Probabilità

1. *Considerazioni introduttive.*

1.1. La probabilità: chi è costei?

Prima di rispondere a tale domanda è certamente opportuno chiedersi: ma davvero «esiste» la probabilità? e cosa mai sarebbe? Io risponderei di no, che non esiste. Qualcuno, cui diedi questa risposta (ribadita, col motto in tutte maiuscole – PROBABILITY DOES NOT EXIST – nella prefazione all'edizione inglese di *Teoria delle probabilità* [1970]), mi chiese ironicamente perché mai, allora, me ne occupo.

Mah! potrei anche dire, viceversa e senza contraddizione, che la probabilità regna ovunque, che è, o almeno dovrebbe essere, la nostra «guida nel pensare e nell'agire», e che per ciò mi interessa. Soltanto, mi sembra improprio, e perciò mi urta, vederla *concretizzata* in un sostantivo, 'probabilità', mentre riterrei meglio accettabile e piú appropriato che si usasse soltanto l'aggettivo, 'probabile', o, meglio ancora, soltanto l'avverbio, 'probabilmente'.

Dire che la probabilità di una certa asserzione vale 40 per cento appare – purtroppo! – come espressione concreta di una verità apodittica. Non pretendo né desidero che tale modo di esprimersi vada bandito, ma certo è che l'asserzione apparirebbe assai piú appropriatamente formulata se la si ammorbidisse dicendo, invece, che quel fatto lo si giudica «probabile al 40 per cento», o, meglio ancora (a parte che suona male), che ci si attende «al 40 per cento – probabilmente» che sia o che risulti vero.

Il guaio è che il realismo (come acutamente osservò Jeffreys) ha il vantaggio che «il linguaggio è stato creato da realisti, e per di piú da realisti molto primitivi», ed è per ciò che «noi abbiamo larghissime possibilità di descrivere le proprietà attribuite agli oggetti, ma scarsissime di descrivere quelle direttamente conosciute come sensazioni» [1939, p. 394].

Da ciò la mania (che forse per altri è invece indizio di saggezza, serietà, acutezza) di assolutizzare, di concretizzare, di oggettivizzare perfino quelle che sono soltanto proprietà dei nostri atteggiamenti soggettivi. Non altrimenti si spiegherebbe lo sforzo di fare della Probabilità qualcosa di *nobler than it is* (sempre parole di Jeffreys), nascondendone la natura soggettiva e gabellandola per oggettiva. Secondo la spiritosa fantasia di Hans Freudenthal si tratterebbe di uno strano pudore per impedire di farci vedere la Probabilità «come Dio l'ha fatta»: occorre «una foglia di fico», e spesso la si riveste tutta di foglie di fico rendendola addirittura invisibile o irriconoscibile.

1.2. Le probabilità: pretesamente oggettive.

Vi sono molti casi (spesso banali, ma anche no) nei quali le valutazioni di probabilità dei vari individui coincidono o tendono a coincidere. È abbastanza «na-

turale» (anche per un profano) attribuire, ad esempio, probabilità 40 per cento all'estrazione di pallina bianca da un'urna che ne contiene 100 di cui 40 bianche, o anche (piú o meno) se, anziché conoscerne la composizione, si sa che in 100 estrazioni con reimbussolamento (e rimescolamento) le estrazioni di palline bianche sono state 40.

Si è detto «naturale» (e tra virgolette) perché si tratta pur sempre di un giudizio probabilistico soggettivo (anche se, di solito, appare naturale e viene accettato da tutti). Non si tratta di sottigliezza sofistica: si tratta del fatto che un'opinione, in quanto tale, è sempre soggettiva, personale; è, cioè, tutt'altra cosa che un dato oggettivo (quali, ad esempio, la vera composizione dell'urna o l'effettiva frequenza osservata). A prescindere poi dal fatto che, per quanto riguarda la frequenza, essa è solo «probabilmente» vicina alla composizione dell'urna, e varia da un gruppo di «prove» a un altro.

È tuttavia un fatto che, in casi siffattamente schematici, piú o meno tutti giungono a valutazioni piú o meno concordanti, considerate per ciò dalla piú parte degli autori come espressione di «probabilità *oggettive*». Ma sarebbe piú appropriato, in tali casi, e verrà qui (ove occorra) seguito, l'uso del termine neutro 'probabilità *pubbliche*', suggerito da Leonard Jimmie Savage (acutissimo pensatore e impareggiabile amico, scomparso, purtroppo, anzi tempo), oppure, come mi è sembrato (ripensandoci) ancor piú appropriato, 'probabilità *consuete*' (conformi a consuetudine): è infatti inutile, ingiustificato e fuorviante attribuir loro qualifiche piú ambiziose.

Vero è, come dato di fatto, che il consenso su certe valutazioni di probabilità è spesso piú o meno generale. E ciò costituisce un fatto concreto, una circostanza che può avere interesse in sé (ed essere utile in quanto favorisce mutua comprensione e consenso). Ma – attenzione! – essa non avrebbe alcun valore, avrebbe anzi un valore *fallace e negativo*, se venisse fraintesa come velleitaria e pretestuosa giustificazione di credenze di tipo superstizioso: la credenza, anzitutto, nella «esistenza» di una fantomatica «probabilità oggettiva», magari camuffata di volta in volta sotto le tradizionali spoglie della dea Fortuna e della strega Scalogna, cui attribuire tutto quel poco o tanto di bene e di male che a ciascuno viene largito.

Queste considerazioni introduttive non pretendono, né potrebbero, fornire fin d'ora indicazioni positive sul senso in cui occorre intendere la nozione di probabilità, precisando e perfezionando l'idea intuitiva che tutti ne abbiamo. Al contrario, sono intese a sgombrare il terreno da troppe idee preconcette, sia grossolane o sofisticate, che tuttora imperversano.

1.3. La presente occasione.

L'occasione che scaturisce dall'iniziativa di questa *Enciclopedia* sembra suscettibile di favorire un costruttivo chiarimento, un sostanziale passo in avanti nell'auspicata direzione.

I due magistrali articoli di Stefan Amsterdamski, «Caso/probabilità» e «Causa/effetto» (vol. II, pp. 668-87 e 823-45), aprono infatti la visuale su di una va-sta e complessa tematica, assai analoga a quella prevista – sia pure come «sottofondo» – per la presente trattazione: una trattazione di carattere piú tecnico ma anche concettuale, la quale risulterà pertanto arricchita e meglio precisata in un puntuale confronto.

Confronto e *non* contrapposizione, direi, in quanto si tratta di proporre e cercar di giustificare una scelta univoca e precisa entro il largo ventaglio delle opzioni prospettate, o almeno non escluse, nei due già citati articoli.

In forma schematica, e approfittando della possibilità di far riferimento all'ampia e approfondita panoramica di Amsterdamski, posso precisare fin d'ora la mia posizione in poche parole dicendo che, delle due interpretazioni della probabilità ivi prospettate (pp. 674-75), escluderei senz'altro la prima secondo la quale «le asserzioni probabilistiche riguarderebbero gli eventi e sarebbero analitiche», mentre potrei accettare – in una versione invero molto radicalizzata – la seconda, riformulandola come segue: «La probabilità, pur essendo sempre una caratteristica dei giudizi, non è mai un concetto logico; le asserzioni contenenti valutazioni probabilistiche non sono mai analitiche in quanto esprimono sempre e soltanto il grado di credenza che, nel suo presente stato d'informazione, il soggetto che giudica attribuisce all'oggetto dell'asserzione. Sinteticamente, essa caratterizza, cioè, l'atteggiamento del soggetto conoscente nei riguardi di una data asserzione».

Per chiarire la situazione in forma piú esplicita basta chiedersi quali risposte può dare una persona interrogata riguardo a un *evento*, cioè a *una data affermazione* (dotata di senso univoco e per lei comprensibile). Evidentemente, le risposte possibili, tra cui ciascuno può scegliere quella che corrisponde allo stato delle sue attuali conoscenze al riguardo, sono, in senso *oggettivo*, tre: «Sí», «No», «Non so». La differenza essenziale fra le tre risposte sta nel fatto che (in qualunque versione) le due estreme: «Sí» (o «Vero», o «Certo») e «No» (o «Falso», o «Impossibile») sono dotate di un senso univoco, di un carattere definitivo e categorico, mentre quella intermedia «Non so» (o «Dubbio», o «Incerto») non ha invece che un carattere provvisorio in quanto esprime solamente il perdurare di una attuale ignoranza o indecisione tra il «Sí» e il «No», che sono le sole due risposte definitivamente concludenti.

In tale situazione di incertezza, ciascuno potrà propendere piú o meno sensibilmente per il «Sí» o per il «No», ed esprimere tale sua propensione dicendo che l'affermazione gli appare piú o meno probabile. Ma frasi del genere sono sempre vaghe, non impegnative, di dubbia interpretazione, magari a volte anche volutamente equivoche, come o quasi come i famigerati responsi della Sibilla, del tipo: «Ibis, redibis / non / morieris in bello», con libertà di immaginare la virgola prima o dopo del «non».

1.4. Come eliminare tale vaghezza?

Chi si limita ad esprimere la propria opinione dicendo che qualcosa è «molto» o «poco» probabile, che la sua probabilità è piú o meno «piccola» o «grande», dice ben poco; comunque, niente di preciso. È però sempre possibile (e, quando l'indicazione vaga non basta, necessario) tradurre il proprio convincimento, il proprio grado di fiducia, in un'indicazione numerica, come 10 per cento, 40 per cento, 75 per cento di probabilità. E non c'è dubbio che ogni persona, anche poco o affatto istruita, sappia esprimere correttamente in tale forma le proprie opinioni e, analogamente, comprendere il significato di quelle esposte da altri. Si vedranno, ad esempio, nei §§ 1.7-1.8, i cenni esplicativi riguardanti i pronostici probabilistici sul calcio.

Semmai, il rischio è quello di esser stati deviati, allontanati, dalla concezione naturale causa una certa moda incomprensibilmente imperversante, favorevole a certe disgraziate concezioni della probabilità, banali, artificiose, e, per sovrappiú, fuorvianti e limitative delle capacità d'intendere di quanti vi si assuefanno.

La sola concezione che (come si spiegherà in seguito) abbia senso, l'unica che comporti una vera comprensione del significato e della validità del ragionamento probabilistico, è quella genuina di un qualunque «uomo della strada»: quella che ci guida in ogni attimo ed azione della nostra vita, anche se inconsciamente, con elaborazioni mentali e sintesi istintive piú rapide di quelle di un qualunque calcolatore elettronico.

Il «calcolo delle probabilità» (in quanto *calcolo*) può servire in casi artificialmente complessi, ma sempre considerandolo come un sussidiario dell'intuizione e non come sostituto (o come possibile sostituto) di essa.

In questo senso, si dovrebbe insistere soprattutto per far considerare il ragionamento probabilistico *non come un sostituto* bensí come *uno strumento integrativo* delle capacità intuitive che tutti (uomini ed altri animali) possediamo. Tali capacità, secondo una felice espressione (di cui mi spiace non ricordare chi ne sia l'autore) costituiscono una *built in machinery* nel nostro cervello (un macchinario innato). Ed è quindi da ciò che si deve partire. Si tratta (si ponga ben attenzione!) di comprensione effettiva, anche se un po' rozza, ed occorrerà soltanto *approfondirla e affinarla*; sarebbe invece un *regresso* sostituire questa comprensione intuitiva e pratica con delle *pseudodefinizioni* (!) della *probabilità*: pseudodefinizioni – purtroppo di *moda*! – che si autodefiniscono *oggettivistiche*. Se ne riparlerà a suo tempo e luogo.

Ma occorre anzitutto indicare (come verrà fatto nei prossimi §§ 1.5-1.6) un procedimento *operativo* atto a *misurare* la probabilità di un evento *E*: ripetiamo (meglio ripeterci fino alla noia pur di evitare il rischio di fraintendimenti) che 'evento' significa «caso unico ben determinato».

Attenzione: abitualmente il termine 'evento' viene invece usato in senso generico, per indicare tutti gli eventi di un certo tipo, detti «prove» di quell'«evento». Ciò comporta molti inconvenienti ed inestricabili confusioni; per evitarli si potrebbe dire che certi eventi (analoghi) sono «prove di un medesimo *fenomeno*» (ma senza intendere con ciò che siano ugualmente probabili o indipendenti od altro salvo che non sia esplicitamente detto). La terminologia attuale è, oltretutto, ambigua, perché a volte si considera anche il caso in cui la probabilità di un «evento» *varia* di prova in prova (ma allora vuol dire che la probabilità si riferisce, anche per gli oggettivisti, non all'«evento» secondo l'accezione oggettivistica, bensí all'evento (caso singolo) secondo la terminologia conforme alla concezione soggettivistica. Ciò dovrebbe bastare per far riconoscere a chiunque – chissà come cosí non è? – che la concezione (e la stessa terminologia) degli *oggettivisti* altro non è che una vuotaggine confusionaria).

1.5. «Previsione» e «scarto» (quadratico medio).

Finora si sono considerati soltanto eventi e loro probabilità, ma si tratta solo di un caso particolare di quello piú generale dei numeri aleatori e della loro previsione. Come caso particolare (già noto) è un numero aleatorio ogni evento, se, con convenzione di cui si vedrà sempre meglio l'appropriatezza, si identifica l'evento E col numero aleatorio che vale 1 se E è vero e 0 se è falso. (Spesso lo si chiama «indicatore di E», ma senza alcun costrutto: una distinzione senza differenza non crea che apparenti ed inutili doppioni di parole, oltre a contravvenire una norma sacrosanta: «Entia non sunt multiplicanda sine necessitate»).

Un numero aleatorio, X, può assumere un numero finito di valori (come i punti da 1 a 6 con un dado o da 2 a 12 con due dadi, o da 1 a 90 alla tombola) o tutti i valori reali entro un intervallo «verosimile» se si tratta ad esempio della «temperatura di domani mattina» a un dato osservatorio meteorologico.

Nel caso *discreto* (valori possibili in numero finito: $x_1, x_2, ..., x_n$) basterà indicare le probabilità $p_1, p_2, ..., p_n$ attribuite a ciascuno di essi, e si avrà una distribuzione *discreta*; altrimenti una distribuzione *continua*, nel caso piú regolare con una densità, $f(x)$ (probabilità $f(x)\,dx$ che X cada tra x e $x+dx$, per dirla in termini comprensibili anche se criticabili). C'è anche un caso intermedio (in certo senso «patologico»): vedansi la figura 1 e relativa didascalia. (Per maggiori informazioni cfr. l'articolo «Distribuzione statistica» in questa stessa *Enciclopedia*).

Il nostro attuale obiettivo è molto limitato ed elementare; tuttavia, il modo di considerarlo è inteso a preparare il terreno per una discussione semplice ma

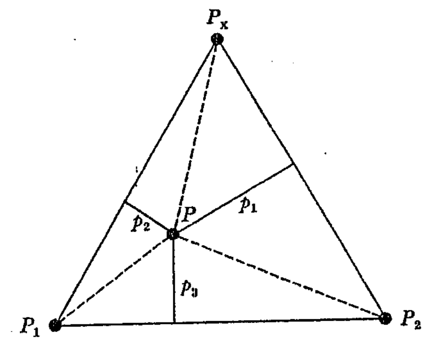

Figura 1.

Triangolo equilatero: per ogni punto interno P la somma delle tre distanze dai lati costante (e precisamente è uguale all'altezza). Perciò, nel caso di tre eventi incompatibil ed esaustivi (come i risultati «1», «x», «2» in una partita di calcio), ogni opinione sull'esi to di una data partita (probabilità p_1, p_x, p_2) è rappresentabile come un punto del triangolc

standard deviation

critica sul modo appropriato di valutare la previsione, $m = \mathbf{P}(X)$, nonché lo scarto (quadratico medio) di X, che si indica con sigma di X, $\sigma(X)$. Il suo quadrato $\sigma^2(X) = \mathbf{P}(X-m)^2$ (con $m = \mathbf{P}(X)$) si chiama «varianza» di X. In parole: σ è la radice della media di $(X-m)^2$, cioè del quadrato degli scostamenti di X dalla media m. Come dice il nome stesso, σ fornisce una misura (inversa) dell'addensamento della distribuzione attorno al valor medio (o, in termini meccanici, baricentro). Si può anche considerare lo scarto quadratico medio da un punto (o valore) x diverso dalla media m; lo si indichi σ_x. È facile vedere che $\sigma_x^2 = \sigma^2 + (x-m)^2$; in termini geometrici, σ_x è l'ipotenusa del triangolo rettangolo di lati σ_x ed $(m-x)$, e quindi il baricentro è il punto rispetto al quale il momento è minimo. (Ed è, del resto, intuitivo che, se l'asse di rotazione passa per il baricentro, la massa vi è ravvicinata e il momento diminuisce). Ed è questa la conclusione che serve: il baricentro, oltre che come punto di equilibrio, è anche caratterizzato dall'essere il punto rispetto al quale il momento è minimo; si dispone pertanto di due metodi per trovare il baricentro di un solido (nel caso che ci interessa: una sbarra): 1) è il punto per cui si deve sospendere la sbarra affinché rimanga in equilibrio; 2) è il punto della sbarra che occorre tener fisso affinché, facendo ruotare la sbarra (beninteso, a parità di velocità angolare), l'energia o (come forse è piú familiare) la «forza viva», sia minima.

Immagino e comprendo lo stupore del lettore: chi mai farebbe tanti tentativi per misurare l'energia per rotazioni con assi diversi fino a individuare il minimo di «forza viva» e quindi il baricentro? Ha ragione, anzi ragioni da vendere...; ma, nel caso che ci interessa, l'aspetto meccanico scompare e rimane per analogia la questione di convenienza tra gli analoghi metodi nel contesto probabilistico.

E qui sta il punto: nel caso della probabilità la misura diretta, anziché essere la piú appropriata come nel caso meccanico, si riduce a profferire una cifra per la probabilità «ad occhio», senza alcun ausilio di controlli o correttivi; invece la procedura indiretta – cioè una «stima», ma collegata ad una «penalizzazione» (appropriata, nel senso di *proper scoring rule*) serve ad affinare la sensibilità degli «stimatori» e a vagliarne l'abilità tenendo conto (mediante i punteggi, *scores*, di ogni stima) dell'abilità dimostrata nel complesso della loro attività in tale campo. Naturalmente, se da una parte occorre buona compenetrazione con lo spirito del procedimento, occorre – e del resto è questa la motivazione del farne uso – una buona competenza e informazione nel campo specifico (nel nostro esempio, (cfr. §§ 1.7 e 1.8), valore e situazione delle squadre del campionato di calcio). Quanto alla regola di penalizzazione quadratica, che già avevo applicata in concorsi probabilistici sul calcio, appresi poi che era già nota (*Brier's rule* 'regola di Brier') ed applicata in America per dare indicazioni probabilistiche per la pioggia nei bollettini meteorologici diffusi ogni mattina da radio, Tv e giornali.

Beninteso, non avrebbe senso pensare che una previsione basata su questo tipo di procedure ed informazioni sia di per sé migliore (sarebbe miracolismo!); sta di fatto, però, che il metodo fornisce un autocontrollo, nonché un controllo comparativo se gli addetti o partecipanti a tali pronostici sono parecchi e possono (a posteriori) confrontare pronostici e risultati di tutti e diagnosticare il perché taluno va piú bene che male e talaltro piú male che bene.

	00	10	20	30	40	50	60	70	80	90	100
$(1-x) =$	100	90	80	70	60	50	40	30	20	10	00
$y=x(1-x) =$	00	09	16	21	24	25	24	21	16	09	00
Ordinate delle tangenti nel punto 0 =	00	01	04	09	16	25	36	49	64	81	100
Ordinate delle tangenti nel punto 100 =	100	81	64	49	36	25	16	09	04	01	00
Dislivello =	+100	+80	+60	+40	+20	00	−20	−40	−60	−80	−100
Retta a =	09	13	17	21	25	29	33	37	41	45	49
Retta b =	64	58	52	46	40	34	28	22	16	10	04

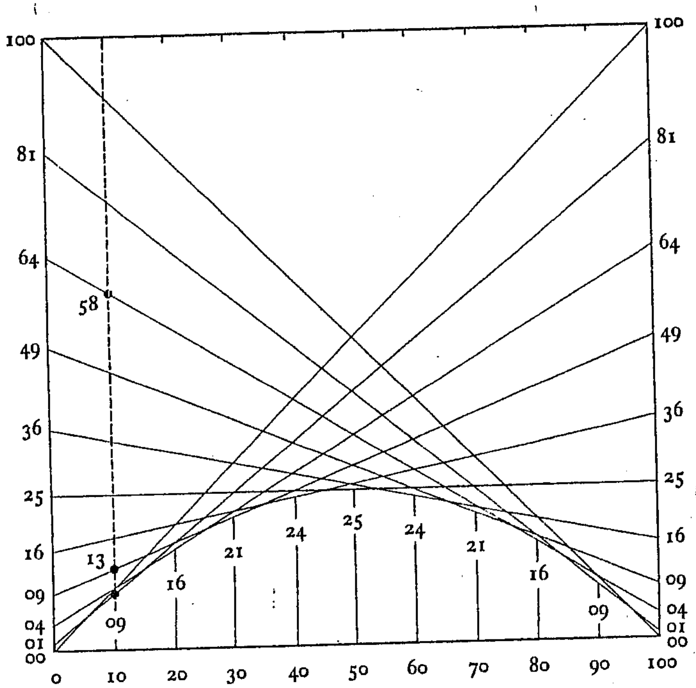

Figura 2.

Le rette corrispondono alle combinazioni di penalizzazione tra cui il metodo consente di scegliere (si può ridurre la penalizzazione in uno dei due casi a spese di un aumento nell'altro: per abbassare l'ordinata in un estremo si alza nell'altro). L'ordinata di una retta nel punto p è la previsione di penalizzazione per chi sceglie quella retta e attribuisce all'evento in questione la probabilità p. In tal caso il minimo ottenibile è dato dall'ordinata della parabola (nessuna retta vi passa al di sotto!) e la scelta ottima è quella della retta tangente alla parabola in corrispondenza all'ascissa p.

1.6. Una presentazione alternativa.

Può riuscire istruttiva e appropriata, sotto vari punti di vista, un'illustrazione anche in forma grafica del senso e del funzionamento delle valutazioni basate sulla minimizzazione del quadrato dell'errore (che, nel gergo statistico, si chiama «regola di Brier»).

La figura 2 mette in evidenza, visivamente, come e *perché* una regola di penalizzazione appropriata obblighi ciascuno, nel suo proprio interesse, a comportarsi in accordo con quanto segue dalla sua valutazione di probabilità e ad esprimere sinceramente tale sua valutazione.

La figura rappresenta un quadrato di lato unitario con l'arco di parabola $y = x(1-x)(0 \leq x \leq 1)$ e le tangenti ad essa per ogni decimo dell'ascissa.

(Per comodità tutti i valori sono indicati moltiplicati per 100, cioè, ad esempio, 100 anziché 1 e 24 anziché 0,24).

1.7. Pronostici probabilistici.

È importante intrattenersi sull'argomento dei pronostici probabilistici per vari motivi.

Il motivo teorico consiste nel mostrare come il concetto informatore della «regola di Brier» si trasporti, con le stesse utili proprietà, dal caso di due sole eventualità (eventi) a quello di tre (o piú).

Il motivo esemplificativo-psicologico consiste nell'illustrare la validità educativa di esercizi sistematici di valutazioni di probabilità, riferendo su concorsi di pronostici probabilistici riguardanti le partite del campionato di calcio.

Ed infine, incidentalmente ma appropriatamente, verrà messa in luce l'antitesi di mentalità di educatività e di moralità (in senso lato) tra i giochi-scommesse in cui si stimola la sciocca «furbizia» del «tirare a indovinare» e quelli in cui si tratta di dare una valutazione quanto piú «obiettiva» e spassionata possibile.

Quanto all'educatività e all'importanza pratica, si vedranno subito dopo (§ 1.9) le analoghe esperienze nel campo (nientemeno!) delle prospezioni petrolifere!

Nel caso del calcio (come per molti altri giochi) i risultati possibili per ogni partita sono tre: «1» = vittoria, «x» = pareggio, «2» = sconfitta (sempre con riferimento alla squadra ospitante). Ogni pronostico probabilistico consiste pertanto nell'indicare le tre probabilità, p_1, p_x, p_2 (di somma = 1, ossia 100 se le indicazioni sono fatte, come è usuale, in percento), ed è opportuno pensarle come masse (o «pesi») collocate nei vertici P_1, P_x, P_2 di un triangolo equilatero. La penalizzazione è il quadrato della distanza tra il punto-pronostico, P, e il punto-risultato: P_1 o P_x o P_2: ovvia estensione della «regola di Brier» a due (e anche, volendo, piú) dimensioni.

Statisticamente, la proporzione dei tre risultati «1» - «x» - «2» è in media 50 : 30 : 20 (50 per cento per vittoria in casa, 30 per cento per pareggio, 20 per cento per vittoria esterna), ma, è chiaro, queste sono indicazioni statistiche globali, mentre ogni caso singolo differisce per molte circostanze da tener presenti e vagliare attentamente: rapporto di valori tra le due squadre che si affrontano, fattore campo, condizioni del tempo, assenze e sostituzioni di giocatori per malattie, incidenti, squalifiche, importanza per la classifica dell'una e/o dell'altra squadra, ecc.: tutte cose che ciascuno dovrà tener presenti, nella misura in cui ne è informato, e vagliare con cura.

È bene sottolineare e ricordare insistentemente, per non cadere in distorsioni di visuale di tipo superstizioso o fatalistico o metafisicheggiante, che non si tratta di «scoprire» un preteso e fantomatico «valore vero» di ascose «probabilità oggettive», bensí di indicare il valore che *ciascuno a suo modo* (come nella commedia di Pirandello) vi attribuisce. Sperabilmente, lo farà previa attenta riflessione sui pro e sui contro, in conformità alla misura in cui ciascuno propende per l'una o l'altra delle tre possibilità. È *questa* la probabilità nell'*unico senso che appare valido*, universalmente valido.

(Esistono però, come si vedrà (§ 2.3), altre sedicenti «definizioni» che non possono venir considerate e accettate come tali, ma soltanto valide come criteri ausiliari per la valutazione – sempre, beninteso, soggettiva – delle probabilità).

1.8. Pronostici e concorsi pronostici.

Concorsi pronostici sulle partite del campionato di calcio, nell'illustrata forma *significativa*, sono stati ripetuti per parecchi anni (presso l'Università di Roma, con partecipazione anche di colleghi e di studenti di altre sedi). In forma «significativa» significa «nel modo già indicato»: significa cioè che si tratta dell'opposto del diseducativo criterio del «tirare a indovinare», del «tentare la fortuna», come al Totocalcio ove si tratta di «predire» il risultato secco (o «1» o «x» o «2»), o come al famigerato gioco del Lotto. Il quale – sia detto per inciso – concorre anche, indirettamente, a perpetuare incorreggibili diffuse idiozie, quali l'attesa con crescente fiducia di numeri «arretrati» o suggeriti da sogni o da astrologi o da «maghi» o dalla «cabala» o da calcoletti cervellotici... e chissà cos'altro!

Nulla vi è in comune, nei pronostici probabilistici, con tali disgustose forme di «predizioni secche», che appena di poco appaiono meno peggiori del sullodato Lotto (in quanto, nel calcio, la scelta fra «1» - «x» - «2» implica almeno un po' di riflessione). Meno male che tali superstiziose fole e scimunitaggini giovano allo Stato, e che molti cittadini, magari evasori fiscali o riluttanti e dispiaciuti nel pagare le debite tasse ed imposte, dimostrano un immenso anche se involontario (e pertanto non meritorio) zelo nel versare abbondantemente denaro per tale «tassa sulla imbecillità».

Il concorso pronostici probabilistico richiede invece ad ogni partecipante di indicare per ogni partita le probabilità che egli attribuisce ai tre risultati possibili, e, trattandosi di una «regola di penalizzazione appropriata», ha convenienza ad esprimerle sinceramente ed esattamente, dopo aver vagliato il valore delle squadre, preso nota di assenze per malattie o squalifiche, dello stato di forma dei giocatori, delle condizioni meteorologiche previste e dell'influenza di tutto ciò

(e di quant'altro ritiene rilevante) sul rendimento dell'una e dell'altra squadra.

Tale attività si è dovuta interrompere da qualche anno causa difficoltà create da scioperi e disservizi nel servizio postale e da analoghe complicazioni entro l'Università. È allo studio la possibilità di riprenderla in altra sede e con partecipazione di pronosticatori piú qualificati (giornalisti sportivi ed altre persone legate all'ambiente calcistico e sportivo).

Comunque, sarebbe auspicabile che tale capacità di esprimere in termini di probabilità il grado di attendibilità o di fiducia che uno attribuisce a risultati possibili di una qualsiasi azione o iniziativa venisse apprezzata e incoraggiata, al fine di venire effettivamente sfruttata, con consapevolezza e coerenza, per vagliare accuratamente il pro e il contro di ogni elemento che influisce sul risultato di ogni possibile decisione.

Particolarmente significativo e istruttivo a tale riguardo risulterà l'esempio che verrà illustrato nel § 1.9, considerate anche le necessarie nozioni di «numeri aleatori» e loro «previsione» introdotte nel § 1.10.

1.9. Ruolo della previsione in decisioni importanti.

Gli esempi finora introdotti riguardavano situazioni piú o meno di carattere ludico, in particolare risultati sportivi, e ciò sembrava utile per avviare e far entrare nello spirito della trattazione senza dover superare – sperabilmente – eccessive riluttanze. Forse, dopo aver appreso e meditato il modo in cui le valutazioni probabilistiche hanno un ruolo essenziale in situazioni di gioco, risulterà però ora – sempre «sperabilmente» – accettabile l'affermazione che i medesimi criteri e procedimenti sono applicabili, come sono stati effettivamente applicati, con risultati significativamente validi in campi ove è altissima l'*importanza pratica* di una attenta e accurata valutazione (da parte di esperti dei diversi rami) dei fattori e delle circostanze che, con le loro probabilità, incidono sulla probabilità da attribuire ad ipotesi di risultati globali piú o meno favorevoli.

Molte di tali questioni sono trattate sotto l'etichetta di «Ricerca operativa» (*Operation Research*), e alcuni esempi semplici, ma utili a scopo illustrativo, si possono vedere nell'articolo «Decisione» di questa *Enciclopedia* (vol. IV, pp. 421-84).

Ma l'esempio piú significativo, e in cui meglio appare la connessione fra tante valutazioni fatte da esperti diversi, è certamente quello relativo alla decisione di intraprendere, e poi di proseguire, e in quale modo, le ricerche petrolifere in una data zona, oppure di abbandonarle.

Presupposto per tale decisione è l'acquisizione di elementi di giudizio (geologici, ecc.) da parte di esperti, di una attenta e accurata valutazione da parte loro delle prospettive di successo o insuccesso – in termini di utile o perdita – di una tale costosissima impresa.

Riguardo ad esperienze su questo particolare ma assai rilevante e istruttivo problema, e al tipo di argomentazioni interconnesse cui conduce, vale la pena di segnalare soprattutto il libro *Decisions under Uncertainty: Drilling Decisions by Oil and Gas Operators* di Grayson jr [1958]. Egli descrive come sia riuscito ad ottenere dagli esperti (geologi, ingegneri, ecc.) di esprimere in *valutazioni probabilistiche* (*numeriche*) i loro giudizi sulle prospettive di successo di ricerche in una data località, anziché usare (come in precedenza era abituale) frasi studiatamente vaghe e ragionamenti sofisticati con comode riserve a titolo cautelativo... quasi ad imitazione della già menzionata Sibilla.

In base ad informazioni probabilistiche dettagliate (cioè, concernenti varie sottoipotesi sulla natura e ricchezza dei presunti giacimenti) diviene possibile anche stabilire, mediante un'analisi delle previsioni probabilistiche relative a diverse circostanze, la convenienza o meno (speranza di risparmio o timore di perdita) per ogni ulteriore esperimento di questo o quel tipo (ad esempio, perforazione di un pozzo di sondaggio o prospezione sismica) atto a consigliare o sconsigliare, a seconda dell'esito, la decisione finale (o, eventualmente, quella di rinviare la decisione procedendo, prima, ad ulteriori indagini).

Probabilmente molti saranno perplessi e troveranno ridicolo fare dei calcoli «campati in aria» (assimilando ad «aria», magari ad «aria fritta», le probabilità soggettive sia pure stimate da esperti); certamente, esse non possono dare alcuna certezza, ma l'indicazione di un grado di probabilità presentato come tale è il massimo grado ottenibile di informazione oggettiva: un'indicazione comunque molto piú dotata di senso di responsabilità, e quindi di attendibilità, che non una «certezza» fasulla, asserita con leggerezza, o un responso «oggettivo» ma ambiguo. Confucio, del resto, non aveva già detto che la parola 'certezza' era una di quelle che si sarebbero dovute abolire?

Tutto ciò appare naturale per chiunque, libero da preconcetti assolutistici, tenga conto del fatto che tutto è incerto, ma che per decidere occorre e basta basarsi su ciò che si sa (con certezza) e su ciò che si ritiene probabile, piú o meno probabile, sulla base di ciò che si sa e di ciò che non si sa. Ed ogni informazione arricchisce questo sfondo sempre incompleto, ma soltanto l'onniscienza potrebbe completarlo: guai a chi, rinunziando ad avvalersi dell'informazione possibile, decide a vanvera o rinunzia a decidere (o decide secondo pregiudizi generici, ritenendo di dovervisi attenere alla cieca, senza vagliarne l'appropriatezza e l'opportunità che variano caso per caso).

1.10. Probabilità, previsione, prezzo.

Ma – ci si potrà obiettare – non è un'inutile complicazione il riferimento a «regole di penalizzazione» dal momento che ciò (come si è visto) equivale all'affermazione banale, chiara per chiunque, che $\mathbf{P}(E)$ (sia ad esempio $\mathbf{P}(E)=0{,}40$) significa che 0,40 Lire è il *prezzo equo* per ricevere una Lira se E si verifica? Si usa 'una Lira' come termine generico: chi sentisse il bisogno di riferirsi a una scala piú attuale potrebbe intendere per 'Lira' una Kilolira (mille Lire) od altro importo a suo piacimento. Meglio però non troppo piccolo da rendere insignificante il risultato né troppo grande per evitare il divario tra valore monetario e *utilità* (cfr. il già citato articolo «Decisione»).

Una siffatta brutale identificazione della probabilità a prezzo avrebbe però il difetto di condurre in una deprecabile situazione di «gioco», nel senso magistralmente esposto da John von Neumann e Oskar Morgenstern nella loro famosa opera *Theory of Games and Economic Behavior* (1947) (cfr. l'articolo «Giochi» in questa stessa *Enciclopedia*).

Una tale situazione di «gioco» dà spesso adito, infatti, ad astuzie, a mercanteggiamenti o tentativi di mercanteggiamento, sicché il prezzo non sarebbe un dato certo e significativo su cui ci si possa basare. Se non si ponesse attenzione a tali inconvenienti la stessa probabilità verrebbe a confondersi con un frutto di patteggiamenti, di un labile compromesso tra chi vorrebbe spendere meno e chi vorrebbe incassare piú di quanto potrebbe venire ragionevolmente stabilito.

Questa critica non inficia tuttavia l'idea di considerare la probabilità come un prezzo: è soltanto necessario ricorrere ad uno «strumento di misura» insensibile ai menzionati fattori di distorsione. E tali strumenti – le «regole di penalizzazione appropriate» – si conoscono già, pur non avendone finora rilevato la proprietà che qui interessa.

A questo punto (per non ripetere due volte lo stesso discorso) conviene introdurre, oltre agli eventi, anche i «numeri aleatori», ad esempio $X=x_1E_1+x_2E_2+\ldots+x_nE_n$ (dove $E_1, E_2, \ldots, E_n$ formano una *partizione*: sono cioè incompatibili ed esaustivi, nel senso che se ne verifica certamente uno e uno solo): X è pertanto (come mostra la scrittura) il numero aleatorio che assume il valore x_1 se si verifica E_1 (e cosí via: $x_2, \ldots, x_n$ se si verificano, rispettivamente, $E_2, \ldots, E_n$).

Naturalmente, si possono considerare anche numeri aleatori con un'infinità (discreta o continua) di valori possibili: ad esempio, pensando ad un numero X (qualunque, o soltanto razionale) scelto «a caso» – cioè con densità uniforme di probabilità – tra 0 e 100, e quindi con probabilità $(x''-x')/100$ di trovarsi in qualunque intervallino (x', x'') contenuto in $(0, 100)$. Ma, per il momento, ci si limita al caso elementare di valori possibili in numero finito per non dover parlare di derivate e integrali.

Per sviluppare l'argomento in termini matematici (pur cercando di evitare discorsi in forma astrusa per non spaventare i profani) è necessario introdurre alcuni concetti e simboli (del resto già usati in casi particolari).

Anzitutto il simbolo $\mathbf{P}$, comodo per indicare indifferentemente sia *probabilità* (nel caso di eventi, ad esempio $\mathbf{P}(E)$), e sia *previsione* (nel caso di numeri aleatori, per esempio $\mathbf{P}(X)$). Però, con una interpretazione unitaria e banale, riferentesi a una scommessa unitaria, $\mathbf{P}(X)$ si può anche dire «*prezzo* di X» (prezzo da pagare per ricevere l'importo incognito X quando sarà noto), e cosí $\mathbf{P}(E)$, prezzo di E (di «una Lira» se si verifica E).

1.11. Probabilità (e previsione): sempre subordinate.

Parlare (come è stato fatto finora, «sic et simpliciter») di eventi e di numeri aleatori, come di enti cui riferire probabilità e rispettivamente previsione, è però un nonsenso. Per giustificare tale colpa occorre dire che era tuttavia utile far cosí per evitare troppe complicazioni tutte d'un colpo e per attirare maggiormente l'attenzione su di esse ora, facendo notare e correggere la provvisoria («calcolata») dimenticanza.

Dire che la probabilità di un dato evento, E, vale, ad esempio, $\mathbf{P}(E)=0{,}40$, non è un'affermazione avente un senso compiuto, a meno che non si pensi sottinteso il secondo *essenziale* fattore: il nostro attuale stato di conoscenza. Lo si indichi con H_0. Allora, a rigore, si dovrebbe indicare la scrittura completa, cioè $\mathbf{P}(E|H_0)$. Se quella che si vuole considerare è la probabilità subordinata all'ulteriore conoscenza o «ipotesi» H, quindi ad HH_0, si avrà $\mathbf{P}(E|HH_0)$ ove H_0 serve per rammentare lo stato di conoscenza attuale, mentre H è l'ipotesi aggiuntiva sotto la quale ci interessa stimare la probabilità di E.

Anziché $\mathbf{P}(E)$ e $\mathbf{P}(E|H)$ *dovremmo* pertanto scrivere sempre $\mathbf{P}(E|H_0)$ oppure $\mathbf{P}(E|HH_0)$, rispettivamente per ricordare e indicare quale sia il nostro stato di conoscenza, oppure, inoltre, quale sia l'ulteriore *ipotetica* circostanza H da aggiungervi, interessando conoscere quale sarebbe detta probabilità condizionandola a tale ampliata conoscenza (o informazione).

Ho detto «dovremmo», e non «dovremo», perché la continua indicazione e ripetizione di H_0 risulterebbe inutilmente ingombrante. Tuttavia, andrà *sempre* tenuto presente che questo «H_0» dovrà *sempre* intendersi *sottinteso*, mai soppresso come cosa superflua. E può essere sottinteso soltanto se dal contesto risulta in modo non dubbio quale sia la situazione (per quanto riguarda le circostanze rilevanti al riguardo).

La probabilità di un evento E dato un H si esprime, in base al «teorema delle probabilità composte»: $\mathbf{P}(EH)=\mathbf{P}(E)\cdot\mathbf{P}(H|E)$ od anche (è ovvia la simmetria) $\mathbf{P}(EH)=\mathbf{P}(H)\cdot\mathbf{P}(E|H)$.

Tenendo conto di tale identità è possibile ricavare, per $\mathbf{P}(E|H)$, l'espressione seguente:

$$\mathbf{P}(E|H)=\frac{\mathbf{P}(EH)}{\mathbf{P}(H)}=\mathbf{P}(E)\frac{\mathbf{P}(H|E)}{\mathbf{P}(H)}.$$

A parole: la probabilità di E, subordinandola ad H, si modifica nel medesimo rapporto in cui si modifica la probabilità di H subordinandola ad E.

È questo il fondamentale *teorema di Bayes*, base del ragionamento induttivo, e in particolare della statistica matematica (quando non venga ridotta a ricettari empirici piú o meno grossolani).

La principale fonte di errori e malintesi, nel campo probabilistico-statistico, consiste proprio nel considerare certi dati come se fossero dotati di senso assoluto, non pensando che esso è sempre relativo a un certo stato di conoscenze.

Eppure ciò sembra difficile da far capire (o «inghiottire»: a molti ripugna)! Quanti non insistono nel sostenere che esistano «probabilità oggettive» (e perché no, allora, anche quadrati circolari?!)

A chiunque parli di probabilità oggettive si dovrebbe dare una risposta drastica: la sola probabilità oggettiva, per un qualunque evento E, è $\mathbf{P}(E|E)=1$ nell'ipotesi che E si verifichi e $\mathbf{P}(E|\tilde{E})=0$ nell'ipotesi che E non si verifichi. (Su questo punto si veda anche il § 3.9).

1.12. Il «punto», dopo le considerazioni introduttive.

Gli argomenti e le considerazioni finora svolti hanno (come già espresso nel titolo) carattere e scopo introduttivo, ma sotto una duplice visuale: l'una di chiarire alcuni aspetti generali del ragionamento probabilistico e del suo significato effettivo, e l'altra di precisarli (quanto piú elementarmente possibile, ma in modo concettualmente preciso) come preparazione alla trattazione matematica (e, necessariamente, piú organica) da svilupparsi a suo tempo.

È quindi opportuno, in questo momento, fare il «punto» della situazione cui si è giunti, riflettendo sinteticamente su ciò che è stato detto e delineando un abbozzo panoramico degli argomenti ed aspetti che andranno sviluppati in seguito. Naturalmente, gli sviluppi comporteranno in genere una trattazione in forma matematica, senza però appesantirla con tecnicismi; fatta – si potrebbe dire – per aiutare a comprendere il «succo», in forma matematica, anche a coloro che sono o si sentono «digiuni» in matematica ma non cadono nell'errore di rifiutare ogni aiuto per capire una spiegazione in forma idonea per chiunque abbia interesse ad afferrare il «succo» usualmente nascosto «sotto il velame delli *sgorbi* strani»: quegli «sgorbi» che sono, per lui, le formule e i simboli che vi compaiono.

In chiusura di questo primo paragrafo è necessario indicare alcuni simbolismi e forme di scrittura che occorreranno in seguito: aiutano alla concisione e al risparmio di spazio, e quindi alla chiarezza. (Purché uno si degni di abituarvisi: è un po' faticoso – specie per coloro che si sentono «profani» o «refrattari» alla matematica – ma vorrei dire loro, per incoraggiarli (ma *con convinzione*, non per illuderli o per ingraziarmeli), che non si tratta né di loro inettitudine né di indigeribilità della matematica, bensí di indigeribilità dell'insegnamento matematico formalistico-mnemonico-astratto nelle scuole; salvo, beninteso, parecchie lodevoli eccezioni).

Oggetto della teoria delle probabilità sono gli *eventi* e i numeri aleatori (potrebbero considerarsi anche punti aleatori, funzioni aleatorie, passeggiate aleatorie, processi aleatori, ecc.). Gli eventi aleatori si indicano in genere con E e indici (E_0, E_1, E_2, ...) oppure altre maiuscole (A, B, C, ...); i numeri aleatori con maiuscole a fine alfabeto (X, Y, Z, ... oppure X_0, X_1, X_2, ...)

Il simbolo $\mathbf{P}$ significa sia *probabilità* sia *previsione*: probabilità se riferito a un evento, ad esempio $\mathbf{P}(E)$; previsione se riferito ad un numero aleatorio, ad esempio $\mathbf{P}(X)$; il simbolo $\boldsymbol{\sigma}$ indica lo scarto quadratico medio (o «scarto standard»): $[\boldsymbol{\sigma}(X)]^2=\mathbf{P}(X-m)^2$ dove $m=\mathbf{P}(X)$.

Ogni evento, E, si identifica col numero aleatorio che vale 1 se E si verifica e 0 se non si verifica. Le operazioni aritmetiche hanno (naturalmente) il medesimo significato per numeri aleatori che nel caso abituale; interessa però aggiungerne altre: col segno $\sim$ («tilde»: il segno che in spagnolo si sovrappone alla lettera n (ñ) per farla pronunziare come «gn» in italiano; ad esempio «giugno») si indica il complemento ad 1: $\sim x=1-x$ (sovrapponendolo: $\tilde{x}$, quando si tratta di una sola lettera); in particolare, per un evento, E, $\sim E$ (o $\tilde{E}$) significa «negazione di E» (infatti, il segno «tilde» scambia vero con falso e viceversa).

Altre operazioni logiche (su eventi ma anche su numeri, aleatori o no) sono quelle di «sup» e «inf», indicate con $\vee$ e $\wedge$: $A \vee B$ indica il maggiore (e, analogamente, $A \wedge B$ il minore) tra i numeri A e B (e lo stesso per piú termini: ad esempio $A \wedge B \wedge C$ ed $A \vee B \vee C$ significano rispettivamente che *tutti* i tre eventi sono veri, o che lo è almeno uno). Per dare un esempio un pochino piú complesso, $(A \vee B) \wedge (C \vee D)$ significa che c'è almeno un evento vero in entrambe le parentesi.

Per dare un esempio relativo a numeri aleatori, basta pensare che gli A, B, C, D del caso precedente siano numeri qualunque (non piú solo 0 o 1): in tal caso il significato è «il minore tra i massimi di ciascuna coppia».

A seconda delle sue preferenze, il lettore potrà cercar di assimilare fin dall'inizio tali concetti e simbolismi, oppure ricordare che può ricorrere a queste pagine ogni qual volta abbia bisogno di decifrare un caso singolo o di rinfrescarsi le idee.

2. *Molteplicità anche di concezioni.*

2.1. Un preambolo pirandelliano.

Parafrasando un brano di Pirandello nel romanzo *Uno, nessuno, centomila* («parafrasandolo» col sostituire «probabilità» a «realtà» e «sento» a «mi do»), il discorso potrebbe iniziare cosí: «Ci fosse fuori di noi, per voi e per me, ci fosse una signora probabilità mia e una signora probabilità vostra, dico per se stesse, e uguali, immutabili. Non c'è. C'è in me e per me una probabilità mia: quella che io sento, e una probabilità vostra in voi: quella che voi sentite; le quali non saranno mai le stesse, né per voi né per me».

Sarebbe stato impossibile, senza l'aiuto di Pirandello, esprimere questo concetto (e, in nuce, l'essenza della nostra tesi) in un modo cosí preciso, completo, efficace; rimane però da chiarire la specifica interpretazione – anzi, le due *opposte* interpretazioni – in cui potrebbe sembrare appropriato intenderlo nel presente contesto.

Questa citazione pirandelliana si presta infatti – nel tentativo qui presentato di suo adattamento in campo probabilistico – a due diverse interpretazioni, esprimenti i due aspetti complementari delle tesi qui contrapposte; quella *soggettivista* dove si ha unicità d'interpretazione e molteplicità di valutazioni, e quella *oggettivista* dove si ha una molteplicità d'interpretazioni ciascuna delle quali si traduce nell'unicità (o pretesa unicità?) della corrispondente valutazione.

Per chiarire un po' meglio fin d'ora le posizioni contrapposte dei soggettivisti e degli oggettivisti si aggiungono le precisazioni che seguono:

Nel campo dei *soggettivisti* si ha un'*unica* concezione ammissibile basata soltanto sul requisito della *coerenza*, e dove la definizione in senso operativo della probabilità si traduce nella «regola di Brier» (o simili; cfr. § 1.6).

Ma è proprio nell'ambito di tale concezione che l'illimitata molteplicità delle valutazioni di probabilità ammissibili (conformi all'opinione di ciascuno: «Ciascuno a suo modo») si presenta come cosa naturale e *necessaria*. Necessaria in

dipendenza del fatto stesso che ogni valutazione (ciascuna delle «una-nessuna-centomila») è, per definizione: *soggettiva*, nel senso che riflette *non* circostanze oggettive, oggettive di per sé, bensí l'opinione che se n'è fatto, sia pure *in base ad esse*, l'individuo che le valuta, e *specifica*, nel senso di riferirsi specificamente alla probabilità di un «evento», inteso sempre come «caso singolo univocamente individuato» nelle date circostanze (e *non* al modo degli oggettivisti che usano «evento» in senso generico e chiamano «prove di tale evento» tutti gli eventi di quel certo tipo).

Nel campo degli *oggettivisti* si ha invece – secondo il gusto di ciascuno: «Ciascuno a suo modo» – una fungaia di (una? nessuna? centomila?) «definizioni» (piú o meno cervellotiche, e che piú appropriatamente, come si vedrà, dovrebbero dirsi «pseudodefinizioni»), le quali – almeno nelle pie intenzioni dei loro fautori – dovrebbero conferire, *motu proprio eorum*, alla probabilità di ogni evento («evento» da interpretarsi – quel che è peggio – come un ammasso incomprensibile e stravagantemente «collettivistico»?!) il diritto a fregiarsi del titolo onorifico di «oggettive». È inutile dire quale assurda confusione ciò possa ingenerare; la miglior prova è data dagli stessi oggettivisti che distinguono – palesemente contraddicendosi! – il caso in cui «tutte le prove» siano «ugualmente probabili» e il caso in cui la probabilità «varia di prova in prova». Accettando sul serio tale formulazione, sarebbe naturale concludere che ogni evento ha probabilità o uno o zero a seconda che si verifichi oppure non si verifichi!

Poiché tentativi prematuri di spiegazioni e chiarimenti riguardo alle molteplici diatribe sul significato delle probabilità (diatribe che si riducono di regola a «dialoghi tra sordi») riuscirebbero oscuri e finirebbero per confondere ancor piú le idee anziché facilitarne la comprensione, sembra consigliabile seguire una via di mezzo: *dapprima* (nel seguito di questo § 2) prospettare il senso delle diverse concezioni e discutere l'appropriatezza di diverse terminologie e notazioni, aggiungere qualche cenno storico al riguardo, discuterne la validità (se esiste, ed entro quali limiti), ma sempre a scopo di preliminare orientamento in vista della trattazione piú approfondita e precisa, che *poi* (nel § 3, ed ultimo) sarà sviluppata un po' piú col necessario rigore (anche matematico). La lettura e comprensione dovrebbe tuttavia risultare facilitata anche ai lettori non troppo agguerriti in fatto di conoscenze matematiche, dato che gli sviluppi di formule e i risultati matematici, nella maggior parte, non saranno che la traduzione in termini precisi di quanto sarà già stato fatto intravedere da varie osservazioni critiche che verranno sviluppate nel seguito del presente secondo paragrafo. E, naturalmente, sarà considerato acquisito quanto premesso nel § 1.

In particolare, e soprattutto, si tenga sempre presente la «regola di Brier» (cfr. § 1.4), che sarà sempre considerata come lo strumento-base per la misura (e, sostanzialmente, per la definizione operativa) della probabilità. Si rammenti, tuttavia, che essa è equivalente a quella banale ($\mathbf{P}(E)$ è il valore di «uno» – una Lira, oppure un Dollaro, se si vuole dare un nome all'unità – da ricevere se E è vero), salvo la situazione di «gioco» («io credo ch'ei credesse ch'io credessi») che potrebbe falsare la decisione e che grazie alle «regole di penalizzazione appropriate» viene eliminata.

2.2. «Eventi»: ambiguità da eliminare.

Per evitare di discutere di probabilità *nel vuoto* o nell'ambiguo (come purtroppo può capitare e spesso capita) è certo opportuno – e direi addirittura (a mio avviso) necessario – introdurre subito alcune precisazioni terminologiche (almeno in parte nuove). Si tratta anzitutto di stabilire in senso univoco il significato di 'evento' e di introdurre i due termini 'fatto' e 'fenomeno' da sostituirsi ad 'evento' nei sensi in cui, per evitare equivoci, non dovrebbe mai piú essere usato.

Il termine 'evento' dovrebbe, a tal fine, venire riservato al senso di «caso *unico* perfettamente *specificato* (in anticipo)»: ad esempio, il pareggio in una ben precisata partita di calcio; l'aumento della percentuale di voti per una data lista nelle prossime elezioni in confronto alle precedenti; la cattura (entro un precisato limite di tempo) di un dato criminale ora evaso, ecc.

Al contrario, indicazioni *generiche* come «la cattura di *un* evaso», «il pareggio in una (non specificata) partita di calcio», «un forte acquazzone», mancano dei requisiti necessari per consentire una risposta univoca e certa, «Sí» o «No», e pertanto *non* costituiscono «eventi» nel senso precisato, ma soltanto «fatti».

Per dare un chiarimento concreto e completo su di un esempio: il fatto che cada (o che sia caduto) un fulmine è un *fatto* (che può «accadere»); il fatto che esso colpisca o abbia colpito un dato edificio causando danni coperti da assicurazione (secondo le condizioni di polizza vigenti e le clausole convenute) è un *evento* (che può «verificarsi»); la «caduta di un fulmine» (in senso generico: dove e quando che sia) è un *fenomeno* (che può «ripetersi», sempre e dovunque).

Piú radicalmente ancora, va scartato l'uso del termine 'evento' in senso generico, conformemente alla locuzione confusionaria disgraziatamente invalsa di «prove di un evento» per dire «eventi» (in genere, piú o meno analoghi) e, quel che è peggio, «probabilità di un evento» (!) per probabilità di ciascuno di tali eventi, detti (in virtú di tale loro intruppamento) «prove di quello stesso evento». Gli stessi oggettivisti, però, si smentiscono, in quanto parlano anche del caso in cui «la probabilità... varia di prova in prova», in contrasto coll'«assioma» che la probabilità debba riguardare, per aver senso, senso *collettivo*, un gran numero (o, secondo i piú raffinati, una successione *infinita* (!)) di «prove».

È penoso, ma doveroso, segnalare queste palesi assurdità (non si potrebbe spiegare, salvo per l'assuefazione, come esse possano non apparire palesi anche al piú sprovveduto mortale che vi ponga il minimo di attenzione!)

L'opportunità di queste precisazioni terminologiche è scaturita da discussioni critiche su tali problemi durante un corso di lezioni all'Istituto nazionale di alta matematica (Roma, marzo-maggio 1979).

2.3. Certe «definizioni».

Anziché di «definizioni» della probabilità sarebbe piú appropriato parlare di «pseudodefinizioni»; esse non sono in genere che dei conati di definizione: così somiglianti a quelli di chi volesse sollevarsi da terra tirando verso l'alto i lacci delle proprie scarpe.

Un'altra immagine – altrettanto bella e del tutto diversa, dovuta all'indimenticabile Leonard Jimmie Savage – ribadisce e arricchisce il medesimo concetto dicendo che «è impossibile fare una omelette probabilistica senza spezzare uova probabilistiche». Fuori di metafora (e si potrà notare e apprezzare sempre piú quanto dette metafore siano appropriate!) è insensato cercar di foggiare una definizione usando il medesimo termine che si vuol definire, o altri che lo presuppongono. Ed è proprio questo, invece, il sistema cui tentano di aggrapparsi gli aspiranti ideatori e gli illusi scopritori della «probabilità oggettiva».

Le «definizioni» correnti si basano poi entrambe sul circolo vizioso di supporre già noto il significato di probabilità, almeno nel senso di saper distinguere se certi dati eventi *sono* o *non sono* «ugualmente probabili». È necessario infatti – per la «definizione classica» – «avere una partizione in n risultati *ugualmente probabili* e sapere che m sono favorevoli a un dato evento E» per dire che $\mathbf{P}(E) = m/n$, oppure – per la «definizione frequentista» – che «su n "prove" se ne sono verificate m» per dire che la frequenza è stata m/n. (Ma in quel singolo caso, e quindi «per caso». Di per sé, il fatto di stimare la loro probabilità in «circa m/n» è frutto di un'illazione infondata, a meno che non si tratti di molti eventi giudicati «scambiabili»: cfr. cenno nel § 2.7 e sviluppi nel § 3).

Come conclusione: in entrambi i casi la frazione m/n può essere una scelta piú o meno ragionevole ma non obbligatoria; è bene riflettere caso per caso senza elevare a teoremi o a dogmi delle semplici norme di buon senso affinate con la familiarità a ragionare sull'incertezza.

Non è detto, naturalmente, che parlando di scommesse si debba pensare ad esempi nel senso piú deplorevole e dannoso del termine (lotto, lotterie, giochi d'azzardo con carte o dadi o roulette, ecc.); fortunatamente rientrano nello schema anche operazioni formalmente analoghe ma di motivazione e direzione opposta, come il risparmio che dà una protezione generale contro ogni rischio imprevisto o piú o meno genericamente prevedibile, e come, piú specificamente, le assicurazioni che coprono ogni genere di rischi, incoraggiando la preveggenza anziché l'incoscienza, l'estraneità anziché la soggezione a certe manie sciocche e difficilmente curabili per chi ne è vittima come il fumo, l'alcool, la droga. Tutte cose, tra l'altro, che, sia pure indirettamente, danneggiano purtroppo anche chi ne è immune ed estraneo.

E conviene rammentare, qui, che ogni valutazione di probabilità è sempre *subordinata*, o condizionata. Anziché $\mathbf{P}(E)$ si dovrebbe a rigore scrivere sempre $\mathbf{P}(E|H_0)$ o $\mathbf{P}(E|H_0H)$ per indicare che la valutazione è fatta nello stato di conoscenza H_0 e, rispettivamente, condizionatamente *anche* all'ipotesi H.

2.4. Certezza, incertezza, probabilità.

Sembra un dannato destino quello di molte scienze che vedono sopravvivere credenze da esse ridicolizzate e smentite dai fatti, e che trovano in tal modo inquinato il loro campo dal tentativo di intrusioni da parte di squallidi residui superstiziosi e cabalistici. Si pensi all'astronomia travisata in supporto di sproloqui astrologici! Eppure, è strano: non siamo forse ben lontani dal medioevo e prossimi al 2000?

Una delle credenze superstiziose piú pericolose e aberranti è quella che ammette ed afferma che «esistono» delle «probabilità oggettive». Tale binomio, oltre che a «quadrato circolare» come già fatto, potrebbe essere abbinato, per la sua contradditorietà, a «ghiaccio bollente», o «luce nera», o «pioggia asciutta». È chiaro che la probabilità oggettiva (volendo dar senso, sia pure artificiosamente e provvisoriamente, a tale locuzione) non potrebbe essere se non il valore (logico) di *verità*, e cioè:

«1» («certezza», o «vero», o «certo») se l'evento si è verificato o si verificherà; ossia, sinteticamente, «Sí»;
«0» («impossibilità» o «falso») o, sinteticamente, «No» nel caso opposto;

al quale (nella logica dell'incertezza, propria della non-onniscienza umana) è necessario aggiungere

«?» («incertezza», o «dubbio», o «incerto») o, sinteticamente, «Non so», od anche, recitando il ben noto detto dannunziano, «Forse che Sí, Forse che No» (cfr. fig. 3).

La situazione cui si è pervenuti nel momento attuale – con tre livelli di conoscenza: «Sí», «No», «Non so» – è la situazione della «logica dell'incerto», con tre «valori di verità». Quello intermedio è quello dell'incertezza, ma essa va qui considerata come una situazione unica senza differenziazioni tra «il *piú* e il *meno* probabile».

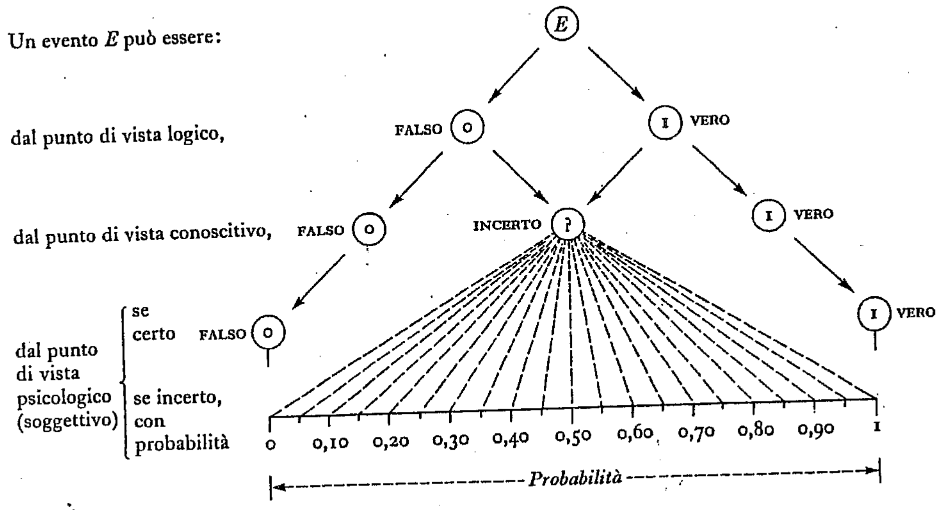

Figura 3.
I tre livelli di conoscenza di un evento.

Questa «logica a tre valori» («1», «0», «?»), che non mi consta sia mai stata teorizzata o utilizzata, può effettivamente risultare significativa per sviluppare espressioni logiche, o matematiche, o logiche e matematiche insieme. Tutto è analogo al «calcolo letterale» dell'algebra elementare, e si possono poi ottenere anche le corrispondenti espressioni in termini di *probabilità* e di *previsione* quando se ne facciano valutazioni ed elaborazioni probabilistiche. (Si tratterebbe del banale procedimento di sostituzione di indicazioni in lettere con le corrispondenti valutazioni numeriche come probabilità e previsioni). Questo non è che un cenno; ma basti darne qui una esemplificazione per rendere chiaro tutto il significato della distinzione tra livello «incertezza» e livello «probabilità».

Si sa che il numero $X = x_1E_1 + x_2E_2 + \ldots + x_nE_n$, supponendo che gli eventi E_h siano incompatibili ed esaustivi (in parole povere: che si debba verificare certamente uno e uno solo di essi), ci darà un guadagno x_1 se si verificherà l'evento E_1 (e cosí per tutti gli altri). Si tratta di un numero (qui, in particolare, di un guadagno) incerto. E non lo si dice «aleatorio»: è utile infatti (anzi, per la presente finalità, essenziale) distinguere nettamente «incertezza» da «probabilità»: si sarebbe detto «aleatorio» se si conoscesse (cioè si fossero stimate, ecc., non importa come) le probabilità degli eventi E_h: allora (ma solo allora) si potrebbe parlare della previsione, $\mathbf{P}(X)$, del guadagno aleatorio X (beninteso, soggettiva, come è soggettiva ogni probabilità).

Il caso piú generale, di n eventi E_h (non disgiunti; $h = 1, 2, \ldots, n$) si riconduce subito al caso precedente di una partizione in N ($N \leq 2^n$) eventi disgiunti, C_i, detti «costituenti». Essi sono ottenibili dal prodotto logico $E_1E_2 \ldots E_n$ cambiando in tutti i 2^n modi possibili alcuni degli E_i nella loro negazione $\tilde{E}_i$; non è detto però, naturalmente, che tutti i 2^n prodotti siano non vuoti; perciò essi sono *al piú* 2^n, non 2^n senz'altro.

2.5. Dall'incertezza alla probabilità.

Le considerazioni che precedono avevano uno scopo assai semplice e meramente preparatorio: intendevano mostrare fin dove si poteva portare avanti il discorso e la trattazione matematica restando nel campo dell'incerto, per poi passare direttamente dal campo della semplice *incertezza* a quello in cui l'incertezza, venendo tradotta in *probabilità* mediante stime dirette o indirette o mediante calcoli piú o meno complessi su di esse basati, fornisce gli elementi necessari e desiderati come base per prendere le decisioni nel modo piú ragionevole e vantaggioso.

La conclusione è ora semplicissima. Si tratta soltanto di sostituire a tutti gli eventi (siano ad esempio $E_1, E_2, E_3, \ldots, E_n$) le rispettive probabilità, ossia applicare l'operatore lineare $\mathbf{P}$ («probabilità» di) ottenendo $\mathbf{P}(E_1)$, $\mathbf{P}(E_2)$, $\mathbf{P}(E_3)$, ..., $\mathbf{P}(E_n)$. Analogamente, applicando $\mathbf{P}$ al numero aleatorio $X = x_1E_1 + x_2E_2 + \ldots + x_nE_n$, si ha per la *previsione* $\mathbf{P}(X) = \mathbf{P}(x_1E_1) + \ldots + \mathbf{P}(x_nE_n) = x_1\mathbf{P}(E_1) + \ldots + x_n\mathbf{P}(E_n) = p_1x_1 + \ldots + p_nx_n$ (ove si pone $\mathbf{P}(E_i) = p_i$).

Quanto al modo di stimare tali probabilità nulla c'è da aggiungere a quanto detto in generale, salvo far presente una circostanza utile come controprova. La somma delle n probabilità deve ovviamente dare 1 ma in pratica la somma delle probabilità stimate se ne discosterà piú o meno. Se la differenza è piccola basta alterare proporzionalmente (moltiplicandole per il necessario coefficiente, poco inferiore o poco superiore ad 1); altrimenti conviene ripensare alle singole probabilità nel dubbio che una (o piú) presenti una deviazione macroscopica.

2.6. Le «pseudodefinizioni» tuttora in voga.

È necessario parlare anche delle «pseudodefinizioni» della probabilità, sia perché sono quelle usualmente (purtroppo!) presentate tuttora come autentiche «definizioni (!)», e sia perché – considerandole non nella mentita veste di definizioni bensí come criteri ausiliari per la valutazione di probabilità in certi tipi di circostanze – possono costituire spesso un valido punto d'appoggio.

Beninteso, come definizione *vera* si considera sempre quella diretta: $\mathbf{P}(E)$ è il «prezzo equo» per una scommessa che faccia vincere l'importo «uno» (una Lira, un Dollaro, quel che altro si voglia) se l'evento E si verifica. Beninteso, «equo» secondo la valutazione dell'interessato. Tuttavia, per eliminare il carattere di «gioco», di «rischio», di «azzardo», conviene fare una di quelle «scommesse col morto» considerate fin dal § 1.5 e chiamate «regole di penalizzazione appropriate» (*proper scoring rules*). Si rammenta che, all'opposto delle usuali scommesse, questa specie di «scommessa col morto» tende a rendere minimo il rischio anziché a crearlo. Si tenga presente, senza ripeterne qui la spiegazione, la «regola di Brier», particolarmente elementare e pertanto chiarificatrice (anche grazie alla sua interpretazione meccanica).

Le altre cosiddette «definizioni» tuttora imperversanti non dovrebbero in alcun modo venir chiamate «definizioni»; escludendo di considerarle tali possono però talvolta, se intese ed usate con discernimento, costituire criteri utili per agevolarci o guidarci, in particolari circostanze, alla valutazione soggettiva delle probabilità.

Quella che pretende di «definire» la probabilità come rapporto, $p = m/n$, tra il numero dei «casi favorevoli» e dei «casi possibili» *supposti... ugualmente probabili* esprime una proprietà esatta ma pressoché tautologica; non è comunque una definizione perché presuppone di saper già cosa significhi «ugualmente probabili». Altrettanto poco «definizione» sarebbe quella di «peso», «volume» «carica elettrica», ecc., che riconducesse ad una altrettanto non definita nozione di «uguaglianza di peso», «di volume», ecc.; ciò fisserebbe la *scala* (lineare, non logaritmica, non ..., ecc.) ma non permetterebbe di distinguere quale sia la grandezza che viene chiamata «peso» e quale venga chiamata «volume», ecc.: eppure è proprio su questa distinzione sostanziale che si sorvola (come se «tutto ciò che l'Autore tace non sapendolo spiegare» potesse, per chissà quale miracolo, riuscire chiaro al Lettore!)

Ma al peggio non si è ancora arrivati: non si è ancora arrivati al gradino peggiore di confusionismo che è quello in cui si cerca di identificare, o almeno assimilare tra loro, due nozioni che richiedono assolutamente, per essere comprese in modo corretto, di venir considerate in certo senso antitetiche e tuttavia legate

da molteplici e reciproci influssi e rapporti: la probabilità e la frequenza. Il peggio è la concezione *frequentista*, che, come l'idra dalle sette teste, presenta sempre nuove varianti di conati di «definizioni frequentistiche»: a mano a mano che quelli precedenti vengono rintuzzati, ecco pullularne versioni sempre piú artificiose e infelici.

Sembrerebbe di dover essere giunti, in questo modo, ad un limite invalicabile, ma l'esperienza in tal campo e un detto popolare romano inducono a prudenza. Secondo tale detto, «il peggio non è mai morto»! E infatti, per liberarsi dall'indeterminatezza della frequenza, la si sostituisce con l'irraggiungibile «frequenza-limite» (conoscibile... dopo la fine dell'eternità!) Comunque, il discorso si disperde, inevitabilmente, in mille rivoli.

2.7. Fraintendimenti: guardarsene!

Tentar di passare sistematicamente in rassegna questi «mille rivoli» sarebbe fatica improba e inutile. Basti soffermarsi a titolo esemplificativo su qualche fraintendimento in cui è facile cadere o che può lasciarci confusi.

I malintesi si possono ricondurre, sempre o quasi sempre, alla tendenza ad interpretare in senso oggettivistico delle considerazioni che sono valide solo in senso soggettivo, o a travisare in senso assoluto dei ragionamenti che sono validi solo in senso relativo.

Esempio tipico del primo malinteso è il confondere «probabilità zero» con «impossibilità» (sarebbe come dire che un insieme di misura nulla, ad esempio un punto o una linea su un piano, *è* l'insieme vuoto!)

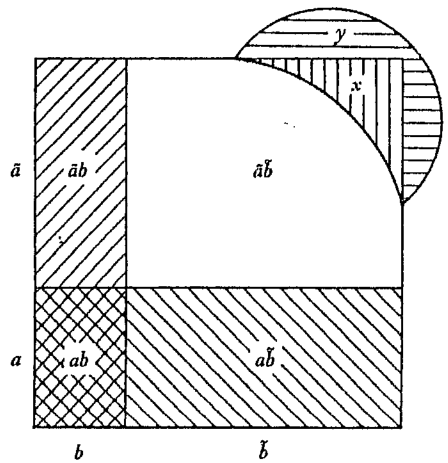

Figura 4.

La figura (quadrato di lato 1 diviso in quattro rettangoli) indica le probabilità dei quattro casi, AB, $A\bar{B}$, $\bar{A}B$, $\bar{A}\bar{B}$, essendo A e B indipendenti. Se però, fermi restando i rettangoli tratteggiati, si altera (allargandolo o restringendolo) il rettangolo bianco $\bar{a}\bar{b}$ (che diverrebbe $\bar{a}\bar{b}-x$ risp. $\bar{a}\bar{b}+y$) l'indipendenza stocastica non sarebbe piú rispettata.

Il secondo malinteso ha luogo, in molti casi, per il fatto di non tener conto della dipendenza di ogni conclusione dallo stato di conoscenza in cui ci si trova. Per dirlo con la notazione già usata, si tratta di pensare all'evento E «nel vuoto» anziché in rapporto al nostro presente stato d'informazione, H_0; di per sé – lo si rammenti – $\mathbf{P}(E)$ non ha senso a meno che si sottintenda H_0, pensando che s'intenda scritto $\mathbf{P}(E|H_0)$.

Ed è chiaro che se, al posto di H_0, si avranno altre «ipotesi» H_0', H_0'', ... diverse, non solo cambieranno valore le probabilità $\mathbf{P}(E)$ ma anche, in genere, conseguentemente, le relazioni tra esse. Per illustrare tale fatto si veda l'esempio rappresentato nella figura 4, osservando dapprima il quadrato (di lato $= 1$) diviso nei quattro rettangoli ab, $a\tilde{b}$, $\tilde{a}b$, $\tilde{a}\tilde{b}$ (dove, si ricordi, la tilde indica negazione, o, aritmeticamente, complemento ad 1 : $\tilde{a} = 1 - a$, $\tilde{b} = 1 - b$), e poi lo stesso quadrato privato del pezzetto x, o invece accresciuto del pezzetto y, in seno al quale a e b risultano *correlati* (risp. positivamente e negativamente).

Ciò mostra che, a rigore, si dovrebbe sempre specificare rispetto a quale *stato di conoscenza* (H_0; H_0H; ...) l'indipendenza stocastica viene affermata: non esiste l'«indipendenza in sé» (come la «cosa in sé» di certi filosofi).

Analogamente, l'indipendenza stocastica può anche sussistere, anziché *tout court*, soltanto «subordinatamente a una data ipotesi» oppure «subordinatamente a ciascuna di certe ipotesi incompatibili»: è il caso che dà luogo, ad esempio, alla «scambiabilità» che verrà presentata – data la necessità di più ampie premesse e di strumenti e ragionamenti alquanto più delicati e complessi – nel § 3.

Il peggiore fraintendimento (e, sembra, il più radicato, tanto che chi ne è immune rischia di venir considerato un idiota o uno squilibrato,... e di ricevere lettere di insulti!) è però quello che induce molti «competenti» o «intenditori» a ritenere molto probabile, al Lotto, l'estrazione di un numero ritardato (cioè che non è uscito da molte settimane su una data ruota, o, occasione ancor più ghiotta per gli «intenditori», su nessuna ruota!) È inutile dire a tali «intenditori» che i numeri «non hanno memoria» e che non c'è quindi alcun motivo di pensare che l'essere stati estratti poche o molte volte più o meno recentemente non sia, come lo è, un fatto passato privo di qualsiasi influenza sulle circostanze in cui l'estrazione avrà luogo. E «acqua passata non macina più».

2.8. Probabilità e frequenza.

Fra i molti equivoci e le molte distorsioni d'interpretazione che riducono spesso le discussioni sulla probabilità a «dialoghi tra sordi», primeggiano indubbiamente quelle che concernono le relazioni tra *probabilità* e *frequenza*. Vi sono addirittura delle scuole che pretenderebbero di identificarle (!), di considerare i due termini come sinonimi, come inutili doppioni l'uno dell'altro.

In tal senso sono orientati specialmente molti statistici, ma anche, sia pure con svariate sfumature e «abbellimenti», molti autori di estrazione filosofica o propensi a filosofanteggiare. Lo scopo, ambito a prezzo di qualsiasi distorsione, è quello di *negare* (in sostanza) l'autentica nozione di probabilità – quella *soggettiva*, quella «naturale» del non abbastanza fuorviato e catechizzato «uomo della strada» – per sostituirla con incredibili frutti di macchinose elucubrazioni.

'Probabilità' e 'frequenza' – secondo certe vedute tuttora abbastanza in voga, specie negli ambienti degli statistici – sono sinonimi o quasi, più o meno intercambiabili, come gemelli identici o forse come un individuo e la sua ombra (per ricollegarsi a una fantasia di Bontempelli).

Solo quest'ultima immagine, però, risulta concettualmente appropriata, nel senso che la statistica rileva oggettivamente i fatti (li enumera, li classifica, li elabora,...) traendone indicazioni significative e certe; mentre la probabilità (*ars conjectandi* 'arte di congetturare', come la chiamò il suo stesso fondatore, Giacomo Bernoulli) fornisce a ciascuno il modo di esprimere il proprio grado di fiducia nelle varie ipotesi di cui si interessa.

Le nostre più o meno istintive valutazioni di probabilità dipendono da una sintesi di esperienze favorevoli e sfavorevoli, vissute o sentite raccontare, di tipo più o meno affine ai casi di cui attualmente ci si preoccupa. In casi schematici e ripetitivi è naturale pensare che le cose andranno in futuro più o meno conformemente alle esperienze del passato, a meno di non ritenere che ci siano motivi di prevedere più o meno sensibili miglioramenti o peggioramenti. Ciò significherebbe rispettivamente aumento o diminuzione della probabilità per fatti desiderabili o viceversa; ecc.

È bene, comunque, cercare di schematizzare un po' le varie situazioni.

Si pensi anzitutto a un certo numero di eventi qualunque, $E_1, E_2, ..., E_n$; beninteso, eventi aleatori. La loro somma, $S_n = E_1 + E_2, ..., E_n$, è il numero dei successi; un numero per noi incognito, quindi aleatorio. Si può però darne la *previsione*, $\mathbf{P}(S_n) = \mathbf{P}(E_1) + \mathbf{P}(E_2) + ... + \mathbf{P}(E_n)$, come somma delle probabilità degli E_h (l'additività vale comunque, siano gli E_h compatibili o incompatibili, logicamente e/o stocasticamente dipendenti o indipendenti tra loro).

Più in generale, ciò vale anche per ogni numero aleatorio $X = x_1E_1 + x_2E_2 + ... + x_nE_n$ (ci si riferisce qui al caso più semplice di numeri aleatori con un numero finito di valori possibili: $x_1, x_2, ..., x_n$ con probabilità $p_1 = \mathbf{P}(E_1)$, $p_2 = \mathbf{P}(E_2)$, ..., $p_n = \mathbf{P}(E_n)$, naturalmente, di somma $= 1$). Concettualmente ciò vale anche se la probabilità (e, se ciò giova a dare un'immagine più intuitiva, la si pensi come massa) è distribuita, anziché su un numero finito di punti, su un'infinità o addirittura con continuità su tutto l'asse da $-\infty$ a $+\infty$ o su una parte qualsiasi di esso.

Ma, per ora, interessa soltanto la previsione della frequenza, $\mathbf{P}(S_n)$, facendo notare anche che $\mathbf{P}(S_n/n) = \mathbf{P}(S_n)/n$ è la probabilità media degli n eventi considerati. D'altra parte, se si indica con $F_0, F_1, F_2, ..., F_n$ gli eventi consistenti nell'aver nessun successo, o uno, o due, ecc., fino a tutti n, si potrebbe esprimere $\mathbf{P}(S_n)$ con una diversa espressione, e cioè $\mathbf{P}(S_n) = \mathbf{P}(F_1) + 2\mathbf{P}(F_2) + 3\mathbf{P}(F_3) + ... + n\mathbf{P}(F_n)$. Nel caso (o nell'ipotesi) che gli n eventi $E_h (h = 1, 2, ..., n)$ siano ugualmente probabili e indipendenti, la probabilità $\omega_h^{(n)}$ $(h = 0, 1, ..., n)$ di h successi su n prove vale $\binom{n}{h} p^h q^{n-h}$ (ove $q = 1 - p$ è la probabilità di insuccesso (in una qualunque prova); si tratta (come è chiaro) dei termini dello sviluppo della potenza n-esima del binomio $(p+q)$; la loro somma è ovviamente $= 1$).

Questo, però, non è che un caso molto particolare. Un caso abbastanza piú generale (benché, in certo senso, differisca poco dal precedente) è quello della *scambiabilità* (che sarà sviluppato nel § 3 con l'attenzione che richiede: si tratta del caso che, con locuzione impropria e contraddittoria, viene comunemente descritto come di «indipendenza con probabilità costante ma incognita»).

Nel presente contesto è però necessario insistere, invece, su ciò che ha validità generale, indipendentemente dalle ipotesi piú comuni che sono indubbiamente idonee in molti casi, ma che sarebbe un errore (o un travisamento? o una superstizione?) considerare come norme valide *in generale* salvo quelle «eccezioni che (con detto comicamente ineffabile) confermano la regola»!

Ma non basta. Occorre sottolineare come sia ingannevole l'idea grossolana, conforme però a certe vedute tuttora abbastanza diffuse (specie nel campo degli statistici), secondo le quali probabilità e frequenza sarebbero da considerarsi sinonimi, o quasi.

È importante, data tale situazione, chiarire e precisare rigorosamente fino a che punto il significato di 'probabilità' e quello di 'frequenza' concordino, e indicare da quale punto e in quale senso le due nozioni risultino contrapposte.

Se si considerano n eventi $E_1, E_2, \ldots, E_n$, la loro somma $X = E_1 + E_2 + \ldots + E_n$ è il numero m dei successi (degli E_h che si verificano, cioè che sono $=1$), ed X/n ne è la *frequenza* (o m/n, come è piú usuale, dato che si pensa già noto il valore m di X). Ma, prima di conoscerne l'esito, cosa è possibile dire di X? Non certo il valore effettivo; se ne può però indicare la *previsione*, $\mathbf{P}(X) = \mathbf{P}(E_1 + E_2 + \ldots + E_n) =$ somma delle probabilità.

Per esprimersi in forma ragioneristica, rendendo piú «palpabile» il significato, basterà dire, concludendo, che la differenza tra probabilità e frequenza, o tra previsione e realizzazione, consiste nella necessaria distinzione tra valutazione *preventiva* (necessariamente piú o meno incerta e soggettiva) e valutazione *consuntiva* (ovviamente certa e oggettiva).

2.9. Valutazioni condizionate.

Vari accorgimenti possono essere spesso d'aiuto per valutare accuratamente la probabilità da attribuire a un dato evento E. Come primo esempio, può a volte aiutare il fare piú valutazioni condizionate a diverse ipotesi, siano $H_1, H_2, \ldots, H_s$, incompatibili ed esaustive (cioè tra le quali una e una sola risulterà essere quella vera). Se uno attribuisce alle s ipotesi le probabilità $q_i = \mathbf{P}(H_i)$, risulterà che, conseguentemente, egli dovrà valutare $p = \mathbf{P}(E) = \sum_i p_i q_i$. In particolare, se si tratta di distinguere solo due ipotesi, H e $\tilde{H}$, si avrà $p = p'\,q + p''\tilde{q}$.

Analogamente, nel valutare la probabilità di un evento E che sia il prodotto di due o piú altri eventi $E = E_1E_2$ (o $E = E_1E_2 \ldots E_n$), conviene confrontare la valutazione soggettiva diretta di $\mathbf{P}(E)$ con quelle indirette come $\mathbf{P}(E) = \mathbf{P}(E_1)\mathbf{P}(E_2|E_1)$ (o, rispettivamente $\mathbf{P}(E) = \mathbf{P}(E_1) \cdot \mathbf{P}(E_2|E_1) \cdot \mathbf{P}(E_3|E_1E_2) \ldots \mathbf{P}(E_n|E_1E_2 \ldots E_{n-1})$, anche cambiando comunque l'ordine degli n eventi). Ci saranno delle discordanze (la coerenza in casi complessi non è visibile di primo acchito) e si dovrà vagliare

quali, fra i ritocchi atti a ristabilire la coerenza, dànno luogo al complesso di valutazioni globalmente soddisfacente come espressione delle proprie opinioni.

Un caso analogo, ma molto piú semplice, è quello in cui si abbiano da valutare le probabilità degli n eventi $E_1, E_2, \ldots, E_n$ costituenti una *partizione* ($E_1 + E_2 + \ldots + E_n \equiv 1$; cioè, se ne deve verificare uno e uno solo). Evidentemente, anche la somma delle probabilità deve dare 1; ma sembra dia maggiore affidamento il procedere alla valutazione delle n probabilità $p_1 = \mathbf{P}(E_1)$, $p_2 = \mathbf{P}(E_2)$, $\ldots$, $p_n = \mathbf{P}(E_n)$ singolarmente (senza le probabili tentazioni di aggiustare via via gli addendi per arrivare ad 1); verificare poi quanto la loro somma si scosti in piú o in meno da 1, e in base a ciò ripensare quali valori sembri ragionevole ritoccare per eliminare tale differenza.

Tutto ciò, beninteso, non può rientrare in forma troppo «ufficiale» nei precetti della teoria delle probabilità secondo il concetto di coloro che ne fanno un'astrazione perfetta, immutabile, apodittica: ma è bene che sia cosí, altrimenti probabilità e teoria delle probabilità cesserebbero di essere creature vive e vitali riducendosi a spoglia imbalsamata o addirittura a nudo scheletro.

Oltre che come immagine descrittiva, il termine 'scheletro' è appropriato per sottolineare l'impostazione astratta e meramente «assiomatica» che molte scuole impongono alla teoria delle probabilità, rinunziando ad ogni scelta (buona o cattiva che sia) di una interpretazione da dare al termine 'probabilità'. Potrebbero abolirlo, e dire 'teoria della misura' (con un qualsiasi aggettivo di loro gradimento), e cesserebbe ogni rischio di confusione ed ogni motivo di recriminazione.

La teoria delle probabilità *si serve* della matematica, ma *non* in astratto, bensí per applicazioni concrete nei problemi di previsione, e non si basa su assiomi artificiosi bensí è essa stessa – come ben disse (salvo errore) Henri Poincaré – «il buon senso ridotto a calcolo».

2.10. Le «certezze» col «quasi».

Per finire (e proprio, purtroppo, anche nel senso dei «per finire» umoristici) occorre segnalare (ma – beninteso! – soltanto per guardarsene) le affermazioni in cui si parla di «certo» e di «impossibile»... col «quasi» (un «quasi» talora espresso ma spesso addirittura sottinteso).

È chiaro che questa voluta imprecisione, intesa a «dire e disdire» nonostante «la contradizion che no l'consente», è particolarmente esiziale se inquina fin dall'inizio il discorso dal quale si pretenderebbe di estrarre lo spunto per le *definizioni*: qualunque cosa si dica in seguito risulterà allora fatalmente ambigua e, a rigore, priva di senso. Eppure le piú tipiche (pseudo)-definizioni «oggettivistiche» della probabilità si sforzano per l'appunto di ridurre il senso di «probabilità» all'esistenza di certi comportamenti pretesamente obbligati dalle «leggi del caso» in (lunghe) successioni di «*prove*».

Il caso piú estremo è quello della cosiddetta «concezione statistica» in cui la probabilità viene addirittura «definita» (!) come la frequenza (cioè la percentuale di successi) su «un grande numero di prove».

La critica alla definizione frequentista (e la spiegazione del perché la si dovrebbe dire semmai «pseudodefinizione») sta nel fatto che tra probabilità e frequenza sussistono parecchi legami in entrambi i sensi che non consentono però di confondere le parti in certo senso *opposte* che giocano nei ragionamenti e nei fatti. In forma di scherzoso apologo si può paragonare il caso di probabilità e statistica a quello dei due gemelli identici di Plauto, che, col loro entrare e uscire dalla scena, provocavano continui equivoci.

Occorre ribadire (cfr. § 2.8) che le probabilità riguardano il «preventivo» di ciò che ci si aspetta accada (o sia accaduto ma non ci sia ancora noto), mentre le frequenze indicano ciò che realmente è accaduto. Nei casi in cui i dati non noti relativi (in genere) al futuro riguardano frequenze, la valutazione di probabilità riguarderà i valori piú o meno ritenuti ragionevoli da attendersi per esse; si tratterà di *previsione* di frequenze.

A tale riguardo, accade spesso che la previsione della frequenza sia ritenuta abbastanza «certa» o «buona», nel senso di attribuire piccola probabilità a scarti sensibili dal previsto.

Per limitarsi al caso piú banale, di Testa e Croce, se la moneta appare non deformata, è naturale pensare che lo scarto fra i risultati Testa o Croce sia piccolo (dell'ordine di grandezza della radice del numero dei colpi); ma non si deve attribuire ciò ad un meccanismo o magia che tende a correggere le deviazioni favorendo la faccia che è in minoranza. Il procedere a sempre piú numerosi colpi non ha alcuna tendenza alla compensazione: l'avvicinamento alla situazione «equa» avviene, ma non per compensazione bensí soltanto per «sommersione»: la differenza si *diluisce* e scompare per il prevalere dei risultati successivi.

La «tendenza» alla compensazione non ha nulla di intenzionale o guidato su 10 colpi a Testa o Croce si può ottenere una qualsiasi delle $2^{10} = 1024$ successioni; non si deve pensare che una successione *data* con 5T e 5C abbia probabilità maggiore di quelle con tutte T o con tutte C; ma la probabilità di una successione *qualunque* con 5T e 5C ha probabilità 1/4 (esattamente 252/1024) perché 5T e 5C si possono ordinare in 252 modi differenti; la probabilità di 6T e 4C (o viceversa) è 210/1024; quindi la probabilità di ottenere parità con al piú uno scarto di 1 è 672/1024 (praticamente, 2/3).

A parte l'utilità di spiegazioni esplicative su esempi semplici, è opportuno (come conclusione del presente § 2) sottolineare ancora – e magari «enfasizzare» (le «voci» riprese dai pavidi puristi sono spesso le piú efficaci) – la distinzione fra i termini che hanno significato oggettivo e quelli che hanno significato soggettivo.

La verità o falsità di un'affermazione, o evento, è un fatto oggettivo, la sua probabilità è un fatto soggettivo; lo stesso vale per il valore (vero) di un numero (numero aleatorio per chi non ne conosce il valore effettivo), e per la sua previsione che è soggettiva. Per due eventi, il fatto che siano logicamente compatibili (in base alle informazioni certe che qualcuno ne ha) è un fatto oggettivo; per chi non ne ha è soggettivo. La frequenza (in un certo gruppo di eventi) è un fatto oggettivo se essa è conosciuta con certezza; nel caso opposto, la sua previsione è soggettiva.

E, per finire, si aggiunga ancora il caso-limite (sia in senso matematico, sia in

senso interpretativo): quello che – a volerlo prendere sul serio e non come una insensatezza – permetterebbe di conoscere le probabilità solo dopo la fine dei tempi. È il caso della cosiddetta «concezione statistica» in cui la probabilità viene addirittura «definita» (!) come la frequenza (cioè la percentuale di successi) su «un gran numero di prove», ed anzi, in una sua versione piú spinta, su... un'infinità di «prove» (probabilmente richiedenti di continuarle fino alla fine dell'eternità). Fra i piú impegnati sostenitori di tale concezione si possono segnalare il matematico Richard von Mises e il filosofo Hans Reichenbach; da menzionare, del primo, *Wahrscheinlichkeitsrechnung und ihre Anwendung* (1931), e, del secondo, *Wahrscheinlichkeitslehre* (1935).

3. *«Ab omni naevo vindicata»?*

3.1. Perché «complicare le cose semplici»?

Il proposito di presentare la probabilità e la teoria delle probabilità in forma «ab omni naevo vindicata» apparirà indubbiamente come uno dei piú ardui, tale è l'inestricabile connessione di vedute e concezioni e terminologie radicalmente disparate e spesso anche intrinsecamente inconsistenti che contraddistinguono il variegato campo dei cultori specifici o collaterali della teoria delle probabilità.

Non si tratta affatto, tuttavia, di difficoltà inerenti alla nozione di probabilità e alla teoria delle probabilità, bensí semplicemente dell'effetto di storture d'interpretazione, di ambiguità di concetto e di linguaggio che sono in voga fra i cultori di versioni artefatte della teoria, e, peggio ancora, delle molteplici superfetazioni che ne derivano. Mai come in questo campo riesce indispensabile l'assillo di cui parla Giovanni Papini quando dice, riferendosi all'amico Mario Calderoni, che «a lui premeva insegnare con quali cautele e quali accorgimenti si possa giungere a ottenere delle proposizioni che abbiano un senso» (*Stroncature*, n. 14). Nel campo della probabilità – finché perdurerà il confusionismo imperversante – sembrerebbe quasi utopistico riuscire a tanto.

Non però perché il ragionare in termini di probabilità sia qualcosa di difficile o astruso, bensí perché tale lo si rende sovrapponendo al significato intuitivo, limpido e genuino, di probabilità, delle deformazioni che lo rendono oscuro e vuoto. Accettando invece, secondo le spiegazioni e indicazioni già date, il *naturale* significato *soggettivo* della probabilità – liberata da contraffazioni e da insulsi camuffamenti oggettivistici – tutto diventa assolutamente chiaro, sia per chi lo voglia accettare e sia per chi (anche senza volerlo accettare) non disdegni di apprendere e comprendere cosa ciò comporti.

È necessario, a tal fine, porre attenzione alle precisazioni terminologiche occorrenti per eliminare e sostituire e correggere locuzioni improprie, confuse, fuorvianti, che risentono delle deviazioni «oggettivistiche» o pretesamente tali. Già gran parte delle osservazioni critiche presentate nei due precedenti paragrafi avevano espressamente l'intento di segnalare e far riconoscere le manchevolezze di fraseologie ambigue e devianti, inquinate di oggettivismo: oggettivismo che inevitabilmente dà luogo ad errori e nonsensi impossibili da correggere con semplici ritocchi.

Tali nonsensi, infatti, come scrisse molto efficacemente e spiritosamente Bernard O. Koopman, «al contrario della Guardia di Napoleone si possono sempre far arretrare, ma mai arrendersi e scomparire».

3.2. Delle impostazioni assiomatiche.

In realtà, di impostazioni «assiomatiche» ne esistono molte e molto diverse, ma la distinzione preliminare e radicale è la fondamentale dicotomia fra le due concezioni in cui la probabilità si riferisce *a*) a un *evento* (nell'accezione qui fissata di «caso singolo univocamente specificato»), e si potrebbe dirla concezione *chiara*; o invece *b*) a una *collettività* di eventi *in un qualche senso* «analoghi» (che, nel gergo oggettivistico, si dicono «prove di uno stesso evento» e che, per evitare ambiguità, si potrebbero chiamare – come qui si propone – «prove di uno stesso fenomeno»). A volte si pensa a collettività numerose ma finite, ma a volte qualcuno pensa addirittura a successioni infinite; comunque tutte queste sovrastrutture non giovano che a «complicare le cose semplici», a recarsi dalla località A alla vicina B, non direttamente, bensí percorrendo tutto un cerchio massimo intorno alla Terra tranne il tratto AB. Non sembra eccessiva cattiveria battezzare tale concezione come *confusionaria*. D'altronde, a chiunque s'interessi a un qualche fatto, o *evento*, premerà valutare la probabilità di *quell'evento* (*Hic Rhodus, hic salta!*) e non avrà scopo, in genere, curarsi di altri eventi, piú o meno analoghi, magari (secondo una fraseologia corrente) «prove dello stesso *evento*» (che andrebbe semmai corretta – lo si ribadisce! – in «eventi» che sono «prove di uno

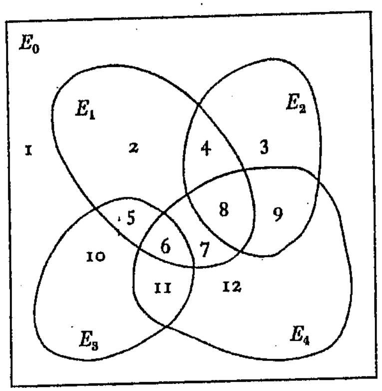

Figura 5.
Il quadrato rappresenta schematicamente tutti i «casi possibili» (punti); le quattro «patate» delimitano gli eventi E_1, E_2, E_3, E_4 e le loro dodici intersezioni (a due a due e a tre a tre): E_0 è la parte del quadrato esterna a tutte le «patate», e indica l'evento «Nessuno degli E_h ($h = 1, 2, 3, 4$) si verifica».

stesso fenomeno»). Può darsi, naturalmente, che questi eventi (o «prove»), specie se in qualche senso «analoghi», vengano considerati ugualmente probabili, o anche (stocasticamente) indipendenti, ecc. (Ma ciò va detto: non è da considerarsi sempre tacitamente affermato se non è esplicitamente negato).

La dicotomia fra concezioni *chiare* e concezioni *confusionarie*, come quella fra *soggettivisti* e *oggettivisti*, non esauriscono la sia pur sommaria rassegna che occorre per dare una efficace anche se pallida idea dell'insieme.

Si può anzi cominciare dagli *astrattisti*: coloro che si occupano sostanzialmente di «teoria della misura» in spazi astratti qualunque, e chiamano «eventi» dei sottoinsiemi e «probabilità» una loro «misura» (additiva, magari completamente additiva) con misura = 1 per tutto lo spazio. È questo l'esempio piú spinto di quel modo di vedere dei formalisti che vantano la matematica come quella scienza in cui «non si sa di cosa si parla né se ciò che si dice è vero o falso». Si veda la figura 5 con quattro eventi rappresentati da «patate» e che, con le loro intersezioni, dànno luogo a una partizione in dodici eventi.

Tutto bene; ma se si cura l'aspetto formale senza badare soprattutto al significato pratico, concreto (e in questo caso «concreto» significa *soggettivo autentico* e non *oggettivo fasullo*!), non è lecito pretendere che le conclusioni valide per convenzione in *quella* teoria astratta debbano valere anche nei problemi concreti e pratici concernenti la probabilità. Per fare un solo esempio, non appare lecito pretendere che valga l'additività completa: nel caso (sia pure considerato inaccettabile da molti autori) di «un intero scelto a caso» ogni intero ha probabilità nulla ma tutti insieme (un'infinità numerabile) hanno probabilità = 1.

Quanto agli oggettivisti del tipo «classico», che si basano sulle suddivisioni in «casi ugualmente probabili», si può riconoscere che c'è modo, spesso comodo, di ricondursi ad esemplificazioni di quel tipo; però il giudizio di «uguale probabilità» è non definibile salvo in senso soggettivo, oppure... con varianti verbali o perifrasi: anziché «ugualmente probabili» dire «ugualmente possibili» (che è peggio!), o addirittura «uguali» (peggio che peggio!) In conclusione, il metodo pretesamente oggettivo od oggettivistico ha senso ed è accettabile ed applicabile se e soltanto se lo si *concreta* in senso *soggettivistico* anziché evocare presunti fantasmi oggettivistici.

Non è un gioco di parole: sembra giusto asserire che è piú oggettiva una cosa soggettiva considerata come tale anziché una cosa che viene considerata come oggettiva mentre tale qualifica non può essere avallata senza riserve.

Passando agli oggettivisti di formazione statistica, si giunge talora a veder addirittura non solo confondere, ma perfino identificare (!) probabilità e frequenza. Ciò significa, in sostanza, scambiare l'attesa di un fatto con la sua realizzazione e constatazione, il «preventivo» col «consuntivo». Questa distorsione è terribilmente grande non solo perché oscura entrambi i concetti, bensí, peggio ancora, perché, identificandoli, li trasforma in un ibrido mostro bicipite. Guardando piú a fondo, l'argomentazione è ancor piú inconcludente; per esprimersi in modo sensato, corretto, si dovrà parlare di «scambiabilità» (cfr. § 3.5) anziché di indipendenza stocastica.

C'è qualche conclusione che si può trarre da tutto ciò?

Probabilmente sí, e precisamente nel senso di riconoscere che nel *fondo* (spesso trascurato o negato o svisato), qualcosa, spesso anche molto, si può salvare e utilizzare di quanto dicono le diverse teorie, ma con un «purché»: purché nell'interpretare tutti i termini e tutte le affermazioni o definizioni o nozioni o conclusioni si abbia sempre cura di vivificarle facendovi scorrere la linfa salutare del soggettivismo.

Senza di ciò – lo si può ben affermare senza esitazione – tutto si ridurrebbe a un vaniloquio: proprio come affermava la *boutade* di Bertrand Russell citata sopra.

3.3. Qualcosa che si dice «eccezionale».

Vi sono molte specie, piú o meno fondamentalmente analoghe, di fraintendimenti che fanno giudicare «eccezionale» il verificarsi di qualche fatto, o circostanza, o situazione, e ritenere «accettabili» come «normali» altri fatti o circostanze o situazioni del tutto analoghi. Il fatto piú tipico a questo riguardo è quello che fa ritenere necessario che in una lunga successione di colpi a Testa e Croce, oppure di lanci di un dado o di due dadi, le 2 facce della moneta (e, rispettivamente, le 6 facce del dado o le 36 coppie di facce dei due dadi, *debbano* presentarsi (prolungando le «prove») circa nella proporzione prevista (1/2, o 1/6, o 1/36), ed inoltre trovarsi *in disordine* (non ad esempio sempre Testa, né sempre alternatamente Testa e Croce, né Testa nella prima metà e Croce nella seconda, né una Testa e poi sempre Croce, né in altre modalità «regolari»)). Tant'è vero che tali risultati si scarterebbero come «non ammissibili», «non regolari» (nel senso di non abbastanza irregolari); eppure di per sé non presentano nulla di anomalo. Essi hanno probabilità $(1/2)^n$, $(1/6)^n$, $(1/36)^n$; è molto piccola se n è grande ma è esattamente la stessa di qualunque altra successione, non importa se piú o meno regolare o irregolare (quale che sia il senso – molto arbitrario! – in cui uno potrebbe interpretare tali distinzioni!)

E allora perché meravigliarsi? Lo stupore può essere giustificato dalla sorpresa, ma non dal fatto che la probabilità sia piccola. Ogni fatto, se lo si precisa con tutti i dettagli, ha probabilità piccolissime, ed anche nulla se la precisione è assoluta (non sbagliare la posizione di un millimetro, non sbagliare l'istante di un microsecondo, ecc.).

Si può notare, d'altra parte, l'utilizzazione che viene fatta dagli statistici sperimentatori di tabelle di «numeri casuali» (*random numbers*) allo scopo di eseguire «scelte a caso» di individui o oggetti od altro onde approfondire certi studi (nell'impossibilità di esaminare tutti gli individui, o oggetti, o avvenimenti, od altro) limitandosi ad esaminare un «campione rappresentativo»; la scelta «a caso» dovrebbe eliminare (o almeno ridurre di molto) il rischio di scelte distorte (ad esempio con sproporzionata rappresentanza di persone del tipo piú abbordabile). La cura di scegliere un campione cercando che risulti rappresentativo è il principale requisito per rendere attendibili le previsioni basate su di esso (ad esempio nei sondaggi).

Guardando nel verso opposto, uno potrebbe dire che si sente sicuro perché

i rischi cui si espone sono minimi: ma se sono parecchi o si ripetono frequentemente la risultante può facilmente essere fatale. Tutte queste ovvie riflessioni, a cosa possono giovare? Dovrebbero giovare a non prendere mai troppo sul serio le impressioni di tranquillità, come è giustificato in parte da una scritta profondamente significativa che campeggia in una trattoria di Trastevere: «Il caso ci protegge – piú che qualsiasi legge». È vero, ma vale anche il viceversa. E perciò «non fidarsene è meglio».

3.4. Indipendenza (stocastica) e correlazione.

È bene premettere che esistono diverse proprietà che si chiamano «indipendenza» (tra eventi, tra numeri aleatori). Si accenna dapprima alla indipendenza (o invece dipendenza) *lineare*: se X e Y sono numeri aleatori (qualunque), $Z = aX + bY + c$ è combinazione lineare di X e Y (la costante c è inessenziale); ovviamente la $\mathbf{P}$ (previsione, o, in particolare, probabilità) è lineare (additiva) cosicché, nell'esempio, sarà $\mathbf{P}(Z) = \mathbf{P}(aX + bY + c) = a\mathbf{P}(X) + b\mathbf{P}(Y) + c$. È questa la piú semplice forma di dipendenza funzionale. Un esempio relativo ad eventi: se A e B sono eventi incompatibili la loro unione (o «evento somma») $A \vee B$ coincide con la somma (aritmetica) $A + B$; se non sono incompatibili (se cioè la loro intersezione non è vuota, e sia $C = AB$), la loro unione $A \vee B$ non è piú $A + B$ bensí $A + B - C$. Bastino questi cenni a titolo informativo.

Piú interessante forse è notare una circostanza che, dopo aver visto un esempio, è ovvia, ma di primo acchito può sembrare incredibile: l'indipendenza stocastica a due a due tra n eventi qualsiansi $E_1, E_2, \ldots, E_n$ non implica la loro indipendenza; cioè, il fatto che per ogni coppia di eventi sia $\mathbf{P}(E_h E_k) = \mathbf{P}(E_h)\mathbf{P}(E_k)$ non implica che debba essere anche $\mathbf{P}(E_1 E_2 \ldots E_n) = \mathbf{P}(E_1)\mathbf{P}(E_2)\ldots\mathbf{P}(E_n)$. Si consideri il piú semplice controesempio, illustrato nella figura 6. I quattro eventi sono rappresentati dai tre rombi E_0+E_1, E_0+E_2, E_0+E_3 di probabilità (= area) 1/2; l'intersezione è il triangolo centrale E_0 di probabilità (= area) 1/4, cioè $1/2 \times 1/2$, come si voleva dimostrare, e non $1/8 = (1/2)^3$ come se sussistesse l'indipendenza fra tutti e tre i rombi e non solo a due a due.

Nel caso di numeri aleatori il significato di indipendenza (stocastica) è sostanzialmente il medesimo; limitandosi, per rimanere al livello elementare, al

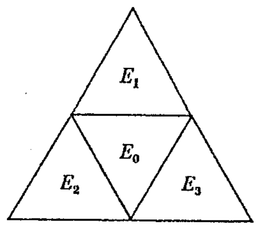

Figura 6.
L'indipendenza stocastica a due a due tra n eventi non implica la loro indipendenza.

caso di numeri aleatori con un numero finito di valori possibili (e siano X ed Y, di valori rispettivamente $x_1, x_2, \ldots, x_n$ con probabilità $p_1, p_2, \ldots, p_n$ ed $y_1, y_2, \ldots, y_m$ con probabilità $q_1, q_2, \ldots, q_m$), esso significa che $\mathbf{P}(XY)=\mathbf{P}(X)\mathbf{P}(Y)=$ somma dei termini $p_h q_k x_h y_k$.

Si può considerare, in certo senso, come una forma piú «debole» di indipendenza un'altra proprietà di cui si farà cenno piú avanti: la «scambiabilità»; si può in certo modo anticiparne il senso dicendo che si tratta di «indipendenza *condizionata* alla conoscenza di dati che al momento *non si conoscono*». E si può approfittare dell'occasione per rilevare anche su questo esempio la necessità logica di certe correzioni a terminologie inveterate ma disgraziatamente fuorvianti causa inconsistenze od ambiguità o facilità di sottintendere precisazioni che non possono essere sottintese.

3.5. La scambiabilità.

Il termine 'scambiabilità' è stato introdotto (dallo scrivente) per sostituire una precedente denominazione inaccettabile poiché di per sé contraddittoria: quella di «eventi indipendenti ed ugualmente probabili con probabilità incognita».

Dicendo «eventi scambiabili» s'intende correggere una delle peggiori assurdità terminologiche: nelle condizioni cui allude la precedente descrizione l'uguale probabilità e l'indipendenza non possono coesistere a meno che la probabilità non sia conosciuta; allora si è nel caso usuale (detto in genere delle «prove ripetute»). Se invece la probabilità (ad esempio il numero di palline bianche e nere) non si conoscesse (si sapesse, ad esempio, che l'urna è stata «scelta a caso» (con uguale probabilità) fra due, di cui una contiene 6 palline bianche e 4 nere, e l'altra viceversa 4 bianche e 6 nere), le estrazioni *non sarebbero indipendenti*. Infatti, a mano a mano che si ripetono delle estrazioni, si sarà giustamente indotti a ritenere che l'urna prescelta sia quella che contiene il maggior numero di palline del colore uscito piú spesso. (Per esercizio – se si trattasse di un testo scolastico – si potrebbero fare esempi numerici, e indicare, per ogni momento (ad esempio dopo 10, o 15, o 20 estrazioni), quali probabilità dovrebbero darsi al fatto che l'urna da cui vengono fatte le estrazioni sia quella con prevalenza di palline dell'uno o dell'altro colore. Qui non interessa fare esercizi, ma basti far notare che, evidentemente, si propenderà per ritenere che l'urna prescelta sia quella con maggior numero di palline del colore che è stato estratto piú spesso).

Quindi, non c'è indipendenza, bensí (in questo caso) spostamento dell'attesa verso il colore presentatosi piú spesso. Sussiste però la scambiabilità, nel senso che tutte le successioni (che variano solo per l'ordine, per esempio con 8 estrazioni di pallina bianca e 10 di pallina nera) hanno la medesima probabilità.

È cosa, in fondo, banale; ma le confusioni che possono derivare dall'uso non sufficientemente inequivocabile del linguaggio sono terribili. Sembrerà forse una affermazione stravagante e assurda sostenere l'importanza di cose che potrebbero sembrare minuzie, ma è proprio dal fatto di considerarle tali e di esprimersi in modi ambigui o sconclusionati che si arriva nel modo piú diretto a crearsi intorno un viluppo inestricabile di concetti confusi e di parole usate a casaccio.

Comunque, la locuzione sbagliata di «indipendenza con probabilità uguale ma incognita» merita di essere ricordata come «memento» per riflettere su tutte le terminologie e accertarsi di capire quali hanno senso e quali sono vuote, o, peggio ancora, contraddittorie.

3.6. Previsione e scarto (quadratico medio).

Il discorso finora svolto aveva carattere e scopo piuttosto orientativo, come sembra particolarmente necessario in un campo dove spesso i veri significati vengono soffocati e sacrificati a favore di astrazioni e formalismi nonché (talvolta) perfino di elucubrazioni piú o meno sconnesse. Si spera che le precedenti considerazioni critiche possano avere per lo meno aiutato a sgomberare un po' il terreno da parecchie delle peggiori insidie che vi allignano.

Ed ora, che fare? Fare una sintesi di tutto ciò che potrebbe costituire la materia di un trattato (o magari trattatello) sarebbe cosa insieme troppo pesante e troppo poco efficace. Nel tentativo di fare un discorso piú appetibile per chi desideri acquisire un po' di familiarità e sicurezza, nonché un po' di competenza, nel ragionare sensatamente in termini di probabilità, dovrebbe giovare maggiormente una sia pur piccola serie di esemplificazioni interessanti e sperabilmente (con appena un po' di sforzo) intuitivamente accessibili.

Non sembri troppo banale (o addirittura offensivo per il lettore) se si inizia dal processo di Testa e Croce; sono molte (e spesso anche elevate e complesse) le considerazioni che scaturiscono dallo studio di questo classico argomento. Il riferimento a tale caso è comunque qui appropriato per considerazioni su *previsione* e *scarto* (scarto quadratico medio: $\sigma(X)=\sqrt{\mathbf{P}(X-m)^2}$, ove $m=\mathbf{P}(X)=$ previsione di X); considerazioni che varranno, sostanzialmente, anche per tutti i casi analoghi piú o meno semplici.

Si indichino con $E_1, E_2, \ldots, E_h, \ldots$ gli eventi «ottenere Testa all'h-esimo colpo a Testa e Croce» (e quindi $\tilde{E}_h$, o $1-E_h$, l'ottenere Croce); piú semplicemente, basta dire che E_h vale 1 oppure 0 a seconda che l'h-esimo colpo dà Testa oppure Croce. Si suppone al solito di ritenere che le probabilità dei due risultati siano uguali (1/2 e 1/2) e che le prove siano indipendenti (cioè che i risultati precedenti non modifichino l'opinione circa le probabilità dei casi successivi).

Ciò equivale a dire, in altri termini, che tutte le diverse possibili successioni di n risultati (Testa o Croce; oppure 0 od 1) vengono giudicate ugualmente probabili; precisamente, ciascuna di probabilità $1/2^n$. (Ad esempio per $n=10$ si avrebbe $2^{10}=1024$, quindi probabilità circa 1/1000, e per $n=20$ circa un milionesimo). Ed è il caso di rammentare, in questa occasione, per evitare di cadervi, quei facili fraintendimenti per cui una probabilità molto *piccola* (o molto *grande*) si *confonde* con *impossibilità* (rispettivamente con *certezza*).

Come già detto (e come è ovvio) il numero di successi su n colpi fissati (il successo sia per esempio Testa) è $X=E_1+E_2+\ldots+E_n$; finché di tali eventi non si conosce l'esito, X è sconosciuto ed è possibile soltanto dirne qualcosa in termini probabilistici. Si potrà esprimere la *previsione* di X, $\mathbf{P}(X)$, come somma delle probabilità $\mathbf{P}(E_h)$; per saper restringere un po' l'indeterminatezza di tale conoscenza sarebbe utile ogni indicazione sull'ordine di grandezza dello scostamento del valore vero di X da quello della detta stima. Il dato consueto (e perciò appunto è chiamato spesso «scarto standard») lo si indica con $\sigma(X)$ ed è definito come *radice della varianza* $\sigma^2(X) = \mathrm{P}(X-m)^2$. Da notare che se si considerassero le distanze da un punto m' diverso dal baricentro m si otterrebbe $\mathrm{P}(X-m')^2$ *maggiore* (e precisamente aumentato di un termine proporzionale al quadrato della distanza $m-m'$ fra il baricentro e il punto scelto in luogo di esso come riferimento). Veniamo allora al nostro caso (di n eventi ugualmente probabili e indipendenti); si è supposto che abbiano probabilità $p=1/2$ (ma si potrebbe considerare il caso di p qualunque, od anche di p_h diverso da evento ad evento).

Comunque, per $p=1/2$, la previsione del numero di successi in n colpi è $n/2$, e lo scarto standard è $\sqrt{n/2}$; si può indicare convenzionalmente (come d'uso) tale fatto dicendo che il numero di successi in n colpi (di probabilità $1/2$, e indipendenti) è $n/2 \pm \sqrt{n/2}$. (Per fare un esempio: su 200 colpi (con probabilità di «successo» $1/2$ ad ogni colpo, e indipendentemente dal risultato degli altri) il numero di successi dev'essere (nel senso detto), $100 \pm \sqrt{100} = 100 \pm 10$, ossia tra 90 e 110). Si badi che tali modi di esprimersi sono utili come indicazione grossolana; non c'è però troppo da preoccuparsene perché nei casi pratici serve piú una *mentalità* intuitivamente allenata a vagliare i rischi e ponderare le decisioni in base a una visione panoramica del pro e del contro. Questo non deve però significare «decidere a vanvera» e senza ponderazione, ma tener presente quanto ci sia di troppo o di troppo poco in una schematizzazione matematica prima di affidarsi ciecamente ai consigli che se ne traggono.

3.7. La tendenza alla distribuzione normale.

Le presenti considerazioni sul caso di Testa e Croce vanno approfondite, sia per l'interesse dell'argomento in sé, sia per i legami con molte delle piú consuete applicazioni. Anziché limitarsi a indicare previsione e scarto (risp. $n/2$ e $\sqrt{n/2}$) si possono indicare le probabilità di ogni singolo valore possibile per la frequenza su un certo numero n di colpi. Esse sono date dalla tabellina in figura 10 dell'articolo «Distribuzione statistica» di questa stessa *Enciclopedia* (vol. IV, p. 1208) fino ad $n=6$: su 64 casi possibili ce n'è 1 con sempre Testa (o sempre Croce), 6 con 5 volte Testa e 1 Croce (o viceversa), 15 con 4 volte Testa e 2 Croce (o viceversa), e infine 20 con 3 volte Testa e altrettante Croce.

Su $n=10$ colpi, tra le $2^{10}=1024$ successioni possibili di 10 tra T e C (e per semplificare si arrotonderà a 1000) ce n'è 1 di tutti T (o tutti C), 10 con 1 C e 9 T, 45 con 2 C e 8 T, 120 con 3 C e 7 T, 210 con 4 C e 6 T, e 252 con 6 C e 6 T (e poi simmetricamente, scambiando T e C). È chiaro che, disegnando l'istogramma, si avrebbe la forma di una collina simmetrica, e si può dire subito (tralasciando la dimostrazione) che la sua forma si avvicinerebbe sempre piú (e, al limite col crescere di n all'infinito, coinciderebbe) col diagramma della *distribuzione normale* (o gaussiana, dal nome di Gauss); (cfr. il già citato «Distribuzione statistica», p. 1210). Può interessare una proprietà geometrica del diagramma di tale distribuzione: la collina che si ottiene facendo ruotare tale curva intorno all'asse centrale ha tutte le sezioni verticali simili (uguali alla sezione centrale, salvo il sempre maggiore appiattimento man mano che ci se ne allontana).

Un processo analogo a quello di Testa e Croce può essere immaginato in due, tre (od anche piú) dimensioni, sia pensando che ogni passo possa essere non solo di *avanti o indietro* su una retta, ma in una delle quattro direzioni (sempre «avanti» o «indietro») nel piano, o delle otto nello spazio a tre dimensioni, o in direzioni qualsiansi (indipendentemente dalla direzione degli assi). Si hanno cosí schemi di «passeggiate aleatorie», che, tra l'altro, possono rappresentare schematizzazioni di moti del tipo browniano (moto disordinato di molte particelle che si urtano procedendo cosí a zigzag). Ed anzi (in una schematizzazione un po' semplificata), varrebbe, nel processo di diffusione, una formula (per la densità) identica a quella di Testa e Croce ($N\omega \pm \sqrt{N\omega}$).

Tornando allo schema di Testa e Croce, va segnalata la semplicità ed efficacia con cui il semplice ma potente «ragionamento di Desiré André» facilita e rende intuitiva la soluzione dei problemi tipo «rovina dei giocatori» (cfr. fig. 7).

Esso si basa sul «principio di rovesciamento»; si ha la medesima probabilità per la traiettoria segnata con linea continua e per quella che (dopo l'intersezione

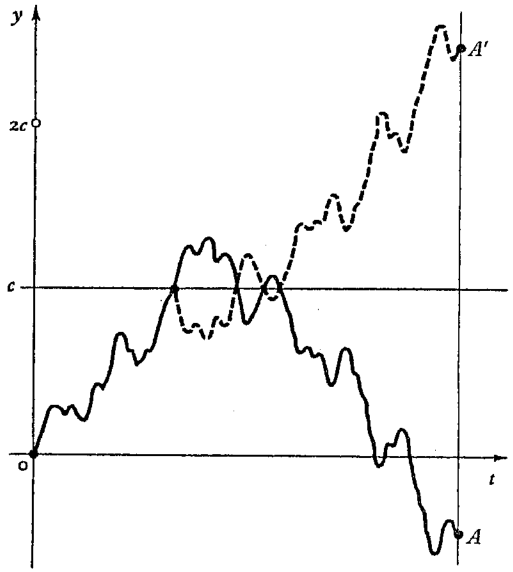

Figura 7.

Ragionamento di Desiré André nel caso di una barriera.

Le traiettorie che, dopo toccato il livello c, ne sono al di sotto alla fine dell'intervallo che interessa (punto A) corrispondono biunivocamente per simmetria a quelle che terminano in A' (simmetrico rispetto alla barriera $y=c$). Di qui (in un processo simmetrico) l'ugual probabilità di terminare in A' o terminare in A dopo toccato il livello c, ed anche, di terminare a un livello $>c$, oppure ad un livello inferiore ma avendo toccato il livello c. Il punto $2c$ sull'asse y è indicato in quanto «sorgente fredda» nel metodo di Lord Kelvin.

col livello c) ne è immagine rovesciata rispetto a detta linea di livello. Nel caso di scommesse eque (probabilità $1/2$ e $1/2$) c'è la stessa probabilità di trovarsi, nell'istante finale, nel punto A o nella sua immagine speculare A'. Ciò significa che c'è ugual probabilità di terminare in A' oppure di *terminare in A dopo aver* toccato il livello c. Problemi del genere sono spesso suscettibili di soluzione elegante con ragionamenti intuitivi di grande efficacia; tale è ad esempio la determinazione della probabilità che un «laccio» (il tratto di traiettoria zigzagante tra due successivi *zeri*) consti di 2, 4, 6, ... (necessariamente numero pari) passi unitari.

Un esempio interessante che fa vedere in modo particolarmente intuitivo la conclusione e l'efficacia del principio di rovesciamento è il cosiddetto «problema dello scrutinio»; si tratta di chiedersi quale sia la probabilità che durante uno scrutinio (supposto che su n voti quelli favorevoli siano in maggioranza in numero di $h>n/2$), essi si siano trovati talvolta in minoranza nel corso dello spoglio. Come la figura 8 mette senz'altro in evidenza, ogni traiettoria che inizi con un passo verso il basso dà luogo ad un'altra che inizi con un passo verso l'alto. Ma il primo passo (come ogni altro) ha probabilità $(n\text{-}h)/n = 1-h/n$ di essere uno degli $n\text{-}h$ passi discendenti. La probabilità di annullamento è doppia, $2h/n$, e quella di non-annullamento è quindi $1\text{-}2h/n$. $\frac{2h}{n} - 1$ $\frac{2(n-h)}{n}$

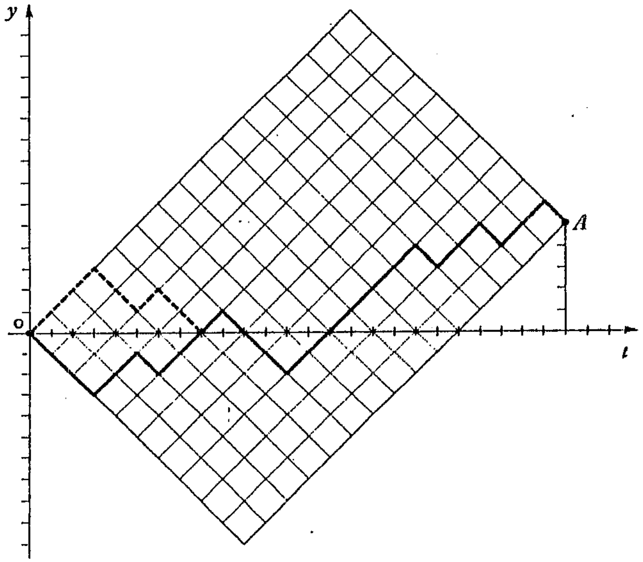

Figura 8.

Ragionamento di Desiré André: problema dello scrutinio (ossia: distribuzione ipergeometrica).

Le traiettorie da O ad A con primo passo discendente corrispondono biunivocamente (per simmetria del tratto fino al primo raggiungimento dell'asse t) a quelle con primo passo ascendente che però toccano l'asse t.

Si accenna ancora ad una conclusione che sembra incredibile: si tratta di chiedere quale sia la previsione della lunghezza di un laccio, cioè del numero di «prove» a Testa e Croce fino al primo ritorno all'equilibrio, cioè a zero. Evidentemente ciò può avvenire dopo 2 colpi (con probabilità $1/2$) o dopo 4, 6, un numero pari qualsiasi,... o anche mai (sebbene tale eternità abbia probabilità zero); ma la previsione è tuttavia infinita (è data dalla somma di una serie che diverge come quella di termini che tendono a zero soltanto come $n^{-1/2}$ (cioè $1/\sqrt{n}$)).

3.8. Riflessioni su presunti paradossi.

I presunti «paradossi» derivano dall'impressione che il calcolo delle probabilità conduca *sempre* ad avvalorare le ipotesi e le previsioni piú appiattite, quelle conformi alla media e alla vetusta massima *in medio stat virtus*. Ma cosí non è: la teoria delle probabilità considera tutti i casi, svariatissimi, rispetto ai quali siamo in condizioni di incertezza, e ciascuno (coll'ausilio di tale teoria o seguendo la sua intuizione, corretta o distorta che sia, nel seguire sia pur istintivamente e grossolanamente certe sue concezioni) dovrà scegliere la decisione e la via che riterrà preferibile, caso per caso.

I casi piú frequenti (o almeno quelli per abitudine considerati «normali») saranno quelli in cui tutti si attendono che permanga piú o meno costante la frequenza osservata nel passato recente (e meglio se anche nel passato meno recente), col «disordine» dovuto al «caso». E probabilmente con eccessiva faciloneria: una sequenza di 10 Teste consecutive ha probabilità assai piccola ($1/2^{10} = 1/1024$, diciamo un millesimo), ma ciò non significa che non si otterrà mai (salvo «miracoli» o... «eccezioni che confermano la regola» secondo un modo di dire stravagante); è da attendersi che il fatto si ripeta in media circa una volta ogni 1000 colpi.

Piú illuminante è l'indicazione data da Willy Feller circa la durata di «permanenza in vantaggio» fra due giocatori che puntano rispettivamente l'uno su Testa e l'altro su Croce ininterrottamente per un anno (un colpo ogni ora, od ogni minuto, od ogni secondo: la numerosità non conta perché sia tanto grande da condurre alle stesse conclusioni del caso limite di infiniti colpi). Per dare un'idea concreta di ciò che è «naturale» prevedere in dette circostanze, si riporta un esempio (di Feller): «Si pensi di giocare continuamente a Testa e Croce per un anno (un colpo ogni ora, o minuto, o secondo: praticamente valgono sempre le conclusioni limite per infiniti colpi); sembrerebbe che, per ragioni di simmetria, i due contendenti dovrebbero trovarsi in vantaggio (complessivamente) ciascuno circa per metà del tempo. Invece: c'è appena probabilità del 30 per cento che entrambi stiano in vantaggio per piú di 100 giorni (circa 28 per cento del tempo totale), mentre c'è probabilità del 50 per cento che uno dei due vi rimanga meno di 54 giorni (15 per cento del tempo), del 20 per cento che vi stia meno di 9 giorni (2,4 per cento del tempo), del 10 per cento che vi stia meno di 2 giorni e mezzo (ossia meno dello 0,6 per cento, e l'avversario piú del 99,4 per cento)». (Da notare poi che l'esser stato piú a lungo in vantaggio non implica maggior probabilità di aver vinto! Infatti tutti i lacci finiscono (per definizio-

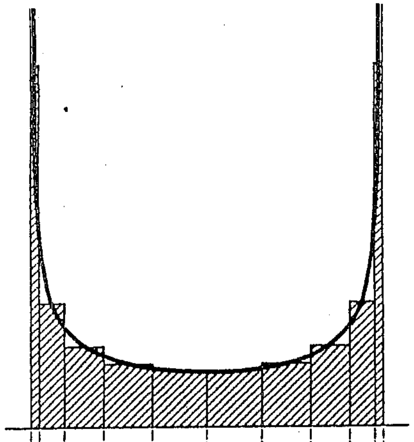

Figura 9.

Densità della distribuzione arcoseno. L'istogramma indica la densità media in ogni intervallo fra i decili. La curva è il diagramma della densità. L'equazione (se l'intervallo-base si assume come $(0,1)$) è $f(x)=k/\sqrt{x(1-x)}$. La densità è infinita negli estremi.

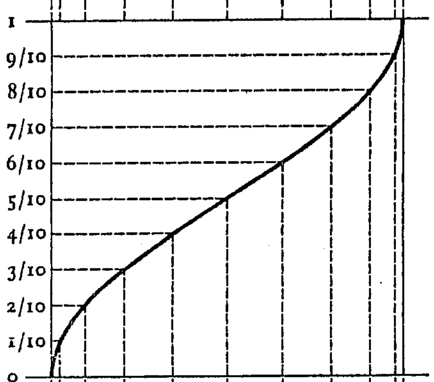

Figura 10.

Funzione di ripartizione della distribuzione arcoseno (ottenibile col modello in fig. 11). Le ascisse segnate sono quelle dei decili come risulta dalle corrispondenti ordinate. I dieci intervalli fra i decili sono ugualmente probabili (probabilità 1/10); notare quanto piú si addensi la probabilità verso gli estremi.

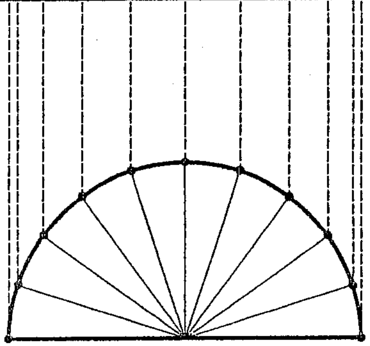

Figura 11.

Si considera la distribuzione di probabilità della proiezione (sul diametro) di un punto «scelto a caso» (densità costante) su una semicirconferenza (o circonferenza). Tale distribuzione si presenta pertanto, ad esempio, se si misura in un istante «a caso» la posizione (o la velocità) di un punto che effettua oscillazioni armoniche. La divisione della semicirconferenza in 10 parti uguali (18°) dà i decili.

ne) in parità, ed è soltanto l'ultimo laccio (anzi frazione di laccio) a decidere la vittoria (o il pareggio)).

Il carattere «squilibrato» di tale processo (probabilità maggiore per squilibri alti anziché modesti) ne fa qualcosa di molto diverso (sotto questo aspetto) dai giochi consueti (tipo Testa e Croce) dove le probabilità maggiori si addensano verso il centro (parità) e svaniscono allontanandosene. Non si tratta però di un caso «eccezionale», ché anzi ne va notata la presenza in fenomeni naturali. L'esempio piú chiaro è quello dei fenomeni periodici (andamento sinusoidale) come ad esempio la proiezione di un punto che percorre una circonferenza a velocità costante, o (approssimativamente, in natura) l'alternarsi dell'alta e bassa marea. È chiaro che l'andamento non è una spezzata a zigzag, in cui si alternerebbero tratti in salita e in discesa rettilinei con passaggio istantaneo e brusco dall'uno all'altro caso, bensí c'è un lento passaggio dalla fase crescente a quella decrescente attraverso un intervallo di quasi stazionarietà. Le figure 9-11 e le relative didascalie bastano a completare la spiegazione.

3.9. Bayes, o del ragionamento induttivo.

È giusto porre ora in testa al titolo il nome di Bayes perché intorno ad esso si è scatenata e tuttora perdura la contrapposizione tra bayesiani e antibayesiani: il suo nome può considerarsi come «segno di contraddizione», come vessillo in una battaglia tra fazioni contrapposte e inconciliabili.

Oggetto della contesa è il fondamento del ragionamento induttivo, del ragionamento che precisa il senso e il modo in cui si fanno, ed è giustificato fare, delle *previsioni*, in termini di *probabilità*, basandosi sull'esperienza, e precisamente – in particolare – sull'osservazione della *frequenza* dei successi in un numero (possibilmente grande) di casi «analoghi» a quello (o quelli) di cui c'interessa prevedere il risultato. In termini piú generali si tratta di vedere il modo in cui «la» (o «le») probabilità in questione vengono modificate in seguito all'acquisizione di ulteriori informazioni, H, che – aggiunte a quelle che costituivano il precedente stato di conoscenza, H_0 – dànno come nuovo stato di conoscenza $H_0'=H\cdot H_0$.

Il teorema di Bayes dice che (come si è già accennato) la probabilità $\mathbf{P}(EH)$ del prodotto di due eventi E ed H è data da $\mathbf{P}(H)\cdot\mathbf{P}(E|H)$, o, simmetricamente, da $\mathbf{P}(E)\cdot\mathbf{P}(H|E)$ da cui risulta che è

$$\mathbf{P}(E|H)=\mathbf{P}(E)\frac{\mathbf{P}(H|E)}{\mathbf{P}(H)} \qquad \mathbf{P}(H|E)=\mathbf{P}(H)\frac{\mathbf{P}(E|H)}{\mathbf{P}(E)};$$

a parole: la probabilità di E, subordinandola ad H, si modifica nello stesso rapporto in cui si modifica la probabilità di H subordinandola ad E.

Le precedenti formulazioni e conclusioni, considerate valide ed anzi ovvie nell'impostazione soggettivistica, sono al contrario ferocemente avversate ed esecrate dagli oggettivisti. Tale contrasto è esploso ripetutamente sia in congressi sia in scambi polemici su riviste statistiche o matematiche, e in particolare al congresso dell'Istituto Internazionale di Statistica a Vienna (1973) dove la difesa del bayesianismo fu affidata a una *contributed paper* (del presente autore). Il titolo rivela di per sé chiaramente l'impostazione: *Bayesianism: its unifying role for both the foundations and the applications of statistics.*

Da parte dei soggettivisti fu fatto notare che la loro posizione è «naturale», e il teorema di Bayes ne è parte in modo ovvio; anzi – come appropriatamente soggiunse Cornfield – è talmente ovvio che è *overly solemn to call it a theorem at all.* I contraddittori non seppero dire nulla che denotasse idee contrastanti ma rispettabili; si limitarono perciò a sfogarsi strappandosi (metaforicamente) le vesti per il sacrilego rifiuto ad adorare le presunte «Probabilità oggettive» come un novello vitello d'oro.

3.10. L'atteggiamento bayesiano.

Sembrerebbe logico, per chiunque abbia una mentalità immunizzata rispetto ad ogni superfetazione oggettivistica, che tutti i suoi ragionamenti e pensieri fossero ispirati (anche inconsapevolmente rispetto a tecnicismi) ad un'attenzione per ogni nuova scoperta o conoscenza tale da farla incorporare in senso bayesiano nel precedente complesso delle sue conoscenze, con tutti gli eventuali conseguenti mutamenti e arricchimenti di esse.

Probabilmente tutti già fanno cosí: fanno un continuo aggiornamento del loro orizzonte globale, riordinandone piú o meno automaticamente tutto il contenuto; dimenticano o mettono in disparte (in una «memoria esterna», per dirlo in termini di informatica) ciò che preme di meno, e collocano quelle per loro piú importanti in posizioni di piú rapido accesso e in piú stretto collegamento con altre zone della memoria.

L'analogia dovrebbe giovare in entrambi i sensi, incoraggiando ad usare le proprie facoltà mentali nel modo migliore per utilizzare (selezionandolo) tutto l'input che ci proviene dall'esterno per aggiornare il complesso delle cose piú o meno ricordate, richiamandole per fondersi coi nuovi apporti. È certo che a tale opera collabora – anche forse a nostra insaputa, ai limiti dell'inconscio – la nostra mente per perfezionarla, la nostra attenzione per precisarla, la nostra fantasia per precorrere le possibilità di farne uso nel modo piú soddisfacente.

Ed è questo, al di là delle applicazioni piú specialistiche e tecniche, il ruolo che ha la probabilità, e in particolar modo il ragionamento bayesiano, per contribuire congiuntamente – nel corso del cammin di nostra vita – a norme oggettive di razionalità e lungimiranza. [B. D. F.].

De Finetti, B.
1970 *Teoria delle probabilità. Sintesi introduttiva con appendice critica*, Einaudi, Torino.

Grayson, C. J. jr
1958 *Decisions under Uncertainty. Drilling Decisions by Oil and Gas Operators*, Harvard University Division of Research, Boston.

Jeffreys, H.
1939 *Theory of Probability*, Clarendon Press, Oxford.

La probabilità si presenta, a prima vista, come un'entità paradossale, perché di essa altrettanto si potrebbe negare l'**esistenza** quanto affermarla come ovunque presente (cfr. anche **essere, idea**). Ma ciò forse inerisce piú che al **concetto** alla sua concretizzazione (cfr. **astratto/concreto**) o piuttosto ancora al suo uso indiscusso e a-problematico, che riversa dalla parte dell'oggetto ciò che invece piú propriamente appartiene al soggetto (cfr. **dato, soggetto/oggetto**). Le necessità scientifiche o pratiche (cfr. **scienza, teoria/pratica**) ove interviene la probabilità impongono naturalmente che in determinati contesti si debbano usare leggi (cfr. **legge**) piú o meno empiriche (cfr. **empiria/esperienza, esperimento**; e anche **convenzione, operatività**) ricorrendo a metodi (cfr. **metodo**) piú o meno attendibili (cfr. **anticipazione, invenzione, ipotesi, modello**); che venga usato un **linguaggio** determinato e specifico (cfr. anche **formalizzazione, logica**) il quale certo impone l'uso di segni determinati o di particolari metafore (cfr. **codice, metafora, segno, significato**); ma è un'istanza **metafisica** ingiustificabile far sí che questo vasto spettro di azioni e di pensieri debba raccogliersi in definizioni pseudo-oggettive e contraddittorie (cfr. anche **opposizione/contraddizione, errore, dicibile/indicibile**). Liberato il concetto di probabilità dai vincoli di una falsa oggettività è possibile suggerirne un uso pertinente nei vari contesti ove esso si dà (cfr. **caso/probabilità, causa/effetto**), legandolo appropriatamente ai concetti e ai procedimenti del **calcolo** dell'**induzione statistica**, alle valutazioni quantitative (cfr. **qualità/quantità**) e in generale collegandolo in modo piú adeguato alle varie esigenze del comportamento umano (cfr. **comportamento e condizionamento, decisione, giochi**).