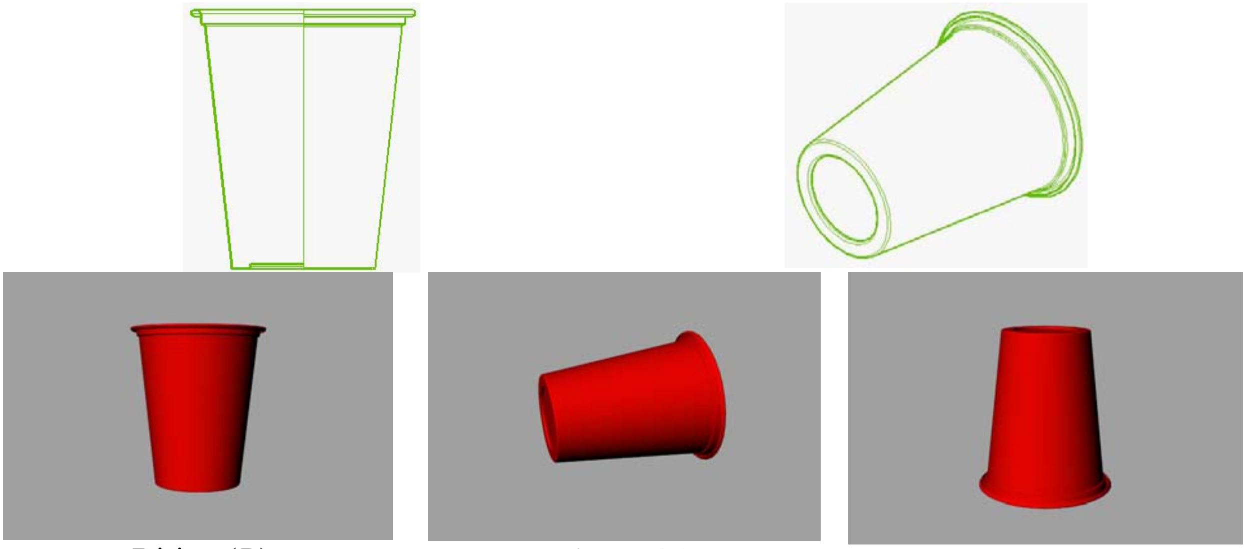
SCOMMESSE, FREQUENZA E PROBABILITÀ (SOGGETTIVA)

Supponiamo di disporre di un bicchiere di plastica (Fig. 1) e di lanciarlo, tenendolo per l'orlo superiore o per il piede, facendolo ruotare. Quanto scommettereste che il bicchiere cada diritto (*D*)? Quanto che cada di fianco (*F*) o capovolto (*C*)?



Diritto (*D*) Fianco (*F*) Capovolto (*C*)

Fig. 1. Il bicchiere di plastica e gli esiti del lancio

Approfondiamo questo semplice esperimento per chiarire alcuni concetti di probabilità e comprendere come siano legati alle scommesse.

Innanzitutto, definite quante volte la posta vorreste guadagnare, relativamente all'esito *D*, per considerare equa la scommessa. Scrivete la Vostra Opinione prima di procedere nella lettura.

Accettereste una scommessa 1 a 2 ossia che il banco paghi due volte la posta da Voi impegnata sull'esito *D*?

Non credo che accettereste perché, pur basandosi il calcolo su di un ragionamento comune, risulta intuitivo ipotizzare, anche prima di dimostrarlo, che raramente il bicchiere cade in piedi. La scommessa, in questo caso, è iniqua. Il banco bara proponendo come base di calcolo consueto lo schema logico del gioco ideale (equiprobabilità degli eventi elementari, v. Erto, caso *a*, p. 13):

$$\Pr\{D\} = \Pr\{F\} = \Pr\{C\} = \frac{1}{3}$$

Si può così spiegare il ragionamento (sbagliato a suo favore) che può fare il banco per giustificare la quota 1 a 2:

$$\frac{\Pr\{D\}}{\Pr\{F \text{ o } C\}} = \frac{1/3}{2/3} = \frac{1}{2}\,.$$

Per definire quali possano essere le quote corrette occorre interrogare il fenomeno fisico, reiterando l'esperimento e raccogliendo dati. Questa strada è propria della Statistica Sperimentale. La Tab. I sinottica riporta sinteticamente le frequenze assolute degli esiti di 6 serie di 100 lanci. Uno scommettitore basandosi sull'esito della prima serie di lanci riterrebbe eque scommesse che diano 1 a 99 l'esito *D*. Addirittura non si accetterebbero scommesse su *D*, basandosi sulle due ultime serie (quando la frequenza osservata di *D* è nulla).

Tab. I. *Tavola sinottica delle frequenze assolute di serie di esperimenti di lancio del bicchiere*

	n_D	n_F	n_C	N
I serie	1	87	12	100
II serie	2	80	18	100
III serie	5	76	19	100
IV serie	2	75	23	100
V serie	0	83	17	100
VI serie	0	70	30	100
Totale	10	471	119	600

Poichè calcoliamo frequenze, mentre per decidere le quote dovremmo conoscere le probabilità solo reiterando un numero praticamente infinito di volte l'esperimento potremmo ottenerle. Questo criterio di calcolo di una probabilità a partire dalle frequenze osservate è noto come frequentista e si esprime con un limite all'infinito (v. Erto, caso *b*, p. 13). Fortunatamente, nelle applicazioni pratiche, può bastare un numero di esperimenti sufficientemente grande, ma fattibile, per valutare, in modo plausibile, come probabilità ciò che misuriamo come frequenza. Ciò significa che, sulla base dei 600 lanci effettuati una plausibile valutazione di probabilità, dai dati raccolti risulti:

$$\Pr\{D\} = \frac{10}{600} = 0.0167 = 1.67\%$$

Da un punto di vista grafico possiamo mostrare come la frequenza vari al crescere delle replicazioni dell'esperimento (Fig. 2) pur tendendo a stabilizzarsi intorno un valore sconosciuto che rappresenta la probabilità. Quest'ultimo è il valore asintotico, che viene approssimato dal valore di frequenza ricavato sulla base di tutte le replicazioni disponibili (nel nostro caso sull'insieme di 600 lanci).

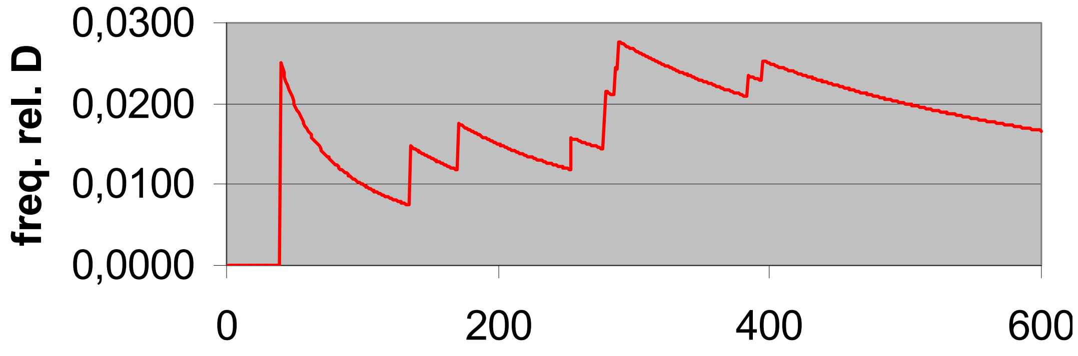


Fig. 2. Fluttuazioni della frequenza relativa al crescere delle replicazioni dell'esperimento

A questo punto, possiamo riprendere le quote precedentemente considerate eque per decidere di scommettere sull'esito *D*. Se le quote da Voi fissate fossero 1 (*a*) a 19 (*b*), il valore di probabilità soggettivo conseguente risulterebbe (v. Erto, caso *c*, p. 13):

$$\Pr\{D\} = \frac{a}{a+b} = \frac{1}{20} = 0.05 = 5\%$$

La valutazione così effettuata può avere il significato di ricordo di esperimenti passati ("*da piccolo lanciai il bicchiere e ricordo che mediamente ogni 20 lanci mi capitava una volta D*") oppure di esperimenti immaginari cui si ricorre per dare corpo ad una valutazione ("*immagino che cosa possa succedere ripetendo il lancio del bicchiere e, data la sua forma ed il tipo di lancio, riterrei plausibile che mediamente una volta su 20 l'esito sia D*"). Questa valutazione viene definita informazione a priori. Grazie alla “macchina” di Bayes (dal nome del reverendo Bayes che la propose), l’informazioni a priori può essere coniugata con i dati sperimentali aggiornando lo stato di conoscenze e le decisioni conseguenti (giungendo così all’informazione a posteriori). Lo schema in Fig. 3 evidenzia il funzionamento (che può essere iterativo) della macchina di Bayes.

La via soggettiva alla valutazione di probabilità è conosciuta come Bayesiana. Lo studioso che più recentemente ne ha dimostrato la coerenza, anche nell’aggiornamento dello stato di conoscenza, è il prof. De Finetti di Roma, cui sono debitori tutti gli statistici bayesiani contemporanei.

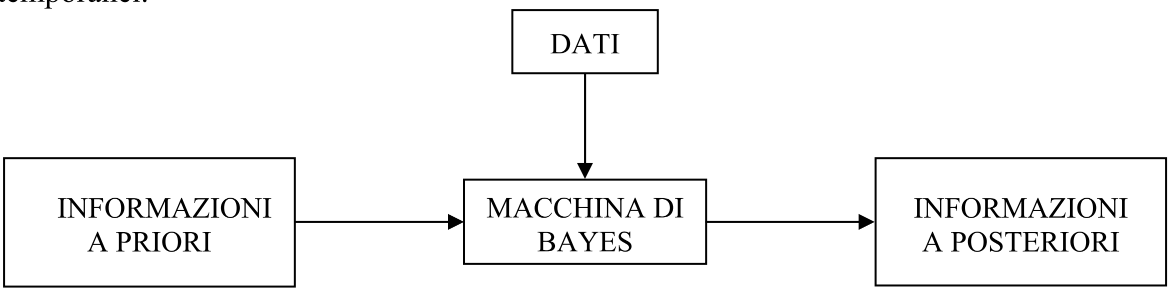


Fig. 3. Schema logico della macchina di Bayes

Ritornando all’esempio del bicchiere, grazie alla macchina di Bayes, a partire dall’informazione a priori, possiamo ottenere la migliore valutazione possibile a posteriori della probabilità dell’esito D condizionata agli esperimenti effettuati:

$$\frac{a+n_D}{a+b+N}=\frac{1+10}{20+600}=0.0177=1.77\%$$

Il risultato mostra come, in questo caso, i dati sperimentali pesino di più rispetto all’informazione a priori, essendo più vicino il risultato finale all’1.67% valutato sperimentalmente che non al 5% anticipato.

Viceversa, una conoscenza a priori espressa come 30 (a) esiti D su 600 esperimenti (ossia sempre del 5% ma su una base di conoscenza più ampia), porterebbe alla seguente valutazione di probabilità:

$$\frac{a+n_D}{a+b+N}=\frac{30+10}{600+600}=0.0333=3.33\%$$

In questo caso, il risultato è medio pesato tra la conoscenza a priori e quella sperimentale. Si mostra, quindi, come la maggiore consapevolezza sull’esperimento possa portare all’esplicitazione (il termine tecnico è *elicitazione*) di un a priori più forte che tende a pesare di più rispetto ai dati sperimentali.

Non sono ancora noti e disponibili (e non lo saranno mai, sic!) strumenti bayesiani in grado di farci vincere sempre le scommesse. Fortunatamente, però, sono disponibili in modo via via crescente strumenti bayesiani di ausilio alle diagnosi in medicina. Infine, è in crescente diffusione, l’approccio bayesiano in campo industriale specificamente aerospaziale ed automobilistico per le valutazioni di affidabilità. In generale, l’approccio bayesiano, consentendo di non sprecare informazioni, è molto utile in tutti quei contesti (industriali e non) dove ci siano reali competenze di esperti e dati (anche pochi) su cui basare le decisioni.

A.L.

Riferimenti essenziali

T. Bayes, *Essay towards solving a problem in the doctrine of chances*, pubblicato in *Philosophical Transactions of the Royal Society of London*, 1764.

B. De Finetti, *Theory of Probabilità*, Wiley, NY, 1990.

P. Erto, *Probabilità e Statistica per le scienze e l’ingegneria*, II ed., McGraw Hill, 2004.

R. Scozzafava, La probabilità soggettiva e le sue applicazioni, editoriale Veschi, 1989.