Test di ingresso sull'intuizione probabilistica (Prodi)

1. Da molti anni il numero 57 non esce sulla ruota del lotto di Cagliari.

Conviene giocarlo? Conviene escluderlo? E' indifferente?

2. Lanciando due monete qual è la probabilità di ottenere:

due volte “testa”; due volte "croce"; una volta "testa" ed una volta "croce"?

3. Due coniugi vogliono ad ogni costo un figlio maschio; se nasce un maschio, smettono, se no ritentano, fino ad un massimo di 6 figli. In media, avranno più maschi o più femmine?

4. Si estraggono le palline della tombola: quale dei due seguenti risultati è più probabile?

1, 2, 3 oppure 9, 84, 37.

5. In ogni clinica di maternità si prende nota dei giorni in cui nasce più del 60% di maschi oppure più del 60% di femmine. In un anno, il numero di questi giorni sarà maggiore:

nelle cliniche più grandi; nelle cliniche medie; nelle cliniche piccole?

L'obiettivo del test è quello di mostrare che, nel calcolo delle probabilità, ci sono casi in cui l'intuizione fornisce la risposta errata.

Prima di cominciare ad affrontare l'argomento, è possibile che tu fornisca le seguenti risposte:

1. conviene giocarlo;
2. 1/3, 1/3, 1/3;
3. più maschi;
4. 9, 84, 37;
5. nelle cliniche più grandi.

Quando avrai terminato lo studio, dovresti capire o verificare che le risposte corrette sono:

1. è indifferente;
2. 1/4, 1/4, 1/2;
3. più femmine;
4. sono ugualmente probabili;
5. nelle cliniche più piccole.

Nel <u>Calcolo della Probabilità</u> ci troviamo in una situazione apparentemente paradossale:

- ci sono delle <u>regole di calcolo</u> che tutti accettano ed utilizzano, in qualunque ambito della scienza;
- ci dovrebbero essere delle <u>definizioni</u> o delle <u>interpretazioni</u> che ci dicano "cosa è la probabilità".

Sull'ultimo punto, però, ci sono forti discussioni. Come afferma uno studioso:

"As to what probability is and how it is connected with statistics, there has seldom been such complete disagreement and breakdown of communication since the Tower of Babel".

Di conseguenza, precisiamo che, mentre le formule che scriveremo sono valide sempre e per tutti, non tutti concordano sull'interpretazione che ne forniremo, che è motivata in buona parte da ragioni didattiche.

Spesso dobbiamo prendere in considerazione una certa affermazione sulla quale abbiamo solo delle informazioni parziali ed incomplete.

Esempi tipici sono quelli associati ai giochi (lanci di dadi, monete, estrarre delle carte da un mazzo).

Alcuni esempi più “concreti” sono:

- “oggi pioverà a Grosseto”, sapendo che “oggi è nuvoloso”;
- “la squadra X vincerà il prossimo campionato di calcio”, conoscendo gli organici delle varie società;
- “la teoria scientifica X deve essere considerata accettabile”, sapendo che ci sono degli esperimenti di laboratorio e dei ragionamenti teorici che la supportano
- un poliziotto cammina per una strada deserta quando sente suonare un allarme; vede una gioielleria con la vetrina in frantumi ed un uomo mascherato che ne asce con una borsa piena di gioielli; il poliziotto decide che si tratta di un ladro.

In casi come i precedenti, <u>non abbiamo sufficienti informazioni per concludere con certezza che la nostra affermazione sia vera o falsa</u>.

Anche nell'ultimo esempio, l'uomo che il poliziotto ha visto potrebbe essere il proprietario della gioielleria ,che stava tornando da una festa in costume, ha visto che una pietra scagliata da un'auto in corsa ha infranto la vetrina, ed ha deciso di mettere in salvo i suoi beni...

In questi casi, vogliamo introdurre una grandezza, che chiamiamo <u>probabilità</u>, la quale misura il <u>grado di fiducia</u> (“degree of belief”) che noi attribuiamo alla validità della nostra affermazione, sulla quale abbiamo a disposizione solo delle informazioni parziali ed incomplete, che in genere non sono sufficienti a concludere se essa è vera o falsa.

Più frequentemente, si parla di probabilità che un determinato <u>evento</u> si realizzi, ovvero che la nostra affermazione sia vera, e che quindi “un determinato fatto avvenga”.

Come casi "estremi", stabiliamo che:

- se riteniamo che un evento sia <u>certo</u>, gli attribuiamo probabilità uguale ad 1;
- se riteniamo che un evento sia <u>impossibile</u>, gli attribuiamo probabilità uguale a 0.

Come possiamo assegnare un valore numerico alla probabilità di un determinato evento nel caso generale in cui questo non sia né certo, né impossibile?

Consideriamo intanto il caso più semplice, e, in pratica, l'unico che ti interessa nello svolgimento di esercizi a livello liceale.

<u>Definizione classica di probabilità</u>.

La probabilità p di un certo evento E è definita come il <u>rapporto tra il numero f di casi favorevoli (ovvero in cui l'evento E viene realizzato) ed il numero totale u di casi possibili</u>, <u>purché essi siano tutti ugualmente possibili</u>:

$$p(E) = \frac{n^\circ \text{casi favorevoli}}{n^\circ \text{casi possibili}} = \frac{f}{u} \ .$$

Esempi

- Lanciando una moneta abbiamo due esiti possibili (testa e croce), per cui le probabilità di ottenere ciascuno di essi sono: $p(T)=p(C)=1/2$.
- Lanciando un dado abbiamo sei esiti possibili, ciascuno dei quali ha probabilità 1/6.

  All'evento $A=\text{"esce un multiplo di 3"}$ corrispondono due esiti (3 e 6), e quindi la sua probabilità è:

  $$p(A)=\frac{2}{6}=\frac{1}{3} .$$
- Estraendo una carta da un mazzo completo, abbiamo 52 esiti possibili. Definiamo gli eventi:

  $A=\text{"esce una figura"}$, $B=\text{"esce una carta di fiori"}$, $C=\text{"esce ua figura di fiori"}$.

  In base alla def. classica, essi hanno probabilità:

  $$p(A)=\frac{12}{52}=\frac{3}{13} \;,\; p(B)=\frac{13}{52}=\frac{1}{4} \;,\; p(C)=\frac{3}{52} .$$

Forse ti sarai reso conto che quella precedente non è una definizione rigorosa.

Affermando che i vari casi debbano essere “ugualmente possibili”, stiamo implicitamente utilizzando proprio quel concetto di probabilità che cercavamo di definire.

D'altra parte, non possiamo eliminare questa precisazione, perché, se ci chiediamo qual è la probabilità che un determinato atleta vinca una gara o che un intervento chirurgico abbia esito positivo, non possiamo “pesare” tutti i possibili esiti allo stesso modo.

Potremmo allora dire che, se un fenomeno può avere diversi esiti, e non abbiamo nessuna informazione che ci permetta di “privilegiare” uno di questi esiti rispetto agli altri, allora dovremo attribuire la stessa probabilità a ciascuno degli esiti.

Quindi, nel lancio di una moneta o di un dado, o nell'estrazione di una carta dal mazzo, è ragionevole applicare la definizione classica di probabilità se non abbiamo informazioni contrarie, che ci dicano che la moneta o il dado o il mazzo di carte sono in qualche modo “truccati”.

Se la condizione precedente non è verificata, ovvero se i diversi esiti non sono "ugualmente possibili", come possiamo attribuire un valore di probabilità ad un determinato evento?

Risposta breve.

Non te ne preoccupare, perché in questo caso tale valore sarà assegnato dal testo del problema.

Ad esempio, ci sarà scritto "la probabilità che una moneta truccata dia croce è $p(C)=0{,}7$ ", oppure "la probabilità che un certo atleta vinca una gara e è $p(V)=0{,}45$ ".

Risposta lunga. Dipende.

Diciamo che ci stiamo occupando di un fenomeno che può essere ripetuto in maniera indefinita, e che avviene sempre in condizioni più o meno simili, e che possiamo controllare.

Questo avviene appunto per il lancio di una moneta o di un dado, per l'estrazione di una carta da un mazzo, per la produzione di un certo prodotto in una fabbrica.

In questo caso, contiamo quante volte avviene l'evento che ci interessa (frequenza assoluta), calcoliamo il rapporto tra la frequenza assoluta del nostro evento ed il numero totale di prove compiute (frequenza relativa), e identifichiamo la probabilità con la frequenza relativa calcolata per un numero “molto grande” di prove.

Questa, espressa meglio, viene chiamata <u>definizione frequentista della probabilità</u>.

Da questo punto di vista, dire che la probabilità che, lanciando una determinata moneta, esca croce è $p(C)=0{,}7$, significa che, se la lanciamo “molte volte”, è “molto probabile” che “circa” il 70% delle volte esca croce (tutte le parole tra virgolette andrebbero definite meglio).

Supponiamo, invece, di studiare un <u>fenomeno che avviene un'unica volta, o che, anche se può essere ripetuto, lo è in condizioni ogni volta diverse e per noi non controllabili</u>.

Tipici esempi possono essere un intervento chirurgico di cui ci interessa l'esito, una gara di atletica, un esame universitario.

Consideriamo in particolare l'esito di una partita di calcio. In questo caso, la definizione classica ci direbbe che ciascuna squadra ha una probabilità del 50% di vittoria, il che ha poco senso, a meno che si tratti di due formazioni per noi completamente sconosciute. La definizione frequentista può dirci che, ad esempio, negli ultimi dieci anni la squadra A ha vinto il 70% delle gare disputate contro la squadra B. Se non abbiamo altre informazioni, possiamo assumere questo numero come stima della probabilità di vittoria di A. Se, però, siamo degli appassionati, e sappiamo che la squadra A ha uno dei suoi migliori giocatori infortunato, e che la squadra B ha acquistato un forte attaccante, dovremo tenere conto di queste nuove informazioni per aggiornare le nostre stime sulla probabilità di vittoria delle due squadre.

Torniamo quindi alla nostra definizione iniziale di probabilità (detta definizione soggettivista di probabilità) come "misura del grado di fiducia che noi attribuiamo, sulla base delle nostre informazioni, al fatto che una certa affermazione sia vera".

Per ottenere un valore numerico, il "grado di fiducia" viene espresso chiedendoci quale sarebbe una scommessa "equa" sull'esito della partita (ma è inutile fornire troppi particolari).

Quando utilizziamo questo punto di vista, stiamo affermando che la probabilità non si trova "negli oggetti", ma "nella nostra mente", e, soprattutto, che la probabilità che attribuiamo al verificarsi di un determinato evento cambia se cambiano le nostre informazioni su di esso.

Probabilità totale (regola della somma)

Supponiamo di conoscere le probabilità che si realizzino due eventi A e B (ovvero che siano vere le affermazioni “avviene A”, “avviene B”).

Ci chiediamo quale probabilità dobbiamo attribuire all'evento $A \vee B$, ovvero all'affermazione “avviene A oppure B” (disgiunzione logica delle due affermazioni precedenti).

Consideriamo prima il caso particolare in cui A e B siano proposizioni incompatibili, ovvero tali che non possano essere contemporaneamente vere.

In questo caso, il numero dei casi in cui si realizza la proposizione $A \vee B$ non è altro che la somma di quelli in cui si realizza la proposizione A e di quelli in cui si realizza la proposizione B, per cui:

$$p(A \vee B) = p(A) + p(B) \ .$$

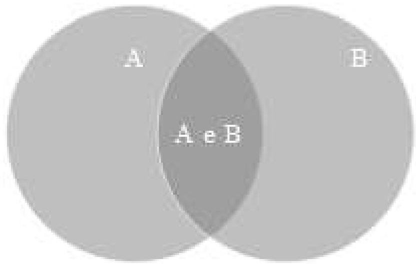
Esempio: in un mazzo di 52 carte abbiamo 12 figure e 4 assi, per cui:

$$p(\text{figura} \vee \text{asso}) = p(\text{figura}) + p(\text{asso}) = \frac{12}{52} + \frac{4}{52} = \frac{16}{52} = \frac{4}{13} \ .$$

In generale, però, le proposizioni A e B saranno <u>compatibili</u>, ovvero ci saranno dei casi in cui esse sono contemporaneamente vere.

Se utilizzassimo la formula precedente, tali casi verrebbero contati due volte; per ottenere la formula corretta, dobbiamo quindi sottrarli:

$$p(A \vee B) = p(A) + p(B) - p(A \wedge B) \quad .$$



In altri termini, stiamo affermando che il numero degli elementi dell'unione di due insiemi A e B è uguale alla somma tra il numero degli elementi di A e quello di B, diminuita del numero di elementi dell'intersezione tra A e B.

Quando gli insiemi A e B sono disgiunti, torniamo al caso precedente delle affermazioni incompatibili.

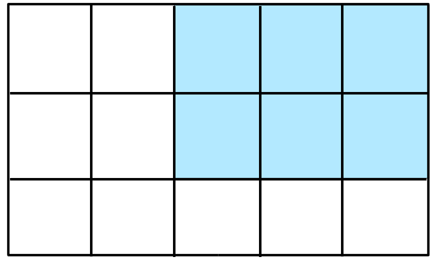
Esempio: in un mazzo di 52 carte abbiamo 13 carte di fiori, 12 figure e 3 figure di fiori:

$$p(\text{figura} \vee \text{fiori}) = p(\text{figura}) + p(\text{fiori}) - p(\text{figura} \wedge \text{fiori}) = \frac{12}{52} + \frac{13}{52} - \frac{3}{52} = \frac{22}{52} = \frac{11}{26} \quad .$$

Probabilità composta (regola del prodotto)

Ci chiediamo quale probabilità attribuire all'evento $A \wedge B$, ovvero all'affermazione “avvengono A e B” (congiunzione logica delle due affermazioni).

Consideriamo prima il caso particolare in cui A e B siano proposizioni indipendenti, ovvero tali che il fatto che una di esse sia vera non influenzi in alcun modo la verità dell'altra.



Aiutiamoci con un esempio concreto: una barretta di cioccolata è divisa in 15 quadretti, e 6 di questi rappresentano i nostri “casi favorevoli”.

La probabilità di scegliere a caso un quadretto dalla zona evidenziata è 6/15.

Potremmo però anche ragionare così: abbiamo una probabilità di 3/5 di scegliere una colonna “favorevole” ed una probabilità di 2/5 di scegliere una riga “favorevole, da cui:

$$p(\text{fav}) = p(\text{riga fav} \wedge \text{colonna fav}) = p(\text{riga fav}) \cdot p(\text{colonna fav}) = \frac{2}{3} \cdot \frac{3}{5} = \frac{6}{15} \quad .$$

Generalizzando l'esempio, possiamo dedurre che, se A e B sono due proposizioni indipendenti, allora:

$$p(A \wedge B) = p(A) \cdot p(B) \quad .$$

Consideriamo ora il caso generale di due proposizioni dipendenti l'una dall'altra, ovvero tali che il fatto che si verifichi una di esse influenza la probabilità che si verifichi l'altra.

Esempio: abbiamo una scatola che contiene 5 palline rosse e 3 palline nere, identiche a parte il colore. Estraiamo dalla scatola prima una pallina e poi un'altra, in maniera casuale, e ci chiediamo quale sia la probabilità di avere estratto due palline rosse.

La probabilità di estrarre una pallina rossa nella prima estrazione è $p(R_1)=5/8$.

La probabilità di estrarre una pallina rossa nella seconda estrazione, però, non è la stessa, ma dipende dall'esito della prima estrazione. In particolare, se nella prima estrazione è uscita una pallina rossa, sono rimaste nella scatola 4 palline rosse e 3 nere, per cui la probabilità di ottenere una pallina rossa nella seconda estrazione sapendo che nella prima è uscita una pallina rossa è $p(R_2/R_1)=4/7$.

In definitiva, abbiamo: $p(R_1 \wedge R_2)=p(R_1)\cdot p(R_2/R_1)=\frac{5}{8}\cdot\frac{4}{7}=\frac{20}{56}$.

Generalizzando l'esempio precedente, definiamo la <u>probabilità condizionata</u> $p(A/B)$ come la probabilità che si verifichi l'evento A sapendo che si è verificato l'evento B.

Quindi, in generale, la probabilità della congiunzione di due eventi A e B è uguale:

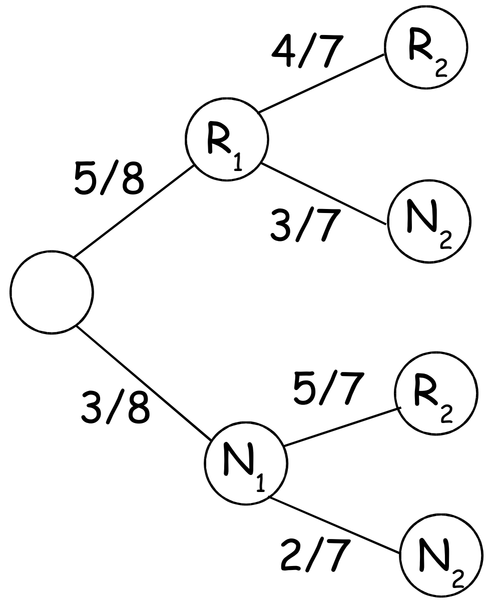
- al prodotto tra la probabilità di A e la probabilità di B condizionata al verificarsi di A,
- oppure al prodotto tra la probabilità di B e la probabilità di A condizionata al verificarsi di B:

$$p(A \wedge B) = p(A) \cdot p(B/A) = p(B) \cdot p(A/B) \ .$$

Nel caso in cui gli eventi A e B siano indipendenti, abbiamo $p(A/B) = p(A)$ e $p(B/A) = p(B)$, per cui ritroviamo il precedente caso particolare.

Spesso un problema di una certa complessità può essere analizzato rappresentandolo tramite un diagramma ad albero, in cui i vari “rami” uscenti da un certo “nodo” rappresentano delle alternative, che vengono pesate con le relative probabilità condizionate.

Diagrammi ad albero 1



Riprendiamo la scatola che contiene 5 palline rosse e tre palline nere e tracciamo il diagramma che rappresenta due estrazioni successive.

Osserva che dopo ogni estrazione <u>non</u> rimettiamo la pallina estratta nella scatola, ovvero, come dicono i matematici, le estrazioni sono <u>senza reimbussolamento</u>; in questo modo, ogni estrazione rappresenta un evento <u>dipendente</u> dalle precedenti estrazioni.

Osserva inoltre che, se l'albero è rappresentato in maniera completa, la somma dei “pesi” (probabilità condizionate) dei rami uscenti da uno stesso nodo deve sempre essere uguale ad 1.

Dalle regole della probabilità totale e composta segue che:

- per calcolare la probabilità di giungere ad un determinato nodo terminale, o “foglia”, dobbiamo partire dalla radice e moltiplicare le probabilità di tutti i rami intermedi;
- per calcolare la probabilità di ottenere un risultato che corrisponde a diversi nodi terminali, dobbiamo sommare le probabilità di giungere a ciascuno dei nodi terminali che ci interessano.

Ad esempio, se cerchiamo la probabilità di estrarre due palline nere, allora l'unico nodo terminale che ci interessa è il quarto, per cui:

$$p(2N)=p(N_1 \wedge N_2)=p(N_1)\cdot p(N_2/N_1)=\frac{3}{8}\cdot\frac{2}{7}=\frac{3}{28} \quad .$$

Se, invece, chiediamo la probabilità di ottenere due palline di colori diversi, indipendentemente dall'ordine, allora costituiscono "casi favorevoli" sia il secondo che il terzo nodo terminale:

$$p(1N \wedge 1R)=p((R_1 \wedge N_2)\vee(N_1 \wedge R_2))=p(R_1)\cdot p(N_2/R_1)+p(N_1)\cdot p(R_2/N_1)=\frac{5}{8}\cdot\frac{3}{7}+\frac{3}{8}\cdot\frac{5}{7}=\frac{15}{28} \quad .$$

Infine, la probabilità di ottenere almeno una pallina rossa, a cui corrisponderebbero i primi tre nodi terminali, si calcola più facilmente dalla probabilità dell'evento complementare:

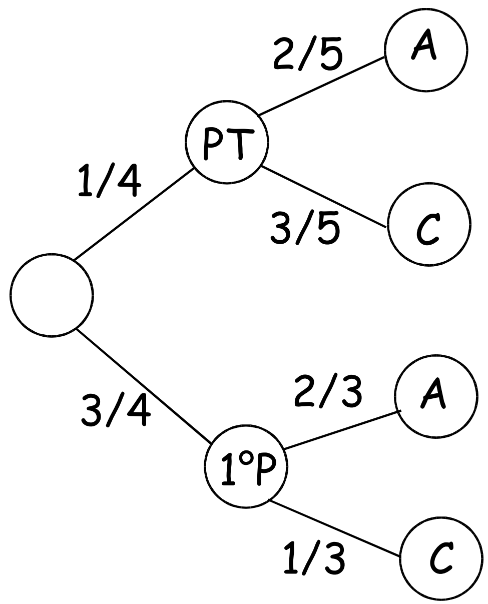
$$p(\text{almeno}1R)=p(\overline{2N})=1-p(2N)=1-\frac{3}{28}=\frac{25}{28} \quad .$$

Diagrammi ad albero 2: problema diretto

Immaginiamo che le macchinette distributrici di bevande di una scuola vengano programmate per funzionare in maniera casuale: quella del piano terra fornisce 2/5 delle volte un'aranciata e 3/5 delle volte una Coca, mentre quella del primo piano fornisce 2/3 delle volte un'aranciata ed 1/3 delle volte una Coca.

Lo studente Pippo si reca, in maniera casuale, 1/4 delle volte alla macchinetta del piano terra e 3/4 delle volte a quella del primo piano.

Rappresentiamo la situazione tramite un diagramma ad albero e calcoliamo la probabilità che Pippo torni in classe con un'aranciata, senza preoccuparci da quale macchinetta essa venga.



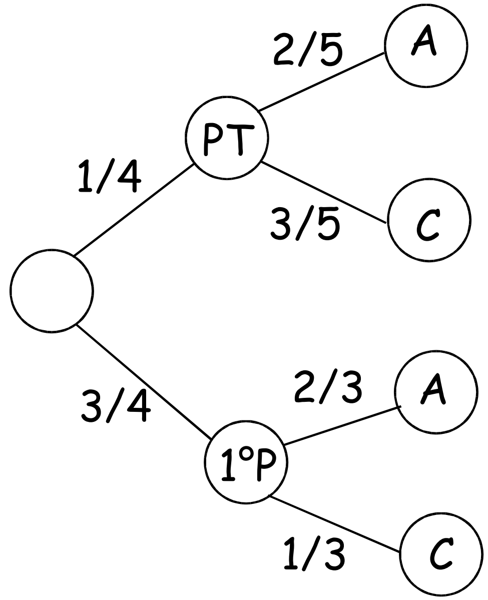
$$p(A)=p((A\wedge PT)\vee(A\wedge 1°P))=p(PT)\cdot p(A/PT)+p(1°P)\cdot p(A/1°P)=$$

$$\frac{1}{4}\cdot\frac{2}{5}+\frac{3}{4}\cdot\frac{2}{3}=\frac{1}{10}+\frac{1}{2}=\frac{3}{5}\ .$$

Diagrammi ad albero 2: problema inverso

Supponiamo che lo studente Pippo dell'esempio precedente entri in classe con un'aranciata.

Qual è la probabilità che l'abbia presa dalla macchinetta del piano terra?



Per rispondere, utilizziamo la definizione classica di probabilità, considerando come “casi possibili” quelli in cui Pippo ottiene un'aranciata (primo ed al terzo ramo del nostro albero, di cui abbiamo calcolato la probabilità nell'esempio precedente), e come “casi favorevoli” quelli in cui l'aranciata viene dal piano terra (solo il primo ramo):

$$p(PT/A)=\frac{p(PT)\cdot p(A/PT)}{p(A)}=\frac{1/4\cdot 2/5}{3/5}=\frac{1}{6}\ .$$

Teorema di Bayes

Generalizziamo l'esempio precedente.

Supponiamo di avere un certo "effetto" (nel nostro esempio, ricevere la bottiglietta di aranciata) che può essere prodotto da varie "cause" (nell'esempio, le due macchinette distributrici).

A differenza di quanto abbiamo fatto in precedenza, non ci chiediamo qual è la probabilità di ottenere un certo effetto, ma, sapendo che questo "effetto" si è verificato, qual è la probabilità che esso sia dovuto ad una tra le varie cause che lo possono determinare.

Indichiamo con B l'effetto, con A_1, A_2, ... , A_n le cause che lo possono produrre, e ci chiediamo qual è la probabilità che l'effetto B sia dovuto alla causa A_1, ovvero la probabilità condizionata $p(A_1/B)$.

Ricordiamo che $p(A_1 \wedge B) = p(A_1) \cdot p(B/A_1)$, ma anche $p(A_1 \wedge B) = p(B) \cdot p(A_1/B)$.

Uguagliando tra loro queste due espressioni, ricaviamo:

$$p(A_1/B) = \frac{p(A_1) \cdot p(B/A_1)}{p(B)} = \frac{p(A_1) \cdot p(B/A_1)}{p(A_1) \cdot p(B/A_1) + p(A_2) \cdot p(B/A_2) + \ldots + p(A_n) \cdot p(B/A_n)}$$

che è l'enunciato del <u>teorema di Bayes</u>.

Esempio

Un insegnante di fisica ha tre scaffali pieni di libri, che indichiamo S_1, S_2, S_3.

Quando ha un po' di tempo libero, va a prendere un libro a caso da uno degli scaffali.

Il libro viene preso 1/2 dello volte da S_1, 1/3 delle volte da S_2 e 1/6 delle volte da S_3.

S_1 contiene il 70% di libri di fisica, S_2 il 50% ed S_3 il 30%.

Qual è la probabilità che l'insegnante prenda un libro di fisica?

$$p(F)=p[(F\wedge S_1)\vee(F\wedge S_2)\vee(F\wedge S_3)]=p(S_1)\cdot p(F/S_1)+p(S_2)\cdot p(F/S_2)+p(S_3)\cdot p(F/3_1)=$$

$$\frac{1}{2}\cdot 0{,}7+\frac{1}{3}\cdot 0{,}5+\frac{1}{6}\cdot 0{,}3=\frac{17}{30}=0{,}5\overline{6}\ .$$

Se l'insegnante ha un libro di fisica, qual è la probabilità che esso provenga dal primo scaffale?

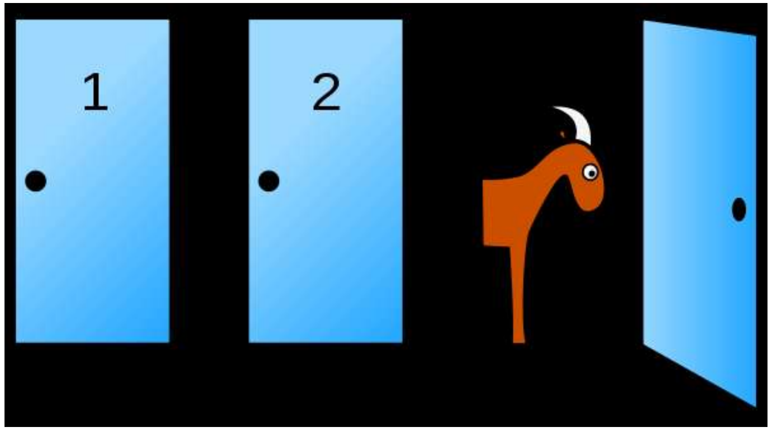
$$p(S_1/F)=\frac{p(S_1)\cdot p(F/S_1)}{p(F)}=\frac{1/2\cdot 0{,}7}{17/30}=\frac{21}{34}\simeq 0{,}618\simeq 61{,}8\%\ .$$

Nelle prossime pagine cerchiamo di dare un rapido accenno ad alcuni problemi di particolare interesse per almeno uno (spesso entrambi) dei seguenti motivi:

1. si tratta di “paradossi” nel senso etimologico del termine, in quanto i loro risultati sono inattesi e vanno contro la nostra intuizione;
2. ci mostrano degli esempi in cui il fatto di conoscere delle nuove informazioni sul problema modifica le nostre assegnazioni dei valori di probabilità.

Trattazioni molto più chiare e dettagliate si trovano facilmente in rete.

Paradosso di Monty Hall



In un gioco televisivo, vengono mostrate ad un concorrente tre porte chiuse; dietro una di esse si trova un'automobile, mentre ognuna delle altre due nasconde una capra. Il giocatore può scegliere una delle tre porte, vincendo il premio corrispondente.

Supponiamo che, dopo che il giocatore ha selezionato una porta, senza averla aperta, il conduttore del gioco apra una delle altre due, rivelando una delle due capre, e offra al giocatore la possibilità di cambiare la propria scelta iniziale, passando ultima porta restante.

Osserviamo che al giocatore conviene cambiare porta.

Infatti, inizialmente egli aveva scelto una porta a caso, e quindi aveva una probabilità uguale ad 1/3 di avere vinto la macchina e 2/3 di avere scelto una capra.

Di conseguenza, cambiando porta avrà una probabilità uguale a 2/3 di vincere la macchina, e solo 1/3 di perderla.

Paradosso dei due bambini

La famiglia Rossi ha due bambini. Sappiamo che almeno uno dei due è un maschio (ma non sappiamo quale). Qual è la probabilità che anche l'altro bambino sia maschio?

Per semplicità, assumiamo che i due sessi siano equiprobabili (in realtà nascono un po' più maschi) e che i sessi dei due figli siano eventi completamente indipendenti (in realtà è un po' più probabile che i due figli siano dello stesso sesso).

Sotto queste ipotesi, e ordinando i bambini ad esempio in base alla data di nascita, avremmo le quattro alternative MM, MF, FM, FF, ciascuna con probabilità 1/4.

Ma, avendo l'informazione che almeno uno dei due figli è maschio, dobbiamo scartare l'ultima possibilità, per cui la probabilità che anche l'altro bimbo sia maschio è:

$$\frac{p(MM)}{p(MM)+p(MF)+p(FM)}=\frac{1/4}{3/4}=\frac{1}{3}\ .$$

Paradosso del compleanno

Una classe comprende 23 studenti. La probabilità che tra di loro ce ne siano due che compiono gli anni nello stesso giorno è maggiore o minore del 50%?

Per semplicità, trascuriamo gli anni bisestili.

In questo caso, è più semplice calcolare la probabilità dell'evento complementare, ovvero che nella classe <u>non</u> ci siano due studenti che compiono gli anni nello stesso giorno.

Questa, a sua volta, è la congiunzione logica di 22 eventi indipendenti: che il secondo studente sia nato in un giorno diverso dal primo, che il terzo sia nato in un giorno diverso dai primi due, etc:

$$p(\overline{C})=\frac{364}{365}\cdot\frac{363}{365}\cdot\ldots\frac{343}{365}\simeq 0{,}4954 \quad .$$

Quindi la probabilità richiesta è $p(C)=1-p(\overline{C})\simeq 0{,}5046$, che, contrariamente all'intuizione, è <u>maggiore</u> del 50%.

Paradosso delle tre scatole

Ci sono tre scatole, di cui la prima contiene due monete d'oro, la seconda due monete d'argento e la terza una d'oro ed una d'argento. Estraendo una moneta a caso da una scatola a caso, ci ritroviamo in mano una moneta d'oro.

Qual è la probabilità che anche l'altra moneta nella scatola sia d'oro?

In tutto, le scatole contengono sei monete, ciascuna delle quali ha una probabilità 1/6 di essere estratta. Ma, avendo l'informazione che la moneta estratta è d'oro, dobbiamo limitarci a tre casi possibili (due monete nella prima scatola e una nella seconda), di cui due sono casi favorevoli (le due monete nella prima scatola).

Quindi la probabilità richiesta è $\frac{2/6}{3/6} = \frac{2}{3}$.

Paradosso di Simpson

In una stanza (1) ci sono due scatole, una rossa e una verde; nella rossa ci sono 90 calzini blu e 10 neri, nella verde 720 calzini blu e 180 neri. In un'altra stanza (2) ci sono altre due scatole, rossa e verde; nella rossa ci sono 160 calzini blu e 640 neri, nella verde 20 blu e 180 neri.

Osserviamo che, nella stanza 1, prendendo un calzino dalla scatola rossa abbiamo il 90% di probabilità di prenderlo blu, mentre dalla verde la stessa probabilità è l'80%. Nella stanza 2 le probabilità di prendere un calzino blu sono il 20% dalla scatola rossa e il 10% dalla scatola verde.

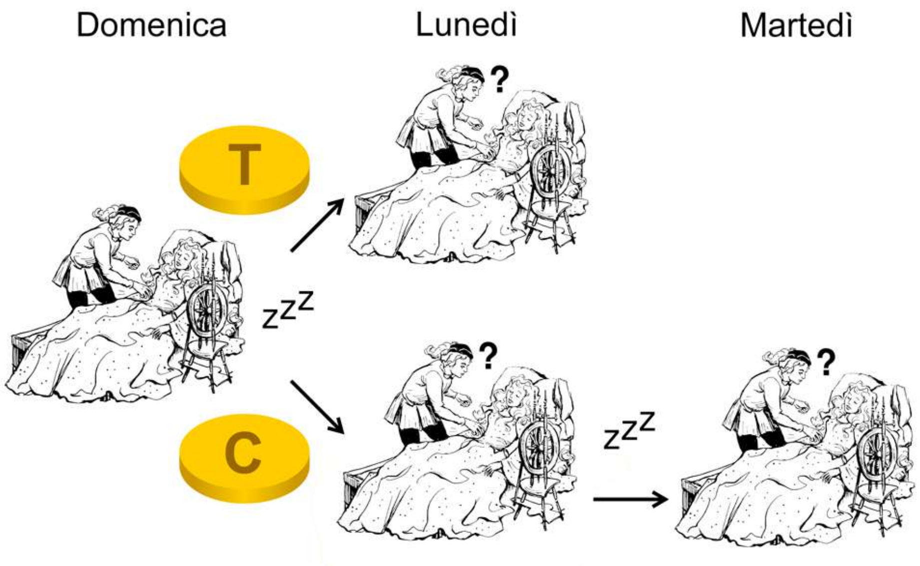
Quindi, sia nella stanza 1 che nella stanza 2, se vogliamo estrarre un calzino blu, ci conviene prenderlo dalla scatola rossa perché la probabilità è maggiore.

Ora prendiamo il contenuto della due scatole rosse e lo mettiamo in una scatola rossa più grande, e facciamo lo stesso con le due scatole verdi.

Questa volta, cosa ci conviene fare per ottenere un calzino blu?

E' più probabile ottenerlo dalla scatola rossa grande o dalla scatola verde grande?

Paradosso della Bella Addormentata



La Bella Addormentata partecipa al seguente esperimento:

- domenica viene addormentata e viene lanciata una moneta;
- se esce testa, viene svegliata lunedì;
- se esce croce, viene svegliata lunedì, le viene dato un farmaco che le fa dimenticare il risveglio, e viene svegliata nuovamente martedì.

Quando viene svegliata, la Bella Addormentata ricorda tutta la procedura dell'esperimento, ma non sa che giorno sia, né se è stata risvegliata in precedenza o meno.

Ad ogni risveglio, le viene chiesto: qual è il tuo grado di fiducia sulla affermazione "è uscita testa"? (ovvero, la probabilità che lei attribuisce soggettivamente a questo evento).

Soluzione A: la Bella Addormentata dovrebbe rispondere che $p(T)=1/2$, in quanto non ha ricevuto nessuna nuova informazione che modifichi la sua conoscenza precedente.

Soluzione B: la Bella Addormentata dovrebbe rispondere che $p(T)=1/3$, in quanto ci sono tre casi possibili (risveglio lunedì se esce testa, risveglio lunedì se esce croce, risveglio martedì se esce croce), di cui solo il primo è favorevole.

Soluzione C: la domanda, così come è posta, è ambigua, e la risposta può essere $p(T)=1/2$ o $p(T)=1/3$ a seconda di quali dettagli aggiuntivi vengono forniti.

Il problema è ancora aperto (anche se la maggior parte degli studiosi sono “terzisti”), e puoi trovare in rete una bibliografia sterminata, ad esempio cercando “Sleeping Beauty problem”.