Probabilità e scommesse

Michele Impedovo

Riassunto Il primo approccio alla probabilità nelle nostre scuole è quello "classico": il rapporto tra il numero di casi *favorevoli* e il numero di casi *possibili*. Questo approccio ha molti pregi; tuttavia se resta l'unico si rischia di confondere la probabilità con la combinatoria e di ridurla a problemi di conteggio, di trattare solo problemi di carte, dadi e monete; invece è importante mostrare che la probabilità ha ambiti di impiego più generali (e più importanti).

L'idea di *scommessa* è precoce nei bambini e potrebbe essere utilizzata per un approccio più generale alla didattica della probabilità. Inoltre l'idea di scommessa è il fondamento dell'approccio *soggettivista*, che rappresenta oggi un modello universale, che contiene tutti gli altri, perché si può applicare a qualunque stima di probabilità.

Abstract In Italian schools the first approach to probability is the "classical" one: the probability of an event is the number of *favorable* outcomes over the total number of outcomes. This approach has many advantages, but, if not taught side by side with other approaches, it may generate confusion between probability and counting, typically focusing only on cards, dice and coins problems. On the contrary, it is important to show that probability has a much wider, and more relevant, scope. Children quickly grasp the concept of a *bet* and this may be used for a more general approach to the teaching of probability. Furthermore, the idea of a bet is the starting point in the *subjective* approach to probability, which is the most comprehensive model, encompassing any assessment of probability.

Michele Impedovo, Università Bocconi di Milano

Com'è noto l'insegnamento della probabilità nella scuola italiana soffre; se ne insegna poca e quasi sempre l'unico approccio è quello "classico": la probabilità è il rapporto tra il numero di casi *favorevoli* e il numero di casi *possibili*. Si tratta di un approccio che ha molti pregi: è relativamente semplice, si può proporre molto precocemente e offre una via facile per partire, ma se resta l'unico si corre il rischio di confondere la probabilità con la combinatoria e di ridurla a problemi di conteggio, di trattare solo problemi di carte, dadi e monete; invece è importante mostrare che la probabilità ha ambiti di impiego più generali: settori interi dell'economia, ad esempio le assicurazioni, si fondano su analisi probabilistiche.

Anche l'idea di *scommessa* è precoce nei bambini e potrebbe ugualmente essere utilizzata per un approccio più generale alla didattica della probabilità. In fondo qualsiasi scambio di somme aleatorie di denaro (titoli, premi assicurativi, pensioni, ipoteche) è una scommessa. Ma c'è di più: l'idea di scommessa è il fondamento dell'approccio *soggettivista* alla probabilità; tale approccio rappresenta oggi un modello universale, che contiene tutti gli altri, perché si può applicare a qualunque stima di probabilità.

Che cos'è la probabilità?

La mia tesi, paradossale e un po' provocatoria,
ma genuina, è che semplicemente la probabilità
non esiste
Bruno de Finetti

Per secoli, fino all'impostazione assiomatica di Andrey Kolmogorov (1908-1987) si è cercata una definizione matematica della probabilità. Oggi accettiamo l'idea di non darne una definizione e di considerarla come un concetto primitivo, al quale chiediamo solo coerenza rispetto agli assiomi:

1) La probabilità è un numero *p* compreso tra 0 e 1.
2) La probabilità dell'evento impossibile è 0 e la probabilità dell'evento certo è 1.
3) Se due eventi A e B sono incompatibili allora la probabilità che si verifichi uno dei due è la somma delle loro probabilità.

Le diverse "definizioni", meglio i diversi approcci alla valutazione della probabilità di un evento corrispondono ad altrettanti modi di intenderla e ad altrettante fasi storiche molto distanti tra loro.

Approccio classico, Pierre Simon Laplace (1749-1827).

$$\textit{probabilità} = \frac{\text{numero di casi favorevoli}}{\text{numero di casi possibili}}$$

1. **O**N a vu dans l'Introduction, que la probabilité d'un événement, est le rapport du nombre des cas qui lui sont favorables, au nombre de tous les cas possibles; lorsque rien ne porte à croire que l'un de ces cas doit arriver plutôt que les autres, ce qui les rend pour nous, également possibles. La juste appréciation de ces cas divers, est un des points les plus délicats de l'analyse des hasards.

"Come si è visto nell'introduzione la probabilità di un evento è il rapporto tra il numero di casi che sono favorevoli ad esso e il numero di tutti i casi possibili; a patto che nulla ci faccia ritenere che un caso debba verificarsi piuttosto che un altro, il che ce li rende tutti ugualmente possibili. La corretta valutazione di questi casi è uno dei punti più delicati nell'analisi del caso."

La genesi è manifestamente quella dei giochi d'azzardo: dadi, carte, monete. La probabilità come rapporto può essere molto utile per consolidare negli alunni le abilità numeriche; poiché le probabilità si sommano e si moltiplicano, questo contesto può essere proficuo per dare alle frazioni e alle loro operazioni una veste semanticamente ricca. Tuttavia i problemi che si possono risolvere sono pochi, sono quelli che presentano simmetria (equiprobabilità) tra i diversi casi possibili. Le critiche a questo approccio sono ben note:

- La "definizione" non è una definizione perché è circolare: i "casi possibili" devono essere equiprobabili.
- Questo approccio funziona solo se i casi possibili hanno tutti la stessa probabilità e, quindi, al di là della circolarità definitoria, funziona solo quando le nostre conoscenze sul fenomeno sono assai elevate (tanto da poter supporre che i casi possibili siano equiprobabili) o, specularmente, talmente basse da essere indotti a considerare equiprobabili i casi elementari in assenza di altre informazioni.

- Se i casi possibili sono infiniti, anche problemi banali diventano inaccessibili: qual è la probabilità che un numero naturale sia pari?
- Anche nel caso finito, se si perde la simmetria si possono ottenere risultati insensati: qual è la probabilità che io sia vivo domani a quest'ora? C'è un caso "favorevole" (molto favorevole: sarò vivo) e due casi possibili (sarò vivo oppure no), dunque la probabilità che domani a quest'ora sia vivo è 1/2?
- L'approccio classico lascia intendere che la probabilità sia oggettiva e viva nelle cose; a noi resta il compito di scovarla e di calcolarla. È davvero così?
- Qual è la probabilità che quest'anno io non provochi incidenti con la mia auto? Qual è la probabilità che io sia in vita tra 10 anni? Qual è la probabilità che il titolo ENI oggi aumenti il proprio valore? Qual è la probabilità che un giovane trovi lavoro entro un anno dalla fine degli studi? Qual è la probabilità che l'Italia raggiunga il pareggio di bilancio entro il prossimo anno? L'approccio classico non ha strumenti di alcun tipo per problemi di questo genere; interi settori della vita economica reale (la gestione pensionistica ed assicurativa, per esempio) presentano problemi per i quali l'approccio classico è impotente.

Approccio frequentista, Richard Von Mises, (1883-1953).

La probabilità è il limite a cui tende la frequenza relativa di un evento quando il numero di esperimenti tende a infinito:

$$\textit{probabilità} = \lim_{N \to \infty} \frac{\text{frequenza di successo}}{\text{numero di esperimenti}}$$

L'approccio frequentista, o statistico, interpreta la probabilità come *frequenza relativa*; ipotizza che un esperimento aleatorio si possa eseguire, *nelle stesse condizioni*, un numero grande di volte, diciamo *N*. Al crescere di *N* il rapporto tra il numero di volte in cui l'esperimento ha dato esito positivo (i cosiddetti "successi") e *N dovrebbe convergere* ad un numero, che chiamiamo probabilità.

La genesi di questa idea è naturalmente quella statistica, viene dal mondo delle assicurazioni, ma anche dalla biologia, dalle scienze sociali, dall'economia.

Si possono risolvere ora tutti i problemi che si potevano risolvere con l'approccio classico (basta ripetere molte volte il corrispondente esperimento) ma il punto di vista è capovolto: mentre nell'approccio classico la probabilità che esca Testa lanciando una moneta è *a priori* 0.5 (in questo senso è un approccio *deduttivo*), un frequentista prima lancia la moneta molte volte, diciamo 1000 volte, e se è uscita Testa circa 500 volte dichiara che la probabilità *a posteriori* è 0.5 (in questo senso è un approccio *induttivo*). Pur partendo da punti di vista differenti l'approccio frequentista e quello classico giungono (circa) alle stesse valutazioni di probabilità.

Un esempio didatticamente interessante è quello del lancio di una puntina da disegno: qual è la probabilità che cadendo rimanga a testa in giù, piuttosto che di lato?

L'approccio classico non può fornire alcun suggerimento; il lancio di una gran numero di puntine potrebbe invece dare una stima ragionevole.

L'approccio frequentista consente di formulare infiniti nuovi problemi: quanti incidenti procurerò con la mia auto nel 2010? Qual è la probabilità che una persona si ammali di influenza il prossimo mese? O che sia mancina? Quale premio mensile devo pagare per un fondo assicurativo che mi garantisca una pensione dignitosa? Che incidenza ha il fumo sul tumore ai polmoni?

Critiche:

- Anche questa "definizione" non è una definizione: chi ha tempo di fare infiniti esperimenti? Chi garantisce che gli esperimenti siano eseguiti nelle stesse condizioni? E soprattutto, perché mai il rapporto tra il numero di successi e il numero di esperimenti dovrebbe convergere? Questo sembra un atto di fede, al solo scopo di mascherare da definizione ciò che altro non è che una stima empirica.

- Ci sono casi in cui la ripetizione dell'esperimento non è possibile: qual è la probabilità che la nazionale italiana vinca il prossimo Sei Nazioni di rugby? Qual è la probabilità che un terremoto distrugga San Francisco entro il 2020? E ancora: qual è la probabilità che io sia vivo domani a quest'ora?
- Qual è la probabilità che lo studente di fronte a me prenda 30 e lode nell'esame di Matematica? L'approccio classico prevederebbe 15 casi possibili {bocciato, 18, 19, …, 30, 30 e lode} e 1 solo caso favorevole; l'approccio statistico vorrebbe che il povero malcapitato ripetesse l'esame un gran numero di volte, e alla fine contasse il numero di volte in cui ha preso 30 e lode.

Nonostante queste critiche, l'approccio statistico ha una notevole valenza didattica, che l'insegnante deve tenere in debito conto perché in molte situazioni una spiegazione fondata sulla frequenza relativa del numero di successi è molto convincente. Sembra che ripetere l'esperimento un gran numero di volte come strumento per validare una stima di probabilità sia nel nostro DNA; anzi a volte l'idea stessa di probabilità è spiegata in termini di frequenza relativa su un gran numero di prove.

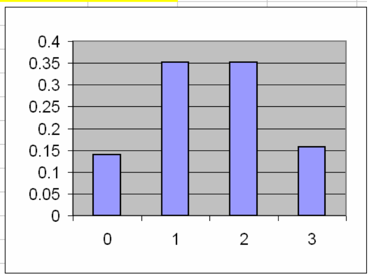
Inoltre l'approccio statistico è in un certo senso connaturato con una verifica *fisica* della probabilità, che consiste nella simulazione (al calcolatore) dell'evento tramite il generatore di numeri casuali, e nella stima empirica della probabilità come frequenza relativa di successo.

Per esempio quando dico, con approccio classico, che la probabilità di avere 2 teste nel lancio di 3 monete regolari è 3/8 (gli eventi possibili ed equiprobabili sono 8: TTT, TTC, TCT, CTT, TCC, CTC, CCT, CCC e 3 di questi sono favorevoli), al tempo stesso suggerisco un esperimento statistico: posso **simulare** il problema lanciando 1000 volte 3 monete e osservando l'esito. L'avvento dei calcolatori ha improvvisamente aperto un orizzonte sterminato di possibilità di verifica di esperimenti, tanto che alcuni problemi, intrinsecamente troppo difficili, vengono risolti solo per mezzo della simulazione. Dal punto di vista didattico il calcolatore ha enormemente facilitato la possibilità di simulare un problema e quindi di formulare congetture sensate a poco prezzo, sulle quali è naturale innestare il gioco razionale dei simboli.

Per esempio con Excel è molto semplice simulare, utilizzando il generatore
=CASUALE(),

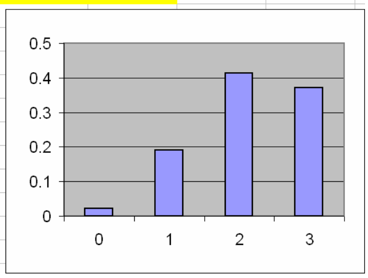
il lancio di 3 monete regolari: basta scrivere, per ciascuna moneta, il comando
=SE(CASUALE()<0.5;"T";"C").

p	moneta 1	moneta 2	moneta 3	n° teste	n° teste	fr.rel.
0.5	C	C	T	1	0	0.139
	C	T	T	2	1	0.352
	C	T	T	2	2	0.352
	T	T	T	3	3	0.157
	T	C	C	1		
	T	T	C	2		
	C	T	C	1		
	T	T	C	2		
	C	T	T	2		
	T	T	C	2		
	C	T	C	1		
	C	T	T	2		
	T	C	T	2		
	T	T	C	2		
	C	C	C	0		



Ma è anche possibile simulare monete di trucco qualsiasi: basta modificare il numero 0.5 nel comando precedente mediante un riferimento assoluto ad una cella (la cella A2, nella figura seguente) che funge da parametro. Ecco per esempio la simulazione del lancio di 3 monete truccate in cui esce testa con probabilità $p = 0.7$.

	A	B	C	D	E	F	G	H	I
1	p	moneta 1	moneta 2	moneta 3	n° teste	n° teste	fr.rel.		
2	0.7	C	C	T	1	0	0.023		
3		T	T	T	3	1	0.191		
4		T	C	C	1	2	0.414		
5		T	T	T	3	3	0.372		
6		T	T	C	2				
7		T	T	T	3				
8		T	T	T	3				
9		T	T	C	2				
10		T	T	T	3				
11		T	T	T	3				
12		T	T	C	2				
13		T	T	T	3				
14		C	T	T	2				
15		T	T	T	3				
16		T	T	T	3				



Approccio soggettivo, Bruno de Finetti (1906-1985)

probabilità = posta che si ritiene equo pagare per una scommessa di valore 1

La probabilità è il *grado di fiducia* (*degree of belief*) di un osservatore nel verificarsi di un evento, ed è misurabile attraverso la seguente richiesta:

qual è l'importo p che si ritiene equo puntare in una scommessa in cui

- se l'evento si verifica si riceve 1 (e dunque si guadagna $1-p$);
- se l'evento non si verifica si perde la posta p.

In una formulazione equivalente più generale, ipotizzando indifferenza al rischio dei giocatori, lo scambio di p contro 1 diventa lo scambio di una qualunque somma pS contro la somma S.

Per esempio, dire che la probabilità di pioggia domani è del 70% significa ritenere equa una scommessa in cui si punta 70 ($= 0.7S$) e si riceve 100 ($= S$) in caso che l'evento si verifichi.

"Ritenere equo" significa in sostanza considerare indifferente ricoprire il ruolo di *scommettitore* oppure di *banco*, e cioè ritenere indifferenti le due posizioni seguenti:

- pagare pS il diritto a riscuotere S nel caso che l'evento si verifichi (ruolo dello scommettitore), oppure
- accettare una scommessa di importo pS e impegnarsi a pagare S nel caso che l'evento si verifichi (ruolo del banco).

Dunque il numero p deve essere consistente con entrambi i ruoli e costringe a mettere in campo correttamente tutte le informazioni di cui si dispone.

Scommesse coerenti

La valutazione di p deve essere *coerente* nel senso che non deve consentire sistemi di puntate che conducano a vincite sicure oppure perdite sicure.

Si può dimostrare che richiedere tale coerenza è equivalente a postulare tutti gli assiomi della probabilità.

Infatti:

1) la posta che ritengo equo pagare per una scommessa di vincita unitaria è necessariamente compresa tra 0 e 1; se fosse minore di 0 incasserei una vincita sicura (la posta) indipendentemente dal verificarsi o meno dell'esperimento; se fosse maggiore di 1, andrei incontro ad una perdita sicura, dato che pagherei più di 1 il diritto a riscuotere, nella migliore delle ipotesi, 1;
2) la probabilità dell'evento certo deve essere 1: qualunque posta minore di 1 mi procurerebbe una vincita sicura; la probabilità dell'evento impossibile deve essere 0: qualunque posta maggiore di 0 mi procurerebbe una perdita sicura;
3) se p è la probabilità assegnata all'evento E allora la probabilità dell'evento complementare E^C deve necessariamente essere $1-p$: infatti il guadagno del sistema di scommesse
    - p su E
    - q su E^C

    risulta essere lo stesso qualunque sia l'esito:
    - $1-p-q$ se si verifica E
    - $-p+1-q$ se si verifica E^C;

    l'unica possibilità affinché tale guadagno non sia una vincita o una perdita sicura è che risulti $1-p-q=0$, cioè $q=1-p$;
4) infine, se ritengo equo pagare rispettivamente p_1 e p_2 per due eventi incompatibili E_1 e E_2 allora l'unico prezzo coerente sull'evento $E=E_1 \cup E_2$ non può che essere p_1+p_2. Infatti, sia q il prezzo di una scommessa su $E_1 \cup E_2$; per quanto abbiamo visto nel punto precedente il prezzo di una scommessa coerente su $(E_1 \cup E_2)^C$ è necessariamente $1-q$; anche in questo caso il guadagno del sistema di scommesse
    - p_1 su E_1
    - p_2 su E_2
    - $1-q$ su $(E_1 \cup E_2)^C$

    risulta essere lo stesso qualunque sia l'esito:
    - $1-p_1-p_2-1+q=q-p_1-p_2$ se si verifica E_1
    - $-p_1+1-p_2-1+q=q-p_1-p_2$ se si verifica E_2
    - $-p_1-p_2+1-(1-q)=q-p_1-p_2$ se non si verifica né E_1 né E_2;

l'unica possibilità affinché tale guadagno non sia una vincita o una perdita sicura è che risulti $q-p_1-p_2 = 0$, cioè $q = p_1+p_2$.

La probabilità è dentro di noi, non fuori di noi

In questo quadro accade un grande cambiamento di prospettiva: la probabilità non vive nell'evento, ma vive nell'osservatore; la probabilità è soggettiva. Valutatori differenti sono disposti a puntare somme diverse sullo stesso evento, ragionevolmente perché dispongono di diverse informazioni.

Ecco dunque, dopo tre secoli di storia della probabilità, una rivoluzione copernicana: la probabilità non è una proprietà della realtà, che devo scoprire o calcolare, è invece una proprietà del soggetto che la valuta ed esprime lo stato attuale delle informazioni che possiede; non solo può cambiare da persona a persona, ma per uno stesso valutatore può cambiare se cambia lo stato di informazioni che possiede. In questo senso de Finetti arriva a dire "*la probabilità non esiste*": non esiste al di fuori dell'osservatore. Quando estraggo una carta da un mazzo dico che la probabilità di pescare una carta di cuori è 1/4; ma se mentre la sto estraendo dal mazzo vedo, riflesso sul tavolo di vetro, una vaga luce rossa, allora la probabilità cambia e diventa 1/2. La probabilità classica ipotizza in sostanza una situazione statica, in cui la simmetria degli eventi elementari non muta, né muta lo stato delle informazioni del valutatore.

La probabilità soggettiva non è in contrasto con gli altri due approcci, ma li contiene come casi particolari. Non richiede simmetria tra gli eventi, né ripetizione dell'evento. È universale, si applica a qualsiasi problema, si accorda bene ai problemi in contesto economico, sociale, biologico ed è particolarmente adatta a prendere decisioni in condizioni di incertezza.

Quando nel bel mezzo della lezione chiedo a bruciapelo a uno studente "Qual è la probabilità che tu prenda 30 e lode nell'esame di Matematica?", lui di solito risponde senza esitazione qualcosa come "5%", oppure "1%", oppure (più spesso) "zero": quale paradigma intuitivo mobilita per dare queste risposte? Non certo quello della probabilità classica (il rapporto tra numero di esiti favorevoli e possibili sarebbe indipendente dalla quantità e qualità del suo studio), e nemmeno quello della probabilità statistica (non pensa certo di ripetere il proprio esame "un gran numero di volte"), bensì quello della probabilità soggettiva: se risponde 1% significa che ritiene equo giocare solo 1 centesimo per ricevere 1 euro nell'eventualità (molto improbabile, secondo lui) di ottenere 30 e lode.

L'Organizzazione Internazionale per la Standardizzazione (ISO), in *Guide to the expression of uncertainty in measurement* (1993)[1], preparata dalle federazioni internazionali di chimica, fisica, biologia, raccomanda di utilizzare la probabilità soggettiva nella valutazione dell'incertezza. Ecco che cosa suggerisce in un celebre passo:

"… an equally valid viewpoint is that probability is a measure of the degree of belief that an event will occur. For example, suppose one has a chance of winning a small sum of money D and one is a rational bettor. One's degree of belief in event A occurring is p = 0.5 if one is indifferent to these two betting choices: (1) receiving D if event A occurs but nothing if it does not occur; (2) receiving D if event A does not occur but nothing if it does occur."

"… un punto di vista altrettanto valido è che la probabilità è una misura del grado di fiducia sul verificarsi di un evento. Ad esempio, supponiamo di avere la possibilità di vincere una piccola somma di denaro D e di essere uno scommettitore razionale. Il grado di fiducia nell'evento A è p=0.5 se si è indifferenti a queste due scelte: (1) ricevere D se l'evento A si verifica, e nulla se non si verifica; (2) ricevere D se l'evento A non si verifica e nulla se si verifica."

Probabilità e scommesse eque

Per decidere chi per primo deve sottoporsi alla tortura quotidiana del bagno, i miei nipotini (8 e 6 anni) giocano a *bimbumbam*: chi perde finisce subito in vasca.

[1] Si può trovare per esempio in http://www.temperatures.ru/pdf/GUM.pdf

Attenti come sono alla *giustizia* ("non è giusto!" urlano spesso), evidentemente ritengono equo questo gioco. Ho provato a chiedere loro quale somma di denaro riterrebbero giusto scommettere sul pari e sul dispari e si sono accordati quasi subito sullo scommettere entrambi la stessa somma, non importa quale.

Poi ho cambiato le regole del bimbumbam: è vietato mettere il pugno chiuso, cioè 0. Abbiamo analizzato i possibili risultati e abbiamo scoperto che in questo caso il pari e il dispari non sono simmetrici: la somma è pari 13 volte su 25, è dispari 12 volte su 25. La differenza è piccola ma loro sono giunti alla conclusione che chi sceglie pari debba pagare una somma un po' maggiore dell'altro; anzi, hanno giudicato corretto che chi scommette sul pari debba pagare 13 centesimi e chi scommette sul dispari 12 centesimi, per bilanciare le differenti possibilità di vittoria. Chi vince intasca comunque 25 centesimi.

Credo che l'idea di *gioco equo* sia molto precoce, e lo sia anche dal punto di vista quantitativo: due giocatori pagano rispettivamente R e S sul verificarsi o meno di un evento di probabilità p. Perché uno dovrebbe scommettere più dell'altro? Ma naturalmente perché la sua probabilità di vincere è maggiore di quella dell'altro.

Chi vince incassa l'intera posta $R+S$. Dunque se l'evento si realizza il primo giocatore vince S e il secondo perde S; se l'evento non si realizza il primo giocatore perde R e il secondo vince R. Fornire una definizione matematica di *gioco equo* non è difficile; l'esempio del bimbumbam senza lo 0 mostra che in una scommessa tra due giocatori è ragionevole chiedere che le poste siano proporzionali alle probabilità di vincere:

$$\frac{R}{p} = \frac{S}{1-p}$$

Dette R e S le poste pagate dai due giocatori, il "guadagno" di ciascuno (eventualmente negativo) è un primo, semplice esempio di *numero aleatorio* (o *variabile aleatoria*, *random variable*), la cui distribuzione di probabilità è descritta dalle seguenti tabelle, che riportano sulla prima riga la somma vinta (se positiva) o persa (se negativa) e sulla seconda riga le corrispondenti probabilità (di somma 1):

Guadagno dello scommettitore 1: $\begin{cases} -R & S \\ 1-p & p \end{cases}$

Guadagno dello scommettitore 2: $\begin{cases} -S & R \\ p & 1-p \end{cases}$

In un certo senso il numero aleatorio rappresenta il superamento della nostra ignoranza sugli eventi futuri. Per esempio: quante TESTE usciranno se lancio una moneta 3 volte? La risposta "*boh?*" non rende giustizia delle nostre conoscenze. La risposta più razionale è un numero aleatorio, la cui distribuzione è la seguente (qui l'approccio soggettivista interviene nella determinazione delle probabilità iniziali, che, comunque, sono assegnate in modo assolutamente coerente con l'approccio classico: in mancanza di informazioni dico 1/2):

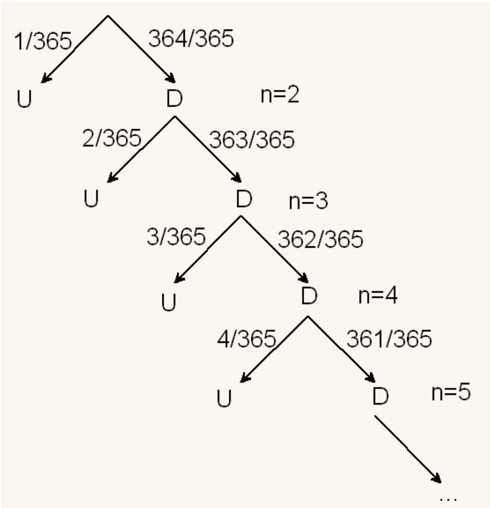
$$\begin{cases} 0 & 1 & 2 & 3 \\ 1/8 & 3/8 & 3/8 & 1/8 \end{cases}$$

La distribuzione di un numero aleatorio, coerentemente con l'impostazione soggettivista, rappresenta dunque lo stato corrente delle nostre conoscenze e delle nostre informazioni.

Mi capita di usare il *problema dei compleanni* con i miei studenti per mostrare questa idea dello *stato informativo*; dico: *Scommetto 100 euro sulla probabilità che in quest'aula ci siano almeno due persone con lo stesso compleanno. Chi scommette 1 euro contro di me?* Il fatto è che in aula ci sono circa 120 studenti, dunque la mia probabilità p di vincere è praticamente … una certezza. Per la precisione, se riteniamo che la data di nascita di quella popolazione sia uniformemente distribuita sui 365 giorni dell'anno, risulta $p \approx 0.9999997$.[2]

[2] Vedi ad esempio in www.matematica.it/paola/curriculum_personale.htm l'articolo di Domingo Paola "Aspetti paradossali in problemi di probabilità".

Infatti: calcoliamo, perché è più facile, la probabilità p_n che n persone abbiano tutte compleanni differenti. La probabilità che almeno due abbiano lo stesso compleanno sarà $1-p_n$. Costruiamo il grafo del problema (U sta per "almeno due compleanni uguali", D sta per "tutti compleanni diversi").



Se $n = 2$ la probabilità di 2 compleanni diversi è

$$p_2 = \frac{364}{365}$$

Se $n = 3$ la probabilità di 3 compleanni diversi è

$$p_3 = \frac{364 \cdot 363}{365 \cdot 365}$$

Se $n = 4$ la probabilità di 4 compleanni diversi è

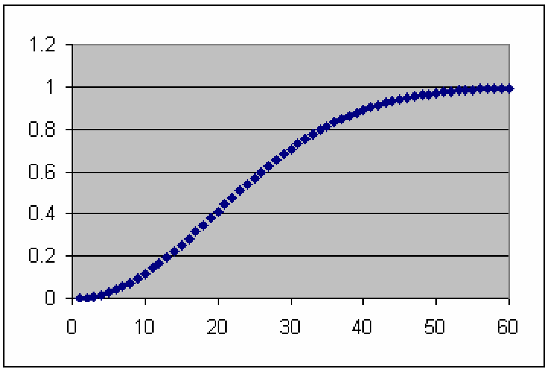
$$p_4 = \frac{364 \cdot 363 \cdot 362}{365 \cdot 365 \cdot 365}$$

e così via.

In generale:

$$p_n = \frac{364 \cdot 363 \cdot \ldots \cdot (366 - n)}{365^{n-1}}$$

Il grafico di $1-p_n$ in funzione di n è il seguente.



Torniamo alla scommessa su un evento di probabilità p e poste R e S. Un altro modo per caratterizzare un gioco equo consiste nel dire che il guadagno medio di ogni giocatore è nullo. Possiamo in prima battuta pensare al *guadagno medio* di un giocatore come il guadagno per ogni giocata in una lunga serie di giocate, diciamo N. Il primo giocatore si aspetta di vincere circa pN volte e di perdere $(1-p)N$ volte; poiché quando vince guadagna S e quando perde "guadagna" $-R$, il guadagno complessivo è

$$pNS-(1-p)RN$$

Il guadagno medio per giocata è dunque

$$\frac{pNS-(1-p)RN}{N}=pS-(1-p)R$$

Osserviamo la distribuzione del guadagno aleatorio del primo giocatore:

$$\begin{cases} -R & S \\ 1-p & p \end{cases}$$

il guadagno medio è dunque la somma dei prodotti (*guadagno*)·(*probabilità*):

$$-(1-p)R+pS$$

ed è nullo se e solo se

$$-(1-p)R+pS=0 \qquad \text{se e solo se} \qquad \frac{R}{p}=\frac{S}{1-p},$$

come ci aspettavamo.

Alla stessa relazione si giunge per il secondo giocatore, il cui guadagno aleatorio è così distribuito:

$$\begin{cases} -S & R \\ p & 1-p \end{cases}$$

Il gioco è equo se e solo se

$-pS + R(1-p) = 0$ se e solo se $\dfrac{R}{p} = \dfrac{S}{1-p}$.

In generale: il "valore medio" di un numero aleatorio X che ha distribuzione

$$\begin{cases} x_1 & x_2 & \dots & x_n \\ p_1 & p_2 & \dots & p_n \end{cases}$$

si chiama *valore atteso* di X, si indica con E(X), e si definisce nel seguente modo[3]:

E(X) = $x_1 p_1 + x_2 p_2 + \ldots + x_n p_n$

Dunque una scommessa è equa se il valore atteso di ogni giocatore è nullo.

Nel caso particolare in cui due giocatori puntano la stessa somma *S* su un evento di probabilità 1/2, il guadagno aleatorio di ciascuno è

$$\begin{cases} -S & S \\ 1/2 & 1/2 \end{cases}$$

e il guadagno atteso è

$$-\frac{1}{2}S + \frac{1}{2}S = 0$$

Sia ben chiaro: "atteso" non significa affatto che ci aspettiamo quel valore nella singola giocata; in ogni giocata o vinciamo *S* o perdiamo *S*. Il valore 0 è "atteso" semplicemente come previsione media nell'ipotesi che si possano scambiare le probabilità con le frequenze relative di successo.

In questo quadro potrebbe essere interessante porre gli studenti di fronte a eventi di probabilità non simmetriche (ma semplici, dal punto di vista classico) e chiedere loro quale somma di danaro sarebbero disposti a

[3] "E" sta per *expectation* in inglese, *esperance* in francese, *erwartung* in tedesco. In italiano non c'è una traduzione equivalente (è stato proposto *speranza matematica*, anche *previsione*) e si preferisce appunto *valore atteso*.

scommettere. Per esempio: se un ragazzo scommette una somma S contro un altro ragazzo sull'uscita del 6 lanciando un dado che ritiene regolare, allora essi dovrebbero ragionevolmente accordarsi sul fatto che il secondo debba puntare la somma $5S$, perché il suo *grado di fiducia* di vincere è 5 volte maggiore.

Con Excel è facile organizzare un sistema di scommesse su eventi E_1, E_2, …, E_n (che costituiscono una partizione di Ω) assegnando rispettive probabilità p_1, p_2, …, p_n. Per esempio, supponiamo
$p_1 = 0.2$, $p_2 = 0.1$, $p_3 = 0.7$.

	A	B	C	D
1	Eventi	Prob.	n° casuale	Esperim.
2	E1	0.2	0.298615032	E2
3	E2	0.1		
4	E3	0.7		

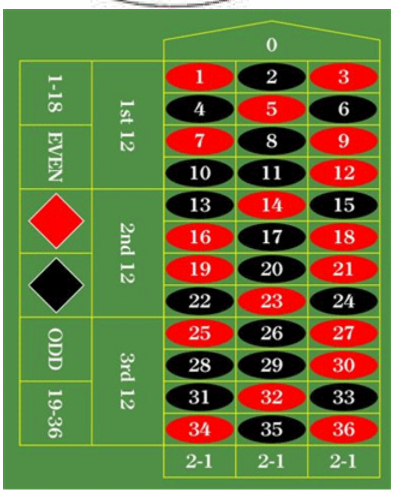
La cella D2 esegue l'esperimento e stabilisce quale evento si è verificato. Il relativo comando è il seguente.
=SE(C2<B2;"E1";SE(C2<B2+B3;"E2";"E3")
Modificando le celle della colonna B si può simulare un esperimento qualsiasi. Premendo F9 si ripete l'esperimento. Assegnando delle quote si può facilmente calcolare l'effettivo "guadagno" di ogni giocatore in un numero arbitrario di puntate.

Scommesse non eque

La richiesta che un gioco sia equo è sensata se ci limitiamo a pensare alle scommesse come contratti privati tra due giocatori che si accordano sulle poste da pagare. Nella realtà uno scommettitore gioca usualmente contro un *banco*, cioè contro un'istituzione che si fa garante del pagamento delle vincite e che quindi, in cambio del servizio offerto, trattiene una parte della posta a titolo di profitto.

Analizziamo un esempio significativo, semplice ma sufficientemente ricco: il gioco della roulette[4], il cui regolamento stabilisce quali sono le vincite di ogni puntata.

[4] Analizzo qui la *roulette francese*, quella con i 37 numeri da 0 a 36 (lo 0 non è né rosso né nero, né pari né dispari, né *manque* né *passe*).



GIOCO 1. Scommetto 1 € sul rosso; se esce un numero rosso guadagno 1 euro (il banco mi restituisce 2 €). Il mio guadagno aleatorio X e il corrispondente valore atteso E(X) sono i seguenti:

$$X \sim \begin{cases} -1 & 1 \\ 19/37 & 18/37 \end{cases} \qquad E(X) = -\frac{19}{37} + \frac{18}{37} = -\frac{1}{37}$$

GIOCO 2. Scommetto 1 € su un singolo numero; se esce quel numero guadagno 35 € (il banco mi restituisce 36 €) . Il mio guadagno aleatorio Y e il corrispondente valore atteso E(Y) sono i seguenti:

$$Y \sim \begin{cases} -1 & 35 \\ 36/37 & 1/37 \end{cases} \qquad E(Y) = -\frac{1}{37}$$

GIOCO 3. Scommetto 1 € su una *terzina* (per esempio 13-14-15); se esce un numero della terzina guadagno 11 € (il banco mi restituisce 12 €). Il mio guadagno aleatorio Z e il corrispondente valore atteso sono i seguenti:

$$Z \sim \begin{cases} -1 & 11 \\ 34/37 & 3/37 \end{cases} \qquad E(Z) = -\frac{1}{37}$$

Il valore atteso è sempre $-1/37$. Dunque non solo questo gioco non è equo (come in ogni gioco contro un banco, il giocatore ha sempre un guadagno atteso negativo), ma si può mostrare che il guadagno atteso del giocatore è $-1/37 \approx -0.027$ qualunque sia la sua puntata unitaria: questo significa in definitiva che per ogni euro puntato il banco trattiene in media 2.7 centesimi.

È ragionevole pensare che al banco non interessi quale sia la puntata, interessa fissare una quota di profitto per ogni euro puntato. Questa quota è esattamente il valore atteso di una qualsiasi puntata unitaria.

Le scommesse e le quote

sabato 7 gennaio 2012

Primera Division (Liga BBVA), Spagna

18.00	Levante	1.95	X	3.30	RCD Maiorca	3.80	+1
18.00	Racing Santander	1.95	X	3.40	Real Saragozza	3.70	+1
18.00	Real Sociedad	2.20	X	3.30	Osasuna	3.10	+1
20.00	Real Madrid	1.08	X	10.50	Granada CF	17.00	+1
22.00	Malaga	2.15	X	3.30	Atletico Madrid	3.20	+1

Un'agenzia di scommesse online assegna alla prossima partita di calcio Real Madrid-Granada le seguenti *quote*:

- vittoria Real Madrid: $q_1 = 1.08$
- pareggio: $q_2 = 10.5$
- vittoria Granada: $q_3 = 17$

L'agenzia svolge il ruolo di banco: incassa le puntate degli scommettitori e, una volta che la partita è finita, paga i vincitori. Il significato di una quota q è ben noto; per ogni euro puntato si ricevono q euro in caso di vittoria (e dunque deve essere necessariamente $q > 1$). Ad esempio, se punto

200 € sulla vittoria del Real Madrid

- ricevo 216 € (e dunque vinco 16 €) se il Real vince;
- perdo 200 € se vince il Granada oppure se le squadre pareggiano.

In generale se punto la somma *S* sull'evento E di quota *q* allora:

- incasso qS (cioè vinco $qS–S$) se E si verifica
- perdo *S* se E non si verifica (cioè se si verifica E^C, l'evento complementare di E).

È chiaro che attraverso la pubblicazione delle quote l'agenzia di scommesse comunica implicitamente quali sono le probabilità assegnate ai diversi eventi.

Come vengono valutate queste probabilità? Non certo con un rapporto tra numero di casi favorevoli e numero di casi possibili (sarebbe sempre 1/3), né attraverso l'analisi statistica di tutte le partite precedenti tra le due squadre (non è ragionevole supporre che il risultato di una partita svolta anni prima possa influenzare il risultato della prossima).

Si tratta evidentemente di una probabilità *soggettiva*, maturata sulla base del rendimento delle due squadre in quel periodo, della forma dei calciatori migliori di entrambe le squadre, e di ogni possibile informazione acquisita. Tali informazioni mutano; se, per esempio, il miglior giocatore del Real Madrid si infortunasse in allenamento e non potesse giocare la partita, la quote sarebbero inevitabilmente riviste. Le quote vengono aggiornate anche in base ad una eventuale evoluzione anomala delle puntate, per esempio nel caso in cui consistenti quote di denaro fossero puntate dagli scommettitori sul Granada.

Quei tre numeri {1.08, 10.5, 17} rappresentano dunque il frutto delle conoscenze attuali dei bookmaker su ogni possibile dato che possa influenzare l'esito della partita.

In una mente naturalmente curiosa, quale può essere quella di un ragazzo, sorgono spontaneamente alcune domande. Per esempio:

- Come vengono stabiliti quei numeri?
- Perché gli allibratori sono disposti a pagare addirittura 17 € in cambio di una puntata di solo 1 € nel caso vinca il Granada?
- Quali sono le probabilità che gli allibratori assegnano ai tre possibili esiti della partita?
- Conviene scommettere?

- È possibile scommettere contemporaneamente su più eventi al fine di realizzare un guadagno certo?

La prima ovvia considerazione è che le quote devono essere tutte maggiori di 1: nessuno scommette S per incassare meno di S e dunque perdere con certezza.

Inoltre, se gli allibratori sono disposti a pagare $17S$ a chi scommette S, è perché ritengono che la vittoria del Granada sia molto *improbabile*.

In modo analogo la vittoria del Real è quotata 1.08, con un guadagno solo dell'8% della somma puntata, perché è ritenuta *molto probabile*.

Queste considerazioni sorgono spontanee in una discussione con i ragazzi: è abbastanza immediato riconoscere che l'ordinamento delle quote è inverso a quello delle probabilità; gli eventi di probabilità maggiore hanno quote minori.

Adottiamo il punto di vista della probabilità soggettiva: la probabilità p di un evento E è il prezzo che riteniamo equo pagare per ricevere 1 in caso che E si realizzi (e 0 in caso contrario). Poiché se puntiamo S riceviamo qS, affinché sia

$qS = 1$

dovrebbe risultare

$S = 1/q$

dunque il prezzo p che riteniamo equo pagare per ricevere 1 dovrebbe essere

$p = 1/q$

Le probabilità degli eventi associate alle quote {1.08,10.5,17} risulterebbero dunque:

$$p_1 = \frac{1}{1.08} \approx 0.926 = 92.6\%$$

$$p_2 = \frac{1}{10.5} \approx 0.095 = 9.5\%$$

$$p_3 = \frac{1}{17} = 0.059 = 5.9\%$$

Ma se così fosse la probabilità totale sarebbe

$$\frac{1}{1.08} + \frac{1}{10.5} + \frac{1}{17} \approx 1.08 = 108\%$$

contro l'assioma irrinunciabile che la probabilità totale sia 1.

Il fatto è che nella relazione

probabilità di E = posta che si ritiene equo puntare per ricevere 1 in caso E si realizzi

si sottointende che la scommessa sia *equa*, e cioè che il ruolo di scommettitore e di banco siano intercambiabili. Nelle scommesse reali non è così: il ruolo di scommettitore e di banco non sono intercambiabili.

È ovvio che nelle scommesse contro un banco reale (il singolo *bookmaker*, oppure l'agenzia di scommesse, oppure il sito di scommesse online), questo, a titolo di profitto per il servizio offerto, trattiene per sé e non redistribuisce una certa percentuale delle puntate degli scommettitori.

Le "probabilità" $1/q_1$, $1/q_2$, $1/q_3$ calcolate prima sono quindi maggiori delle probabilità effettive assegnate dai bookmaker e vanno ridimensionate mantenendo la scala: si tratta di trovare tre numeri proporzionali a

$$\frac{1}{1.08},\ \frac{1}{10.5},\ \frac{1}{17}$$

e di somma 1; è allora sufficiente *normalizzare* i valori $1/q_1$, $1/q_2$, $1/q_3$, dividendo ciascuno di essi per la loro somma:

$$p_1 = \frac{1/1.08}{1/1.08 + 1/10.5 + 1/17} \approx 0.857$$

$$p_2 = \frac{1/10.5}{1/1.08 + 1/10.5 + 1/17} \approx 0.088$$

$$p_3 = \frac{1/17}{1/1.08 + 1/10.5 + 1/17} \approx 0.055$$

Oppure, più semplicemente: le probabilità e le quote sono inversamente proporzionali, il che significa

$p_1 q_1 = p_2 q_2 = \ldots$

ma la *costante di proporzionalità* non è 1; se indichiamo con C tale costante allora

$$p_1 = \frac{C}{1.08}, \qquad p_2 = \frac{C}{10.5}, \qquad p_3 = \frac{C}{17}$$

Imponiamo ora che la somma delle probabilità sia 1:

$$1 = p_1 + p_2 + p_3 = \frac{C}{1.08} + \frac{C}{10.5} + \frac{C}{17} = C\left(\frac{1}{1.08} + \frac{1}{10.5} + \frac{1}{17}\right)$$

da cui

$$C = \frac{1}{\frac{1}{1.08}+\frac{1}{10.5}+\frac{1}{17}} \approx 0.926$$

e dunque

$$p_1 = \frac{C}{1.08} = 85.7\%, \qquad p_2 = \frac{C}{10.5} = 8.8\%, \qquad p_3 = \frac{C}{17} = 5.5\%.$$

La struttura soggettiva della stima di probabilità non è invalidata: semplicemente tiene conto del fatto che tra i due poli della scommessa non c'è simmetria.

La generalizzazione è immediata: se indichiamo con

$E_1, \ldots, E_n$

n eventi che costituiscono una *partizione* dell'insieme Ω dei possibili esiti della scommessa, e cioè:

- $E_i \cap E_j = \emptyset$ per ogni $i \neq j$
- $E_1 \cup \ldots \cup E_n = \Omega$

e con

$q_1, \ldots, q_n$

le corrispondenti quote delle scommesse, allora posto

$$C = \frac{1}{1/q_1 + \ldots + 1/q_n}$$

le probabilità attribuite agli *n* eventi dalle quote sono

$$p_1 = \frac{C}{q_1}, \ \ldots, \ p_n = \frac{C}{q_n}$$

oppure, senza usare *C*:

$$p_1 = \frac{1/q_1}{1/q_1 + \ldots + 1/q_n}, \ \ldots, \ p_n = \frac{1/q_n}{1/q_1 + \ldots + 1/q_n}$$

Adottiamo lo stesso criterio di analisi che abbiamo usato per la roulette alla scommessa sulla partita Real Madrid-Granada: se punto 1 sulla vittoria del Real il mio guadagno aleatorio A_1 è 0.08 (=1.08–1) con probabilità 0.857 e –1 (perdo la posta) con probabilità 0.143 (=1–0.857):

$$A_1 \sim \begin{cases} -1 & 0.08 \\ 0.143 & 0.857 \end{cases}$$

Il valore atteso non è nullo, ovviamente:

$E(A_1) = -1 \cdot 0.143 + 0.08 \cdot 0.857 = -0.074063$

Se punto sul pareggio:

$$A_X \sim \begin{cases} -1 & 9.5 \\ 0.912 & 0.088 \end{cases}$$

$E(A_X) = -1 \cdot 0.912 + 9.5 \cdot 0.088 = -0.074063$

Se punto sulla vittoria del Granada:

$$A_2 \sim \begin{cases} -1 & 16 \\ 0.946 & 0.054 \end{cases}$$

$E(A_2) = -1 \cdot 0.946 + 16 \cdot 0.054 = -0.074063$

Come si vede il valore atteso di qualunque puntata è sempre esattamente lo stesso, come alla roulette: lo scommettitore perde mediamente ad ogni puntata (il banco vince mediamente ad ogni puntata) circa 0.074 per ogni euro puntato, cioè circa il 7.4% della posta.

In generale:

$$X_i \sim \begin{cases} -1 & q_i - 1 \\ 1 - p_i & p_i \end{cases}$$

$E(X_i) = -1 \cdot (1-p_i) + (q_i - 1) \cdot p_i = -1 + p_i + p_i q_i - p_i = p_i q_i - 1$

Ma poiché

$p_i q_i = C$

qualunque sia la puntata risulta sempre

$E(X_i) = C - 1$

Abbiamo così scoperto che la costante

$$C = \frac{1}{1/q_1 + \ldots + 1/q_n}$$

è una misura di quanto il gioco si discosti da un gioco equo, cioè a valore atteso 0.

Si può verificare che qualunque sia il sistema di scommesse proposto (cioè n eventi tali che uno ed uno solo di essi si verifichi), la costante C è sempre minore di 1: la differenza $C-1$ è sempre negativa e rappresenta il "guadagno" atteso dello scommettitore contro il banco. Come dire: alla lunga si perde sempre.

Nel caso delle partita Real-Granada
$C = 0.926$
$C-1 = -0.074$

Se analizziamo altre partite scopriamo che C è circa sempre lo stesso; per esempio:

Levante-Majorca: $\{1.95, 3.3, 3.8\}$ $\quad C = \dfrac{1}{\frac{1}{1.95}+\frac{1}{3.3}+\frac{1}{3.8}} \approx 0.927$

Espanyol-Barcellona: $\{11.5, 5.5, 1.23\}$ $\quad C = \dfrac{1}{\frac{1}{11.5}+\frac{1}{5.5}+\frac{1}{1.23}} \approx 0.924$

Real Sociedad-Osasuna: $\{2.2, 3.3, 3.1\}$ $\quad C = \dfrac{1}{\frac{1}{2.2}+\frac{1}{3.3}+\frac{1}{3.1}} \approx 0.926$

Anche se cambiamo sport e guardiamo le quote della prossima sfida di tennis tra Nadal e Federer, in cui non è previsto il pareggio, le cose non cambiano:

Federer-Nadal: $\{1.5, 2.4\}$ $\quad C = \dfrac{1}{\frac{1}{1.5}+\frac{1}{2.4}} \approx 0.923$

Viene il ragionevole sospetto che i bookmaker decidano a priori, su ogni scommessa, una percentuale di profitto medio, e calcolino poi le quote. Per esempio, se il bookmaker ritiene che le probabilità degli esiti di una partita di calcio siano rispettivamente

$p_1 = 0.5 \qquad p_2 = 0.3 \qquad p_3 = 0.2$

e se stabilisce $C = 0.9$ (cioè decide di intascare il 10% delle puntate), allora poiché

$$q_i = \frac{C}{p_i}$$

fissa le quote nel seguente modo:

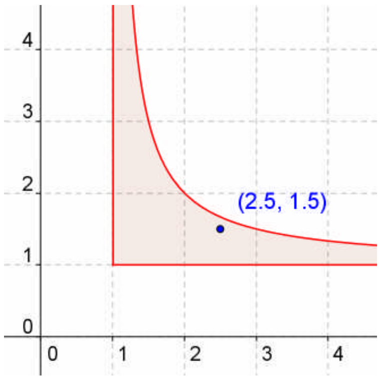
$$q_1 = \frac{0.9}{0.5} = 1.8 \qquad q_2 = \frac{0.9}{0.3} = 3 \qquad q_3 = \frac{0.9}{0.2} = 4.5$$

Nel caso di scommesse a due soli esiti E, E^C è possibile dare un'interpretazione geometrica delle quote x, y di valore atteso negativo per lo scommettitore (di valore atteso positivo per il banco):

$$C-1 < 0 \Leftrightarrow \frac{1}{\frac{1}{x}+\frac{1}{y}} < 1 \Leftrightarrow \frac{1}{x}+\frac{1}{y} > 1 \Leftrightarrow \frac{1}{y} > 1-\frac{1}{x} \Leftrightarrow \frac{1}{y} > \frac{x-1}{x} \Leftrightarrow y < \frac{x}{x-1}$$

(con $x > 1$ e $y > 1$).

L'insieme dei punti (x, y) che rappresentano le quote $\{x, y\}$ per le quali il banco ha un valore atteso positivo è rappresentato in colore nella figura seguente.



Al di sopra della curva di equazione $y = x/(x-1)$ le quote $\{x, y\}$ sono favorevoli allo scommettitore. Sulla curva stanno i punti le cui quote caratterizzano un gioco equo. Per esempio una scommessa di quote

$$q_1 = 1.5 \qquad q_2 = \frac{1.5}{1.5-1} = 3 \qquad \Rightarrow \qquad C = \frac{1}{1/1.5+1/3} = 1$$

è equa. Se $q_2 < 3$ la scommessa è favorevole al banco, se $q_2 > 3$ è favorevole allo scommettitore.

Per esempio, se $q_1 = 1.5$, $q_2 = 2.5$, ricaviamo

$$C = \frac{1}{2/3+2/5} = \frac{15}{16} < 1.$$

Si può dimostrare che se $C < 1$, cioè se la scommessa è a valore atteso negativo per lo scommettitore, allora sono possibili infiniti sistemi di puntate che conducono ad una perdita certa per lo scommettitore (e a una vincita certa per il banco), qualunque sia l'esito della partita.

Usando $q_1 = 1.5$ e $q_2 = 2.5$ come nell'ultimo esempio, supponiamo che alla chiusura delle scommesse il banco abbia incassato dagli scommettitori complessivamente 80, di cui 50 puntati su E e 30 su E^C:

- se si verifica E il banco paga $1.5 \cdot 50 = 75$ e dunque vince $80-75 = 5$;
- se si verifica E^C il banco paga $2.5 \cdot 30 = 75$ e dunque vince $80-75 = 5$.

Il banco vince 5 (circa il 6% della somma puntata) qualunque sia l'esito dell'esperimento aleatorio.

Il banco vince sempre

Si osservi che nell'esempio precedente sono state scelte delle puntate (50 su E e 30 su E^C) che sono inversamente proporzionali alle quote $q_1 = 1.5$ e $q_2 = 2.5$ e che sono quindi proporzionali alle probabilità degli eventi $p_1 = 5/8$ e $p_2 = 3/8$: si dimostra che con questa distribuzione (sempre nel caso $C < 1$) il banco ha un guadagno positivo qualunque sia l'esito della partita.

Dimostriamolo in generale: siano

- $q_1, \ldots, q_n$ le quote di n eventi $E_1, \ldots, E_n$ che costituiscono una partizione di Ω (per esempio gli esiti 1, X, 2 di una partita di calcio);
- $p_1, \ldots, p_n$ le corrispondenti probabilità, con $p_1+\ldots+p_n = 1$;
- $s_1, \ldots, s_n$ le somme complessivamente puntate dagli scommettitori su $E_1, \ldots, E_n$.

I flussi di cassa del banco sono:

- $s_1+\ldots+s_n-q_1s_1$ se si verifica E_1
- …
- $s_1+\ldots+s_n-q_ns_n$ se si verifica E_n

Supponiamo ora che $s_1, \ldots, s_n$ siano proporzionali a $p_1, \ldots, p_n$: $s_1 = kp_1, \ldots, s_n = kp_n$; i flussi di cassa del banco diventano:

- $kp_1+\ldots+kp_n-kq_1p_1$ se si verifica E_1
- …
- $kp_1+\ldots+kp_n-kq_np_n$ se si verifica E_n

Ma

$p_1+\ldots+p_n = 1$

e

$p_1q_1 = \ldots = p_nq_n = C$

dunque il flusso di cassa dello scommettitore è, qualunque evento si realizzi:

$k - kC = k(1-C)$

Poiché $C < 1$ risulta sempre $k(1-C) > 0$: se le puntate sono direttamente proporzionali alle probabilità, il banco, qualunque evento E_i si realizzi, incassa la quantità positiva $k(1-C)$.

Sorprendentemente (ma forse neanche tanto) si verifica che nella realtà accade proprio questo; il popolo degli scommettitori si orienta in generale per puntare proporzionalmente di più su eventi più probabili e proporzionalmente di meno su eventi meno probabili, facendo così proprio il gioco migliore per il banco, che si assicura una vittoria certa qualunque sia l'esito.

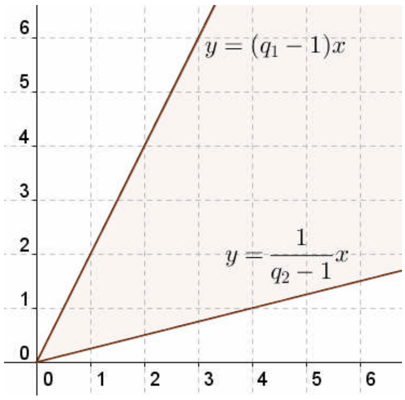
Invece, sempre nel caso $C < 1$, non esiste un'analoga possibilità per lo scommettitore: non esiste cioè un sistema di puntate su più eventi che assicuri allo scommettitore una vincita sicura qualunque sia l'esito dell'esperimento. Dimostriamolo nel caso semplice di due eventi complementari E e E^C di quote q_1 e q_2; supponiamo di scommettere x su E e y su E^C. Il guadagno risulta essere:

- $q_1x - x - y$ se si verifica E_1
- $-x + q_2y - y$ se si verifica E_2

Risolviamo il sistema

$$\begin{cases} q_1 x - x - y \ge 0 \\ -x + q_2 y - y \ge 0 \end{cases} \qquad \begin{cases} (q_1 - 1)x - y \ge 0 \\ -x + (q_2 - 1)y \ge 0 \end{cases} \qquad \begin{cases} y \le (q_1 - 1)x \\ y \ge \dfrac{1}{q_2 - 1}x \end{cases}$$

Le puntate x e y che assicurano una vincita certa allo scommettitore dovrebbero dunque essere comprese (nel I quadrante) tra le rette di pendenza $1/(q_2-1)$ e (q_1-1).



Ma se $C < 1$ allora risulta

$$C = \frac{1}{1/q_1 + 1/q_2} < 1 \Leftrightarrow \frac{1}{q_1} + \frac{1}{q_2} > 1 \Leftrightarrow \frac{q_1 + q_2}{q_1 q_2} > 1 \Leftrightarrow q_1 + q_2 > q_1 q_2 \Leftrightarrow$$

$$\Leftrightarrow q_1 q_2 - q_1 - q_2 < 0 \Leftrightarrow q_1 q_2 - q_1 - q_2 + 1 < 1 \Leftrightarrow (q_1 - 1)(q_2 - 1) < 1 \Leftrightarrow$$

$$\Leftrightarrow q_1 - 1 < \frac{1}{q_2 - 1}$$

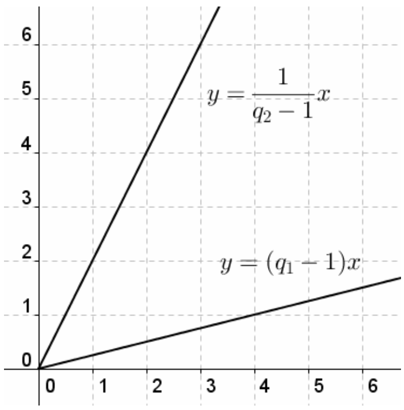
cioè la retta

$$y = (q_1 - 1)x$$

ha pendenza minore della retta

$$y = \frac{1}{q_2 - 1}x$$

e dunque l'insieme dei punti del I quadrante compreso tra le due semirette è vuoto.



Naturalmente il guadagno positivo del banco non è sempre sicuro: può accadere anzi che il banco, in particolari occasioni, vada incontro a perdite consistenti. Per esempio, supponiamo che alla chiusura delle scommesse con quote $q_1 = 1.5$, $q_2 = 2.5$, il banco abbia incassato 300, interamente puntati sull'evento E^C (quello di quota massima e di probabilità minima). Se si verifica E^C, il banco deve pagare agli scommettitori

$$2.5 \cdot 300 = 750$$

e dunque perde 450. Questo spiega perché le somme puntate sono soggette a controlli periodici e se sono eccessivamente sbilanciate i bookmaker modificano le quote in gioco. Tra l'altro, diverse inchieste sul cosiddetto *calcio-scommesse* italiano sono partite proprio da insoliti volumi di danaro puntato su eventi poco probabili.

Conclusioni

Spero di aver mostrato come sia possibile un approccio didattico alla probabilità che non faccia uso esclusivo né del rapporto tra numero di casi favorevoli e numero di casi possibili, né della frequenza relativa di successo in un gran numero di esperimenti, ma parta direttamente da una stima soggettiva della probabilità e, in modo coerente con essa, sviluppi i primi passi. A pensarci bene, anche l'approccio classico usa (senza dichiararlo, forse inconsciamente) una certa dose di soggettività, per esempio quando stabilisce che nel lancio di un dado gli esiti possibili ed equiprobabili sono gli elementi dell'insieme

$$\Omega = \{1,2,3,4,5,6\}.$$

Ma già questa impostazione è il risultato di una forte ipotesi soggettiva: il dado, una volta lanciato, potrebbe cadere in un tombino, smaterializzarsi in una nube di particelle alfa, essere mangiato dal mio gatto che sta saltando proprio in quel momento …

In sostanza il numero $p(E)$ non può essere visto come una caratteristica intrinseca, oggettiva dell'evento E, perché dipende in modo essenziale dalle informazioni che si possiedono su di esso. Supponiamo che una moneta abbia "testa" su entrambe le facce; la probabilità che lanciandola esca testa sarebbe

- 1, per chi è a conoscenza del trucco;
- 1/2, per chi non ha alcuna informazione;

- 0, per chi ha invece avuto l'informazione, per lui attendibile, che la moneta è truccata con due croci.

Dunque?

L'idea di scommessa, che de Finetti pone a fondamento semantico della probabilità, ha un impatto semplice, diretto e precocemente riconosciuto dai ragazzi; inoltre la sola ragionevole richiesta che la scommessa sia coerente è equivalente all'intero impianto assiomatico di Kolmogorov.

L'approccio classico non può che rimanere inesorabilmente legato a problemi sui giochi d'azzardo e quindi si consuma tra dadi, monete, carte, urne: sono modelli lontani dalle reali situazioni di incertezza con cui i ragazzi si confrontano e la loro valenza didattica non è forse tra le più accattivanti.

L'approccio frequentista è più generale e può essere collegato a simulazioni al calcolatore e quindi può essere didatticamente proficuo nella formulazione di congetture.

L'approccio soggettivo può spaziare ovunque: come è ben noto nel mondo anglosassone, si può scommettere su qualunque cosa!

E naturalmente non è necessario scommettere davvero, è sufficiente ipotizzare una scommessa per giungere ad una stima di probabilità.

Scoprire e verificare analiticamente che la scommessa contro un banco (chiunque lo rappresenti) non è mai equa ed è strutturalmente a sfavore dello scommettitore può forse essere utile, in tempi in cui pare che il gioco d'azzardo si avvicini ad essere una vera e propria piaga sociale (anche a livello giovanile), a far sì che i nostri ragazzi abbiano qualche anticorpo razionale in più.

Bibliografia

Kolmogorov A., 1933, Foundations of the theory of probability

Laplace P. S., 1812, Théorie Analytique des Probabilités

Barra M., 2003, Difficoltà nascoste nella didattica dei primi elementi di calcolo delle probabilità, *Progetto Alice* n° 10

Barra M., 2003, Aspetti storici e pedagogici relativi al calcolo combinatorio, *Progetto Alice* n° 11

Barra M., 2005, Alcuni problemi dell'insegnamento del calcolo delle probabilità, *Progetto Alice* n° 18

Barra M., 2006, La probabilità è nata in Italia, *Progetto Alice* n° 19

Barra M., 2011, Argomenti e termini nell'insegnamento del Calcolo delle probabilità, *Progetto Alice* n° 34

Invernizzi S., Rinaldi C., Sgarro A., 2000, Moduli di matematica e statistica, Zanichelli

Prodi G., 1992, Metodi matematici e statistici, McGraw-Hill

Scozzafava R., 2001, Incertezza e probabilità, Zanichelli

Scozzafava R., 1997, Probabilità soggettiva. Significato, valutazione, applicazioni, Masson

Ekeland I., 1997, Il caos, Il Saggiatore

Ekeland I., 1992, A caso, Bollati Boringhieri

de Finetti B., 1995, Filosofia della probabilità, Il Saggiatore

de Finetti B., 1989, La logica dell'incerto, Il Saggiatore

de Finetti B., 1970, Teoria delle probabilità, Il Saggiatore

Bucciarelli A., 2010, L'insegnamento della matematica al liceo: de Finetti vs Laplace, *Bollettino dei Docenti di Matematica* n° 61, CH

Eisenring M., 2011, Un'introduzione al calcolo delle probabilità, *Bollettino dei Docenti di Matematica* n° 62, CH

Negrini P., Ragagni M., 2005, La probabilità, Carocci Faber

Impedovo M., 2003, Un problema di dadi, *Progetto Alice* n. 9

Impedovo M., Probabilità e Bridge, *Progetto Alice* n. 11

Impedovo M., 2004, Variabili aleatorie continue e simulazioni, *Progetto Alice* n. 15

Impedovo M., 2011, Somma di numeri aleatori, *Progetto Alice* n. 36

Michele Impedovo