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.

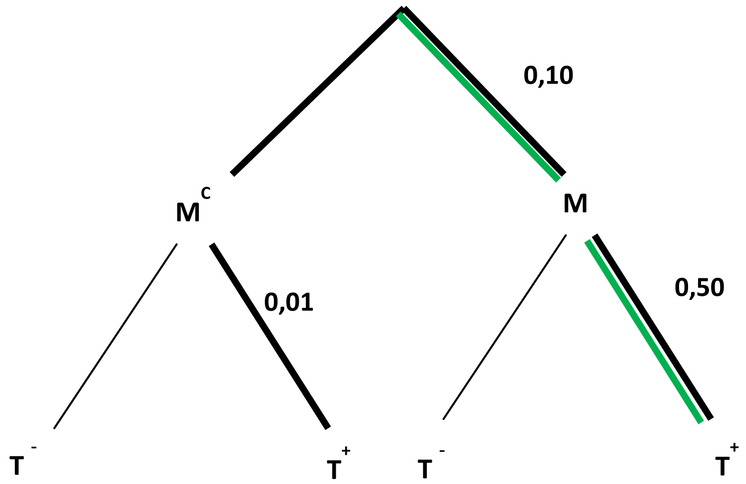
Test antidoping

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

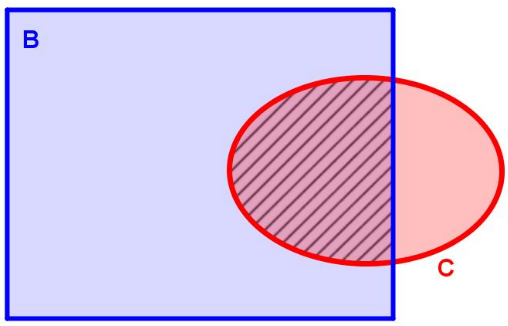
b) Studi successivi: tra le donne percosse dal compagno **e uccise**, il 90% è uccisa da lui

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

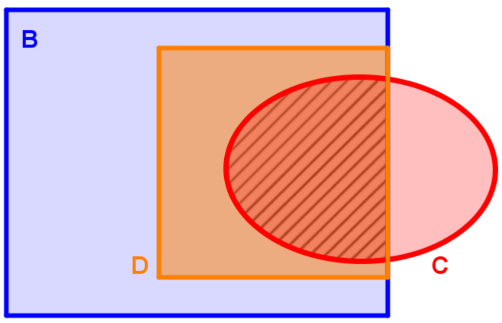
- Rappresenta con diagrammi di Venn le due situazioni a), b) ora descritte.
- Esprimi ciascuna situazione mediante la probabilità condizionata .
- Quale tra le due valutazioni di probabilità ti sembra più adeguata? Argomenta in dettaglio il tuo punto di vista.

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:

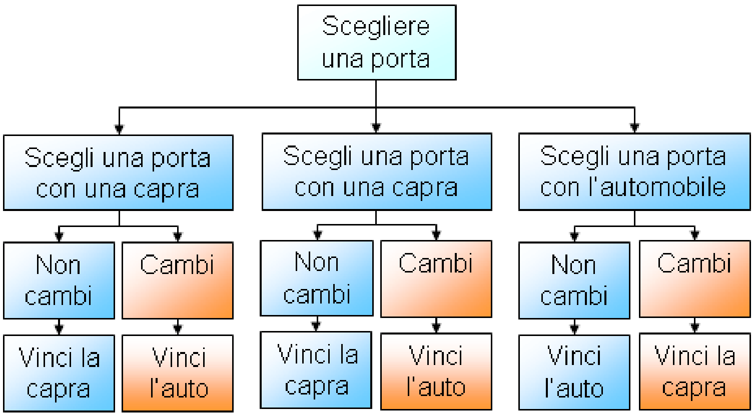
$$p_B(C) = 0{,}04\% \qquad p_D(C) = 90\%$$

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