La Probabilità condizionata

JON HOLMEN, UNO STUDENTE DELL'ILLINOIS STATE UNIVERSITY, RACCONTAVA DI AVERE AVUTO MOLTI INCIDENTI D'AUTO MENTRE VIVEVA A CASA CON I SUOI GENITORI.

UN GIORNO, IL SUO PROFESSORE DI STATISTICA SPIEGÒ CHE L'83% DEGLI INCIDENTI D'AUTO ACCADE A MENO DI 20 MIGLIA DA CASA.

IL GIORNO DOPO JON ANDÒ A VIVERE A 22 MIGLIA DAI SUOI CARI.

La Probabilità e la Statistica possono giocare "brutti scherzi": la Probabilità condizionata ancor di più.

Introduciamola e poi facciamo alcune considerazioni.

La probabilità di un evento A spesso deve essere modificata dopo che sono state ottenute informazioni su un altro evento H che è già accaduto (o che potrebbe verificarsi) e che potrebbe essere legato al nostro evento A. Le informazioni su aspetti particolari di alcuni risultati sperimentali potrebbero quindi rendere necessaria una riconsiderazione della probabilità di un evento che riguarda lo stesso fenomeno. La probabilità, così eventualmente modificata, prende il nome di **probabilità condizionata** dell'evento A, avendo informazioni sull'evento H.

In genere, ogni (nuova) informazione dovrebbe alterare le probabilità che noi assegniamo ad altri eventi. Vediamo anzitutto ciò attraverso alcuni Esempi.

Iniziamo con due casi tipici, uno legato all'attualità ed uno ai giochi: campionario Ω.

Esempio 1 *i) – Luigi, studente di Ingegneria, ha già sostenuto* 10 *esami ed ha la media del* 25

1) Qual è la probabilità che superi il prossimo esame (di Calcolo delle Probabilità e Statistica) ?

2) Qual è la probabilità che superi il prossimo esame (di Calcolo delle Probabilità e Statistica), sapendo che ha ottenuto 30 *nel compito scritto ?*

ii) – Da un mazzo di 40 *carte da gioco (francesi) si estrae una carta. Ci si chiede:*

1) Qual è la probabilità che essa sia un "Re"?

2) Qual è la probabilità che essa sia un "Re" se, nel frattempo, sono venuto a sapere che la carta estratta è una figura ?

Svolgimento Anche senza aver mai sentito parlare di "probabilità condizionata" si può rispondere subito dicendo

1) Parlare di una probabilità intorno al 70% sembrerebbe abbastanza ragionevole.

2) Adesso, dopo questa informazione anche il 95% sembrerebbe ragionevole.

1) $\mathbb{P} = \dfrac{4}{40} = \dfrac{1}{10}$.

2) $\mathbb{P} = \dfrac{4}{12} = \dfrac{1}{3}$. (Adesso i casi possibili sono solo $12 \cdots$.)

Come si vede abbiamo affrontato il caso della assegnazione di probabilità di un evento **supponendo** di sapere che un certo (altro) evento H si è verificato (con probabilità positiva). Vediamo un altro Esempio.

Esempio 2 *Supponiamo di scegliere a caso una persona (cittadino italiano) e chiamiamo questo l'evento H. Sia poi A l'evento che la persona scelta è un tifoso della Fiorentina e sia B l'evento che la persona sia un abitante di Firenze.*

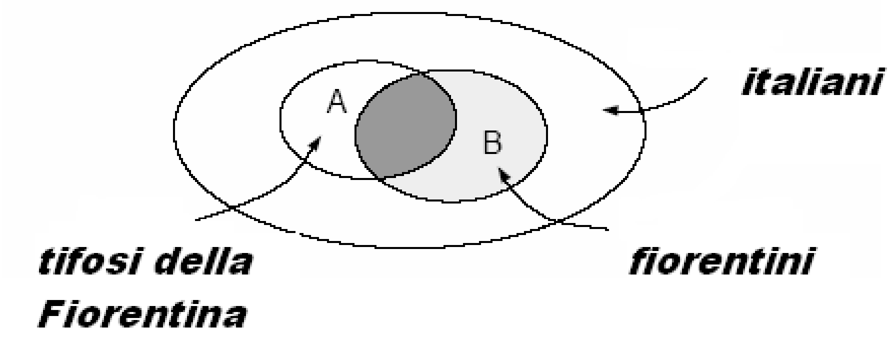

Figura 1

Quali sono le probabilità dei due eventi considerati ? (Si veda la Figura 1) Intuitivamente, dovremo scegliere (a caso) un punto dell'insieme universo e chiederci con quale probabilità esso appartenga all'insieme A oppure all'insieme B.

La maggior parte degli italiani non vive a Firenze e non è tifoso della Fiorentina: quindi gli eventi A e B hanno probabilità "piccola". Ma, se ci chiediamo quale sia la probabilità che una persona che viva a Firenze sia anche "tifoso della Fiorentina"; tale probabilità, certamente, non è adesso "piccola".

Tenendo allora conto dell'informazione data dall'evento (condizionante) H possiamo dire che il nostro evento A si verifica se e solo se si verificano sia A sia H; in tal modo lo spazio campionario si riduce ad H. Quindi, la probabilità di A, dopo la conoscenza di H, dovrà essere "collegata" al valore di $\mathbb{P}(A \cap H)$.

Nota Il fatto che $\mathbb{P}(H)$ sia ipotizzato diverso da zero segue subito dal fatto, ovvio, che non possiamo subordinare alcunché ad un evento di probabilità nulla !

Possiamo allora dare la seguente definizione:

Definizione. Siano A e H eventi di un esperimento casuale con $\mathbb{P}(H) > 0$. La probabilità condizionata di A dato H è definita come

$$\mathbb{P}(A/H) = \frac{\mathbb{P}(A \cap H)}{\mathbb{P}(H)}.$$

Vediamo ora questo concetto anche alla luce dell'assegnazione classica o frequentista.

Supponiamo, ad es., di ripetere più volte l'esperimento. Per un certo evento H, sia $n(H)$ il numero di volte che H si è verificato nelle prime n prove.

Se $n(H)$) è grande, la probabilità condizionata che A si sia verificato, dato il verificarsi di H, dev'essere vicina alla frequenza relativa condizionata di A dato H, ovvero la frequenza relativa di A per le prove in cui H si verificato:

$$n(A \cap H)/n(H).$$

Ma questa è, tenendo conto del concetto di frequenza relativa, ciò che abbiamo scritto sopra.

In alcuni casi, le probabilità condizionate possono essere calcolate direttamente, riducendo in modo opportuno lo spazio campionario.

Vediamo ancora un Esempio.

Esempio 3 *Consideriamo una classe di* 100 *studenti dei quali conosciamo pesi (in ordinata) ed altezza (in ascissa).*

==	160	165	170	175	180	≥ 185	*Tot.*
≤ 55	–	4	7	2	–	–	13
60	–	3	18	2	2	–	25
65	1	2	11	3	5	1	23
70	1	–	2	4	9	4	20
75	–	–	–	6	4	2	12
80	–	–	–	2	1	2	5
≥ 85	–	–	–	–	2	–	2
Tot	2	9	38	19	23	9	100

E ci chiediamo "Qual è la probabilità che uno studente pesi 70 *kg sapendo che è alto* 170 *cm.*

Svolgimento Dobbiamo calcolare

$$\mathbb{P}(p = 70/a = 170) = \frac{\mathbb{P}(p = 70) \cap (a = 170)}{\mathbb{P}(a = 170)}.$$

Dalla tabella abbiamo
$\mathbb{P}(a = 170) = 0.37; \quad \mathbb{P}((p = 70) \cap (a = 170)) = \frac{2}{37} \approx 0.054.$

Esempio 4 *Se $H = \Omega$ abbiamo* $\mathbb{P}(A/\Omega) = \dfrac{\mathbb{P}(A \cap \Omega)}{\mathbb{P}(\Omega)} = \dfrac{\mathbb{P}(A))}{1)} = \mathbb{P}(A).$

Da questa osservazione seguono anche altre conseguenze. Ad es., $\mathbb{P}(A^c/B) = 1 - \mathbb{P}(A/B)$*. mentre NON vale in generale che* $\mathbb{P}(A/B^c) = 1 - \mathbb{P}(A/B)$*.*

Esercizio 1 *Mostrare, attraverso Esempi, che valgono le due affermazioni dell'Esempio precedente.*

Esempio 5 *Il Prof. Baruzzo ha due figli. Uno dei due è Maschio. Che probabilità c'è che siano entrambi Maschi ?*

Il Prof. Baruzzo ha due figli. Il primogenito è è Maschio. Che probabilità c'è che siano entrambi Maschi ?

Svolgimento Le risposte, forse non in favore della nostra intuizione, sono..... diverse. Nel primo caso si ha un caso favorevole (M-M) su 3 (F-M, M-F, F-F), mentre nel secondo caso ce n'è uno su 2 (F-M, F-F).

Esempio 6 *Quattro giocatori pescano ciascuno una carta da un mazzo di* 40 *carte. Qual è la probabilità di pescare tutti "semi" diversi ?*

Svolgimento: Definiamo i seguenti eventi:

$A = \{$ le quattro carte sono di seme diverso $\}$;
$A_1\{$ si estrae (a caso) la prima carta $\}$;
$A_2\{$ si estrae (a caso) la seconda carta di seme diverso dalla prima $\}$;
$A_3\{$ si estrae (a caso) la terza carta di seme diverso dalle prime due $\}$;
$A_4\{$ si estrae (a caso) la quarta carta di seme diverso dalle prime tre$\}$.

Allora

$$\mathbb{P}(A) = \mathbb{P}(A_1 \cap A_2 \cap A_3 \cap A_4) =$$

$$\mathbb{P}(A_1)\mathbb{P}(A_2/A_1)\mathbb{P}(A_3/(A_1 \cap A_2)\mathbb{P}(A_4/capA_1 \cap A_2 \cap A_3) =$$

$$\frac{40}{40} \cdot \frac{40}{40} \cdot \frac{30}{39} \cdot \frac{20}{38} \cdot \frac{10}{37} \approx 0.109.$$

Ovviamente, la probabilità condizionata, dato il verificarsi di H dovrà essere sempre una probabilità, ovvero deve soddisfare le proprietà della probabilità, (riassunte ad es., dagli assiomi di Kolmogorov). Possimao pensare ad una certa normalizzazione attraverso una una costante di proporzionalità: essa potrà essere $\frac{1}{\mathbb{P}(H)}$. Questo ed altri fatti saranno considerati nel seguente esercizio.

Esercizio 2 *Dimostrare, utilizzando le proprietà precedenti, che:*

1) $\mathbb{P}(A/H)$, *in funzione di* A *e per* H *dato, è una misura di probabilità.*

2) se $H \subset A$ *allora* $\mathbb{P}(A/H) = 1$;

3) se A *e* H *sono incompatibili, allora* $\mathbb{P}(A/H) = 0$;

4) $\mathbb{P}(A/H) < \mathbb{P}(A)$ *se e solo se* $\mathbb{P}(H/A) < \mathbb{P}(H)$;

5 $\mathbb{P}((A \cup B/)C) = \mathbb{P}(A/C) + \mathbb{P}(B/C) - \mathbb{P}((A \cap B)/C)$;

6) $\mathbb{P}(A_1 \cap A_2 \cap \cdots A_n) = \mathbb{P}(A_1)\mathbb{P}(A_2/A_1)\mathbb{P}(A_3/A_1 \cap A_2) \cdots \mathbb{P}(A_n/(A_1 \cap A_2) \cdots A_{n-1})$.

Esempio 7 *Un'urna contiene* 10 *palline Rosse,* 5 *Viola e* 7 *Gialle. Qual è la probabilità di estrarne successivamente, senza restituzione. E Tre di colore diverso ?*

Svolgimento Usiamo la proprietà 6) dell'Esercizio precedente. Abbiamo, indicando con A_1, A_2, A_3 gli eventi che indicano rispettivamente l'estrazione di una pallina Rossa, Viola e Gialla, si ha:

$$\mathbb{P}(A_1) = \frac{10}{22}, \quad \mathbb{P}(A_2/A_1) = \frac{5}{21}, \quad \mathbb{P}(A_3/A_1 \cap A_2) = \frac{7}{20}.$$

Dunque $\mathbb{P}(A_1 \cap A_2 \cap A_3) = \frac{10}{22}\frac{5}{21}\frac{7}{20} = \frac{5}{132}$.

Se supponiamo che gli eventi A_1, A_2, A_3 siano tali che $\mathbb{P}(A_2/A_1) = \mathbb{P}(A_2)$ e $\mathbb{P}(A_3/A_1 \cap A_2) = \mathbb{P}(A_3)$: ciò significa che l'accadere di un evento NON è condizionato dall'accadere (o dal non accadere) di uno (o da una qualsiasi "combinazione") degli altri eventi. In parole la probabilità del verificarsi dell'evento A_1 è la stessa della probabilità del verificarsi dell'evento "A_2 sapendo che si è verificato A_1".

In tali ipotesi la proprietà 6) dell' Esercizio precedente può essere scritta come:

$$\mathbb{P}(A_1 \cap A_2 \cap \cdots A_n) = \mathbb{P}(A_1)\mathbb{P}(A_2)\cdots\mathbb{P}(A_n).$$

Ciò ci porta al concetto di **eventi indipendenti.**

Se abbiamo Due eventi la scrittura precedente assume la forma:

$$\mathbb{P}(A_1 \cap A_2) = \mathbb{P}(A_1)\mathbb{P}(A_2).$$

Se ne abbiamo Tre dobbiamo tener conto del fatto che, oltre alla prevedibile richiesta:

$$\mathbb{P}(A_1 \cap A_2 \cap A_3) = \mathbb{P}(A_1)\mathbb{P}(A_2)\mathbb{P}(A_3),$$

deve verificarsi anche che

$$\mathbb{P}(A_1 \cap A_2) = \mathbb{P}(A_1)\mathbb{P}(A_2), \quad \mathbb{P}(A_1 \cap A_3) = \mathbb{P}(A_1)\mathbb{P}(A_3), \mathbb{P}(A_3 \cap A_2) = \mathbb{P}(A_3)\mathbb{P}(A_2).$$

In altre parole, quando abbiamo più eventi, l'indipendenza deve essere "verificata" a due a due, a tre a tre, ecc. La definizione formale quindi è la seguente:

Definizione.N eventi sono indipendenti se $\forall k = 2, 3, \cdots, N$ si ha

$$\mathbb{P}(E_{i_1} \cap E_{i_2} \cap \cdots \cap E_{i_k}) = \mathbb{P}(E_{i_1})\mathbb{P}(E_{i_2})\cdots\mathbb{P}(E_{i_k})$$

per ogni scelta degli indici $i_1, \cdots, i_k$.

Esempio 8 *Un sistema per la generazione di corrente elettrica è formato da tre generatori indipendenti. La potenza richiesta dal sistema può essere effettivamente fornita se almeno due dei tre generatori funzionano correttamente. Sia $G = \{$il sistema eroga la potenza richiesta$\}$ un evento d'interesse: se ne calcoli la probabilità.*

Svolgimento Indichiamo con A, B e C gli eventi che indicano il corretto funzionamento del primo, secondo e terzo generatore e poniamo $\mathbb{P}(A) = \mathbb{P}(B) = \mathbb{P}(C) = p$.

I casi possibili sono $2^3 = 8$. Essi sono dati da:

$$(A \cap B \cap C), (A \cap B \cap C^c), (A \cap B^c \cap C), (A \cap B^c \cap C^c),$$
$$(A^c \cap B \cap C), (A^c \cap B \cap C^c), (A^c \cap B^c \cap C), (A^c \cap B^c \cap C^c).$$

Di questi, quelli in cui almeno due generatori sono attivi sono quattro: il caso in cui funzionano tutti e tre e i tre casi in cui si ha il cattivo funzionamente di uno solo. La loro unione è data da:

$$G = \{(A \cap B \cap C) \cup (A \cap B \cap C^c) \cup (A \cap B^c \cap C) \cup (A^c \cap B \cap C)$$

Trattandosi di eventi incompatibili, la probabilità dell'unione è data dalla somma delle probabilità. Dunque si ha:

$$\mathbb{P}(G) = p^3 + 3p^2(1-p) = 3p^2 - 2p^3.$$

Il seguente è un risultato teorico di notevole importanza:

Teorema 1 *: Per ogni coppia di eventi A e B, di probabilità non nulla, si ha*

$$\mathbb{P}(A) = \mathbb{P}(A/B)\mathbb{P}(B) + \mathbb{P}(A/B^c)\mathbb{P}(B^c).$$

(L'uguaglianza vale anche se scambiamo A con B.)

Dimostrazione Infatti, da $A = A \cup \Omega = A \cap (B \cup B^c)$ segue

$$\mathbb{P}(A) = \mathbb{P}(A \cap B) + \mathbb{P}(A \cap B^c) = \mathbb{P}(A/B)\mathbb{P}(B) + \mathbb{P}(A/B^c)\mathbb{P}(B^c),$$

avendo usato direttamente le definizione di probabilità condizionata.

Esempio 9 *Un commerciante di televisori si serve principalmente di tre industrie produttrici diverse Anvedi, Bellavista e Chiaroschermo (che indicheremo con A, B, C). La sua conoscenza delle ditte lo porta a sapere che A produce metà dei televisori di B ed il 20% dei televisori prodotti è difettoso, B il 10% di difettosi e C, che produce il 50% di televisori più di quanti ne produce B, ne produce il 5% difettosi. Un giorno gli arriva al negozio un suo ordinativo di 600 televisori.*

1) qual è la probabilità che, scegliendone uno a caso, esso sia difettoso ?

2) qual è la probabilità che, sapendolo difettoso, esso sia stato prodotto da A ?

Svolgimento Dai dati del problema si ha subito che, indicando con D l'evento "televisore difettoso":

— $\mathbb{P}(A) = \frac{1}{6}$; $\mathbb{P}(B) = \frac{1}{3}$; $\mathbb{P}(C) = \frac{1}{2}$;

— $\mathbb{P}(D/A) = \frac{1}{5}$; $\quad \mathbb{P}(D/B) = \frac{1}{10}$; $\quad \mathbb{P}(D/C) = \frac{1}{20}$.

Si ha allora:

1) $\mathbb{P}(D) = \mathbb{P}(D/A)\mathbb{P}(A) + \mathbb{P}(D/B)\mathbb{P}(B) + \mathbb{P}(D/C)\mathbb{P}(C) =$
$= \frac{1}{5}\frac{1}{6} + \frac{1}{3}\frac{1}{10} + \frac{1}{2}\frac{1}{20} = \frac{11}{120}$.

2) $\mathbb{P}(A/D) = \frac{\mathbb{P}(A \cap D)}{\mathbb{P}(D)} = \frac{\mathbb{P}(D/A)\mathbb{P}(A)}{\mathbb{P}(D)} = \frac{1}{5}\frac{\frac{1}{6}}{\frac{11}{120}} = \frac{4}{11}$.

Il Teorema precedente è solo un risultato utile al più importante:

Teorema 2 (di Bayes) ([1]) *Dato un evento A ed una classe completa di eventi incompatibili H_j, $j = 1...N$, si ha*

$$\mathbb{P}(H_j/A) = \frac{\mathbb{P}(A/H_j)\mathbb{P}(H_j)}{\sum_j \mathbb{P}(A/H_j)\mathbb{P}(H_j)}.$$

[1]Thomas Bayes (Londra, 1702 Tunbridge Wells, 1761) è stato un matematico e ministro presbiteriano britannico. È noto soprattutto per questo risultato, pubblicato postumo nel 1763.

Abbiamo

$$\mathbb{P}(H_j/A) = \frac{\mathbb{P}(H_j) \cap A}{\mathbb{P}(A)},$$

da cui

$$\mathbb{P}(H_j/A) = \frac{(\mathbb{P}(H_j) \cap A)}{\mathbb{P}(\sum_j A/H_j)\mathbb{P}(H_j)}.$$

Esempio 10 *Un sistema di comunicazione telematica invia segnali attraverso i bit "0" oppure "1". A causa di interferenze il segnale trasmesso è spesso ricevuto male. Calcolare la probabilità di ricevere* 0 *avendo trasmesso* 1 *e la probabilità di un errore di trasmissione dopo dieci mesi di attvità del servizio.*

Svolgimento: Sapendo che il servizio è attivato da 10 mesi abbiamo dati stastistici archiviati che ci permettono di mettere in atto alcune informazioni.

Supponiamo che $T_k = \{$ il bit k è trasmesso $\}$, e $E_k = \{$ il bit k è ricevuto, per $k = 0, 1.\}$ Supponiamo allora che dai dati si abbia:

$$\mathbb{P}(R_0/T_0) = 0.7, \quad \mathbb{P}(R_1/T_1) = 0.8, \quad \mathbb{P}(T_0) = 0.6.$$

(l'ultima informazione ci dice ad es., che il bit 0 è inviato mediamente il 60% delle volte).

Abbiamo allora

$$\mathbb{P}(T_0/R_1) = \frac{\mathbb{P}(R_1/T_0)\mathbb{P}(T_0)}{\mathbb{P}(R_1/T_0)\mathbb{P}(T_0) + \mathbb{P}(R_1/T_1)\mathbb{P}(T_1)} =$$

$$\frac{0.3) \cdot 0.6}{0.3 \cdot 0.6 + 0.8} \cdot 0.4 \approx 0.36.$$

L'errore di trasmissione ha probabilità $\mathbb{P}(T_0 \cap R_1) + \mathbb{P}(T_1 \cap R_0) =$

$$\mathbb{P}(R_1/T_0)\mathbb{P}(T_0) + \mathbb{P}(R_0/T_1)\mathbb{P}(T_1) = 0.3 \cdot 0.6 + 0.2 \cdot 0.4 \approx 0.26.$$

Qualche confusione può essere fatta fra l'interpretazione degli eventi

$$\mathbb{P}(A/B) \quad \text{e} \quad \mathbb{P}(B/A).$$

Vediamone alcuni Esempi elementari (e meno) utilizzando anche il Teorema di Bayes.

Esempio 11 *Consideriamo gli eventi (nell'anno calcistico 2006-2007):*

- $A = \{$ *un uomo, abitante a Firenze, è capocannoniere in Serie A* $\}$;
- $F = \{$ *un uomo, capocannoniere in Serie A, è un abitante di Firenze A* $\}$.

Calcoliamo $\mathbb{P}(A/F)$ e $\mathbb{P}(F/A)$. Abbiamo:

$\mathbb{P}(A/F) \approx \dfrac{1}{400000} \approx 2.5 \times 10^{-6}$;
$\mathbb{P}(F/A) = 1.$ [2]

Esempio 12 *In un Test medico una malattia, presente nella popolazione all'*1%, *è accurato nella misura dell'*80% *ed ha un "falso positivo" del* 10%. *In un'inchiesta fra futuri medici venne fuori che la probabilità di quella malattia, in chi ha avuto Test positivo, era circa del* 75%. *È questo una risposta "credibile" ?*

Svolgimento: Applichiamo il Teorema di Bayes agli eventi:
$A = \{$ il paziente è malato $\}$; $H = \{$ il Test è positivo $\}$.

Abbiamo

$$\mathbb{P}(A/B) = \frac{\mathbb{P}(B/A)\mathbb{P}(A)}{\mathbb{P}(B/A)\mathbb{P}(A) + \mathbb{P}(B/A^c)\mathbb{P}(A^c)} = 0.075.$$

La differenza fra $\mathbb{P}(A/B) = 0.075$ e $\mathbb{P}(B/A) = 0.80$ anche questa volta è notevole !!

Qui *l'errore* sta nel fatto che si pensa che le probabilità di malattia siano "proporzionali" alla percentuale di positività del Test.

Nota: Questo è un "errore" molto comune anche in altri campi. Ne citiamo alcuni:

- Molti incidenti accadono entro 20 km da casa: quindi siamo più sicuri quando guidiamo la macchina lontano da casa;
- Molti tossicomani fumano spinelli: chi fuma spinelli prima o poi diventerà tossicomane.

[2] In quell'anno calcistico il capocannoniere è stato Luca Toni, giocatore della Fiorentina.

- ...

La probabilità condizionata è lo strumento "giusto" per affrontare le questioni che trattano lo studio dell'incertezza. Il fatto è che la realtà condizionata dal nostro modo di descriverla e che sta a noi decidere, per fare ciò, quale metodo impiegare: in un certo senso noi non scopriamo propriamente, ma "determiniamo" come il mondo è.

Per gli eventi che appaiono verosimili, la differenza dei dati che ogni uomo ha su di essi è una delle cause principali della diversità di opinioni relative ai medesimi oggetti.

Supponiamo, ad esempio, che si abbiano tre urne A, B, C, di cui una contenga solo una pallina Nera, mentre le altre due solo palline Rosse: ci si chiede se, estrando una pallina dall'urna C quale sia la probabilità che essa sia nera. Se si ignora quale delle tre urne contenga solo pallinee Nere, le tre ipotesi possibili sembreranno ugualmente probabili, e poiché una pallina Nera può essere estratta solo dalla prima ipotesi, la probabilità di estrazione è uguale ad un terzo. Se abbiamo altre informazioni, ad es., se è noto che l'urna A contiene solo palline Rosse, l'indecisione verte allora solo sulle urne B e C, e la probabilità che la pallina estratta dall'urna C sia nera è del 50%. Naturalmente si ha probabilità 1 se abbiamo avuto (da qualche informatore) la certezza che le urne A e B contengono solo palline Rosse. Analogamente, il medesimo giudizio, esposto di fronte ad una folta assemblea, viene accettato a diversi livelli, secondo le profondità delle conoscenze degli ascoltatori. Se colui che lo espone ne è intimamente persuaso e se, per la sua posizione e per il suo carattere ispira una grande fiducia, la sua esposizione, per quanto straordinaria, avrà, per gli ascoltatori privi di altre fonti, lo stesso grado di verosimiglianza di un fatto ordinario, e gli potranno concedere una fiducia completa.

Talvolta ciò accade ai buoni insegnanti.

Esempio 13 (Tale Esempio ricalca il "paradosso del carceriere",) *[ben presente in molti libri di Calcolo delle Probabilità.]*

Al master di Matematica Sublime vengono ammessi 5 *studenti ogni anno: nel 2006 si sono presentati* 9 *candidati, fra i quali il giovane Carlo. Due giorni dopo la prova Carlo incontra il Direttore della Scuola e gli chiede come è andata: il Direttore risponde che la commissione è ancora al lavoro: gli può dire comunque che il candidato Adelmo ha superato la prova. È possibile determinare:*

a) la $\mathbb{P}(C) = \mathbb{P}\{$ *Carlo venga ammesso* $\}$.

b) la $\mathbb{P}(C/A) = \mathbb{P}\{$ *Carlo venga ammesso sapendo che Adelmo viene ammesso* $\}$.

c) la $\mathbb{P}(C/D) = \mathbb{P}\{$ *Carlo venga ammesso dopo aver ricevuto l'informazione dal Direttore*$\}$.

Svolgimento

a) L'unica ipotesi che possiamo fare è che i 9 candidati siano "equibravi" e dunque

$$\mathbb{P}(C) = \mathbb{P}\{ \text{ Carlo venga ammesso } \} = \frac{5}{9}.$$

Nota: Lo stesso risultato si ottiene osservando che i casi favorevoli sono $\binom{8}{4}$, mentre i casi possibili sono $\binom{9}{5}$.

b)

$$\mathbb{P}(C/A) = \mathbb{P}\{ \text{ Carlo venga ammesso sapendo che Adelmo viene ammesso } \} =$$

$$\frac{\binom{7}{3}}{\binom{8}{4}} = \frac{1}{2}.$$

c) Notiamo anzitutto che $D \subset A$ in quanto Adelmo può essere assunto anche se il Direttore avesse fatto un altro nome. Da questa osservazione segue

$$\mathbb{P}(C/D) = \frac{\mathbb{P}(C \cap D)}{\mathbb{P}(D)} = \frac{\mathbb{P}(C \cap D \cap A)}{\mathbb{P}(D \cap A)} =$$

$$\frac{\mathbb{P}(A)\mathbb{P}(C/A)}{\mathbb{P}(D/A \cap C)}\mathbb{P}(A)\mathbb{P}(C/A)\mathbb{P}(D/A \cap C) + \mathbb{P}(A)\mathbb{P}(C^c/A)\mathbb{P}(D/A \cap C^c).$$

Per valutare questa espressione, osserviamo che:

- Poniamo $P_A = \mathbb{P}((D/A \cap C)$ ovvero la probabilità che il Direttore ricordi il candidato Adelmo, sapendo che sia Carlo che Adelmo vengono ammessi; $P_{A^c} = \mathbb{P}(D/A \cap C^c)$ ovvero la probabilità che il Direttore ricordi il candidato Adelmo, sapendo che che Adelmo è ammesso ma Carlo no;

- $\mathbb{P}(C^c/D) = 1 - \mathbb{P}(C/D) = \frac{1}{2}$;
- $\mathbb{P}(A) = \mathbb{P}(C) = \frac{5}{9}$. (per l'"equibravura".)

Abbiamo allora

$$\mathbb{P}(C/D) = \frac{\frac{5}{9}\frac{1}{2}P_A}{\frac{5}{9}\frac{1}{2}P_A + \frac{5}{9}\frac{1}{2}P_{A^c}}$$

Cosa possimao dire sui valori P_A e P_{A^c} ?

In genere non possiamo dire che gli eventi D e C sono indipendenti (e tutto dipenderà da situazioni di scelta della correzione da parte della commissione, ecc.). Se comunque la commissione avesse iniziato a caso a correggere gli elaborati, avremmo

$$P_A = \frac{1}{4}, \quad P_{A^c} = \frac{1}{5}$$

(la prima nell'ipotesi che il Direttore sappia che anche Carlo è stato ammesso, la seconda nel caso in cui sappia che Carlo non è stato ammesso).

Avremo allora

$$\mathbb{P}(C/D) = \frac{\frac{5}{9}\frac{1}{2}\frac{1}{4}}{\frac{5}{9}\frac{1}{2}\frac{1}{4} + \frac{5}{9}\frac{1}{2}\frac{1}{5}} = \frac{5}{9}.$$

Vediamo, come forse ci si poteva aspettare, che in questo caso gli eventi C e D sono eventi indipendenti, avendo noi postulato l'"equibravura" dei candidati; peraltro se i candidati non sono "equibravi" dobbiamo avere un'informazione sulla loro "bravura" e ciò determina la probabilità necessaria.

Alcuni fatti paradossali

a) *Droghe leggere e droghe pesanti*

Quando su televisioni o giornali si parla di problemi collegati all'uso di stupefacenti gli aspetti emotivi prevalgono di gran lunga su quelli razionali, ovvero si parla in modo "ignorante", ignorando cioè il modo per poter fare affermazioni corroborate da un supporto scientifico. Indagando sulle storie individuali si chiede al consumatore di droghe pesanti (eroina, cocaina, ...) se "prima" (di diventare tossicodipendente) fumava spinelli (o una qualsiasi droga leggera. Alla risposta, generalmente affermativa, se ne deduce (rigorosamente ??) che si deve proibire l'uso di droghe leggere per evitare, poi, l'uso di quelle pesanti.

Nello stesso modo, se gli si chiedeva se al semaforo si fermava col Rosso, dopo al risposta – anche adesso generalmente affermativa – si dovrebbe proibire di fermarsi al Rosso per non diventare tossicodipendente.

Dal punto di vista della probabilità condizionata qual è la differenza fra $\mathbb{P}(A/B)$ e $\mathbb{P}(B/A)$?

La differenza può essere massima, come mostra il seguente Esempio:

Esempio 14 *Siano dati gli eventi:*
$A\{$ *X è un uomo che vive a Roma* $\}$
$B\{$ *X è il Presidente della Repubblica Italiana* $\}$

Si vede facilmente che:

$$\mathbb{P}(A/B) = 1, \qquad \mathbb{P}(B/A) \approx 0.$$

(Essendo circa 2 *milioni gli uomini che vivono a Roma, la probabilità che uno a caso sia il Presidente della Repubblica è praticamente nulla, mentre il primo evento è certo.)*

In un articolo sulla rivista *Induzioni (n. 6, 1993, pag. 45)* Carla Rossi e Romano Scozzafava hanno mostrato che la probabilità che la probabilità di passaggio da droghe leggere a droghe pesanti è di circa 6.75% Tale probabilit, solo per dare un'idea di una valutazione *razionale*, è dello stesso ordine della probabilità della vittoria del Montenegro nella sfida calcistica con la nazionale italiana (15 ottobre 2008).

b) **Il paradosso di Simpson**

Terminiamo la parte relativa alla Probabilità condizionata con quello che, noto come **Paradosso di Simpson**, è un argomento molto sottile che richiede sempre un'analisi accurata dei dati statistici.

In una scuola, in cui insegnano i prof. Anichini e Brunelli, si hanno, per quanto riguarda le promozioni all'anno successivo i seguenti dati:

	0	1	2
Tutti	Promossi	Non Promossi	% di promossi
Anichini	20	20	50%
Brunelli	24	16	60%

Il prof. Brunelli è un professore meno severo del prof. Anichini? Sembrerebbe di sì

Ma vediamo i dati da un'altra ottica, dividendo gli studenti fra maschi e femmine.

	0	1	2
Maschi	Promossi	Non Promossi	% di promossi
Anichini	12	18	40%
Brunelli	3	7	30%

	0	1	2
Femmine	Promossi	Non Promossi	% di promossi
Anichini	8	2	80%
Brunelli	21	9	70%

La seconda e la terza tabella ci dicono che il prof. Anichini è meno severo del Prof. Brunelli, sia nei confronti degli studenti Maschi sia delle studentesse. E questo fatto non dipende di "piccoli" numeri (nel senso che posso moltiplicare tutti i dati per 100 o per 900 e niente cambierebbe).

Sembra dunque che la prima Tabella sia in completa contraddizione con i dati delle altre due. la situazione in cui una relazione tra due fenomeni viene apparentemente modificata o persino invertita dai dati in possesso a causa di altri fenomeni non presi in considerazione nell'analisi.

L'apparente paradosso nasce dal fatto che quando mettiamo insieme i dati stiamo supponendo di di fare la media delle proporzionalità: in effetti il mettere insieme i dati crea una "media" ma sitratta di una media "pesata" (rispetto alla numerosità del gruppo osservato).

Esso dice in particolare che le seguenti disuguaglianze, dati 3 eventi A, B, H possono anche verificarsi contemporaneamente.

$$\begin{gathered}
\mathbb{P}(H/A \cap B) > \mathbb{P}(H/A \cup B^c), \\
\mathbb{P}(H/A^c \cap B) > \mathbb{P}(H/A^c \cup B^c), \\
\mathbb{P}(H/B) < \mathbb{P}(H/B^c).
\end{gathered}$$

È sufficiente osservare (Esercizio !) che se $\mathbb{P}(A/B) = \mathbb{P}(A/B^c)$ ciò non accade.

Non è questo un fatto facilmente eliminabile: talvolta una variabile anomala o nascosta (religione, età, nazionalità, ...) può sempre reintrodurlo.