www.sce.unimore.it

Scienze della Comunicazione
e dell'Economia

INTRODUZIONE ALLA PROBABILITA

Legacy Edition
Copyright 27 febbraio 2018

Luca La Rocca
luca.larocca@unimore.it

UNIVERSITÀ DEGLI STUDI DI MODENA E REGGIO EMILIA

Il calcolo delle probabilità è una forma di logica dell'incerto sulle cui regole vi è, in buona sostanza, ampio consenso:

> *una formulazione assiomatica del calcolo delle probabilità è stata data da Kolmogorov (1933).*

Vi sono, al contrario, diversi punti di vista sul significato da dare al concetto di probabilità:

> *si assume qui il punto di vista soggettivo (de Finetti, 1970) in contrapposizione con quello frequentista (von Mises, 1928) e in continuità con quello classico (Laplace, 1812).*

L'origine del calcolo delle probabilità viene comunemente fatta risalire al 1654, anno in cui Antoine Gombaud, Chevalier de Méré, sottopose al matematico francese Blaise Pascal alcuni problemi relativi all'ambito dei giochi d'azzardo; Pascal scrisse al suo collega Pierre de Fermat per comunicargli le sue soluzioni ai problemi del Cavalier de Méré e questo diede origine a uno scambio epistolare che può considerarsi il primo approccio razionale sistematico a problemi nel dominio dell'incerto.

Precedentemente Girolamo Cardano (1501–1576) e Galileo Galilei (1564–1642) si erano interessati, in modo estemporaneo, a problemi concernenti il lancio di (tre) dadi.

Il Cavalier de Méré si chiedeva…

- è più facile per me ottenere almeno un ⚅ lanciando 4 volte un dado (a sei facce) oppure ottenere almeno un ⚅ ⚅ lanciando 24 volte due dadi (a sei facce)? non è forse ugualmente facile, visto che nel primo caso mi è favorevole 1 di 6 esiti possibili, mentre nel secondo caso 1 di 36 esiti possibili (e $4 : 6 = 24 : 36$)?
- se gioco contro un'altra persona al meglio di cinque partite, ne ho già vinte due e il mio avversario ne ha già vinta una, come dovremo dividerci la posta in palio nel caso in cui decidessimo di interrompere il gioco?

Il secondo è noto come problema "dei punti" o "della divisione della posta".

Per alcuni l'ambito dei giochi d'azzardo è di interesse pratico, per altri è fonte di curiosità intellettuale, per altri ancora è provvisto di scarso fascino. . . in ogni caso, quali che siano i gusti personali, si tratta di un buon punto di partenza per lo studio del calcolo delle probabilità:

- i problemi formulati nell'ambito dei giochi d'azzardo sono relativi a un contesto "sperimentale" dove tutto è noto (es. si conosce la composizione di un'urna da cui si estraggono delle biglie);
- i problemi formulati nell'ambito dei giochi d'azzardo sono spesso utili come modelli per situazioni "osservazionali" (es. l'estrazione di biglie da un'urna è un modello per il campionamento casuale).

Chiameremo esperimento aleatorio un qualsiasi processo il cui esito non sia prevedibile con certezza. . .

. . . per esempio il lancio di una moneta da due euro, avendo cura che dopo il lancio la moneta mostri inequivocabilmente una delle due facce (es. non affondi nella sabbia, né tantomento rotoli in un tombino. . .) mentre durante il lancio le due facce non siano fra loro distinguibili (in conseguenza del veloce ruotare della moneta). . .

. . . oppure l'estrazione di due biglie da un'urna contenente 5 biglie di cui 2 rosse e 3 bianche (senza reinserimento), avendo cura di non guardare durante l'estrazione (usando un'urna opaca) e di mantenere tutte le biglie alla stessa temperatura. . .

. . . sorvolando, nel seguito, su "ovvi" dettagli del processo.

Un evento è un ente logico il cui valore di verità (VERO o FALSO) dipende dall'esito di un esperimento aleatorio; diremo che

- l'evento E si verifica quando E è VERO,
- l'evento E non si verifica quando E è FALSO.

Nell'ambito di un (ben definito) esperimento aleatorio, definiremo un evento mediante una frase non ambigua (che permetta di concordare una scommessa sul verificarsi dell'evento); per esempio definiremo

$$E = \text{“ottengo testa”}$$

nell'esperimento del lancio di una moneta da due euro, chiamando “testa” il lato della moneta con Dante Alighieri (e “croce” l'altro)...

. . . oppure definiremo gli eventi

$$E = \text{“le due biglie estratte hanno lo stesso colore”}$$
$$F = \text{“almeno una delle biglie estratte è rossa”}$$

nell'esperimento dell'estrazione di due biglie da un'urna contenente 5 biglie di cui 2 rosse e 3 bianche.

Si noti che non stiamo richiedendo la “ripetibilità nelle stesse condizioni” dell'esperimento aleatorio (necessaria dal punto di vista frequentista).

Altri esempi di eventi (esperimenti aleatori) sono:

E_1 = "ottengo ⚅" (lancio di un dado a sei facce);

E_2 = "lo scarto dalla media è inferiore a un'unità standard" (misura dell'IQ di un individuo estratto a caso da una popolazione con media e deviazione standard note);

E_3 = "il Barcelona vince la coppa" (Champions League 2011/12).

La Champions League 2011/12 fornisce un esempio di esperimento non ripetibile nelle stesse condizioni; ne segue che E_3 non sarebbe un evento valido, se adottassimo il punto di vista frequentista.

Due eventi speciali (definiti per ogni esperimento aleatorio) sono:

- l'evento impossibile $\emptyset$, sempre FALSO;
- l'evento certo Ω, sempre VERO.

Diverse formulazioni sono possibili (nell'ambito di uno stesso esperimento aleatorio):

$$\begin{aligned}
\emptyset &= \text{“la prima biglia estratta è nera” (urna)} \\
\emptyset &= \text{“la prima biglia estratta è gialla” (urna)} \\
\Omega &= \text{“ottengo un punteggio minore di sette” (dado)} \\
\Omega &= \text{“ottengo un punteggio maggiore di zero” (dado)}
\end{aligned}$$

Si noti che la nozione di evento certo/impossibile dipende in generale dallo stato delle conoscenze del soggetto che considera l'evento.

Per esempio (Scozzafava, 2000) dell'evento

$$E = \text{"L'anno della morte di Garibaldi è il 1882"},$$

nell'esperimento "storia dell'umanità" si può dire che

- nel 1870 non era né l'evento certo, né l'evento impossibile (era, come si dice, un evento possibile);
- oggi è l'evento certo (per chi sa che, in effetti, Garibaldi è morto nel 1882).

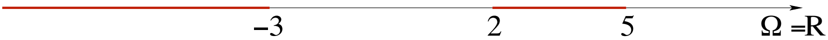
Per un dato esperimento aleatorio, identificheremo l'evento certo con un opportuno insieme ambiente i cui elementi rappresentino i possibili esiti elementari dell'esperimento, di modo che ogni evento di interesse sia a sua volta identificato da una parte (sottoinsieme) di Ω (quella formata dagli esiti elementari che lo rendono VERO); troveremo inoltre utile precisare una famiglia di eventi $\mathcal{A}$ come oggetto della nostra attenzione.

Per esempio, se lanciamo un dado a sei facce, possiamo prendere $\Omega = \{1, 2, 3, 4, 5, 6\}$ e $\mathcal{A}$ sarà formata dalle $2^6 = 64$ parti di Ω:

$$\begin{aligned} \text{DISPARI} &= \{1, 3, 5\} \\ \text{UNO} &= \{1\} \\ \emptyset &= \{\} \qquad \text{(insieme vuoto)} \\ &\cdots \end{aligned}$$

Se misuriamo l'IQ (scarto in unità standard dalla media di popolazione) di un individuo preso a caso, possiamo prendere $\Omega = \mathbb{R}$ (la retta reale) e definire $\mathcal{A}$ come la famiglia dei plurintervalli (unioni finite di intervalli):

$$E_1 \quad = \quad (-\infty, -3) \cup (2, 5)$$



$$E_2 = (-1, 1)$$

$$\ldots$$

Ci si limita ai plurintervalli perché la retta reale è troppo "ricca" di parti.

Nell'esperimento della Champions League 2011/12, a valle degli ottavi di finale e a monte dei quarti di finale (es. alla data del 15 marzo 2012), le squadre ancora in gara sono. . .

APOEL (CYP)
Barcelona (ESP)
Bayern (GER)
Benfica (POR)
Chelsea (ENG)
Marseille (FRA)
Milan (ITA)
Real Madrid (ESP)

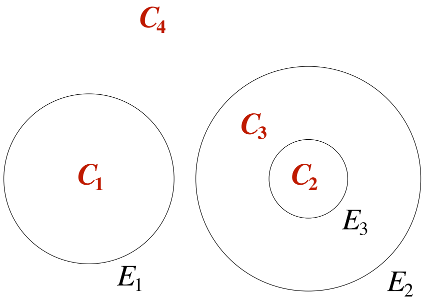
. . . e il loro elenco può fare da insieme ambiente. . .

…oppure, se interessano solo gli eventi

$$
\begin{aligned}
E_1 &= \text{“una squadra italiana vince la coppa”} \\
E_2 &= \text{“una squadra spagnola vince la coppa”} \\
E_3 &= \text{“il Barcelona vince la coppa”},
\end{aligned}
$$

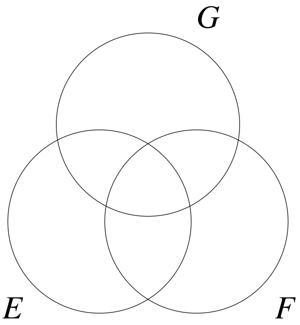
possiamo prendere come insieme ambiente il piano euclideo (per comodità ma senza un'interpretazione diretta) e identificare ogni evento di interesse con una sua regione:

> *$\mathcal{A}$ sarà la famiglia di tutte le unioni finite dei loro costituenti (intersezioni non vuote di queste regioni e dei loro complementi al piano Ω) rappresentati nel seguente diagramma di Venn…*



C_1 = "italiana", C_2 = "Barcelona", C_3 = "altra spagnola", C_4 = "altro paese"
$(E_1 = C_1,\ E_2 = C_2 \cup C_3,\ E_3 = C_2)$

Tre eventi (E, F e G) logicamente indipendenti



In generale 3 eventi danno luogo a un massimo di $2^3 = 8$ costituenti, nel qual caso si dicono logicamente indipendenti (questa situazione è raffigurata nel diagramma di Venn precedente).

Non è chiaramente il caso della Champions League 2011/12, visto che abbiamo solo 4 costituenti:

- E_1 ed E_2 sono mutuamente esclusivi (è impossibile che la squadra vincente sia allo stesso tempo italiana e spagnola);
- E_3 implica E_2 (se il Barcelona vince, allora ha vinto una squadra spagnola).

Secondo la definizione classica (nata nell'ambito dei giochi d'azzardo) la probabilità di un evento E è il rapporto tra il numero di casi favorevoli a E (esiti elementari per i quali E è VERO) e il numero di casi totali (esiti elementari in Ω):

$$\mathbb{P}(E) = \frac{\#E}{\#\Omega}.$$

Per esempio, con riferimento al lancio di un dado a sei facce, si trova

$$\begin{aligned} \mathbb{P}(DUE) &= \mathbb{P}\{2\} = 1/6 \simeq 17\% \\ \mathbb{P}(ALTO) &= \mathbb{P}\{4,5,6\} = 3/6 = 50\%. \end{aligned}$$

Inoltre (in qualsiasi esperimento) si ha $\mathbb{P}(\Omega) = 1$ e $\mathbb{P}(\emptyset) = 0$, mentre in generale si ottengono valori intermedi (cioè compresi tra zero e uno).

Similmente, se estraiamo una carta da un mazzo francese, troveremo

$$\mathbb{P}(\text{“carta di cuori”}) = \frac{13}{52} = \frac{1}{4} = 25\%$$

perché ci sono in tutto 52 carte nel mazzo, di cui 13 di cuori, e

$$\mathbb{P}(\text{“carta vestita”}) = \frac{16}{52} = \frac{4}{13} \simeq 31\%$$

se per carta vestita intendiamo fante, donna, re o asso (ce ne sono quattro per ognuno dei quattro semi).

D’altra parte, se invece di estrarre una carta da un mazzo estraiamo un individuo da una popolazione, possiamo valutare empiricamente la probabilità di osservare una certa caratteristica mediante la frequenza relativa (percentuale) di tale caratteristica nella popolazione da cui effettuiamo l’estrazione.

Per esempio, se il 60% degli abbonati a una certa rivista è donna, questa sarà anche la probabilità di selezionare una donna, quando si scelga a caso un lettore dalla lista degli abbonati a tale rivista.

La definizione classica di probabilità dipende dalla scelta dell'evento certo: se per esempio, con riferimento al lancio di un dado a sei facce, ci viene detto che è uscito un numero pari, allora prenderemo $\Omega = \{2, 4, 6\}$ e troveremo $\mathbb{P}\{2\} = 1/3$.

Più in generale, per evitare di cambiare Ω ogni volta che cambiano le informazioni disponibili, si definisce la probabilità condizionata di un evento F dato l'evento possibile E (un evento E tale che $\#E > 0$) come

$$\mathbb{P}(F|E) = \frac{\#F \cap E}{\#E}$$

ritrovando la probabilità incondizionata come probabilità condizionata all'evento certo (fissato una volta per tutte): $\mathbb{P}(F) = \mathbb{P}(F|\Omega)$.

Per esempio (con riferimento al lancio di un dado a sei facce) troveremo

$$\mathbb{P}(DUE|PARI) = \frac{\#\{2\} \cap \{2,4,6\}}{\#\{2,4,6\}} = \frac{\#\{2\}}{\#\{2,4,6\}} = 1/3$$

per la probabilità di *DUE* sapendo che si è verificato *PARI* (come già visto cambiando l'evento certo) e

$$\mathbb{P}(ALTO|PARI) = \frac{\#\{4,5,6\} \cap \{2,4,6\}}{\#\{2,4,6\}} = \frac{\#\{4,6\}}{\#\{2,4,6\}} = 2/3$$

per la probabilità di *ALTO* sapendo che si è verificato *PARI*; si noti che il primo caso è meno generale perché $DUE \subset PARI$.

Quali relazioni sussistono (nell'ambito della definizione classica) tra le probabilità (condizionate) dei diversi eventi che formano la famiglia di interesse $\mathcal{A}$ (in un dato esperimento)?

Risponderemo mediante l'introduzione di alcune fondamentali operazioni su eventi, giungendo a formulare le regole del calcolo delle probabilità. . .

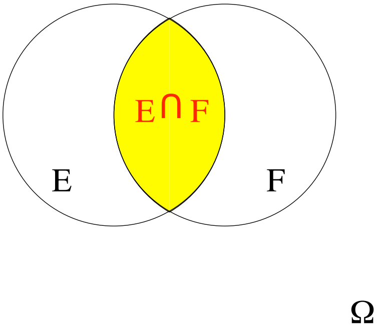
Il prodotto logico di E ed F è l'evento $E \cap F$ che è VERO se e solo se sia E che F sono VERO (congiunzione "e", intersezione insiemistica).

Es. nella misura dell'IQ di un individuo preso caso si ha

$$(-\infty, 1) \cap (-1, \infty) \quad = \quad (-1, 1)$$

e nel lancio di un dado a sei facce si trova

$$\begin{aligned} \mathit{ALTO} \cap \mathit{DUE} &= \{4,5,6\} \cap \{2\} = \{\} = \emptyset \\ \mathit{ALTO} \cap \mathit{DISPARI} &= \{4,5,6\} \cap \{1,3,5\} = \{5\}. \end{aligned}$$



E	F	$E \cap F$
VERO	VERO	VERO
VERO	FALSO	FALSO
FALSO	VERO	FALSO
FALSO	FALSO	FALSO

La probabilità condizionata (definita classicamente) soddisfa la regola del prodotto:

$$\mathbb{P}(E \cap F) = \mathbb{P}(F|E) \cdot \mathbb{P}(E)$$

come si verifica direttamente dalla definizione (moltipicandone entrambi i membri per $\#E/\#\Omega$).

Si noti che quando E è l'evento impossibile ($\#E = 0$) la probabilità condizionata di F dato E non è definita, ma la regola del prodotto può considerarsi ugualmente soddisfatta (in modo banale) perché sarebbe soddisfatta da qualunque valore di $\mathbb{P}(F|E)$.

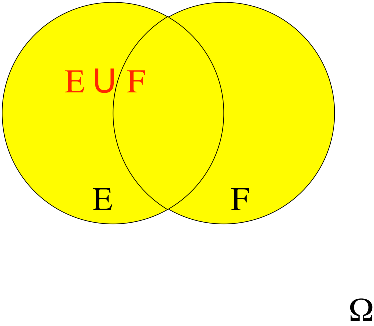
La somma logica di E ed F è l'evento $E \cup F$ che è VERO se e solo se almeno uno tra E ed F è VERO ("o" inclusivo, unione insiemistica).

Es. nella misura dell'IQ di un individuo preso caso si ha

$$(-\infty, -1) \cup (-1, 1) \quad = \quad (-\infty, 1)$$

e nel lancio di un dado a sei facce si trova

$$\begin{aligned} \mathit{ALTO} \cup \mathit{DUE} &= \{4,5,6\} \cup \{2\} &= \{2,4,5,6\} \\ \mathit{ALTO} \cup \mathit{DISPARI} &= \{4,5,6\} \cup \{1,3,5\} &= \{1,3,4,5,6\}. \end{aligned}$$



E	F	$E \cup F$
VERO	VERO	VERO
VERO	FALSO	VERO
FALSO	VERO	VERO
FALSO	FALSO	FALSO

Se E ed F sono mutuamente esclusivi ($E \cap F = \emptyset$) non vi sono esiti elementari che rendono VERO entrambi, quindi $\#(E \cup F) = \#E + \#F$ e $\mathbb{P}(E \cup F) = \mathbb{P}(E) + \mathbb{P}(F)$: la regola della somma.

Per esempio, nell'esperimento del dado a sei facce, troviamo

$$\mathbb{P}\{2,4,5,6\} = \frac{4}{6} = \frac{1}{6} + \frac{3}{6} = \mathbb{P}\{2\} + \mathbb{P}\{4,5,6\}$$

mentre, nell'esempio del mazzo francese, si ha

$$\mathbb{P}(\text{“cuori o vestita”}) = \frac{25}{52} \neq \frac{13}{52} + \frac{16}{52} = \mathbb{P}(\text{“cuori”}) + \mathbb{P}(\text{“vestita”})$$

perché tra le 16 carte vestite ce ne sono 4 di cuori.

In generale ($E \cap F \neq \emptyset$) troviamo $\#(E \cup F) = \#E + \#F - \#E \cap F$ e quindi

$$\mathbb{P}(E \cup F) = \mathbb{P}(E) + \mathbb{P}(F) - \mathbb{P}(E \cap F).$$

Per esempio, nell'esperimento dell'estrazione di una carta da un mazzo francese, troviamo

$$\begin{aligned}
\mathbb{P}(\text{“cuori o vestita”}) &= \frac{25}{52} \\
&= \frac{13}{52} + \frac{16}{52} - \frac{4}{52} \\
&= \mathbb{P}(\text{“cuori”}) + \mathbb{P}(\text{“vestita”}) - \mathbb{P}(\text{“cuori e vestita”}).
\end{aligned}$$

La negazione di E è l'evento E^c, o $\overline{E}$, che è VERO quando E è FALSO e FALSO quando E è VERO ("non", complementazione insiemistica).

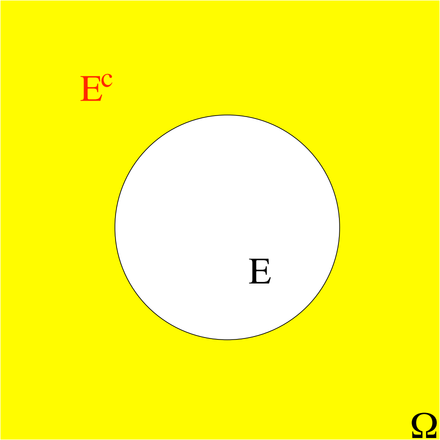
Es. nella misura dell'IQ di un individuo preso caso si ha

$$(-1,1)^c = (-\infty,-1)\cup(1,\infty)$$

e nel lancio di un dado a sei facce si trova

$$\begin{aligned}\overline{DISPARI} &= \{1,3,5\}^c &= \{2,4,6\} &= PARI\\ \overline{UNO} &= \{1\}^c &= \{2,3,4,5,6\} &= NONUNO\\ \overline{\emptyset} &= \{\}^c &= \{1,2,3,4,5,6\} &= \Omega.\end{aligned}$$

Si noti che la negazione è idempotente: $(E^c)^c = E$.



E	E^c
VERO	FALSO
FALSO	VERO

Poiché $\mathbb{P}(\Omega) = 1$ e per qualunque evento E

$$\begin{aligned} E \cap E^c &= \emptyset \\ E \cup E^c &= \Omega \end{aligned}$$

la regola della somma fornisce $\mathbb{P}(E) + \mathbb{P}(E^c) = 1$ e dunque ha come corollario la regola della negazione:

$$\mathbb{P}(E^c) = 1 - \mathbb{P}(E).$$

Per esempio, oltre che banalmente $\mathbb{P}(\emptyset) = 0 = 1 - 1 = 1 - \mathbb{P}(\Omega)$, si ha

$$\begin{aligned} \mathbb{P}(\text{PARI}) &= \frac{1}{2} = 1 - \frac{1}{2} = 1 - \mathbb{P}(DISPARI) \\ \mathbb{P}(\text{NONUNO}) &= \frac{5}{6} = 1 - \frac{1}{6} = 1 - \mathbb{P}(UNO). \end{aligned}$$

La definizione classica ci fornisce dunque dei valori di probabiltà tra 0 (evento impossibile) e 1 (evento certo) che misurano, alla luce delle informazioni disponibili, la propensione degli eventi a verificarsi, soddisfacendo alcune regole: possiamo usarla in ogni esperimento?

No, come ci accingiamo a verificare, ma possiamo “salvare” le regole (approccio assiomatico) e anche motivarle (definizione soggettiva); questa è la strada che seguiremo e che ci porterà a considerare la definizione classica come un caso particolare...

Preliminarmente conviene supporre che operando con gli eventi di $\mathcal{A}$ si ottengano ancora eventi di $\mathcal{A}$ (ovvero che $\mathcal{A}$ sia un'algebra):

in particolare $\mathcal{A}$ risulta essere un'algebra quando

- *$\mathcal{A}$ è la famiglia delle parti di un insieme finito*
- *$\mathcal{A}$ è la famiglia dei plurintervalli della retta reale*
- *$\mathcal{A}$ è la famiglia delle unioni finite dei costituenti di alcuni eventi (supposti in numero finito)*

e dunque in tutti e tre i casi sinora considerati.

Vale la pena osservare che l'algebra delle unioni finite dei costituenti di alcuni eventi (supposti in numero finito) è la più piccola algebra che contenga tali eventi e si dice algebra generata dagli eventi in questione.

Se misuriamo l'IQ di un individuo preso a caso:

- abbiamo un infinità di esiti elementari in $\Omega = \mathbb{R}$ (stiamo supponendo di conoscere solo la media e la deviazione standard di popolazione, ignorando i valori individuali);
- non sembra ragionevole che gli eventi elementari $\{\omega\}$, $\omega \in \mathbb{R}$, siano equiprobabili, mentre la definizione classica comporta che lo siano, avendo tutti probabilità $\mathbb{P}\{\omega\} = 1/\#\Omega$ (es. $\mathbb{P}\{1\} = \mathbb{P}\{2\} = \mathbb{P}\{3\} = \mathbb{P}\{4\} = \mathbb{P}\{5\} = \mathbb{P}\{6\} = 1/6$ nel caso del lancio di un dado a sei facce).

Se consideriamo la Champions League 2011/12, i costituenti possono vedersi come un numero finito di eventi elementari, ma

- non sembra ragionevole ritenerli equiprobabili (es. $\mathbb{P}$(“altro paese”) $=$ $\mathbb{P}$(“Barcelona”) significa che il Barcelona vale da solo tutte le squadre né italiane né spagnole in gara);
- non sembra ragionevole che diversi soggetti debbano necessariamente concordare su un’unica valutazione di probabilità (es. qualcuno potrebbe ritenere che il Barcelona vinca con probabilità 20%, qualcun altro che vinca con probabilità 10%).

Si noti che persino sul semplice lancio di una moneta può esserci disaccordo: es. $\mathbb{P}$(“testa”) $= 1/2$ esprime la mia convinzione che la moneta sia onesta, laddove $\mathbb{P}$(“testa”) $= 1$ esprime la vostra consapevolezza che la moneta ha due teste…

. . . si giunge così al punto di vista soggettivo, secondo il quale la probabilità di un evento misura il grado di fiducia di un soggetto nel verificarsi dell'evento (non la propensione dell'evento a verificarsi).

Nel seguito, assumendo il punto di vista soggettivo, si darà una definizione operativa di probabilità, in termini di scommesse, tale da fornire un significato alla nozione di probabilità (cosa che la definizione classica non fa) e da determinarne le regole (che abbiamo già visto e su cui comunque c'è accordo).

Ritroveremo la definizione classica di probabilità come strumento per assegnare in modo univoco dei valori di probabilità quando via sia un numero finito di eventi elementari equiprobabili; si noti che per giudicarli tali dobbiamo già possedere il concetto di (equi)probabilità.

Secondo la definizione soggettiva di probabilità, data un'algebra di eventi $\mathcal{A}$, diremo che $\mathbb{P}$ è la mia valutazione di probabilità su $\mathcal{A}$ se sono pronto a scommettere contro di voi (o chiunque altro) su tutti gli eventi $E \in \mathcal{A}$ adottando le seguenti regole:

1. voi scegliete un importo S e mi pagate $\mathbb{P}(E) \cdot S$;
2. io vi pago S se e solo se E si verifica.

N.B. Potete anche scegliere $S < 0$, in modo da scambiare i ruoli (io pago con certezza, voi pagate se e solo se E si verifica).

Per esempio, nell'esperimento della Champions League 2011/12, potrei valutare $\mathbb{P}(\text{“vincitrice italiana”}) = 15\%$: potreste allora anticiparmi 15 euro per il diritto a riceverne 100 in caso di vittoria italiana, oppure chiedermi 15 euro certi a fronte della vostra disponibilità a pagarne 100 incerti.

Se non voglio regalarvi denaro, la mia valutazione di probabilità deve essere coerente: qualunque combinazione di scommesse scegliate, deve esservi impossibile ottenere un guadagno certo; ne seguono le regole del calcolo delle probabilità.

Se scommettete per un importo S sull'evento $E \in \mathcal{A}$ e $\mathbb{P}$ è la mia valutazione di probabilità su $\mathcal{A}$, il vostro guadagno sarà

$$G(E, S, \mathbb{P}) = \begin{cases} (1 - \mathbb{P}(E)) \cdot S & \text{se } E \text{ è VERO} \\ -\mathbb{P}(E) \cdot S & \text{se } E \text{ è FALSO} \end{cases}$$

di modo che, per coerenza, deve aversi $-\mathbb{P}(E) \cdot (1 - \mathbb{P}(E)) \cdot S^2 \leq 0$ e quindi

$$0 \leq \mathbb{P}(E) \leq 1.$$

Resta dunque stabilito che la probabilità di un evento è un numero tra zero e uno; in particolare coerenza vuole che

$$\begin{aligned} \mathbb{P}(\Omega) &= 1 \\ \mathbb{P}(\emptyset) &= 0 \end{aligned}$$

perché Ω si verifica sempre, mentre $\emptyset$ non si verifica mai (in entrambi i casi il guadagno è certo e deve annullarsi).

La probabilità di un evento è dunque la quota dell'importo in palio che dovete contribuire per scommettere contro di me sull'evento (tanto maggiore quanto più io ritengo che l'evento si verificherà).

Supponete ora di scommettere per un importo S su due eventi mutuamente esclusivi di $\mathcal{A}$: $E \in \mathcal{A}$ e $F \in \mathcal{A}$ con $E \cap F = \emptyset$.

Poiché E ed F non possono verificarsi assieme, state pagando $\mathbb{P}(E) \cdot S + \mathbb{P}(F) \cdot S$ per ottenere S se uno tra E ed F si verifica.

Questo per voi equivale a scommettere per l'importo S su $E \cup F$, nel qual caso paghereste $\mathbb{P}(E \cup F) \cdot S$.

Quindi, per coerenza, devo porre

$$\mathbb{P}(E \cup F) = \mathbb{P}(E) + \mathbb{P}(F)$$

ovvero la mia $\mathbb{P}$ deve soddisfare la regola della somma.

Per esempio, con riferimento alla Champions League 2011/12, se valuto

$$\begin{aligned} \mathbb{P}(E_1) &= \mathbb{P}(\text{“vincitrice italiana”}) &= 15\% \\ \mathbb{P}(E_2) &= \mathbb{P}(\text{“vincitrice spagnola”}) &= 40\% \end{aligned}$$

allora devo necessariamente valutare $\mathbb{P}(E_1 \cup E_2) = 55\%$; altrimenti, se per fissare le idee valuto

$$\mathbb{P}(E_1 \cup E_2) = \mathbb{P}(\text{“vincitrice italiana o spagnola”}) = 60\%,$$

voi potete scommettere per un importo $S = 100$ euro su E_1 ed E_2 (pagando $15 + 40 = 55$ euro) e per un importo $S = -100$ su $E_1 \cup E_2$ (incassando 60 euro) in modo da <u>guadagnare certamente</u> 5 euro.

Si dimostra che, data un'algebra di eventi $\mathcal{A}$, qualsiasi $\mathbb{P}$ tale che

(I) $\mathbb{P}(E) \geq 0$, per ogni $E \in \mathcal{A}$

(II) $\mathbb{P}(\Omega) = 1$

(III) $\mathbb{P}(E \cup F) = \mathbb{P}(E) + \mathbb{P}(F)$, se $E \cap F = \emptyset$

è una valutazione di probabilità coerente su $\mathcal{A}$.

Si noti che $\mathbb{P}(E) \leq 1$ e $\mathbb{P}(\emptyset) = 0$ seguono da (I) e (II) una volta che si sia ricavata la regola della negazione $\mathbb{P}(E) = 1 - \mathbb{P}(E^c)$ da (II) e (III).

Quindi <u>la regola della somma e l'avere zero e uno come valori limite</u> è tutto quanto si richiede a una valutazione di probabilità coerente…

…rinviando per il momento la discussione del condizionamento.

Le proprietà (I), (II) e (III) precedentemente elencate sono, in buona sostanza, gli assiomi della probabilità di Kolmogorov (1933).

L'unica differenza è che per Kolomogorov (1933) la famglia $\mathcal{A}$ deve essere una σ-algebra (deve contenere le unioni numerabili di suoi elementi) e la probabilità $\mathbb{P}$ deve essere σ-additiva (la regola della somma deve valere anche per unioni numerabili di eventi).

La σ-additività (additività completa) è un'ipotesi di continuità che non trova giustificazione nella definizione soggettiva di probabilità, ma permette di sviluppare una teoria matematica più ricca.

Come si può assegnare una valutazione di probabilità coerente su una data algebra di eventi?

Se Ω è identificato da un insieme finito, una valutazione di probabilità coerente sull'algebra delle parti di Ω può ottenersi assegnando le probabilità degli eventi elementari (singoletti):

- qualsiasi scelta di numeri positivi che sommino a uno (la probabilità dell'evento certo) garantirà la coerenza;
- la regola della somma darà la probabilità di ogni $E \in \mathcal{A}$ (somma logica di un certo numero di eventi elementari).

Per esempio, con riferimento al lancio di un dado, si ha

$$\begin{aligned} \Omega &= \{1\} \cup \{2\} \cup \{3\} \cup \{4\} \cup \{5\} \cup \{6\} \\ \mathit{PARI} &= \{2\} \cup \{4\} \cup \{6\} \end{aligned}$$

di modo che, <u>scelti</u> $\mathbb{P}\{1\}$, $\mathbb{P}\{2\}$, $\mathbb{P}\{3\}$, $\mathbb{P}\{4\}$, $\mathbb{P}\{5\}$, $\mathbb{P}\{6\}$ positivi con somma $\mathbb{P}(\Omega) = 1$, risulta $\mathbb{P}(\mathit{PARI}) = \mathbb{P}\{2\} + \mathbb{P}\{4\} + \mathbb{P}\{6\}$.

Nell'ipotesi di dado bilanciato, sembra ragionevole supporre gli eventi elementari equiprobabili e porre

$$\mathbb{P}\{1\} = \mathbb{P}\{2\} = \mathbb{P}\{3\} = \mathbb{P}\{4\} = \mathbb{P}\{5\} = \mathbb{P}\{6\} = 1/6$$

ritrovando la valutazione classica come regola di calcolo.

Se invece, per esempio, sappiamo che si tratta di un dado caricato in modo da mostrare più spesso la faccia ⚅, potremmo porre

$$
\begin{aligned}
\mathbb{P}\{1\} &= 6\% \\
\mathbb{P}\{2\} = \mathbb{P}\{3\} = \mathbb{P}\{4\} = \mathbb{P}\{5\} &= 11\% \\
\mathbb{P}\{6\} &= 50\%
\end{aligned}
$$

e trovare, per esempio,

$$\mathbb{P}\{2, 4, 6\} = \mathbb{P}\{2\} + \mathbb{P}\{4\} + \mathbb{P}\{6\} = 0.11 + 0.11 + 0.50 = 0.72.$$

Sia questa valutazione che quella classica sono coerenti: quale sia quella “giusta” (quella che, scommettendo, non fa perdere denaro) dipende dal dado che stiamo lanciando.

Se interessa un numero finito di eventi, una valutazione di probabilità coerente sull'algebra da questi generata (la più piccola algebra che li contiene) può ottenersi assegnando le probabilità dei loro costituenti; anche in questo caso

- qualsiasi scelta di numeri positivi che sommino a uno (la probabilità dell'evento certo) garantirà la coerenza;
- la regola della somma darà la probabilità di ogni $E \in \mathcal{A}$ (somma logica di un certo numero di costituenti).

In pratica è la stessa situazione dell'insieme ambiente finito, con i costituenti nel ruolo degli eventi elementari (in modo da trattare situazioni più generali).

Per esempio, con riferimento alla Champions League 2011/12, si può porre

$$\begin{array}{rcll} \mathbb{P}(C_1) & = & \mathbb{P}(\text{“italiana”}) & = & 15\% \\ \mathbb{P}(C_2) & = & \mathbb{P}(\text{“Barcelona”}) & = & 20\% \\ \mathbb{P}(C_3) & = & \mathbb{P}(\text{“altra spagnola”}) & = & 20\% \\ \mathbb{P}(C_4) & = & \mathbb{P}(\text{“altro paese”}) & = & 45\% \end{array}$$

di modo che, per l'evento “vittoria spagnola”, si trova

$$\begin{array}{rcccl} \mathbb{P}(E_2) & = & \mathbb{P}(C_2 \cup C_3) & = & \mathbb{P}(C_2) + \mathbb{P}(C_3) \\ & & & = & 0.20 + 0.20 \quad = \quad 0.40; \end{array}$$

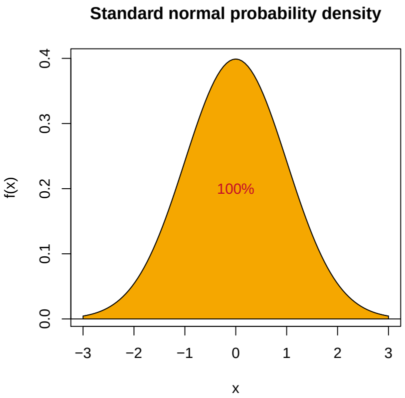
si noti che in questo caso non c'è una valutazione convenzionale (come invece accadeva nel caso del dado a sei facce).

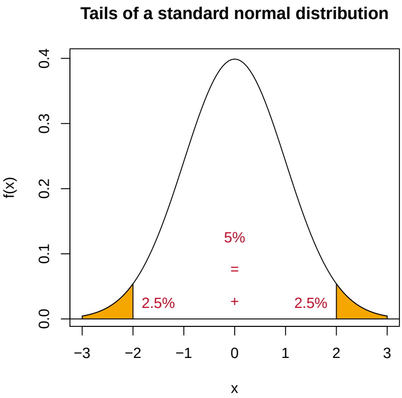
Infine, se Ω è identificato dalla retta reale, ovvero $\Omega = \mathbb{R}$, una valutazione di probabilità coerente sull'algebra dei plurintervalli di $\mathbb{R}$ può ottenersi assegnando una funzione di densità:

- qualsiasi funzione $f : \mathbb{R} \to \mathbb{R}$ tale che
  (i) $f(x) \geq 0$, per ogni $x \in \mathbb{R}$
  (ii) $\int_{-\infty}^{\infty} f(x)\, dx = 1$

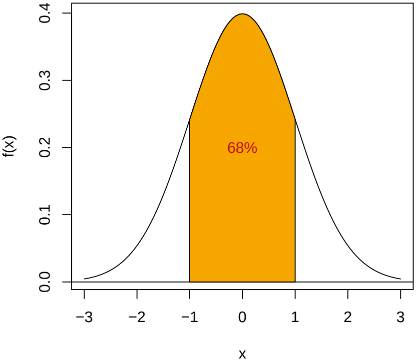
  garantirà la coerenza;
- per ogni intervallo (a, b) di $\mathbb{R}$ si troverà $\mathbb{P}(a, b) = \int_a^b f(x)\, dx$, pari all'area sottesa a f tra a e b, e la regola della somma darà la probabilità di ogni plurintervallo.

Questo approccio consente di trattare il caso dell'IQ (in unità standard) di un individuo preso a caso da una popolazione di cui si conoscano media e deviazione standard. . .


Standard normal probability density
f(x)
0.0
0.1
0.2
0.3
0.4
100%
−3
−2
−1
0
1
2
3
x


Tails of a standard normal distribution
f(x)
0.0
0.1
0.2
0.3
0.4
5%
=
2.5%
+
2.5%
−3
−2
−1
0
1
2
3
x

$$\mathbb{P}(-1,+1) = \int_{-1}^{+1} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx \simeq 0.6827 = 68.27\%$$



Una valutazione di probabilità, $\mathbb{P}$, sui plurintervalli della retta reale, $\mathbb{R}$, definita da una funzione di densità, f, dà probabilità nulla a tutti gli eventi elementari (anche se uno di essi, alla fine, si verificherà):

$$\mathbb{P}\{a\} = \int_a^a f(x)\,dx = 0, \qquad \text{per ogni } a \in \mathbb{R}.$$

Tuttavia, se $h > 0$ è "piccolo", si ha

$$\mathbb{P}(a-h, a+h) = \int_{a-h}^{a+h} f(x)\,dx \simeq 2hf(a)$$

di modo che, se per es. $f(b) > f(a)$, l'esito dell'esperimento è più probabilmente in un intorno di b che in un intorno di a (a parità di lunghezza dell'intorno): f è una probabilità per unità di lunghezza.

Da un punto di vista soggettivo la mia probabilità condizionata di F dato E esprime il mio grado di fiducia nel verificarsi di F al verificarsi di E ed è definita da una procedura di scommessa condizionata:

1. voi scegliete un importo S e mi pagate $\mathbb{P}(F|E) \cdot S$;
2. se E non si verifica, vi restituisco $\mathbb{P}(F|E) \cdot S$;
3. se E si verifica, vi pago S se e solo se F si verifica.

Per esempio, con riferimento all'esperimento della Champions League 2011/12, se $\mathbb{P}(\text{“Barcelona”}|\text{“Spagnola”}) = 50\%$, potete puntare 50 euro sul Barcelona, riaverli indietro se non vince una squadra spagnola e riceverne 100 se il Barcelona vince la coppa, ma li perderete se la coppa è vinta da un'altra squadra spagnola; oppure ($S < 0$) potete chiedermi di invertire i ruoli.

Coerenza vuole che, per E fissato, la valutazione $\mathbb{P}(F|E)$, $F \in \mathcal{A}$, soddisfi gli assiomi della probabilità:

(I) $\mathbb{P}(F|E) \geq 0$;

(II) $\mathbb{P}(\Omega|E) = 1$;

(III) $\mathbb{P}(F_1 \cup F_2|E) = \mathbb{P}(F_1|E) + \mathbb{P}(F_2|E)$, se $F_1 \cap F_2 = \emptyset$.

Inoltre, una volta che mi avete pagato $\mathbb{P}(F|E) \cdot S$, si verificherà uno e uno solo dei seguenti eventi:

E^c, nel qual caso vi pagherò $\mathbb{P}(F|E) \cdot S$;

$E \cap F$, nel qual caso vi pagherò S;

$E \cap F^c$, nel qual caso non vi pagherò affatto.

Potete pertanto mettervi nelle stesse condizioni scommettendo su $E \cap F$ per un importo S e, allo stesso tempo, su E^c per un importo $\mathbb{P}(F|E) \cdot S$; quindi, per coerenza, devo offrirvi lo stesso prezzo d'ingresso:

$$\mathbb{P}(F|E) \cdot S \quad = \quad \mathbb{P}(E \cap F) \cdot S + \mathbb{P}(E^c) \cdot \mathbb{P}(F|E) \cdot S.$$

Semplificando S e sostituendo $\mathbb{P}(E^c) = 1 - \mathbb{P}(E)$ si trova

$$\mathbb{P}(F|E) \quad = \quad \mathbb{P}(E \cap F) + \mathbb{P}(F|E) - \mathbb{P}(E) \cdot \mathbb{P}(F|E)$$

e quindi la regola del prodotto

$$\mathbb{P}(E \cap F) \quad = \quad \mathbb{P}(F|E) \cdot \mathbb{P}(E).$$

Si noti che dalla regola del prodotto segue l'identità $\mathbb{P}(F|\Omega) = \mathbb{P}(F)$, analogamente a quanto accadeva in ambito classico.

Se $\mathbb{P}(E) > 0$, si può riscrivere la regola del prodotto come

$$\mathbb{P}(F|E) = \frac{\mathbb{P}(E \cap F)}{\mathbb{P}(E)}$$

ricavando la probabilità condizionata da quella incondizionata.

Nell'approccio assiomatico questa riscrittura viene presa come definizione di probabilità condizionata (e la regola del prodotto ne consegue); si noti che in questo caso gli assiomi della probabilità (come funzione di F) sono automaticamente soddisfatti.

Vi sono due possibili usi della regola del prodotto:

- per determinare la probabilità condizionata $\mathbb{P}(F|E)$ dalla congiunta $\mathbb{P}(E \cap F)$ e dalla marginale $\mathbb{P}(E) > 0$;
- per determinare la probabilità congiunta $\mathbb{P}(E \cap F)$ dalla marginale $\mathbb{P}(E)$ e dalla condizionata $\mathbb{P}(F|E)$.

Il primo uso si adatta al caso in cui interessino diversi aspetti di una singola "estrazione"; per esempio troveremo

$$\mathbb{P}((-1,1)|(0,\infty)) = \frac{\mathbb{P}(0,1)}{\mathbb{P}(0,\infty)} = \frac{0.34}{0.50} = 0.68$$

se $\mathbb{P}$ è assegnata mediante una funzione di densità normale standard.

Il secondo uso si adatta invece al caso di più "estrazioni" successive...

Con riferimento all'estrazione senza reinserimento di due biglie da un'urna contenente 5 biglie, di cui 2 rosse e 3 bianche, si considerino gli eventi

$$\begin{aligned} R_1 &= \text{“la prima biglia estratta è rossa”} \\ R_2 &= \text{“la seconda biglia estratta è rossa”.} \end{aligned}$$

Valutando classicamente prima l'urna iniziale e poi l'urna residua (supponendo di avere estratto una biglia rossa) assegneremo

$$\begin{aligned} \mathbb{P}(R_1) &= 2/5 = 40\% \\ \mathbb{P}(R_2|R_1) &= 1/4 = 25\%. \end{aligned}$$

Possiamo allora usare la regola del prodotto per calcolare la probabilità del prodotto logico $R_1 \cap R_2 =$ "entrambe rosse":

$$\mathbb{P}(R_1 \cap R_2) = \mathbb{P}(R_2 | R_1) \cdot \mathbb{P}(R_1) = \frac{1}{4} \times \frac{2}{5} = \frac{1}{10}.$$

Analogamente troviamo

$$\begin{array}{rcccccc} \mathbb{P}(R_1 \cap R_2^c) & = & \mathbb{P}(R_2^c | R_1) \cdot \mathbb{P}(R_1) & = & \frac{3}{4} \times \frac{2}{5} & = & \frac{3}{10} \\ \mathbb{P}(R_1^c \cap R_2) & = & \mathbb{P}(R_2 | R_1^c) \cdot \mathbb{P}(R_1^c) & = & \frac{2}{4} \times \frac{3}{5} & = & \frac{3}{10} \\ \mathbb{P}(R_1^c \cap R_2^c) & = & \mathbb{P}(R_2^c | R_1^c) \cdot \mathbb{P}(R_1^c) & = & \frac{2}{4} \times \frac{3}{5} & = & \frac{3}{10} \end{array}$$

ovvero le probabilità degli altri costituenti...

…dell'algebra generata da R_1 ed R_2 (algebra dagli eventi esprimibili in termini di R_1 ed R_2 eventualmente negati).

La regola del prodotto garantisce la coerenza dell'assegnazione di probabilità ottenuta (infatti le probabilità dei costituenti sono positive e sommano a uno) una volta che sia stata garantita la coerenza delle assegnazioni relative alle singole estrazioni (qui ottenute mediante valutazioni classiche).

A partire dalle probabilità dei costituenti possiamo calcolare le probabilità di tutti gli eventi dell'algebra…

...per esempio potrebbero interessarci gli eventi

$$
\begin{aligned}
E &= \text{“le due biglie estratte hanno lo stesso colore”,} \\
F &= \text{“almeno una delle biglie estratte è rossa”.}
\end{aligned}
$$

Ricordando che le probabilità dei costituenti valgono

$$
\begin{aligned}
\mathbb{P}(R_1 \cap R_2) &= 10\% \\
\mathbb{P}(R_1 \cap R_2^c) &= 30\% \\
\mathbb{P}(R_1^c \cap R_2) &= 30\% \\
\mathbb{P}(R_1^c \cap R_2^c) &= 30\%
\end{aligned}
$$

troviamo immediatamente...

$$\begin{aligned}\mathbb{P}(E) &= \mathbb{P}(R_1 \cap R_2) + \mathbb{P}(R_1^c \cap R_2^c) &= \\ & \qquad 10\% + 30\% &= 40\%\end{aligned}$$

per la probabilità di “due biglie dello stesso colore” e

$$\begin{aligned}\mathbb{P}(F) &= \mathbb{P}(R_1 \cap R_2) + \mathbb{P}(R_1 \cap R_2^c) + \mathbb{P}(R_1^c \cap R_2) &= \\ & \qquad 10\% + 30\% + 30\% &= 70\%\end{aligned}$$

per la probabilità di “almeno una biglia rossa”.

Si noti che nel secondo caso si poteva trovare il risultato <u>più speditamente</u> mediante la regola della negazione:

$$\mathbb{P}(F) = 1 - \mathbb{P}(F^c) = 1 - \mathbb{P}(R_1^c \cap R_2^c) = 1 - 0.3 = 70\%.$$

Similmente possiamo calcolare la probabilità di R_2 in assenza di informazioni sul colore della prima biglia estratta:

$$\mathbb{P}(R_2) = \mathbb{P}(R_1 \cap R_2) + \mathbb{P}(R_1^c \cap R_2) = 10\% + 30\% = 40\%$$

verificando che $\mathbb{P}(R_2) = \mathbb{P}(R_1)$ come suggerisce l'intuizione.

In generale, se si vuole ricavare $\mathbb{P}(F)$ da $\mathbb{P}(F|E)$, $\mathbb{P}(F|E^c)$ e $\mathbb{P}(E)$ si può direttamente applicare la regola delle probabilità totali:

$$\mathbb{P}(F) = \mathbb{P}(F|E) \cdot \mathbb{P}(E) + \mathbb{P}(F|E^c) \cdot (1 - \mathbb{P}(E))$$

Avremo modo di farlo quando ricaveremo la formula di Bayes, combinando i due usi della regola del prodotto per passare da $P(F)$ a $P(F|E)$ quando si osservi E.

Dati due eventi E ed F tali che $\mathbb{P}(E) > 0$ e $\mathbb{P}(F) > 0$, diremo che F è stocasticamente indipendente da E, secondo la probabilità $\mathbb{P}$, se il verificarsi di E non fornisce informazioni sul verificarsi di F:

$$\mathbb{P}(F|E) \quad = \quad \mathbb{P}(F)$$

ovvero, in virtù della regola del prodotto, se

$$\mathbb{P}(E \cap F) \quad = \quad \mathbb{P}(E) \cdot \mathbb{P}(F)$$

di modo che l'indipendenza stocastica risulta essere una relazione simmetrica.

Se $\mathbb{P}(E) = 0$ o $\mathbb{P}(F) = 0$, la fattorizzazione sopra esposta vale banalmente e diremo ugualmente che E e F sono indipendenti.

In corrispondenza dei due diversi usi della regola del prodotto, l'indipendenza può essere imposta o scoperta.

Per esempio, con riferimento al lancio di due dadi a sei facce, è prassi assumere che un evento relativo al primo lancio non porti informazione su un evento relativo al secondo lancio...

...questo conduce alla valutazione

$$\begin{aligned}\mathbb{P}(\text{“doppio sei”}) &= \mathbb{P}(\text{“primo sei”}) \cdot \mathbb{P}(\text{“secondo sei”}) \\ &= \frac{1}{6} \times \frac{1}{6} \simeq 3\%\end{aligned}$$

nell'ipotesi di dadi bilanciati.

Analogamente se da un'urna si estraggono due biglie con reinserimento (prima si è considerato il caso senza reinserimento).

Per contro, se da un'urna contenente 8 biglie colorate, di cui 2 gialle, 2 rosse, 2 verdi e 2 nere, se ne estrae una e si considerano gli eventi

$$\begin{aligned} G &= \text{“la biglia estratta è gialla”} \\ R &= \text{“la biglia estratta è rossa”} \\ V &= \text{“la biglia estratta è verde”} \end{aligned}$$

troviamo che, secondo la valutazione classica, $G \cup R$ ed $R \cup V$ sono stocasticamente indipendenti:

$$\begin{aligned} \mathbb{P}(G \cup R) &= 4/8 &= 1/2 \\ \mathbb{P}(R \cup V) &= 4/8 &= 1/2 \\ \mathbb{P}(R) &= 2/8 &= 1/4. \end{aligned}$$

Si noti come l'indipendenza stocastica dipenda dalla valutazione di probabilità (composizione dell'urna): se l'urna contenesse 10 biglie colorate, di cui 2 gialle, 2 rosse, 2 verdi e 4 nere, avremmo

$$
\begin{aligned}
\mathbb{P}'(G \cup R) &= 4/10 = 2/5 \\
\mathbb{P}'(R \cup V) &= 4/10 = 2/5 \\
\mathbb{P}'(R) &= 2/10 = 1/5
\end{aligned}
$$

e gli eventi $G \cup R$ ed $R \cup V$ sarebbero stocasticamente dipendenti.

Si noti la differenza rispetto all'indipendenza logica: questa è una proprietà dei soli eventi ed è condizione necessaria, ma non sufficiente, per l'indipendenza stocastica di due eventi (con probabilità non nulla).

Nel seguito, come è prassi fare, ci riferiremo all'indipendenza stocastica semplicemente come indipendenza, mentre l'indipendenza logica sarà sempre qualificata come tale.

La nozione di indipendenza si estende a famiglie finite di eventi: $E_1, E_2, \ldots, E_n$, $n \geq 3$, sono congiuntamente indipendenti se

$$\mathbb{P}(E_1 \cap E_2 \cap \ldots \cap E_k) = \mathbb{P}(E_1) \cdot \mathbb{P}(E_2) \cdot \ldots \cdot \mathbb{P}(E_k),$$

per ogni $k = 2, \ldots, n$, con il significato che nessuna sottofamiglia fornisce informazioni sui restanti eventi; per esempio

$$\mathbb{P}(E_3 | E_1 \cap E_2^c) = \mathbb{P}(E_3).$$

Con riferimento all'estrazione con reinserimento di tre biglie da un'urna che ne contiene 5, di cui 2 rosse e 3 bianche, è prassi assumere che gli eventi R_i = "i-esima biglia rossa", $i = 1, 2, 3$, siano congiuntamente indipendenti; questo fornisce la valutazione

$$\mathbb{P}(\text{"tutte rosse"}) = \mathbb{P}(R_1 \cap R_2 \cap R_3) = \frac{2}{5} \times \frac{2}{5} \times \frac{2}{5} = \left(\frac{2}{5}\right)^3 \simeq 6\%$$

Si noti che:

- lo stesso evento avrebbe evidentemente probabilità nulla, se effettuassimo delle estrazioni senza reinserimento;
- valutare la probabilità dell'evento "tutte rosse" è anche il modo più facile per calcolare $\mathbb{P}(\text{"almeno 1 bianca"}) = 1 - (2/5)^3 \simeq 94\%$.

D'altra parte, nell'esempio dell'urna con 8 biglie colorate, troviamo che $G \cup R$, $R \cup V$ e $G \cup V$ sono congiuntamente dipendenti, in quanto sono addirittura logicamente dipendenti:

> *es. se sia $G \cup R$ che $R \cup V$ sono VERO, allora è stata estratta una biglia rossa e quindi necessariamente $G \cup V$ è FALSO.*

Si noti come gli stessi eventi siano indipendenti a coppie (lo abbiamo visto per $G \cup R$ e $R \cup V$, analogamente potremmo verificarlo per le altre due possibile coppie); evidentemente l'indipendenza a coppie risulta essere una proprietà più debole di quella congiunta.

Si supponga di avere osservato il verificarsi dell'evento

$$E = \text{“il mio test è positivo”}$$

nell'ambito di una campagna diagnostica per una malattia rara.

Poiché il test non è infallibile, vi sono due possibili spiegazioni:

$$\begin{aligned} H_0 &= \text{“sono sano”} \\ H_1 &= \text{“sono malato”} \end{aligned}$$

Si noti che H_0 e H_1 sono mutuamente esclusive ($H_0 \cap H_1 = \emptyset$) ed esaustive ($H_0 \cup H_1 = \Omega$).

Le specifiche del test diagnostico (ottenute mediante esperimenti su pazienti il cui stato di salute è noto e riportate nel foglio illustrativo) ci dicono che

$$\begin{aligned} \mathbb{P}(E|H_1) &= 0.999 \quad \text{sensitività} \\ \mathbb{P}(E^c|H_0) &= 0.975 \quad \text{specificità} \end{aligned}$$

ovvero ci danno la verosimiglianza delle due possibili spiegazioni:

$$\begin{aligned} \mathbb{P}(E|H_1) &= 0.999 \\ \mathbb{P}(E|H_0) &= 0.025 \end{aligned}$$

A noi tuttavia interessa discriminare tra le due possibili spiegazioni non sulla base della loro verosimiglianza, ma sulla base delle loro probabilità finali (a posteriori) data l'evidenza osservata:

$$\mathbb{P}(H_0|E) = ? \qquad \& \qquad \mathbb{P}(H_1|E) = 1 - \mathbb{P}(H_0|E).$$

A tal fine dobbiamo preliminarmente valutare le probabilità iniziali (a priori) delle due spiegazioni:

$$\begin{aligned} \mathbb{P}(H_1) &= 0.02 \qquad \text{prevalenza} \\ \mathbb{P}(H_0) &= 1 - 0.02 = 0.98 \end{aligned}$$

in base a studi epidemiologici condotti sulla "mia" popolazione (quella su cui viene condotta la campagna diagnostica e di cui io faccio parte).

A questo punto, applicando due volte la regola del prodotto, troviamo

$$\mathbb{P}(H_0|E) = \frac{\mathbb{P}(E \cap H_0)}{\mathbb{P}(E)} = \frac{\mathbb{P}(E|H_0) \cdot \mathbb{P}(H_0)}{\mathbb{P}(E)}$$

e analogamente per H_1; quindi, calcolando la verosimiglianza marginale $\mathbb{P}(E)$ con la regola delle probabilità totali, otteniamo la formula di Bayes

$$\mathbb{P}(H_0|E) = \frac{\mathbb{P}(E|H_0) \cdot \mathbb{P}(H_0)}{\mathbb{P}(E|H_0) \cdot \mathbb{P}(H_0) + \mathbb{P}(E|H_1) \cdot \mathbb{P}(H_1)}$$

e analogamente per H_1: abbiamo espresso le probabilità finali in funzione di quelle iniziali e della verosimiglianza...

...nell'esempio della campagna diagnostica troviamo

$$\mathbb{P}(H_0|E) = \frac{0.025 \times 0.98}{0.025 \times 0.98 + 0.999 \times 0.02} \simeq 0.55$$

e quindi

$$\mathbb{P}(H_1|E) = 1 - \mathbb{P}(H_0|E) = 0.45$$

di modo che gli odds per scommettere contro la malattia sono scesi da 98 : 2 a 55 : 45 in virtù del test positivo; sebbene il test abbia delle buone prestazioni (alta sensitività e specificità) la rarità della malattia rende il risultato non conclusivo (un approfondimento è necessario).

Ignaccolo R. (2009). Allarmarsi per un risultato positivo al test sull'HIV? Non subito..., SIS Magazine, `http://www.sis-statistica.it/magazine/spip.php?article118`.

Prologo

Eventi

Probabilità

Epilogo

Armati dei rudimenti del calcolo delle probabilità, proviamo a risolvere i problemi sollevati nel XVII secolo dal Cavalier de Méré.

Per quanto riguarda il primo problema, se si lancia 4 volte un dado, l'ipotesi di indipendenza per i successivi lanci fornisce la valutazione

$$\mathbb{P}(\text{“4 punteggi} \neq ⚅\text{”}) = \frac{5}{6} \cdot \frac{5}{6} \cdot \frac{5}{6} \cdot \frac{5}{6} = \left(\frac{5}{6}\right)^4 \simeq 48.23\%$$

di modo che, passando alla negazione, si trova

$$\mathbb{P}(\text{“almeno 1 punteggio} = ⚅\text{”}) \quad \simeq \quad 51.77\%.$$

Similmente, posto che lanciando due dadi l'ipotesi di indipendenza per i due lanci ci dà

$$\mathbb{P}(\text{“doppio sei”}) = \frac{1}{6} \times \frac{1}{6} = \frac{1}{36},$$

se lanciamo 24 volte due dadi troviamo

$$\mathbb{P}(\text{“almeno 1 ⚅ ⚅ su 24 lanci”}) = 1 - \left(\frac{35}{36}\right)^{24} \simeq 49.14\%$$

di modo che, a parità di quota, la prima scommessa è preferibile.

Come mai, allora, il Cavalier de Méré pensava che le due scommesse fossero equivalenti?

I 4 lanci per il singolo dado e i 24 lanci per la coppia di dadi sono calibrati per rendere la probabilità di vittoria circa pari al 50%:

l'equazione

$$0.5 = (1 - p)^n \simeq e^{-np}$$

per $p_1 = 1/6$ *e* $p_2 = 1/36$ *ha soluzioni approssimate non intere*

$$n_1 = 6 \log 2 = 4.16$$

e

$$n_2 = 36 \log 2 = 24.95$$

tali che $n_2 = 6n_1$*; in pratica* n_1 *ed* n_2 *devono essere interi e questo (assieme all'approssimazione) fa la differenza.*

Il secondo problema del Cavalier de Méré chiede la probabilità di

$$E = \text{“vittoria al meglio delle cinque partite”}$$

prendendo come evento certo

$$\Omega = \text{“in vantaggio 2 a 1 dope le prime tre partite”}.$$

La posta in palio andrà infatti divisa in parti proporzionali a

$$\mathbb{P}(E) \qquad \text{e} \qquad 1 - \mathbb{P}(E)$$

ovvero assegnando a ogni giocatore la sua quota nella scommessa (al momento di interrompere la partita).

Conviene introdurre l'evento

$$G = \text{“vittoria della quarta partita”}$$

e utilizzare la regola delle probabilità totali per ottenere

$$\begin{aligned}\mathbb{P}(E) &= \mathbb{P}(E|G)\cdot\mathbb{P}(G)+\mathbb{P}(E|G^c)\cdot\mathbb{P}(G^c)\\ &= 1\cdot\frac{1}{2}+\frac{1}{2}\cdot\frac{1}{2}\\ &= 3/4\end{aligned}$$

cosicché di 100 euro in palio 75 spettanno al giocatore in vantaggio (abbiamo supposto che i due giocatori abbiano la stessa probabilità di vittoria in ogni singola partita . . . es. punteggio più alto con un dado).

BORRA, S. & DI CIACCIO, A. (2008).
Statistica: Metodologie per le Scienze Economiche e Sociali (Seconda Edizione).
McGraw-Hill, Milano.

DE FINETTI, B. (1970).
Teoria delle probabilità.
Einaudi, Torino.

KOLMOGOROV, A. N. (1933).
Grundbegriffe der Wahrscheinlichkeitsrechnung.
Springer, Berlin.
Traduzione inglese: Foundations of the Theory of Probability, Second Edition, Chelsea, 1956.

LAPLACE, P. S. (1812).
Théorie analytique des probabilités.
Courcier, Paris.

VON MISES, R. (1928).
Wahrscheinlichkeit, Statistik und Wahrheit.
Springer, Wien.
Traduzione inglese ampliata: Probability, Statistics and Truth, Second Edition, Allen and Unwin, 1957.

SCOZZAFAVA, R. (2000).
Primi Passi in Probabilità e Statistica.
Decibel, Padova.