\mathbf{100}} \cdot \frac{\mathbf{99,7}}{\mathbf{100}} + \frac{\mathbf{99,9}}{\mathbf{100}} \cdot \frac{\mathbf{0,3}}{\mathbf{100}} = \mathbf{0,997 \cdot 0,002 + 0,003 \cdot 0,999}$$

Sostituendo i valori così ottenuti, concludiamo che

$$\frac{p(\boldsymbol{M^c} \cap \boldsymbol{T^+})}{p(\boldsymbol{T^+})} = \frac{\mathbf{0,997 \cdot 0,002}}{\mathbf{0,997 \cdot 0,002 + 0,003 \cdot 0,999}} \approx 0{,}40$$

Abbiamo così ottenuto la stessa espressione numerica a cui eravamo giunti mediante la schematizzazione con il grafo ad albero.

<u>Osservazione</u>

Anche relativamente a questo approccio possiamo *ripercorrere* il procedimento risolutivo *mediante il solo linguaggio degli insiemi.* Facendo riferimento alla rappresentazione grafica, abbiamo suddiviso l'insieme dei T^+ nel modo seguente:

$$T^+ = (M^c \cap T^+) \cup (M \cap T^+)$$

E abbiamo così espresso la probabilità richiesta:

$$p_{T^+}(M^c) = \frac{p(\boldsymbol{M^c} \cap \boldsymbol{T^+})}{p(\boldsymbol{T^+})} = \frac{p(\boldsymbol{M^c} \cap \boldsymbol{T^+})}{p(\boldsymbol{M^c} \cap \boldsymbol{T^+}) + p(\boldsymbol{M} \cap \boldsymbol{T^+})} = \frac{p(\mathbf{M^c}) \cdot p_{M^c}(\boldsymbol{T^+})}{p(\mathbf{M^c}) \cdot p_{M^c}(\mathbf{T^+}) + p(\mathbf{M}) \cdot p_M(\boldsymbol{T^+})}$$

Anch'essa è identica all'espressione mediante il linguaggio degli insiemi che abbiamo ottenuto interpretando il procedimento sul grafo ad albero.

---

[24] Siamo interessati al valore di probabilità dell'insieme colorato in verde **rispetto all'insieme dei casi possibili iniziali.** Sappiamo che il valore $\mathbf{0,2}\%$ è la probabilità dell'insieme colorato in verde *rispetto all'insieme M^c,* non rispetto all'insieme richiesto. Come fare? Possiamo sfruttare l'informazione ulteriore che M^c rappresenta il $\mathbf{97}\%$ dei casi possibili iniziali. Pertanto la probabilità sarà ***0,2% del 97%*** ossia $\frac{\mathbf{0,2}}{\mathbf{100}} \cdot \frac{\mathbf{99,7}}{\mathbf{100}}$

[25] Osserviamo innanzitutto che conviene pensare l'insieme T^+ come *unione* dei due rettangoli che costituiscono la colonna destra in figura. La probabilità dell'insieme T^+ sarà allora la *somma* delle probabilità relative a questi due rettangoli.

Ora, la probabilità dell'insieme colorato in verde è già stata calcolata. Pertanto resta da determinare solo quella relativa al rettangolo in basso nelle figura. Analogamente a quanto osservato nella nota precedente, tale probabilità è

99,9% dello 0,3% ossia $\frac{\mathbf{99,9}}{\mathbf{100}} \cdot \frac{\mathbf{0,3}}{\mathbf{100}}$

c) Interpretazione e giustificazione del risultato. Che fine ha fatto la formula di Bayes?

Il problema che ci ha accompagnato fin dall'inizio del Percorso è stato finalmente risolto e ci ha premesso di sviluppare strumenti metodologici efficaci ed espressivi. Perché dunque non impiegarli subito per affrontare questioni analoghe?
In effetti, quella di fermarsi al risultato, è una forte tentazione, che serpeggia tra gli studenti … e tra gli stessi docenti. Ma tale scelta (è opportuno pensarla in questi termini, visto che in effetti di scelta si tratta, anche se più o meno consapevole) priva gli studenti della possibilità di sviluppare abilità fondamentali, quali la previsione e il controllo dei procedimenti e dei risultati. Non riflettere su tali aspetti può portare acriticamente gli studenti a conclusioni del tipo "L'aereo ha impiegato 24 ore per effettuare il decollo" o "Mario pesa … 1326 kg", candidamente fornite da alcuni studenti.
Di qui l'esigenza di attuare una didattica che preveda anche un esame critico dei risultati e dei procedimenti. Ecco alcuni spunti al riguardo.

1. **Interpretazione del risultato**
   Si è concluso che la probabilità di falso positivo su un test risultato positivo non è trascurabile, visto che vale $\mathrm{p}_{\mathrm{T}^+}(\mathrm{M}^{\mathrm{C}}) \approx 40\%$. Dato che c'è questo rischio, come si comunica tale risultato al paziente? Simmetricamente qual è la probabilità di avere un falso negativo sapendo che il test è negativo? È importante disporre di questa valutazione?

2. **Giustificazione del risultato**
   Il test "Elisa" ha sensibilità e specificità "grandi" (rispettivamente 99,9% e 99,8%). Perché allora non è "piccola" la probabilità che il test positivo fornisca informazioni errate (è circa il 40%)?

3. **Che fine ha fatto la formula di Bayes?**
   Il problema si può risolvere anche ricorrendo direttamente alla formula di Bayes. Del resto, tale risultato compare esplicitamente nelle Indicazioni nazionali per i licei scientifici. Perciò la questione merita di non venir liquidata mediante una risposta secca, anche se la nostra posizione al riguardo si può intuire da quanto finora espresso.

Di seguito proponiamo una possibile risposta a tali questioni.

1. Interpretazione del risultato

- Abbiamo visto che $p_{T^+}(M^c) \approx 40\%$. Questo significa che se il test "Elisa" è positivo vi è comunque una probabilità non nulla che l'individuo sia sano.
  Perciò, prima di comunicare l'esito del test al paziente, il dato viene **confermato mediante un altro test** chiamato "Western Blot".

- Invece qual è la probabilità di avere un **falso negativo sapendo che il test è negativo**? Cioè quanto vale $p_{T^-}(M)$?
  Procediamo in modo analogo al calcolo relativo ai falsi positivi, facendo riferimento al modello del grafo ad albero. Utilizzando i valori di probabilità che troviamo già indicati accanto ai rami[26], possiamo concludere che
  $$p_{T^-}(M) = \frac{0{,}003 \cdot 0{,}001}{0{,}997 \cdot 0{,}998 + 0{,}003 \cdot 0{,}001} \approx 3 \cdot 10^{-6}$$
  Un valore praticamente trascurabile. Se così non fosse si correrebbe il rischio, ben più grave di quello relativo ai falsi positivi, di indicare come sani dei pazienti in realtà malati.

[26] Naturalmente i cammini che ora rappresentano i casi possibili e quelli che rappresentano i casi favorevoli sono diversi da quelli relativi alla richiesta dei falsi positivi. Anzi, sono i loro "complementari".

2. Una giustificazione del risultato

Perché se il test ha sensibilità e specificità così “grandi”, non è “piccola” la probabilità che il test positivo sia errato (è circa il 40%)?

La motivazione si può articolare in tre passi:

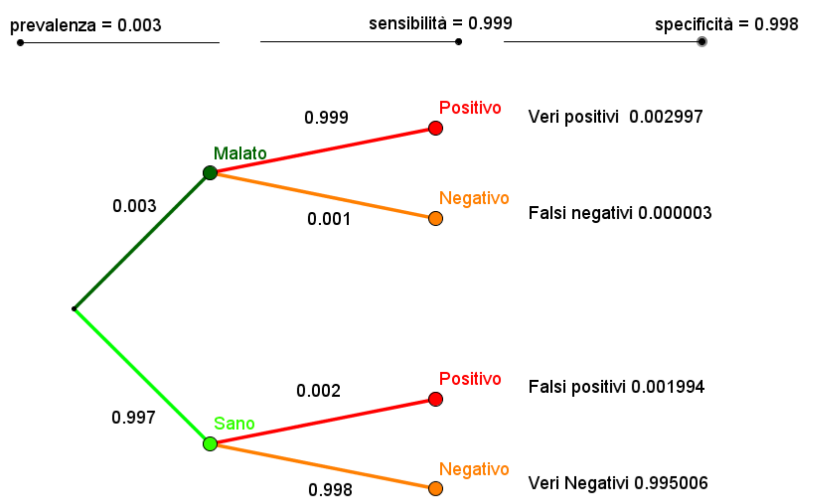
- la malattia ha **bassa prevalenza**, pertanto ci sono “molti” sani
- la probabilità di falso sui sani è “bassa” ma è applicata a “molti”
- quindi possono esserci “non pochi” falsi

Per fissare le idee, esaminiamo un esempio numerico. Supponiamo che la popolazione sia costituita da 1.000.000 di individui. E assumiamo, per semplicità, che i valori di probabilità in ipotesi (sensibilità e specificità) rappresentino delle frequenze relative[27].
In tal caso vi sono:

- 997.000 sani; tra essi i test **positivi** sono lo 0,2%, ossia **1994 falsi**
- 3.000 malati; tra essi i test **positivi** sono il 99,9%, ossia **2997 veri**

Si conclude così che, tra i test positivi (1994+2997), i falsi non sono “pochi”.

Più in generale, ha senso chiedersi come varia la probabilità richiesta al variare dei valori in ipotesi. Immergiamoci dunque in <u>un’attività esplorativa</u>, con il supporto del file Geogebra TestClinici_ARitroso.ggb[28].



	A	B	C
1		Positivi	Negativi
2	Malati	0.002997	0.000003
3	Sani	0.001994	0.995006
4		0.004991	0.995009
5			
6			
7	**P (sano)**	**tra i positivi**	**0.3995191344**
8			
9			
10			
11			
12			
13			
14			
15			

[27] Abbiamo già osservato in altri contesti che se la popolazione è “numerosa” ha senso approssimare i valori di probabilità con le frequenza relative.

[28] Il materiale è stato ideato dalla collega Francesca Arrigoni. Variando gli slider "prevalenza", "sensibilità" e "specificità", si può osservare sull'albero e sulla tabella come variano le cardinalità dei quattro insiemi in cui è stata partizionata la popolazione. Inoltre è evidenziata come varia la probabilità richiesta nel problema che stiamo esaminando.

3. Che fine ha fatto la formula di Bayes?

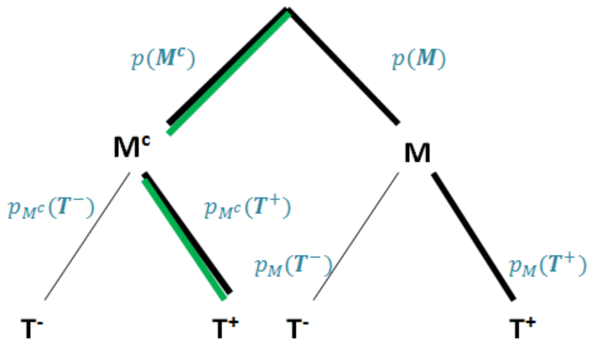
Il problema in esame si può risolvere anche ricorrendo direttamente la **formula di Bayes**[29]:

$$p\,(M^c|\,T^+) = \frac{p(M^c)\cdot p(T^+|M^c)}{p(M^c)\cdot p(T^+|M^c) + p(M)\cdot p(T^+|M)}$$

Sostituendo alle probabilità indicate i corrispondenti valori numerici, si ottiene immediatamente

$$p\,(M^c|\,T^+) = \frac{0{,}997\cdot\ 0{,}002}{0{,}997\cdot 0{,}002 + 0{,}003\cdot 0{,}999}$$

Ma questa è proprio l'espressione numerica a cui eravamo giunti seguendo il nostro approccio di tipo grafico! Anzi, a ben guardare, il nostro procedimento permette addirittura di ricavare la stessa formula di Bayes. Basta infatti considerare il modello ad albero e indicare le probabilità relative ai rami mediante le notazioni formali, come nella figura seguente:



Non resta dunque che ripercorrere sul grafo il nostro procedimento risolutivo, per ottenere subito la formula di Bayes.

Già dalle considerazioni fin qui esposte nella sezione dovrebbe trapelare la nostra posizione sull'utilizzo didattico della formula di Bayes. In realtà la questione non è di poco conto, visto che il termine compare esplicitamente nelle Indicazioni nazionali per il secondo biennio dei licei[30]. Vediamo perciò di precisarla.
Cominciamo con l'osservare che sia la formula di Bayes che il nostro procedimento risolutivo concorrono allo stesso scopo pratico: fornire la probabilità (a posteriori) delle "cause"[31], a partire dalle probabilità degli "effetti".
Ma se si opta per un'applicazione **diretta** (per non dire, brutale) della formula di Bayes, si rischia di ridurre il procedimento risolutivo alla pura manipolazione algebrica di simboli. La stessa dimostrazione della formula è presentata su vari testi in adozione nella scuola secondaria, come semplice manipolazione algebrica[32].
Invece, l'utilizzo di schemi grafici simili a quelli mostrati, è ricco di valenze didattiche: permette di cogliere in profondità il significato di ogni passo del procedimento risolutivo, di controllarlo, nonché di **ricostruirlo volta per volta**, anche a lungo termine.

[29] La formula è volutamente espressa nella notazione adottata da molti libri di testo.
[30] "[…] Studierà la probabilità condizionata e composta, la formula di Bayes e le sue applicazioni, nonché gli elementi di base del calcolo combinatorio. […]"
[31] Per essere precisi, non si dovrebbe parlare di "causa" ma di evento condizionante. Esso infatti non influenza sempre l'evento condizionato, come abbiamo avuto modo di chiarire a più riprese. Comunque il ricorso a tale terminologia può servire a facilitare l'intuizione, purché sia attuato nella consapevolezza dei limiti indicati.
[32] Un approccio formale ha senso nell'ambito di una trattazione assiomatica della probabilità. Ma è opportuno che lo studente di scuola secondaria vi arrivi gradualmente: così può contare su una base di concetti a partire dai quali costruire i significati degli oggetti matematici che incontra.

6.6 Letture e attività

La probabilità condizionata e il suo impiego congiunto con i grafi ad albero hanno permesso di esaminare efficacemente un problema significativo riguardante i test clinici.
Perché non sfruttare il potente approccio discusso nella sezione precedente, per affrontare questioni analoghe, ossia determinare la probabilità (a posteriori) delle "cause" a partire dalle probabilità degli "effetti"? Potremmo riprendere in considerazione alcune domande-stimolo poste all'inizio del Percorso ed esaminarle in modo razionale mediante gli strumenti matematici così introdotti, che consentono in particolare di esprimere sinteticamente le richieste attraverso specifici valori di probabilità.
Tuttavia, presteremo attenzione a non assolutizzare i risultati ottenuti. In particolare eviteremo di confondere il valore di probabilità a cui porta il procedimento, con un valore di verità. Infatti il modello probabilistico si limita a fornire un numero; il significato che tale numero assume resta una nostra scelta e non esprime una presunta verità oggettiva, come abbiamo più volte osservato.

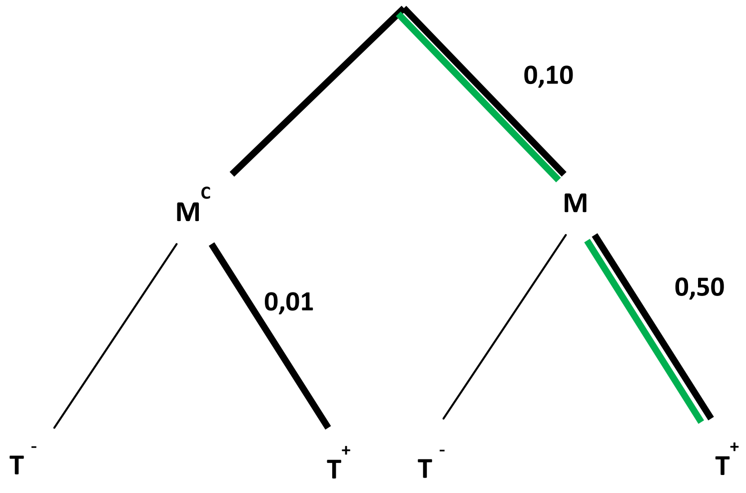
Test antidoping

Pensiamo sia più chiaro riformulare la questione nei seguenti termini.

Qual è la probabilità che l'atleta positivo al test antidoping sia effettivamente dopato?
Assumi che la probabilità di risultare positivo per il non dopato sia dell'1%, quella di essere positivo per il dopato sia del 50%, e che i dopati siano il 10% degli atleti.

Come si vede, il problema è analogo a quello esaminato nella sezione 5 sui test clinici. Perciò ha senso utilizzare addirittura le stesse notazioni, con l'avvertenza di intendere ora con M l'evento "l'individuo è dopato".

È analoga anche la modellizzazione della situazione mediante il grafo ad albero:



Dunque, per ottenere la probabilità richiesta non resta che interpretare adeguatamente la figura e leggere il processo di calcolo direttamente sul grafo:

$$p_{T^+}(M) = \frac{p(M \cap T^+)}{p(T^+)} = \frac{0{,}1 \cdot 0{,}5}{0{,}9 \cdot 0{,}01 + 0{,}1 \cdot 0{,}5} \approx 0{,}847$$

Per facilitare la comprensione della questione, come già suggerito nel capitolo 5 a proposito dei test clinici, può essere didatticamente significativo iniziare con una versione semplificata del problema. Partire cioè da una popolazione di 1.000 atleti e assumere che le probabilità in ipotesi rappresentino delle frequenze relative. Ciò significa ad esempio che, tra gli atleti non dopati, esattamente l'1% è positivo al test e che i dopati sono 100. Il passo successivo consiste nello stabilire il numero di elementi di ciascuno dei 4 sottoinsiemi in cui si può suddividere la popolazione in base agli esiti del test e all'aver o meno fatto uso di doping: dopati positivi, dopati negati, non dopati positivi, non dopati negativi. In questa fase può essere utile anche rappresentare la situazione mediante una tabella a doppia entrata, le cui due voci sono appunto "dopato (si/no)", "esito del test".
A questo punto dovrebbe risultare agevole individuare la consistenza numerica dell'insieme dei casi favorevoli e di quello dei casi possibili, e concludere così con il valore di probabilità richiesto.

Processo ad O.J. Simpson

Già si è visto nel primo capitolo come si possono condensare nel linguaggio naturale le posizioni opposte sull'intera questione.

a) Difesa: tra le donne percosse dal compagno, solo lo 0,04% è uccisa da lui

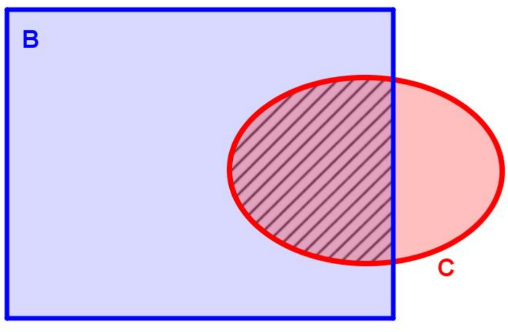
b) Studi successivi: tra le donne percosse dal compagno **e uccise**, il 90% è uccisa da lui

Ora si può richiedere agli studenti di esaminare razionalmente la situazione, mediante i nuovi strumenti matematici introdotti. La richiesta può essere formulata nei termini seguenti

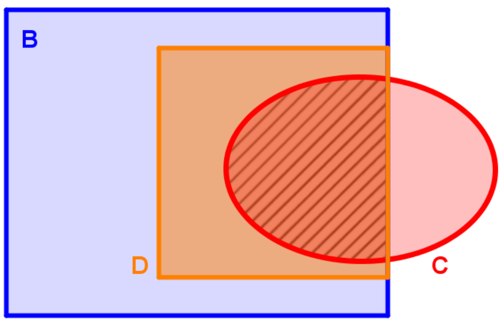
- Rappresenta con diagrammi di Venn le due situazioni a), b) ora descritte.
- Esprimi ciascuna situazione mediante la probabilità condizionata .
- Quale tra le due valutazioni di probabilità ti sembra più adeguata? Argomenta in dettaglio il tuo punto di vista.

Ecco le due rappresentazioni richieste.

a)



b)



$B = \{picchiate\ dal\ compagno\}$

$C = \{uccise\ dal\ compagno\}$

$D = \{picchiate\ dal\ compagno\ e\ uccise\}$

Utilizzando le notazioni della probabilità condizionata, le due situazioni si esprimono in modo sintetico ed espressivo, rispettivamente come di seguito:

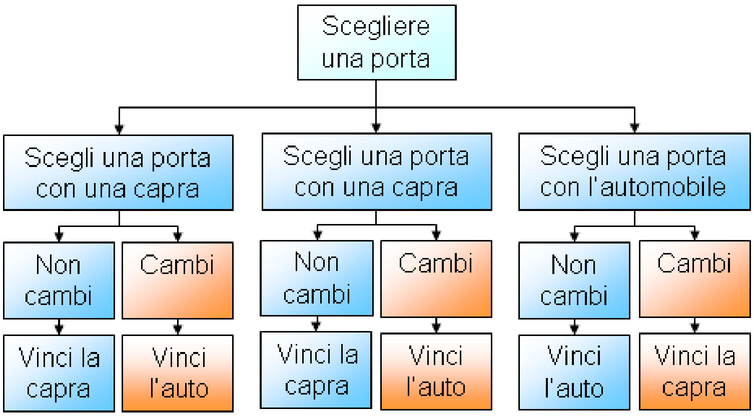
$$p_B(C) = 0{,}04\% \qquad p_D(C) = 90\%$$

Filtri anti-spam

Ora gli studenti dispongono degli strumenti per comprendere ed interpretare, eventualmente in modo autonomo, la lettura proposta all'inizio del Percorso.

Monty Hall

Un problema diventato classico è quello legato al gioco "televisivo" Monty Hall, in cui il giocatore vince un'automobile se indovina dietro quale delle tre porte essa si nasconde; dietro ciascuna delle altre due invece c'è una capra. Più precisamente, il concorrente sceglie una porta e il presentatore apre una delle due rimanenti, dietro la quale sa che si trova certamente una capra. A questo punto il giocatore deve decidere se confermare la sua scelta oppure modificarla. La questione è: cosa gli conviene fare? Ossia è meglio confermare la scelta oppure cambiarla?
La questione è diventata un caso che ha coinvolto molti matematici e diversi sono state gli approcci risolutivi proposti nel tempo.
Uno che ci sembra particolarmente efficace è quello che si basa sulla rappresentazione dei casi possibili mediante un diagramma analogo al seguente[33]:



In definitiva, se il giocatore cambia porta vince 2 volte su 3, altrimenti vince solamente 1 volta su 3, dunque conviene cambiare.

[33] Da https://it.wikipedia.org/wiki/Problema_di_Monty_Hall

APPENDICE

A1 Attività che preparano o consolidano l'uguaglianza $\mathbf{p_B(A)} = \frac{\mathbf{p(A \cap B)}}{\mathbf{p(B)}}$ (*)

Vediamo alcune attività che permettono di comprendere meglio il significato della formula in esame. Esse possono essere affrontate in vari momenti del percorso che precedono l'introduzione della probabilità condizionata, oppure come consolidamento.

Probabilità come misura

- Consideriamo, ad esempio, il problema discusso in dettaglio nel capitolo 2:

  determinare il punteggio più probabile nel lancio di due dadi[34]

  Come ampiamente discusso, possiamo modellizzare la situazione mediante una tabella a doppia entrata.
  In tale contesto, abbiamo ottenuto la probabilità richiesta come rapporto tra il numero dei casi favorevoli all'evento "il punteggio è 7" e il numero dei casi che si possono presentare: $\frac{6}{36}$ ossia $\frac{1}{6}$. In sostanza, abbiamo contato i quadratini verdi e i quadratini bianchi.

dado blu

6	7	8	9	10	11	12
5	6	7	8	9	10	11
4	5	6	7	8	9	10
3	4	5	6	7	8	9
2	3	4	5	6	7	8
1	2	3	4	5	6	7
"+"	1	2	3	4	5	6

dado nero

  Ora osserviamo che la **probabilità** richiesta si può interpretare anche in un modo diverso e che risulterà utile anche in altri contesti: la probabilità è l'**area dell'insieme che rappresenta l'evento** "esce il 7", rispetto al quadrato in cui è contenuto[35].

- Un altro esempio, anch'esso discusso nel capitolo 2:

  determinare la probabilità che il punteggio sia pari

  Come già discusso, possiamo modellizzare la situazione mediante la tabella "contratta" che rappresentiamo a lato, in cui sono raggruppati tra loro i primi tre numeri pari e i primi tre numeri dispari.
  Analogamente all'esempio precedente, la **probabilità** richiesta è uguale all'**area dell'insieme verde**, rispetto al quadrato in cui tale insieme è contenuto. Pertanto, per valutarla non serve contare le celle verdi e bianche, ma basta osservare che le celle verdi occupano esattamente la metà dell'intero quadrato che rappresenta la tabella. Dunque essa vale ½.

dado blu

d		
p		
"+"	p	d

dado nero

[34] Somma dei due numeri usciti su ciascun dado.
[35] Ossia la probabilità è l'area dell'insieme verde, misurata assumendo come **unità di misura** il quadrato che rappresenta l'intera tabella.

In definitiva, in entrambi gli esempi esaminati, la probabilità richiesta è la **misura**[36] di un insieme, rispetto ad un insieme opportuno che lo contiene; più precisamente, rispetto a quello in cui si decide siano contenuti gli eventi in esame.

- Vediamo ancora un ultimo esempio. In diverse situazioni si richiede di

  valutare la ***probabilità*** *come rapporto tra il* ***numero di elementi*** *di un insieme e quello dell'insieme che lo contiene.*

  In tali contesti, perché non cogliere l'occasione di notare che il numero di elementi di un insieme finito è una **misura** di tale insieme? Così come è una misura il numero di elementi dell'insieme rispetto al numero di elementi di un opportuno insieme che lo contiene, ossia è una misura il rapporto tra questi due numeri.
  Quello che a noi interessa è che, ancora una volta, possiamo interpretare la probabilità come misura di un insieme.

L'uguaglianza (*)

a) *Esaminiamo prima la situazione in cui A e B, insieme non vuoto, sono costituiti da un numero finito di elementi*[37]*. E supponiamo che l'insieme U abbia n elementi (i casi possibili iniziali), dove n è un numero naturale diverso da 0.*

   Per il significato attribuito a $\mathrm{p_B(A)}$ come la probabilità che si verifichi l'evento A qualora si sappia che si è verificato B, si ha che l'insieme dei nuovi "casi possibili" **=** B e l'insieme dei "casi favorevoli" = A∩B.
   Pertanto, valutando tale probabilità mediante l'approccio classico, vale[38]

$$\boldsymbol{p_B(A)} = \frac{\#(A \cap B)}{\#B}$$

   Ora, dividendo numeratore e denominatore per *n*, possiamo dedurre che

$$\boldsymbol{p_B(A)} = \frac{\#(A \cap B)}{\#B} = \frac{\frac{\#(A \cap B)}{n}}{\frac{\#B}{n}} = \frac{\boldsymbol{p(A \cap B)}}{\boldsymbol{p(B)}}$$

   È importante che lo studente intenda tale procedimento non come una pura manipolazione algebrica. Ma, interpretando la probabilità come misura di un insieme, riesca a leggere il passaggio intermedio come un cambio della misura con cui si valutano gli eventi $A \cap B$ e B.

b) *L'uguaglianza in esame (*) continua comunque a valere anche se A e B non hanno un numero finito di elementi*[39]*.* Basta ragionare in modo esattamente analogo al punto a), considerando non più il *numero* di elementi di ogni insieme, ma un'opportuna **misura dell'insieme**[40]*.*

---

[36] Facciamo qui riferimento all'idea **intuitiva** che lo studente può avere di misura di un insieme. Eventualmente possiamo precisarla un po', rileggendo le situazioni note costituite dalla lunghezza dei segmenti e dall'area di figure piane: la lunghezza di un segmento o l'area di una figura piana sono numeri (positivi) che si associano a tali insiemi; tale corrispondenza deve verificare opportune condizioni, la cui precisazione non rientra però negli intenti di queste note.
[37] Esempi specifici di questo tipo possono essere proposti allo studente *in più occasioni* nel percorso prima di affrontare la probabilità condizionata oppure dopo aver visto la formula (*).
[38] Dato un insieme E che ha un numero finito di elementi, indichiamo con il simbolo #E il numero di elementi dell'insieme E.
[39] Come, ad esempio, nel caso in cui rappresentano i punti di un cerchio o di un cubo.
[40] Ad esempio **una** misura del cerchio di raggio r è la sua area πr^2. Una misura di un cubo di lato l nello spazio è il suo volume l^3.