LEZIONE 2
La definizione classica generale di Probabilità e il paradosso di Bertrand. Le tre scuole: assiomatica, frequentistica, soggettivistica. Chi ha ragione?

La definizione classica originaria di probabilità di un evento come rapporto fra numero di casi favore-voli e numero di casi possibili, purchè tutti equiprobabili, è carente sotto vari punti di vista.
Esaminiamo tre obiezioni di cui solo le ultime due sono significative.

1) La definizione classica sottintende un circolo vizioso. Infatti nel definire la probabilità di un evento si invoca il concetto stesso che si vuole definire, nell'istante in cui si parla di eventi equi*probabili*. In realtà questa obiezione è sofistica e accademica. E' vero che il termine equiprobabile richiede già una certa conoscenza del concetto di probabilità. Ma si può facilmente rispondere con un postulato preliminare: N eventi diconsi equiprobabili se non esiste alcun motivo, né fattuale né di principio, per pensare che qualcuno di essi si verifichi con una frequenza maggiore di qualsiasi altro. Come si vede, in questa definizione non si fa ricorso al concetto di probabilità, bensì a quello di frequenza. Il suddetto postulato è generalmente noto come "principio di indifferenza". A ben riflettere, anche il concetto di temperatura viene definito presupponendo che si sappia almeno riconoscere in anticipo se due corpi hanno la stessa temperatura. Nell'apparato logico della termodinamica la definizione di corpi isotermici viene presupposta e precede la definizione operativa di temperatura, quella cioè che mette in grado il fisico di associare un numero alla grandezza stessa. La verifica che due corpi siano isotermici si può fare infatti senza misurarne la temperatura, ma semplicemente mettendoli in contatto termico e controllando che non vi sia passaggio di calore fra loro né in un senso né nell'altro.
2) La definizione classica di probabilità ha un ambito di applicabilità troppo angusto. Per esempio, non sarebbe possibile valutare la probabilità che un ago, libero di ruotare in un cerchio, si fermi casualmente all'interno di un qualsiasi assegnato settore del cerchio. Qui il problema nasce dal fatto che sia i casi possibili, sia casi i favorevoli sono infiniti, tanti quante sono le diverse possibili posizioni finali dell'ago all'interno di un angolo prescelto comunque piccolo. Per giunta tali infinità sono infinità con la potenza del continuo. Ma anche l'intervento di infinità di tipo numerabile vanifica l'applicabilità della definizione classica. Qual è la probabilità che un intero positivo, scelto a caso, sia multiplo di 5? Ognuno sa che tale probabilità è 1/5, ma tale risposta, in sé banale, non può ricavarsi usando semplicemente la definizione classica originaria di probabilità, perché sia la totalità dei numeri interi (i casi possibili) che la totalità dei multipli di 5 (i casi favorevoli) sono un'infinità di numeri. Come si possa superare questa limitazione si dirà più avanti, il che ci consentirà di arrivare alla definizione classica estesa ed anche al paradosso di Bertrand. Per ora, esaminiamo la terza obiezione, il che ci darà modo di discutere il punto di vista della scuola frequentistica.
3) Un altro limite troppo restrittivo insito nella definizione classica di probabilità è la richiesta che gli eventi elementari usati come casi possibili siano tutti equiprobabili. Facciamo qualche esempio. Qual è la probabilità che durante il 2002 qualche fulmine cada all'interno del Grande Raccordo Anulare di Roma? Possiamo dare una risposta convincente a questa domanda solo esaminando le statistiche per esempio dell'ultimo secolo. Se dal 1901 al 2000 sono stati 25 gli anni nei quali si è registrato almeno un fulmine caduto nell'area delimitata dal GRA, allora è ragionevole pensare che la probabilità richiesta sia ¼ , cioè 25%, dato su cui si baseranno le compagnie di assicurazione per stipulare le loro polizze di assicurazione che coprono i casi di infortunio e di incendio. La definizione classica invece non ci è di alcun aiuto per determinare il suddetto valore di probabilità. Infatti non è possibile per un evento del genere individuare a priori dei casi possibili, né dei casi favorevoli, contando i quali valutare la probabilità richiesta. E' vero che il valore ottenuto si presenta ancora come il rapporto fra due numeri: il numero di anni in cui è caduto qualche fulmine, e il numero totale di anni presi in esame. Ma questi numeri non possono considerarsi come numero di casi favorevoli e numero di casi possibili secondo la definizione classica, perché questi andrebbero determinati a priori, indipendentemente dai fatti osservati. Quello che invece si è fatto è calcolare semplicemente una frequenza relativa su un numero elevato di fatti osservati e assimilarla alla probabilità. Difatti una frequenza relativa per un dato evento è definita proprio come il rapporto fra il numero di volte in cui un evento si è effettivamente verificato e il numero totale di osservazioni fatte. Nel caso di un dado il procedimento equivale a rinunciare a prendere in considerazione la simmetria delle facce del dado, lanciare il dado un gran numero N di volte, registrare quante di queste volte compare il 6 sulla faccia superiore, diciamo n, e identificare la probabilità col rapporto n/N, rapporto che in realtà è la frequenza relativa osservata del 6. Questo procedimento di identificare frequenza relativa di un evento con la sua probabilità (punto di vista che è stato proposto e sviluppato a partire dagli anni venti da von Mises e Reichenbach, capiscuola della così detta **scuola frequentistica**) ha diversi difetti, se adottato in modo radicale. Primo, è un procedimento tautologico, perché il valore della definizione di probabilità consiste nel fatto che, pur essendo la valutazione di una probabilità indipendente in linea di principio dall'osservazione fattuale, il suo valore è ben approssimato da tutti i valori di frequenza relativa per lo stesso evento che si possono desumere dall'esperienza. Se invece postuliamo che la probabilità coincide con la frequenza relativa, tale identità risulta banalmente vera in virtù di una *petitio principii*, il che svuota completamente di significato

l'impegno, insito nella definizione classica, di predire alcune proprietà dei fatti osservati facendo ricorso esclusivamente ad un'analisi speculativa a priori basata sul riconoscimento delle simmetrie implicite nel fenomeno in esame. Secondo difetto, se proviamo a calcolare la probabilità come una frequenza, sorge il problema di quanta statistica vada usata per calcolare la frequenza. In effetti, nel problema del fulmine dentro il GRA, si poteva prendere come campione di riferimento non un secolo, ma gli ultimi 50 anni, oppure due secoli, e le risposte sarebbero state sicuramente diverse: avremmo facilmente ottenuto 28% o 22%, invece che 25%. Allora qual è il vero valore della probabilità in questo problema? Mentre la frequenza di un evento calcolata su campioni di taglia diversa può differire da caso a caso, il concetto di probabilità dovrebbe misurare una proprietà costante dell'evento, una sua caratteristica intrinseca, indipendente cioè dal numero di volte che noi siamo disposti ad osservarlo. Altrimenti non sarebbe sensata l'introduzione stessa del concetto di probabilità e potremmo benissimo accontentarci di parlare solo di frequenze. La scuola frequentistica ha cercato di risolvere questo problema, ma lo ha fatto nel modo più cretino possibile: la soluzione proposta è che, mentre il calcolo di una frequenza richiede un numero finito di osservazioni, per calcolare la probabilità dell'evento ne occorre un numero infinito! Infatti la probabilità viene definita dai frequentisti come il limite della frequenza relativa di un evento quando il numero di osservazioni tende a infinito! Questa scelta è il suicidio stesso del punto di vista frequentistico. Questa scuola di pensiero, partita da un'impostazione pragmatica e strettamente empirica, in antitesi alla pretesa impostazione *idealistica* insita nella definizione classica, finisce in un gesto utopistico disperato, cioè l'introduzione di un'irrealizzabile infinità di atti osservativi, che preclude per principio un vero controllo basato sui fatti osservabili. Nata insomma con la pretesa di fornire una definizione concreta e operativa di probabilità, la scuola frequentistica approda, per risolvere le sue difficoltà interne, ad una definizione che per sua stessa struttura è antitetica a tutti i principi dell'operazionismo. Purtuttavia l'approccio frequentistico è tuttora molto diffuso, soprattutto nelle facoltà di statistica e fra gli statistici professionisti e gli attuari, i quali nel loro lavoro hanno molto più spesso a che fare con problemi come quello del fulmine, piuttosto che col lancio di dadi. In effetti, bisogna riconoscere che, quando è impossibile, esaminando il fenomeno, rintracciare un numero sufficiente di simmetrie nelle sue modalità di accadere, e individuare quindi sia un insieme di casi possibili tutti equiprobabili, sia, al suo interno, il sottoinsieme di casi favorevoli all'evento preso in considerazione, il punto di vista frequentistico indica, in mancanza di meglio, una via praticabile per dare una valutazione della probabilità dell'evento. Ma in tal caso occorre essere consapevoli che tutto quello che si ha è una stima più o meno approssimata della probabilità che si cerca, non il valore vero della probabilità. Abbiamo però acquisito comunque un risultato importante. Così facendo, siamo in grado di estendere il concetto di probabilità e di usare tutto l'armamentario del calcolo delle probabilità per tentare predizioni su eventi sui quali non avremmo altrimenti nulla di quantitativo da dire. Purtroppo un frequentista puro è costretto ad andare oltre e, mentre guadagna un ambito di fenomeni alla sua capacità di analisi, ne perde uno altrettanto vasto. Infatti, se gli si presenta un dado nuovo di zecca, anche se è stato lui stesso a costruirlo con le sue mani, ponendo la massima attenzione ad evitare di privilegiare in alcun modo nessuna delle sei facce, e se gli si chiede qual è la probabilità che lanciandolo esca il 6, egli non può che rifiutarsi di rispondere e ammettere la propria assoluta ignoranza in merito. Solo dopo aver osservato migliaia di lanci di quel dado, potrà egli calcolare la frequenza relativa del 6 ed arrivare ad una stima della probabilità in questione. Un altro tipico problema davanti al quale un frequentista viene colto da paralisi afasica è il seguente: qual è la probabilità che prendendo un pianeta a caso nella galassia esso sia abitato oggi da esseri intelligenti con una civiltà tecnologica? C'è un modo di ragionare speculativo (cfr. ad esempio P. Angela, 1981: "Nel cosmo alla ricerca della vita", 2.a ediz.,Garzanti,Cap. 6), molto vicino al punto di vista classico, che porta a valutare questa probabilità come prossima ad uno su un milione (e quindi a un numero enorme di civiltà tecnologiche attualmente esistenti nella nostra galassia). Il frequentista, dal canto suo, per rispondere alla stessa domanda cercherà nell'unica statistica per ora disponibile. Si conoscono solo 9 pianeti di cui uno abitato da esseri intelligenti (intelligenti ?). Quindi l'unica frequenza disponibile è 1/9, cioè circa l'11%. Ma il campione è così poco numeroso (e così poco indicativo) che egli non crederà assolutamente a questi numeri e si vedrà costretto a tacere del tutto. In più considererà priva di qualsiasi fondamento la stima fatta con qualsiasi altro metodo, il che ci lascia senza uno straccio di risposta. Anche nel caso in cui si chieda ad un frequentista di valutare la probabilità che un intero positivo sia multiplo di 5, egli dovrà rinunciare a rispondere, a causa dell'enorme difficoltà di procurarsi un campione casuale rappresentativo dell'insieme dei numeri naturali. E invece la risposta al quesito è di una banalità disarmante, 20%. Ebbene, a mio avviso, siamo qui davanti ad un ottuso eccesso di scrupolo. Non ogni conoscenza ci viene dall'esperienza, come hanno chiarito le indagini degli strutturalisti intorno alla metà del Novecento. A volte l'informazione proviene addirittura da una completa ignoranza. Se non si ha nessun motivo per privilegiare alcuna di N diverse ipotesi, il cosiddetto principio di indifferenza autorizza un soggetto razionale ad assegnare a priori a ciascuna di queste ipotesi una probabilità pari a 1/N, senza necessità di condurre nessuna osservazione pratica. Certo, un dado può essere truccato, ed allora, se lo sospettiamo, non c'è motivo di considerare equiprobabili le sei facce. In tal caso il principio di indifferenza non può essere usato. I casi sono sì ancora 6, ma non sono più equiprobabili e quindi la definizione classica di probabilità non è di alcun aiuto. Allora sì che si è costretti ad adottare la cautela dell'approccio frequentistico, cioè ad aspettare di aver fatto molti esperimenti prima di pronunciarsi sul valore della probabilità dell'evento. Adottare questa cautela quando non v'è motivo, è d'altro canto una auto-limitazione troppo drastica, che invece di arricchire ed estendere il concetto di probabilità lo impoverisce declassandolo in tutti i casi al rango di una frequenza relativa, per giunta mai misurabile esattamente per sua stessa definizione.

Torniamo ora all'obiezione n. 2. Come valutare una probabilità quando i casi possibili non sono in numero finito? Se i casi favorevoli sono in numero finito, il problema non esiste. Un numero finito diviso un numero infinito dà zero. Per esempio la probabilità, scegliendo a caso un intero positivo, che si tratti di un numero inferiore a 1000, è zero. Ma che fare se anche i casi favorevoli sono in numero infinito? L'esempio dei multipli di 5 in realtà comporta qualche sottigliezza. Il numero dei casi favorevoli è infinito. Ora si potrebbe sostenere che la probabilità in questione è 1, cioe 100%, perché i casi favorevoli sono dello stesso numero dei casi possibili. Difatti, se si prendono tutti i multipli di 5 e si divide ciascuno per 5, si riottiene la successione di tutti i numeri naturali: 1, 2, 3, 4, 5, 6, 7, ….
Questo ovviamente non è un caso isolato. Esiste un teorema che dice che ogni insieme infinito può mettersi in corrispondenza biunivoca con qualche suo sottoinsieme. A volte questa viene assunta come la definizione di insieme infinito. Il modo corretto per rispondere alla domanda posta è però un altro. Consideriamo una stringa finita fatta dei primi N numeri naturali: 1, 2, 3, 4, 5, 6,…., N. Abbiamo N casi possibili e circa N/5 casi favorevoli, l'approssimazione essendo tanto più precisa quanto più N è grande. Al tendere di N all'infinito il rapporto fra casi favorevoli e casi possibili tende esattamente a 1/5, numero che esprime il valore della densità dei multipli di 5 nell'insieme dei numeri naturali. Questo è allo stesso tempo il valore della probabilità dell'evento considerato, stando all'unico senso ragionevole che può ad essa attribuirsi. Consideriamo ora il problema dell'ago libero di ruotare nel cerchio (si intende che un'estremità dell'ago sia imperniata nel centro del cerchio). Qual è la probabilità che l'ago si fermi in un settore qualunque, ad esempio ampio 23 gradi? Qui, se prendessimo per casi possibili ed equiprobabili tutti i punti della circonferenza e per casi favorevoli tutti i punti dell'arco che delimita il settore dato, avremmo un'infinità di casi possibili e un'infinità di casi favorevoli. Ora risulta intuitivo superare questa difficoltà facendo ricorso al concetto di misura di un insieme fatto di infiniti elementi. La misura dell'intero cerchio, intesa come angolo, è 360 gradi, mentre la misura del settore è di 23 gradi. Inoltre tutti i settori comunque piccoli ma di uguale ampiezza sono equiprobabili, almeno fino a prova contraria. Quindi viene spontaneo valutare la probabilità richiesta col rapporto 23/360=6.39%. Ciò equivale in realtà ad estendere la definizione classica come segue:

Probabilità = misura della totalità dei casi favorevoli / misura della totalità dei casi possibili

a patto che, se due insiemi (anche infiniti) di casi hanno la stessa misura, comunque piccola, gli eventi corrispondenti siano equiprobabili. Questa estensione, che fa intervenire il concetto di misura di una totalità di casi o, il che è lo stesso, di un insieme di punti (contenuto in un universo ben definito), è l'idea basilare a fondamento della generalizzazione del concetto di probabilità che caratterizza l'approccio della cosiddetta **scuola assiomatica**. Questo punto di vista è stato presentato per la prima volta in forma coerente e sistematica da Kolmogorov nel 1933 e resta tuttora l'approccio deduttivo più soddisfacente alla definizione e all'uso del concetto di probabilità. Esso parte da tre definizioni e due assiomi, dai quali è possibile dedurre tutti gli altri risultati, introdurre tutte le altre definizioni utili e calcolare ogni altro valore di probabilità cui si possa essere interessati. Innanzitutto si definisce *prova* l'esecuzione di un esperimento aleatorio dotato di un carattere di ripetibilità, nel senso che esso può essere eseguito o osservato ripetutamente senza che si alterino le condizioni che ne determinano l'esito. Una prova è ad esempio il lancio di un dado o l'estrazione di una pallina da un'urna o la rotazione casuale dell'ago dentro il cerchio fino al suo arresto. Una prova genera un insieme di esiti possibili. La totalità di questi esiti definisce l'*universo* ambiente. Questo va inteso come un insieme o uno spazio i cui punti sono tutti e soli gli esiti possibili di una prova, e sono detti anche eventi elementari. Di solito, un esito può essere caratterizzato associandogli uno o più numeri. Per esempio se si lanciano 3 dadi, ci sono 6x6x6 esiti possibili, quindi l'universo è costituito da 216 punti, ognuno dei quali è caratterizzabile in modo biunivoco con una terna di numeri interi (a,b,c), ciascuno compreso fra 1 e 6, dove a indica il numero uscito sul primo dado, b quello uscito sul secondo, c quello sul terzo. A questo punto si definisce *evento* un qualsiasi sottoinsieme dell'universo (in realtà per universi infiniti occorre limitarsi a una certa collezione di sottoinsiemi, che sia dotata della proprietà di costituire una sigma-algebra, ma qui questi dettagli non ci interessano). Infine si introduce una funzione d'insieme che ha le proprietà di una *misura*, si introduce cioè una regola per associare ad ogni sottoinsieme E (ben fatto) dell'universo U un numero P(E), non negativo, che goda delle seguenti proprietà (si pensi all'area di figure su un piano):

1) la misura dell'universo è 1: $P(U)=1$
2) se due eventi sono incompatibili (insiemi disgiunti), la misura della loro unione è la somma delle rispettive misure: $A \cap B = \varnothing \;\Rightarrow\; P(A \cup B) = P(A)+P(B)$, e lo stesso vale per un'infinità di eventi a due a due incompatibili.

La probabilità di un evento E viene definita ora come la misura dell'insieme corrispondente, cioè P(E). Da queste definizioni e assiomi si può derivare tutto ciò che serve. La prima cosa che si dimostra, per esempio, è che la misura dell'insieme vuoto è zero (cioè una prova, per sua definizione, un qualche esito lo deve pur avere), e, più in generale, che la misura dell'insieme complementare a E è $1 - P(E)$, vale a dire, se un evento ha probabilità 30%, l'evento contrario ha probabilità 70%.
Il successo della teoria assiomatica risiede nella possibilità di assimilare gli eventi a insiemi misurabili. Questo consente di usare l'armamentario matematico già sviluppato per la teoria generale degli insiemi per dedurre elegantemente e rapidamente tutte le proprietà implicate dal concetto di probabilità. L'assimilazione di eventi a insiemi si può fare perché tutte le relazioni significative fra eventi che ci interessano sono traducibili come relazioni fra insiemi e viceversa. Per

esempio due eventi incompatibili corrispondono a due insiemi che non hanno punti in comune, cioè disgiunti; il verificarsi simultaneo di due eventi è un evento ben descritto dall'insieme intersezione dei due insiemi corrispondenti; il verificarsi dell'uno o dell'altro di due eventi è un evento ben descritto dall'insieme unione dei due insiemi corrispondenti; il fatto che un insieme A sia contenuto in un altro B equivale al fatto che il verificarsi dell'evento A comporta automatica-mente il verificarsi di B (ma non necessariamente il viceversa), e così via.
Si noterà che la definizione assiomatica di probabilità, come ogni buona teoria matematica, prescinde totalmente dal significato del termine probabilità. A questo proposito è bene ricordare la famosa definizione che Bertrand Russell diede di tutta la matematica come "la scienza in cui non si sa di che si parla nè se quel che si dice è vero". La definizione assiomatica di probabilità lascia a voi la responsabilità di assegnare in modo sensato la misura ai vari eventi. Se tale assegnazione rispecchia la realtà, tutte le conclusioni che si traggono dalla teoria saranno vere, altrimenti non serviranno a nulla. In ogni caso, il controllo di validità non è compito del calcolo delle probabilità. Questo compito appartiene alla statistica, che combina un uso continuo e pesante del calcolo delle probabilità con l'analisi dei fatti osservati, al fine di accertare se questi ultimi avvengono conformemente alle assunzioni fatte circa la misura (o probabilità) introdotta nell'universo degli eventi. Il calcolo della probabilità è una teoria deduttiva. La statistica è una scienza induttiva. Senza il primo, la seconda sarebbe disarmata, ma senza quest'ultima il primo si ridurrebbe a uno sterile esercizio.

Esaminiamo ora due aspetti di questo approccio assiomatico. L'esperimento di cui si parla nella definizione assiomatica di probabilità, da cui origina il concetto di prova e, quindi, di universo degli eventi, si è pattuito che sia aleatorio e ripetibile. Il paradosso di Bertrand nasce da una critica all'uso ambiguo dell'attributo "aleatorio". La scuola soggettivistica nasce da una critica molto più sostanziale alla necessità dell'attributo "ripetibile".

Vediamo prima il **paradosso di Bertrand**. Quando si dice che un esperimento ha un esito casuale, possiamo voler dire cose diverse. Un'accezione, la più debole, implica che, comunque si esaminino le circostanze che lo accompagnano, risulta impossibile prevedere con certezza quale dei suoi vari esiti possibili si produrrà. Un'accezione più forte consiste nell'aggiungere la condizione che ogni esito dell'esperimento è tanto probabile quanto qualsiasi altro. Se estraiamo "a caso" una carta da un mazzo, con tutta probabilità intendiamo l'accezione forte. Se estraiamo "a caso" una lettera dell'alfabeto da un testo di italiano, non presupponiamo invece che la lettera "e" sia altrettanto probabile della lettera "z". E' bene quindi precisare, quando si usa l'espressione "a caso" o una simile, chiarire esattamente le operazioni che si compiono per eseguire l'esperimento aleatorio. Difatti, due soggetti, pur avendo entrambi in mente l'accezione forte del termine "casuale", potrebbero avere in mente procedimenti diversi che portano a stime di probabilità differenti per lo stesso evento. Un esempio che chiarisce bene questa sottile ambiguità insita nel termine "casuale" è il seguente problema che apparentemente può essere risolto in diversi modi, ognuno impeccabile, ma ognuno con una risposta diversa.
"Si traccia a caso una corda in un cerchio di raggio r. Qual è la probabilità che la corda sia più lunga del lato del triangolo equilatero inscritto nel cerchio dato?"
Si possono esibire almeno quattro modi (e ce ne sono anche altri) per tracciare a caso una corda, e ognuno risponde alla percezione intuitiva che abbiamo circa il significato della locuzione "traccia-mento casuale di una corda".
Un primo modo è scegliere un punto qualsiasi sulla circonferenza e, fissato questo come primo estremo della corda, scegliere il secondo estremo facendo ruotare casualmente un ago imperniato al centro del cerchio, che al suo arresto indicherà un altro punto della circonferenza, da prendere come il secondo estremo della corda.
Un secondo modo è, fermo restando il modo di scegliere il primo estremo della corda, scegliere il secondo estremo facendo ruotare l'ago non attorno al centro del cerchio ma attorno al primo estremo. Quando l'ago si sarà fermato, la retta che contiene l'ago intersecherà la circonferenza data in un solo punto, che sarà il secondo estremo della corda.
Un terzo modo consiste nel prendere a caso (accezione forte) un numero L compreso fra zero e la lunghezza del raggio e nel far ruotare un ago di lunghezza L attorno al centro del cerchio. La punta libera dell'ago si fermerà in un punto del cerchio, per il quale si condurrà la retta perpendicolare all'ago. Questa individuerà sul cerchio la corda casuale cercata.
Un quarto modo consiste nello scegliere a caso (accezione forte) un numero A compreso fra zero e l'area del cerchio e nel tracciare un cerchio di area A concentrico a quello dato. Si traccia poi la retta tangente al cerchio di area A in un punto qualsiasi della sua circonferenza. Tale retta taglierà una corda sul cerchio originario.
Ora, se si fanno i conti nei quattro casi, nei primi due la risposta viene uguale a 1/3, nel terzo viene uguale a 1/2, nel quarto viene uguale a 1/4. Questa pluralità di risposte è paradossale. Solo una dovrebbe essere quella giusta. Eppure il procedimento è ineccepibile in ognuno dei quattro casi. Dov'è l'errore? L'errore sta a monte, cioè nel pensare che basti dire "estrarre a caso" perché ogni soggetto razionale interpreti tale espressione in modo univoco. L'errore sta insomma nell'ambiguità insita nella formulazione linguistica del problema. Occorre specificare con maggiore precisione il modo con cui si deve procedere nell'estrazione casuale. Ognuno dei quattro modi indicati è perfettamente legittimo come metodo di estrazione casuale di una corda e, a seconda di quale metodo si usi, la risposta trovata è quella giusta. Il paradosso di Bertrand non ha quindi una soluzione univoca. Esso illustra il pericolo di un uso troppo leggero dei termini e mostra in modo esemplare come sia privo di senso parlare di probabilità di un evento, qualora l'esperimento aleatorio di cui si parla nella definizione assiomatica non sia stato chiaramente specificato.

Esaminiamo infine la critica della scuola **soggettivistica** alle altre definizioni di probabilità. La nascita di questo approccio si deve al lavoro svolto indipendentemente dal logico inglese Frank Ramsey e dal matematico italiano Bruno de Finetti. Ammettiamo di chiederci la probabilità che:

a) domani piova
b) l'Ulivo vinca le elezioni del 13 maggio
c) la Roma vinca lo scudetto dell'attuale campionato
d) il prossimo papa non sia europeo
e) le quotazioni della azioni FIAT a fine anno valgano la metà di adesso

A tutte queste domande non si riesce ad associare un esperimento aleatorio che sia **ripetibile**. Quindi il metodo assiomatico non è in grado dare risposte. Per lo stesso motivo fallisce il metodo frequentistico, a meno di farne un uso completamente scorretto (per es.: in passato la Roma ha vinto solo 3 scudetti sui 60 campionati di calcio cui ha partecipato, quindi ha solo una probabilità del 5% di vincere quest'anno lo scudetto! Oppure: non c'è mai stato un papa che non sia europeo, quindi la probabilità che il successore di Woytila sia non europeo è nulla).
La conclusione di chi aderisca alla scuola assiomatica o alla scuola frequentistica è pertanto la stessa, e cioè che per i problemi (a)-(e) non si possa parlare di probabilità, se mai di plausibilità di un evento, e che tale plausibilità non possa essere misurata quantitativamente senza cadere in seri errori.
Al contrario i sostenitori del punto di vista soggettivistico definiscono provocatoriamente :

"Probabilità di un evento è la massima somma di denaro che un soggetto razionale è disposto a scommettere a fronte di una vincita lorda unitaria".

Per esempio, se si lancia un dado, un soggetto razionale cui si prospetti una vincita lorda di un milione nel caso che esca il 6, non dovrebbe accettare di versare una posta maggiore della sesta parte di un milione per entrare nella scommessa, quindi 1/6 è la probabilità dell'evento dato. Questa definizione alquanto spregiudicata di probabilità in realtà non è in contrasto con le due precedenti ma le supera e le sussume entrambe. In sostanza essa non pone limiti, oltre alla richiesta di razionalità del soggetto, sul modo in cui si perviene ad un'assegnazione di probabilità agli eventi elementari. In particolare non richiede che l'evento in questione sia dotato di ripetibilità. E' ovvio che nei casi in cui si applica il metodo classico (cioè speculativo) o frequentistico, la valutazione sarà la stessa. Ma quando entrambi falliscono, per esempio a causa della non ripetibilità dell'evento in questione, talora l'informazione disponibile sulle circostanze che precedono l'evento, consente al soggettivista ugualmente di esprimere quantitativamente, sebbene con un grado di approssimazione più o meno alto, il grado di fiducia che ripone nel verificarsi del dato evento. A riprova di ciò si noti che tutti e cinque gli eventi su elencati, tranne il primo, sono oggetto di scommesse presso i bookmakers professionisti che da vari decenni esercitano la loro attività traendo lauti profitti dalle loro stime soggettivistiche di probabilità. L'approccio soggettivistico non è in realtà alternativo all'approccio assiomatico: una volta assegnate le probabilità basilari agli eventi elementari, il resto segue automaticamente come nel metodo assiomatico. L'unica differenza sta nel fatto che nei casi in cui la natura dell'esperimento aleatorio esclude la ripetibilità, l'universo degli eventi è, per così dire, puramente ipotetico, invece che concretamente generabile dagli esiti delle prove ripetute. Di tali prove se ne potrà fare in realtà solo una nel corso di tutta l'esistenza del cosmo, ma ciò non toglie la possibilità che un soggetto razionale, sulla scorta di un certa quantità di informazione acquisita non importa come, possa trovare vantaggiosa una scommessa che in caso di successo paghi 1 a fronte di una posta per esempio di 0.34.
Quanto poi al controllo degli assunti fatti, qualunque sia il punto di vista che si adotta, resta sempre imperativo il dovere di confrontare le proprie previsioni con le osservazioni fattuali future. La parola finale spetterà dunque in ogni caso all'analisi statistica . Allora perché accettare a priori inutili restrizioni sulla natura dell'esperimento aleatorio o più in generale sul metodo di assumere la distribuzione iniziale di probabilità per gli eventi elementari? Anche se l'esperimento aleatorio non fosse ripetibile, un soggetto razionale può trovarsi nelle condizioni di poter fornire una buona stima quantitativa della probabilità dell'evento. In realtà, ognuno di noi ogni giorno fa decine di valutazioni del genere, per esempio quando decide se prendere o no l'ombrello uscendo di casa, se prendere il raccordo o tagliare per il centro, se investire i risparmi in buoni del tesoro o in azioni, se ritenere colpevole o innocente l'imputato di un processo, a quale ora uscire per arrivare puntuale ad un appuntamento, quale partito votare, quale carriera di lavoro intraprendere o a che tipo di scuola iscrivere i propri figli. Tutte queste scelte riguardano eventi unici, irripetibili, eppure l'uomo medio non se la cava poi così male nel prendere decisioni ragionevoli in tutti questi casi (tranne forse nel caso del gioco in Borsa). Potremmo dire quindi che, se non altro grazie alla pressione selettiva subita durante l'evoluzione, il nostro cervello è stato costretto ad elaborare dei metodi efficaci per valutare la probabilità di eventi futuri da cui può dipendere la nostra sopravvivenza. Se questa capacità esiste, perché vietarsene per principio l'uso? Tanto vale metterla alla prova, quando non si ha niente di meglio, ed affidare alle verifiche a posteriori la conferma della nostra presunta efficienza nel fare previsioni. Tutto sommato, una conoscenza approssimata e soggettiva dei fatti è in genere meglio che il non conoscerli affatto, specialmente nei casi (e sono tanti) in cui si scopre che i risultati finali del calcolo (cioè le probabilità derivate, che serviranno a prendere le decisioni) sono in fin dei conti poco influenzati dai valori iniziali che il soggetto ha dovuto assegnare alle probabilità elementari, da cui il calcolo stesso dipende. In questi casi, casi che sarebbero stati respinti *tout cours* come non matemazzibili nell'ambito degli altri approcci, l'approccio soggettivistico ha avuto il merito (e il coraggio) di applicare ugualmente i metodi di calcolo quantitativi, così consentendo per un verso scoperte interessanti che non si sarebbero altrimenti mai fatte, e per l'altro verso fondando a posteriori le ragioni della propria validità e fecondità.

BIBLIOGRAFIA

Feller W. , 1957 : *An introduction to probability theory and its applications - Vol. 1*. J.Wiley & Sons, 3rd edition.
Von Mises R., 1936 : *Wahrsceinlichkeit, Statistik und Wahrheit*, Springer-Verlag, Vienna.
Papoulis A., 1965: *Probability, random variables and stochastic processes,* McGraw-Hill (traduz. italiana di M. Bozzini e R. Tirani, Boringhieri, 1973), Capitolo 1.
Davenport W.B., 1970 : *Probability and random processes*. McGraw-Hill, 542 pp. (Capp. 1-3).
Kolmogorov A.N., 1956 : Foundations of the theory of probability. Chelsea Pub. Co. , New York (traduz. in inglese di un articolo apparso nel 1933 sulla rivista tedesca Ergebn. Math, vol. 2)
Ramsey, F.P. , 1931 : in *The foundations of mathematics and other logical essays* (a cura di R.B. Braitwaite). Routledge & Keagan, London. Traduzione italiana: *Verità e probabilità: I fondamenti della matematica e altri scritti di logica*. Feltrinelli (Milano, 1964)
De Finetti B., 1989 : *La logica dell'incerto*. Mondadori, Il Saggiatore (Milano)