

Fondamenti e metodi per l'analisi empirica nelle scienze sociali

Corso di laurea in Teorie, culture e tecniche per il Servizio sociale

Enciclopedia delle Scienze sociali

Voce

Probabilità

1. Misura della probabilità di *Italo Scardovi*
2. Significato della probabilità di *Giorgio Dall'Aglia*

Misura della probabilità

SOMMARIO: 1. La probabilità come numero reale. □
 2. La probabilità come proprietà statistica. □ 3. La probabilità come attesa di un evento unico. □ 4. Quale probabilità? Un distinguo irrinunciabile. □
 Bibliografia.

1. La probabilità come numero reale.

Nel parlar comune, 'probabilità' è parola che esprime incertezza, ora per alludere al 'grado di credibilità' di una proposizione della quale non si può affermare o negare in assoluto la verità, ora per intendere il 'grado di avverabilità' di un accadimento non certo. Concetti diversi, ambigualmente confusi, nella lingua italiana, in un medesimo vocabolo. Che tuttavia riflette, in entrambe le fattispecie, un'istanza quantitativa: nel senso di una graduazione, o di una comparazione. S'usa dire, infatti, aggettivando, che un assunto, o un evento, è molto (o poco) probabile; o che è più (o meno) probabile di un altro.

L'incertezza può attenersi a interrogativi di vario ordine, conoscitivo o strategico, e si risolve in una sintesi, conscia e inconscia, di opinioni e di informazioni. E già il senso comune spinge ad ancorare la valutazione di una probabilità a elementi oggettivi. Ciò soprattutto nella determinazione di una probabilità intesa quale misura dell'aspettativa di un evento aleatorio. Nell'*Ars conjectandi* (Basileæ, 1715) di Jakob Bernoulli — una delle prime trattazioni teoriche del 'calcolo delle probabilità' — è data una definizione icastica e universale: "*Probabilitas est gradus*

certitudinis et ab hac differt ut pars a toto'. L'enunciato suggerisce una valutazione quantitativa delle conoscenze, la cui limitatezza è il presupposto psicologico dell'idea stessa di probabilità.

Ma come determinare il grado della certezza, come porzionare la parte al tutto? Si deve distinguere, anche per evitare equivoci che sono all'origine di un inesausto dibattito: una cosa è la veridicità di un'ipotesi, di una tesi; altra cosa è l'avverabilità di un fatto del mondo, che può accadere e può non accadere. Una dualità che assume particolare significato nel sapere scientifico, dove è altresì essenziale il distinguo tra probabilità come espressione di incertezza soggettiva per difetto di informazione e probabilità come espressione dell'oggettiva indeterminazione dei fenomeni. Da quando alla ricerca si è rivelata l'immanente casualità di tutta una realtà naturale, il pensiero scientifico si è dato una nuova logica: la logica di una scienza che ha fatto posto al 'caso' e ha sostituito la probabilità alla necessità.

Canoni per la graduazione quantitativa dell'ammissibilità di un assunto non possono essere addotti in generale, se non nel senso di distinguere e soppesare le circostanze ritenute rilevanti e di tradurre il giudizio nella scala dei numeri reali da 0 a 1. I valori estremi, 0 e 1, indicano, rispettivamente, la 'falsità' e la 'verità' dell'assunto, nell'ambito di una logica a due valori (vero, falso), due essendo gli stati possibili. Un qualunque valore intermedio esprime uno stato personale di incertezza: dipende dalle conoscenze di cui si dispone e dall'uso che se ne fa; e può cambiare al cambiare di chi valuta. Una misura della probabilità, dunque, in cui prevalgono gli argomenti 'soggettivi'.

E invece possibile addurre criteri empirici e principi formali per misurare la probabilità dell'avverarsi di un evento aleatorio, oggettivandola in un numero reale p compreso tra 0 e 1, estremi inclusi; questi designano, rispettivamente, l'impossibilità ($p=0$) e la 'certezza' ($p=1$) dell'evento. Un valore p nell'intervallo $[0, 1]$ corrisponde ad un grado di avverabilità dell'evento. L'evento potrà, poi, presentarsi o non presentarsi, così da proporsi, *a posteriori*, in una logica a due valori; e tuttavia la probabilità p può essere ritenuta come intrinseca all'evento e tale da assumere valori crescenti all'aumentare della sua possibilità di accadere: un dato naturale, prima e più che la sensazione di un osservatore. Il quale può valutare l'accadibilità dell'evento affidandosi a criteri di simmetria o di analogia. Una misura di probabilità in cui prevalgono gli elementi 'oggettivi'.

A siffatti criteri si sono ispirate le prime regole empiriche per razionalizzare l'attesa di un evento aleatorio, prendendo lo spunto dai giochi d'azzardo. Questi hanno offerto modelli essenziali alla riflessione sulla probabilità, suscitata appunto, sul finire del Rinascimento, dai quesiti sorti intorno ai tavoli da gioco e divenuti un divertimento intellettuale per i maggiori matematici dell'epoca (da G. Galilei a P. Fermat, a B. Pascal, a C. Huygens, a G.W. Leibniz, ad A. de Moivre, a J. Bernoulli, ...). Trattandosi di realtà enumerabili (il numero delle facce di un dado, il numero dei lanci del dado...), divenne subito naturale ridurre la misura della probabilità di un

evento aleatorio a una quantità direttamente proporzionale al numero degli esiti favorevoli all'evento e inversamente proporzionale al numero di tutti gli esiti possibili, assegnando a ognuno di questi, se ritenuti ugualmente possibili, una uguale frazione della probabilità ($p=1$) dell'evento certo'. Se m sono le eventualità equipossibili, è attribuita *a priori* a ciascuna di esse la probabilità $p=1/m$.

L'avverabilità di un evento aleatorio resta così misurata da una grandezza $0 \leq p \leq 1$ ottenuta come quoziente del rapporto del numero degli esiti favorevoli al numero degli esiti (ugualmente) possibili. Ed è $p=0$ quando nessun esito produce l'evento (l'evento è 'impossibile') e $p=1$ quando tutti gli esiti possibili lo realizzano (l'evento è 'certo'). Ad esempio, se l'evento aleatorio è l'apparizione di cifra dispari nel lancio di un dado dalle facce contrassegnate con le cifre da 1 a 6, tre sono gli esiti favorevoli all'evento — le cifre 1, 3, 5 — e sei gli esiti possibili: le sei facce del dado; se queste sono ritenute ugualmente possibili, ad ognuna di esse viene attribuita una uguale frazione — cioè $1/6$ — della probabilità dell'evento certo; lasciando cadere un cubo su un piano orizzontale si dà per certo che una delle sei facce resterà rivolta verso l'alto. La probabilità dell'evento 'cifra dispari nel lancio di un dado' è allora fissata nel quoziente ($p=0,5$) del rapporto di parte al tutto $3:6$.

Già delineata nel cinquecentesco *De ludo aleae* di Gerolamo Cardano, questa regola ha preso il nome di 'definizione classica' della probabilità. È piuttosto un criterio di misura della probabilità di eventi aleatori in contesti enumerabili entro sistemi simmetrici di struttura conosciuta. Postulando l'egual probabilità degli esiti possibili, quella 'definizione' sembra dare per noto ciò che intende definire. Per ovviare a questa circolarità, Bernoulli enunciava, rovesciando un canone leibniziano, il "principio di ragione non sufficiente", detto anche, in seguito, "principio di indifferenza". Pierre Simon de Laplace — al quale si deve una fondamentale trattazione teorica della probabilità, intesa come "buon senso ridotto a calcolo" (*Théorie analytique des probabilités*, Paris 1812) — affermerà l'esigenza di "ridurre tutti gli eventi dello stesso genere a un numero di casi ugualmente possibili, cioè tali che siamo ugualmente indecisi sulla loro esistenza", così da identificare la probabilità nel "rapporto tra il numero dei casi favorevoli e quello di tutti i casi possibili, a condizione che siano ugualmente possibili".

Una definizione destinata a durare nel tempo quale paradigma di tutti i fenomeni riducibili a sistemi aleatori composti da un numero finito di alternative equivalenti: il dado del giocatore come gli insiemi di particelle della "statistica di Maxwell-Boltzmann". Allo stesso modo di Galileo che, enumerando i costituenti elementari, aveva saputo rispondere al quesito di uno scommettitore al 'gioco del passadièci', dimostrando (*Sopra le scoperte de i dadi*) che nel lancio di tre dadi possono darsi ventisette terne di cifre a somma dieci e undici, e venticinque a somma nove e dodici, Ludwig Boltzmann (*Vorlesungen über Gastheorie*, Wien 1895), enumerando i 'microstati' e assumendoli come equiprobabili, dedurrà la probabilità dei conseguenti 'macrostati', dando ragione (una ra-

gione statistica) della tendenza dei processi termodinamici a volgere la freccia del tempo verso il macrostato più probabile, secondo le stesse regole che intervengono quando si mischia un mazzo di carte da gioco.

Ma così si venivano mischiando in profondo le carte della metodologia scientifica, e la probabilità si avviava a diventare il linguaggio di una nuova razionalità e lo strumento euristico di una nuova scienza. Una probabilità non più da intendersi come espressione dell'incapacità umana a ricondurre l'essere e il divenire all'obbedienza del determinismo universale eretto sul modello meccanico di Newton-Laplace.

2. La probabilità come proprietà statistica.

Quando i principi di simmetria e di indifferenza nei confronti delle alternative possibili non sono postulabili (nell'esempio, che ha ovviamente valore allegorico: il dado non è bilanciato, essendo il suo centro di gravità spostato rispetto al centro geometrico dell'esaedro regolare), la probabilità non è più determinabile nel modo premesso. La si può tuttavia indurre empiricamente enumerando i risultati di una serie di osservazioni eseguite nelle stesse condizioni (nell'esempio: lanci ripetuti del dado). È acquisito all'esperienza che la proporzione in cui il numero i delle uscite di un evento sta al numero n delle osservazioni effettuate si approssima alla probabilità determinabile *a priori* in un sistema simmetrico e che l'approssimazione tende ad aumentare con il numero delle osservazioni.

Si può allora ravvisare una misura del grado di avverabilità dell'evento nel rapporto $f_n = i/n$ (essendo $i \leq n$), detto 'frequenza relativa': una 'probabilità statistica', che assume stati di grandezza nell'intervallo $[0,1]$ e vale zero per $i=0$ e uno per $i=n$. La probabilità così determinata tende a variare con il numero n delle prove, e tende pure a variare ripetendo altre serie di n prove (è il cosiddetto 'errore casuale di campionamento'); ma, all'aumentare del numero delle prove, tale frequenza si approssima, seppur irregolarmente, a un valore costante; e ciò per una proprietà statistica detta "legge empirica del caso". "Anche la più stupida delle persone — scriveva Bernoulli a Leibniz in una lettera dell'aprile 1703 — [...] sa che, più cresce il numero delle osservazioni, minore è il pericolo di allontanarsi dal vero". Non si tratta, tuttavia, di una tendenza al limite nel senso dell'analisi infinitesimale. Matematicamente si può solo dimostrare ('teorema di Bernoulli') che la probabilità P di una differenza tra f_n e p inferiore ad un valore ε , arbitrariamente piccolo, tende alla certezza all'aumentare indefinito del numero n delle osservazioni. In simboli:

$$\lim_{n \rightarrow \infty} P\{|f_n - p| < \varepsilon\} = 1.$$

La 'probabilità statistica' si addice a tutte le realtà fenomeniche in cui sia possibile ridurre i fatti, per astrazione classificatoria, a replicazioni di un medesimo evento, così da renderli enumerabili e matematizzabili: il numero che ne conta il ripetersi è il dato quantitativo più immediato per determinare un va-

lore di probabilità; ed è tanto più stabile quanto più il numero delle osservazioni è consistente, essendo l'apporto informativo di ogni nuova osservazione omogenea tanto meno rilevante quanto più ampio è l'osservato. La necessità di frequenze relative emergenti da un sufficiente numero n di osservazioni ha indotto Richard von Mises (*Wahrscheinlichkeit, Statistik und Wahrheit*, Wien 1928) a interpretare la probabilità come un limite: il limite a cui tende la frequenza relativa in una successione illimitata ('*Kollektiv*') di prove aleatorie. È la cosiddetta 'definizione frequentista' della probabilità, già intuita da John Venn (1866).

Sulle probabilità statistiche si sono fondate e si fondano intere scienze; e prosperano gli istituti assicurativi. L'assicuratore valuta la probabilità di sopravvivenza di un individuo rifacendosi anzitutto alle 'tavole di eliminazione' di un gruppo omogeneo di suoi coetanei, come il fisico trae la probabilità di decadimento di un nucleo atomico dalla legge dell'estinguersi di insiemi di atomi radioattivi: l'uno interessato alla vita media di una collettività, l'altro al tempo medio di dimezzamento di un aggregato di particelle. Ancor più si è avvalsa e si avvale di probabilità statistiche la genetica, che da Johannes (Gregor) Mendel in poi ha visto tramandarsi i caratteri biologici delle specie attraverso le generazioni come in un gioco di dadi. Classificando ed enumerando i risultati di molte serie di ibridazioni di papilionacee, Mendel riuscì a svelare il meccanismo della trasmissione ereditaria dei caratteri e a darne le leggi: le leggi (statistiche) di una natura sorpresa a giocare a 'testa o croce'.

"Dipende semplicemente dal caso — scriveva nei suoi *Versuche über Pflanzenhybriden*, Brünn 1866 — quale dei due tipi di polline si unirà con ciascuna cellula uovo. Tuttavia, per la legge delle probabilità, nella media di molti casi avverrà sempre che ogni forma di polline si unirà con egual frequenza con ciascuna delle forme di cellule uovo [...]. I veri rapporti numerici possono risultare soltanto dalla media tratta dal maggior numero possibile di casi individuali; più grande è il loro numero, più facilmente si eliminano le perturbazioni accidentali". Non altrimenti in meccanica statistica: "Con ciò vediamo — scriveva Schrödinger, riferendosi al secondo principio della termodinamica — che una legge naturale [...] perde tanto più la propria validità quanto minore diventa il numero dei singoli processi intervenuti" (*Was ist ein Naturgesetz?*, Zürich 1921). E Max Born: "Ormai l'ordine delle idee si è capovolto: il caso è diventato la nozione principale, la meccanica un'espressione delle sue leggi quantitative, e la schiacciante evidenza della causalità con tutti i suoi attributi nel campo dell'esperienza ordinaria viene spiegata dalle leggi statistiche dei grandi numeri" (*Natural philosophy of cause and chance*, Oxford 1949).

È il linguaggio di una nuova filosofia naturale, di un sapere che è andato scoprendo negli algoritmi della natura gli algoritmi della probabilità. Più ancora che la critica settecentesca di David Hume al 'principio di causa', è stata la grande rivoluzione scientifica avviata verso la metà del secolo diciannovesimo a insidiare il dogma della causalità deterministica, a da-

re connotati ontologici alla casualità, a fare della probabilità qualcosa di più e di diverso da un vago soccorso epistemico all'umana conoscenza nei confronti di un inafferrabile mondo deterministico. Scrive Henry Poincaré: "Bisogna dunque che il caso sia altra cosa dal nome che diamo alla nostra ignoranza" (*Science et méthode*, Paris 1908).

Quanto più la filosofia si è interrogata sulla natura della probabilità, tanto più la scienza è andata svelando la probabilità della natura. E ne ha date le leggi: leggi che trascendono il costituente elementare, che esprimono proprietà di insiemi. Leggi tanto più attendibili quanto più gli insiemi osservati (i 'campioni statistici') sono empiricamente stabili. Se, per gli eventi assimilabili allo schema classico dei giochi d'azzardo, ove si ammette l'invarianza del sistema, si tratta soltanto di ingrandire il numero delle osservazioni, per ogni altra categoria fenomenica sorge il problema dell'identificazione della classe di riferimento e della delimitazione della base empirica. Una scelta soggettiva, che si risolve in un dato oggettivo.

Le leggi di questa scienza codificano realtà collettive: sciame di molecole, aggregati di atomi, sequenze di geni, popolazioni di viventi. Anche se la traiettoria di una molecola, la sorte di un atomo, la vicenda di un gene, la storia di un vivente fossero conosciute, esse sarebbero inessenziali. È certamente anche per questo che, da Reichenbach a von Mises e a tanti altri, si è negata ogni rilevanza scientifica alla probabilità del fatto singolo. Ha scritto Bertrand Russell (*The scientific outlook*, London 1931): "Qualsiasi conoscenza noi abbiamo, o è conoscenza di fatti particolari, o è conoscenza scientifica". Certo, ogni fatto è un *unicum*. Ma il sapere è avanzato riducendo i fatti in classi, quali simboli di categorie concettuali (dove il nome comune, il numero naturale), cercando le leggi emergenti dalla variabilità individuale: leggi che oltrepassano la singolarità irripetibile.

I canoni metodologici della scienza che ha scoperto l'irriducibilità di tanti fenomeni al paradigma deterministico non rientrano in una logica a due valori. Non si riconoscono in proposizioni categoriche individuali della forma "Ogni A è B ", o in asserzioni ipotetiche univoche del tipo "Se A , allora B ", bensì in enunciati statistici quali: "Una frazione p degli A è B "; "Se A , allora B con probabilità p " (che per $p=1$ identificano gli enunciati precedenti). L'asserzione ipotetica di questa nuova logica è plurivoca, in uno spazio probabilizzato: "Se A , allora B_1 con probabilità p_1 , B_2 con probabilità p_2 , ..., B_j con probabilità p_j , ..., B_k con probabilità p_k , tali che $p_1 + p_2 + ... + p_j + ... + p_k = 1$ ". Dunque, una pluralità esaustiva e disgiuntiva di 'conseguenti' dell' 'antecedente' A : essi descrivono una variabile sulla quale è ripartita la probabilità dell'evento certo. Prima e più che di insufficienza della mente, tale pluralità è espressione di indeterminazione della realtà. Cambiano, in questi contesti non deterministici, anche i canoni della confutazione e svanisce la tradizionale asimmetria tra 'falsificazione' e 'verificazione'.

Quale è allora il significato di una probabilità statistica? Nell'ambito del paradigma deterministico-causale la probabilità, comunque determinata, ha

sempre un implicito significato soggettivo: è un utile sussidio al pensiero umano nei confronti di una realtà che può essere attinta soltanto da "[...] un'intelligenza che in un dato istante conoscesse tutte le forze da cui la natura è animata e la situazione rispettiva degli esseri che la compongono", come è scritto nelle prime battute del laplaciano *Essai philosophique sur les probabilités* (Paris 1814). Ogni altra immagine della probabilità è esclusa da quel rassicurante "sistema del mondo", che sembrava non lasciare spazio alcuno al libero arbitrio. Anche le regolarità empiriche accertate da Adolphe Quételet — un astronomo che non aveva esitato a cercare negli eventi dell'uomo, nei suoi caratteri fisici, nelle sue vicende sociali, regolarità appena assimilabili a quelle osservate nel cielo — erano intese come tracce epifenomeniche di un determinismo profondo e nascosto.

Nel clima rigidamente deterministico della scienza dominante, che respingeva le scienze sociali ai margini del sapere razionale, quelle regolarità collettive parevano estendere, *more statistico*, il dominio del causalismo meccanicistico ai fenomeni dell'umano: una meccanica sociale, per così dire. Di una 'fisica sociale' si era fatto propugnatore Auguste Comte, caposcuola del positivismo e fondatore della sociologia, che tuttavia rifiutava ogni analisi probabilistica dei fenomeni — "una vergognosa aberrazione scientifica" — perché vi intravedeva un'insidia al determinismo. Eppure, Quételet aveva ricondotto a una sorta di superiore ragione fenomeni che vi si ribellavano, a cominciare da quelli più spontanei; ad esempio, i reati. "L'uomo paga al crimine" — scriveva nelle *Lettres sur la théorie des probabilités appliquée aux sciences morales et politiques* (Bruxelles 1846) e nella *Physique sociale* (Bruxelles 1869) — un tributo ancora più costante di quello che deve pagare alla natura o al tesoro dello Stato".

Una sorta di fatalità statistica, che subordina la casualità alla causalità, la probabilità alla necessità e detta i destini dei singoli nel gioco magico dei grandi numeri: la logica e l'etica di un arcano fato collettivo inteso a decidere le categorie, ma indifferente alle sorti degli individui, lasciate al caso, così da aggirare l'alternativa kantiana tra libertà umana come autodeterminazione e determinismo fenomenico come razionalità della natura. E se di fronte a fenomeni non riconducibili a una causalità immediata si faceva appello all'accidentalità, era solo — rassicurava Laplace — per "l'ignoranza dei legami che li uniscono all'intero sistema dell'universo". La probabilità, dunque, come sfida al caso, come artificio di una ragione incapace di cogliere le più riposte ma indubitabili causalità del mondo.

Altro è il canone di lettura delle scienze imbattute nella casualità immanente, nell'indeterminismo irriducibile di tutta una realtà naturale (le scienze dove la probabilità ha avuto un ruolo costitutivo o addirittura fondante): un macrofenomeno è regolare perché vi concorrono innumerevoli microfenomeni irregolari; non sono questi ad accadere per l'imperio di quello: è quello ad avverarsi per l'apporto di questi.

Nella probabilità come frequenza statistica si può vedere il riflesso empirico di una tendenza effettuale. Una tendenza che ha suggerito a K.R. Popper

(1957) e a M. Bunge (1973) l'idea di una "propensione" del singolo evento: un'espressione suggestiva, che sembra evocare un passo di Galileo sulla caduta dei gravi, intesa come "propensione naturale dei corpi". Un'immagine che può riflettere, a dir vero, anche processi naturali 'statistici': dai fenomeni biologici che evolvono nel 'tempo di Darwin' per il gioco del caso con la necessità, agli eventi molecolari che s'inoltrano nel 'tempo di Boltzmann' sospinti dalla probabilità. Se questa non fosse una componente statistica di quei fenomeni, come dar ragione dell'irreversibilità temporale di macroprocessi composti da innumerevoli microprocessi tutti reversibili? Come spiegare altrimenti la tendenza entropica di una popolazione di particelle, ciascuna delle quali risponde a leggi meccaniche simmetriche rispetto al tempo? Un'intuizione, questa, che ha profondamente trasformato la gnoseologia naturalistica.

Un'intuizione già balenata fra le pagine di alcuni filosofi dell'età dei Lumi'. Se non in Condorcet (1785), che pure non aveva esitato a trattare in chiave probabilistica delicate questioni sociali, a cominciare dal problema della composizione delle giurie e dell'attendibilità delle testimonianze, facendo della probabilità un espediente per sottrarsi alle malie della superstizione, certamente in Diderot (1769), che aveva azzardato un'interpretazione accidentalistica della natura e della vita, e in Maupertuis (1753), che era giunto a fare di un prodotto di probabilità lo strumento euristico per un'indagine d'avanguardia su filogenesi familiari.

3. La probabilità come attesa di un evento unico.

Trascurando, perché estranea alla scienza, la probabilità dell'evento unico, il frequentismo esclude dal dominio della probabilità i problemi strategici tipici delle decisioni in condizioni di incertezza e, in genere, di tutte le situazioni non riconducibili a classi di eventi. "L'identificazione della probabilità con le frequenze" — scriveva John Maynard Keynes in *A treatise on probability* (London 1921) — esclude un grande numero di giudizi, che si ritiene attengano alla probabilità". Sulle orme di A. de Morgan (1847) e di C. Boole (1854), prese le distanze dal frequentismo, Keynes ha delineato una concezione 'logicista' della probabilità, intesa come 'credenza razionale' ed espressa da una relazione logica tra due asserzioni, una relativa all'accadibilità di un evento, l'altra alla conoscenza di cui si dispone. Una relazione formale tra un'ipotesi e un'evidenza, riducibile a una misura univoca non appena si identifichino, con L. Wittgenstein (1921), strutture elementari linguisticamente omogenee da enumerare e comporre in rapporti di parte al tutto sulla falsariga della definizione classica, della quale la 'probabilità logica' può ritenersi una riformulazione. Keynes, tuttavia, non riteneva sempre possibile assegnare un grado univoco di fiducia, dovendosi spesso ripiegare sul confronto tra due o più gradi e limitatamente alle situazioni in cui ricorrono i requisiti della confrontabilità.

La visione keynesiana acquisisce rigore sistematico nell'opera di Rudolf Carnap (*Logical foundations of*

probability, Chicago 1950), ove sono addotti due distinti concetti della probabilità: la 'probabilità logica' (probabilità-1) e la 'probabilità empirica' (probabilità-2). La prima, riferita a eventi singoli su di un'alternativa a due valori, è intesa come "grado di conferma $c(h, e)$ di un'ipotesi h sulla base di un'evidenza empirica e "; la seconda si identifica nelle frequenze statistiche, che Carnap definisce "le probabilità della scienza" e lascia fuori dal proprio sistema logico-formale. Una dualità non nuova. Risale a S.D. Poisson (1837) e ad A. Cournot (1843), i quali usavano opportunamente la parola '*probabilité*' per intendere il grado di credenza e la parola '*chance*' per intendere l'attitudine di un evento a prodursi: l'attitudine che Laplace chiamava '*facilité*'.

All'evento unico, a null'altro che all'evento unico, si riferisce la moderna concezione 'soggettivista' della probabilità (Frank P. Ramsey, Bruno de Finetti, Leonard Savage, Dennis V. Lindley), che non ammette distinzioni di sorta e si propone quale teoria unitaria della probabilità, definita come "grado di fiducia di un dato soggetto, in un dato istante e con un dato insieme di informazioni, riguardo al verificarsi di un evento" (v. de Finetti, 1970). Una probabilità ispirata alle situazioni di incertezza tipiche dei contesti decisionali e misurata in termini di scommessa personale nei riguardi di un evento singolo. La probabilità è allora il prezzo da pagare per ricevere 1 se l'evento accade e 0 se non accade. Un solo vincolo: la scommessa non deve essere né sicuramente vincente né sicuramente perdente.

Alla metrica della scommessa si rifaceva già Immanuel Kant. Sotto il titolo "Dell'opinione, della scienza e della fede" si legge nella sezione terza della *Dottrina trascendentale del metodo*, libro secondo della *Kritik der reinen Vernunft* (Königsberg 1781): "L'ordinaria pietra di paragone per vedere se qualche cosa, che uno afferma, sia una semplice persuasione, o almeno una convinzione soggettiva [...] è la scommessa. Spesso uno enuncia le sue proposizioni con una risolutezza così sicura e irriducibile da parere abbia interamente deposto ogni tema d'errore. Una scommessa lo fa adombrare. A volte si vede che egli possiede bensì una persuasione da poter essere apprezzata per un ducato, ma non per dieci. Infatti, egli rischia il primo, ma, di fronte a dieci, comincia ad avvedersi di ciò che prima non avvertiva, essere cioè possibilissimo che si sia sbagliato. Se si immagina di doversi scommettere la felicità di tutta la vita, se ne va il nostro giudizio trionfale, diventiamo timidi, e cominciamo a scoprire che la nostra fede non va tanto in là. Così la fede pragmatica ha soltanto un certo grado, che, secondo la differenza dell'interesse che vi è in gioco, può essere grande e può essere piccolo".

Anche il principio utilitaristico della scelta di una probabilità in termini di vantaggi e svantaggi viene dal passato: da Blaise Pascal, uno dei primi a cimentarsi con la matematica dei giochi di sorte e a trarre dalla ragione delle scommesse le scommesse della ragione. Il famoso "*pari de Pascal*" invita infatti a puntare sull'esistenza del trascendente, poiché se si vince il guadagno è infinito, e se si perde non si perde nulla. Una visione della probabilità che si iscrive

nella categoria della convenienza, non in quella della conoscenza; che si misura sulla scala dell'utilità, non su quella della verità.

In tale scommessa si può riconoscere il protoschema di una concezione delle probabilità approdata, nella seconda metà del XX secolo, al decisionalismo utilitaristico, quale criterio e prassi della scelta e dell'azione. Una concezione in cui s'avverte un riflesso del pragmatismo filosofico, alla Dewey. Essa ha trovato il suo più naturale contesto logico ed etico nelle strategie economiche e finanziarie, dove l'incertezza è un elemento del sistema, e il comportamento razionale consiste nel comparare le conseguenze delle azioni possibili, in termini di utilità. È di T. Haavelmo (1958) l'auspicio a una teoria probabilistica capace di metodi di valutazione e di previsione economica attraverso probabilità soggettive, in quanto "realtà della mente"; ed è di E. Malinvaud (1964) l'invito a costruire modelli aleatori dell'econometria impostati su probabilità soggettive.

La 'probabilità soggettiva', ancorché stato della mente, non ignora i dati dell'esperienza. Se ne avvale, invece, anche attraverso un algoritmo del calcolo delle probabilità che risale a T. Bayes (1763) e a P.S. Laplace (1774). È l'algoritmo delle inferenze induttive, del comporsi di informazioni e opinioni, e del modificarsi di queste alla prova dei fatti. Data un'ipotesi H e un'evidenza empirica E , la probabilità di H subordinata a E è espressa dalla formula:

$$P\{H|E\} = P\{H\}P\{E|H\}/P\{E\}$$

Dunque, una probabilità 'iniziale' $P\{H\}$, che può essere una sensazione personale, si compone con un 'fattore di verosimiglianza' $P\{E|H\}$ e si risolve nella probabilità 'finale' $P\{H|E\}$. Si può tuttavia dimostrare — così von Mises (1931) — che con l'accrescersi delle esperienze (ossia con l'aumentare del numero n delle osservazioni) i valori delle probabilità $P\{H|E\}$ e $P\{E|H\}$ tendono a convergere (purché $P\{H\}$ sia diversa da 0 e da 1). L'evidenza sperimentale, il valore oggettivo dei fatti, vince, allora, sui giudizi aprioristici (e fa giustizia di tanto dibattere). È solo questione di ampiezza del contesto empirico.

4. Quale probabilità? Un distinguo irrinunciabile.

Quella soggettivistica si propone come concezione onnicomprensiva della probabilità. E tale è, in certo modo. Una teoria delle probabilità diventata 'teoria delle decisioni'. Non si vede però quanto possano giovare della scommessa le scienze che si sono imbatute nell'accidentalità combinatoria dei processi intimi della vita, dei costituenti elementari della materia: le scienze che hanno addotto leggi (statistiche) delle pluralità a lato, o al posto, di leggi (deterministiche) delle singolarità. La scommessa 'conveniente', alla Pascal, è estranea al pensiero scientifico. Non è estranea, invece, alle attività operative ispirate al pragmatismo utilitaristico. Il soppesare guadagni e perdite è essenziale alle scelte il cui criterio guida è la massimizzazione dell'utilità attesa. È il criterio dello stratega, sia esso un giocatore, un *manager*, un militare. Non è, in genere, quello dello scienziato. Per-

ché una cosa è strategia, altra cosa è conoscenza: canoni diversi e diversi intenti.

E diverso è pure il significato delle parole, a cominciare dalla nozione di 'evento', che nella strategia decisionale è un 'particolare' e nella ricerca scientifica è un 'universale'. Non per questo, le probabilità della scienza si escludono alla valutazione dell'attesa di un evento singolo. Valga un esempio. L'incrocio sessuale di due individui eterozigoti per un carattere a due alleli, di cui uno dominante, può dar luogo a tre differenti genotipi — omozigote dominante, eterozigote, omozigote recessivo — con probabilità stabilite dalla prima legge di Mendel (una legge statistica): rispettivamente $1/4$, $1/2$, $1/4$. Sono le stesse probabilità dei tre eventi possibili nel lancio di due monete: gli alleli si combinano a caso come le 'teste' e le 'croci'. Così, a due coniugi microcitemici (eterozigoti per la talassemia) che si interrogano sulla possibilità che un loro nascituro sia omozigote sano — un evento unico — la genetica risponde assegnando a tale evento una probabilità: $p = 1/4 = 0,25$.

Una qualunque induzione diagnostica, per se stessa mirata all'individuale, può avvalersi di probabilità statistiche. Nelle rispettive scienze, esse hanno un valore oggettivo (nomico ed euristico), e tuttavia nelle scelte individuali possono anch'esse suggerire quote di scommessa. Nondimeno, l'idea di 'probabilità oggettive', sulle quali sono cresciute intere scienze della natura, appare, in casa soggettivistica, "superstiziosa", "pericolosa", "aberrante". "La probabilità non esiste", — è lo slogan degli odierni probabilisti soggettivisti.

Per superare i tanti dubbi semantici, per liberarsi dalle dispute definitorie, il 'calcolo delle probabilità', quale corpo di dottrina, si è sviluppato adducendo la probabilità come 'concetto primitivo' e formulando regole assiomatiche per dedurne — attraverso i principi di somma e di prodotto codificati da Laplace — le probabilità di eventi complessi, indipendentemente dai significati intesi e sottintesi e dai criteri adottati per tradurre un'attesa in un numero reale da zero a uno. Il tentativo più completo in questo senso è dovuto ad A.N. Kolmogorov (1933) (ma sono da ricordare i contributi di S. Bernstein, 1912, J.M. Keynes, 1921, H. Reichenbach, 1932): una 'teoria matematica della misura', dove la probabilità è un numero che soddisfa assiomi relativi a eventi intesi come sottoinsiemi di uno spazio a cui è assegnata una probabilità uguale a uno.

Ma, nel momento di dare un valore a una probabilità elementare, ritornano i problemi di contenuto e si ripropone l'antitesi di sempre. Eppure, non sembrano esistere, in concreto, probabilità oggettive e probabilità soggettive, ma piuttosto induzioni in cui prevalgono gli elementi oggettivi, seppur filtrati dalla soggettività di certe astrazioni, e decisioni in cui prevalgono gli elementi soggettivi, seppur tarati su dati oggettivi. E le scelte cambiano in ragione dei contesti e degli intenti; e dell'idea stessa di casualità: realtà fenomenica improntata alla variabilità naturale immanente, o espressione impropria di remote ragioni causali. Nella visione oggettivistica v'è un'istanza al valor semantico dei fatti, all'incontro, per così dire galileiano, della ragion critica con la ragion sperimentale.

tale. Nella visione soggettivista — quand'essa non cada in una sorta di esasperato solipsismo — s'avverte un recupero protagoreo e, più ancora, socratico dell'uomo come misura di tutte le cose.

Due modi di intendere il dialogo eterno tra *logos* ed *empeiria*, di giustapporre l'idea al dato, la ragione all'esperienza. E non deve sorprendere che in questi tempi di fine millennio, segnati dal trionfo dell'«economicismo», si pretenda di sostituire ovunque strategia a ricerca, scommessa a ipotesi. Ma il distinguo resta. E passa tra conoscenza e convenienza, tra verità e utilità. E non è soltanto una questione di misura.

(V. anche GIOCHI, TEORIA DEI; METODO E TECNICHE NELLE SCIENZE SOCIALI; PREVISIONE; **STATISTICA APPLICATA ALLE SCIENZE SOCIALI**).

Bibliografia

- BUNGE, M., *Four concepts of probability*, in "Applied mathematical modelling", V, 5, 1981.
 CASTELNUOVO, G., *Calcolo delle probabilità*, Bologna 1948.
 COSTANTINI, D., *Fondamenti del calcolo delle probabilità*, Milano 1970.
 DALL'AGLIO, G., *Calcolo delle probabilità*, Bologna 1987.
 FELLER, W., *An introduction to probability theory and its applications*, 2 voll., New York-London-Sydney 1965.
 FINETTI, B. DE, *Teoria delle probabilità*, 2 voll., Torino 1970.
 LINDLEY, D.V., *Introduction to probability and statistics from a Bayesian viewpoint*, 2 voll., Cambridge 1965.
 SCARDOV, I., *Ambiguous uses of probability*, in *Probability in the sciences*, Dordrecht-Boston-London 1988.

Significato della probabilità

SOMMARIO: 1. Introduzione. □ 2. Cenni storici. □ 3. La concezione classica della probabilità. □ 4. La definizione frequentista. □ 5. L'impostazione soggettiva. □ 6. Un confronto tra le diverse concezioni. □ 7. L'impostazione assiomatica. □ 8. La probabilità come logica dell'incerto. □ 9. La media. □ 10. Giochi d'azzardo e assicurazioni. □ 11. Il cammino della probabilità. □ 12. La probabilità e le scienze sociali. □ Bibliografia.

1. Introduzione.

La probabilità, con il calcolo delle probabilità a essa collegato, si è affermata come componente essenziale della ricerca scientifica e come strumento indispensabile per la scienza e per la tecnica, ed è entrata da tempo nel linguaggio e nell'uso comune. Il suo significato però, al di fuori del campo degli specialisti, resta vago e imprecisato; né potrebbe essere altrimenti date le varie sfaccettature che l'intuizione fa emergere e l'esperienza fa constatare.

Intuitivamente la probabilità si presenta come 'propensione' o 'attitudine' di un evento a verificarsi o manifestarsi. C'è poi il collegamento con la *frequenza (relativa) dei successi*, che è il rapporto, in una successione di prove, tra il numero delle prove in cui un certo risultato si presenta e il loro numero totale: ci si aspetta che a lungo andare la frequenza si avvicini alla probabilità.

Sono appunto gli aspetti intuitivi, e in particolare quelli citati, che hanno fornito la base di partenza per l'elaborazione del concetto di probabilità, del modo di arrivare alle valutazioni quantitative e delle corrispondenti regole di calcolo. L'elaborazione del concetto ha dato origine a diverse concezioni della probabilità, spesso con contrapposizioni polemiche.

Gli aspetti quantitativi sono in generale quelli che più interessano; è necessario perciò dire qualcosa sul calcolo delle probabilità. Esso si può definire come un ramo della matematica che studia il concetto intuitivo di probabilità, per approfondirlo e svilupparlo ai fini delle applicazioni. Si tratta di calcolare le probabilità del verificarsi di eventi complessi a partire da quelle del verificarsi di eventi più semplici, per le quali si pensa di poter dare direttamente una valutazione.

Un caso tipico è quello del gioco del lotto; si è largamente d'accordo che in ogni estrazione tutti i numeri presenti nell'urna hanno la stessa probabilità di essere estratti (1/90 per il primo estratto, 1/89 per il secondo estratto, dopo che dall'urna è stata tolta una pallina, e così via): su questa base si possono calcolare le probabilità che esca un certo terno, o che un numero accumuli un ritardo prefissato, e ancora, giocando ogni settimana su un ambo, che dopo cinque settimane si sia vinto almeno una volta, che dopo nove settimane si sia perso tutto il capitale inizialmente a disposizione, e così via.

Come valutare le probabilità di partenza? E quali regole di calcolo vanno seguite? Per quanto si è detto le risposte derivano dallo studio del significato della probabilità, che d'altra parte le risposte stesse contribuiscono a chiarire. Questa esposizione ha lo scopo di seguire tale percorso, arrivando anche alle nozioni più elementari di calcolo (per gli sviluppi matematici v. Daboni, 1970; v. Dall'Aglio, 1987).

2. Cenni storici.

Qualche cenno sulla nascita del calcolo delle probabilità, oltre che interessante in sé, può essere utile per illuminarne la natura (v. David, 1962; v. Hacking, 1975).

Nel mondo antico la probabilità era studiata nel suo aspetto filosofico, ma erano praticamente assenti le valutazioni quantitative. I pochi cenni a noi noti si trovano nel *Talmud*, dove si pone il problema se sia lecito mangiare della carne trovata per strada, non sapendo se è *kasher*, cioè conforme alle norme rituali, e lo si risolve dicendo che, se in città ci sono 9 negozi che vendono carne *kasher* e uno che vende carne impura, la carne trovata si può mangiare. Appare evidente un riferimento alla probabilità, e l'evidenza è rafforzata da altri passi, in particolare da uno che a proposito di oggetti utilizzati per il culto degli idoli pone limiti molto più severi. E, con riferimento ai sorteggi, usati ritualmente in determinate circostanze, vengono accennate alcune elaborazioni che, pur essendo poco chiare, sono indubbiamente dei tentativi di calcolo.

Lo studio effettivo del calcolo delle probabilità ebbe inizio solo con l'avvento del metodo sperimentale, che ponendo maggiore attenzione alla raccolta dei

dati permise di cogliere, attraverso le frequenze, delle regolarità e dei valori numerici nell'imprevedibilità dei risultati: ci si accorse insomma che si poteva misurare il caso.

L'origine del calcolo delle probabilità è legata al gioco dei dadi, e ciò è dovuto soprattutto alla semplicità, in questo gioco, del meccanismo aleatorio e delle regole di base, che possono essere immediatamente tradotte in termini matematici. Parecchi giochi permettono di costruire un modello matematico che risulta molto semplice e allo stesso tempo ben aderente alla realtà. Va detto anche che l'attenzione dei giocatori ai risultati ha sempre fornito un notevole campo di sperimentazione.

Naturalmente all'avvento dello studio della probabilità ha contribuito l'interesse per altri meno futili campi di applicazione, come le assicurazioni e più in generale l'economia, o l'attendibilità delle opinioni e la validità delle prove, in particolare nel campo teologico (si pensi a Pascal) e giuridico. Ma i primi tentativi non diedero, né potevano dare, risultati significativi per la complessità dei problemi affrontati, che in quella fase iniziale presentavano difficoltà insormontabili.

Gerolamo Cardano nel suo libro *De ludo aleae*, scritto forse intorno al 1526, studia il lancio di tre dadi calcolando la probabilità dei diversi risultati possibili. Lo stesso fa Galileo, con molta maggior chiarezza, in un suo scritto sul gioco dei dadi apparso intorno al 1620.

Ma la nascita del calcolo delle probabilità viene abitualmente attribuita all'interesse di Blaise Pascal e viene fissata nella corrispondenza tra lui e Pierre Fermat, iniziata nel 1654. L'attenzione di Pascal fu suscitata dal cavalier de Méré, un accanito giocatore con una buona conoscenza della matematica, che si lamentava che la matematica lo facesse perdere ai dadi, perché le frequenze dei risultati non corrispondevano alle probabilità da lui calcolate, sulle quali basava le sue scommesse. E soprattutto egli pose a Pascal il problema di come ripartire la posta in palio nel caso in cui due giocatori debbano interrompere il gioco avendo punteggi diversi. Questo problema era stato già molto studiato senza risultati convincenti, perché, come oggi si intuisce chiaramente, solo il calcolo delle probabilità poteva risolverlo. Fu su tale problema che si sviluppò la corrispondenza con Fermat di cui si è detto.

Ben presto altri matematici si interessarono a questi argomenti; tra i primi l'olandese Christiaan Huygens che con il suo scritto *De ratiociniis in ludo aleae* (1656) ne diede una prima trattazione organica. Seguirono poi i matematici svizzeri Bernoulli, tra cui Jakob, l'autore di *Ars conjectandi* (apparso postumo nel 1715) al cui nome è legato un fondamentale modello matematico. Con l'inglese Abraham de Moivre e la sua *Doctrine of chances* (1718) il calcolo delle probabilità consolida le sue basi elementari e raggiunge la sua affermazione come nuova disciplina.

Qualcuno contesta la paternità assegnata a Pascal. In effetti, come si è visto, egli non fu il primo in assoluto, né i suoi contributi si elevano al di sopra di quelli di alcuni suoi contemporanei, come Fermat e Huygens. Sembra però che a lui vada attribuito il

merito di aver intrapreso e stimolato uno studio organico dei problemi di probabilità. D'altra parte ciò corrisponde alla sua posizione filosofica, in particolare nei confronti di Cartesio, che lo precedeva di poche decine di anni. Come in filosofia alle "idee chiare e distinte" di Cartesio egli contrapponeva "le ragioni del cuore, che la ragione non comprende", così, in contrasto con la sistematica e stabile costruzione della geometria cartesiana, egli dava inizio, certo non consapevolmente, al processo dirompente del calcolo delle probabilità.

L'interesse di Pascal per la probabilità, motivato da ragioni che vanno ben al di là dei giochi d'azzardo, è confermato dagli studi del circolo di Port Royal, di cui egli fu autorevole esponente. Nel trattato *La logique, ou l'art de penser* (1662) viene introdotto il ragionamento probabilistico, anche dal punto di vista quantitativo. Conformemente a quanto si è detto, il paradigma è quello dei giochi, e l'applicazione ad altre situazioni (probabilità di morte per fulmine, validità di un contratto, attendibilità dei miracoli) presenta solo esercizi con valutazioni quantitative ipotetiche. A questi tentativi va collegata la famosa grande scommessa di Pascal sull'esistenza di Dio.

La posizione di Cartesio è alla base dello sviluppo deterministico della scienza, culminante nell'affermazione che, se si conosce con precisione lo stato dell'universo in un dato istante, si dovrebbe poter calcolare la sua evoluzione in tutti gli istanti successivi. La posizione di Pascal rappresenta invece la lontana origine della moderna concezione della scienza, specialmente della fisica, secondo cui tutto quello che possiamo fare è costruire dei modelli matematici che descrivono, con una validità approssimata e provvisoria, i fenomeni che osserviamo; e i modelli deterministici, sufficienti solo per una rappresentazione largamente approssimativa, vanno sostituiti, per una comprensione più approfondita, con modelli basati su leggi probabilistiche.

3. La concezione classica della probabilità.

L'episodio dell'incontro tra Pascal e de Méré permette di vedere come si andava configurando nei protagonisti del dibattito il significato di probabilità; va chiarito che qui se ne parla in termini moderni, diversi da quelli usati allora: in particolare nella fase iniziale degli studi si affermava il concetto di media piuttosto che quello di probabilità.

De Méré (senza enunciarlo esplicitamente) aveva calcolato la probabilità come rapporto tra il numero dei casi favorevoli e quello dei casi possibili; egli coglieva così uno degli aspetti intuitivi della probabilità, che oggi viene enunciato come *definizione classica*, specificando che i casi considerati devono essere ugualmente possibili.

È questa una definizione che presta il fianco a diverse critiche: intanto si riferisce a una situazione di **casi ugualmente possibili**, che non sempre si verifica (dove sono i casi ugualmente possibili in una partita di calcio? e allora non si può parlare di probabilità per i fatidici risultati 1, 2, X?), e poi il concetto di **uguale possibilità non è chiaro**; sembrerebbe più significativo parlare di uguale probabilità, ma ciò renderebbe tautologica la definizione.

Prescindendo dalla ricerca di un impossibile rigore assoluto, va tuttavia sottolineata l'efficacia di tale definizione, che permette di ricavare facilmente alcune regole per la probabilità. Essa è chiaramente un numero compreso tra 0 e 1, che sarà 0 se l'evento è impossibile (nessun caso favorevole all'evento) e 1 se l'evento è certo (tutti i casi possibili sono anche favorevoli). E poi si vede facilmente che, se due eventi sono incompatibili (cioè non si possono presentare insieme; in altre parole, non c'è nessun caso favorevole a entrambi), la probabilità che si presenti uno dei due è uguale alla somma delle probabilità dei singoli eventi. Insomma, indicando con $P(A)$ la probabilità di un evento A abbiamo:

$$\begin{aligned} 0 &\leq P(A) \leq 1 \\ P(\text{evento impossibile}) &= 0 \\ P(\text{evento certo}) &= 1 \\ \text{Se } A \text{ e } B \text{ sono incompatibili,} \\ P(A \text{ o } B) &= P(A) + P(B). \end{aligned} \quad (I)$$

Le leggi così enunciate hanno una portata che va molto al di là della loro semplicità: in effetti è praticamente su di esse che si basa l'intera teoria matematica della probabilità.

4. La definizione frequentista.

Una volta calcolata la probabilità, il cavalier de Méré la confrontava con i risultati che otteneva effettivamente nel gioco, aspettandosi che la *frequenza relativa* dei successi, cioè il rapporto tra il numero delle volte in cui l'evento si verificava e il numero delle prove effettuate, si avvicinasse alla probabilità calcolata. E Pascal, correggendo l'errore nel calcolo di de Méré, rafforzava questa aspettativa.

Aspettativa intuitivamente naturale, che veniva poi affermata esplicitamente con la *legge empirica del caso*, secondo cui in un gran numero di prove fatte nelle stesse condizioni la frequenza relativa dei successi si avvicina alla probabilità, e l'approssimazione in genere migliora con l'aumentare del numero delle prove. È un'affermazione necessariamente vaga, perché l'aleatorietà delle prove non permette di precisare entro quanto tempo ed entro quali limiti l'approssimazione è valida, restando sempre possibile, anche se con probabilità bassissima, che ad esempio 100 lanci di un dado diano tutti il 6.

La legge empirica del caso ha una controparte teorica, la *legge dei grandi numeri*: essa afferma che la frequenza dei successi converge in probabilità verso la probabilità p del singolo evento, dove la 'convergenza in probabilità' e le condizioni per la validità vengono chiarite mediante precise definizioni matematiche. Va altresì sottolineato che questo risultato rimane nell'ambito teorico, ed è quindi concettualmente e concretamente diverso dalla legge empirica del caso, che si riferisce a quanto avviene nell'attuazione pratica.

L'aspetto frequentista si presenta come un'altra faccia del concetto di probabilità, tanto che alcuni studiosi insoddisfatti dell'impostazione classica si basavano proprio su di esso per definire la probabilità come *il limite della frequenza relativa dei successi*, in una successione di prove fatte nelle stesse condizioni, o, più

approssimativamente, come la frequenza relativa dei successi in un gran numero di prove. In altre parole, accettando l'ipotesi, implicita nella legge empirica del caso, che la frequenza relativa si vada stabilizzando intorno a un certo numero, è proprio questo numero che va preso come probabilità, senza alcuna valutazione preventiva.

Si può vedere che la definizione frequentista dà luogo per la probabilità alle stesse leggi (I) già trovate sopra. In effetti la frequenza, attraverso la quale ora si definisce la probabilità, può essere vista come il rapporto tra il numero delle prove favorevoli e il numero totale delle prove, in stretta analogia con la definizione classica.

Anche la definizione frequentista ha aspetti criticabili, come la presenza del limite, che, mentre ha un chiaro significato per una successione definita matematicamente, diventa alquanto nebulosa per una serie, ovviamente finita, di osservazioni effettive. E soprattutto rimane vago il concetto di prove fatte nelle stesse condizioni. A rigore, anche in una situazione abbastanza appropriata come i lanci successivi di un dado, cambia, sia pure di poco, l'ambiente esterno (per esempio la pressione atmosferica) e cambia il dado, che per l'impatto nella caduta modifica, sia pure impercettibilmente, la sua forma.

Si tratta di variazioni che impediscono di dare una piena validità alla definizione teorica, ma sul piano pratico le differenze tra una prova e l'altra possono essere poco rilevanti. Ciò va visto anche in relazione alla notevole diversità delle situazioni in cui la definizione può avere concreta applicazione. Si pensi alle polizze di una compagnia di assicurazione, che possono essere considerate (e lo sono effettivamente nella teoria matematica) come prove ripetute, cosicché la probabilità che nell'anno una polizza registri un sinistro viene valutata mediante la frequenza registrata nel passato. La valutazione sarà tanto più valida quanto più omogenee sono le prove del gruppo considerato; intervengono qui complessi problemi che vengono affrontati dalla statistica.

5. L'impostazione soggettiva.

Restano però molte situazioni nelle quali l'evento di cui si vuole valutare la probabilità non può essere inserito in una successione di prove ripetute. Si pensi ancora a una partita di calcio: gli incontri del passato tra le due squadre ben difficilmente si possono ritenere come avvenuti nelle stesse condizioni. Questa difficoltà è rilevante anche nello studio dei fenomeni economico-sociali, e si impone quindi la considerazione dell'evento singolo a cui va riferita la probabilità. Da questa esigenza è nata la concezione soggettiva, sviluppata soprattutto da Bruno de Finetti.

In questa impostazione la probabilità è il grado di fiducia nel verificarsi dell'evento. Essa dipende allora dalla persona che la valuta, ma anche dal momento specifico, e in particolare dalle informazioni disponibili, dando così ragione della denominazione adottata.

È necessaria una definizione più precisa. Uno dei modi per arrivarci è attraverso lo studio delle scom-

messe, definendo la probabilità come il prezzo equo da pagare per ricevere 1 se l'evento si verifica (e niente nel caso contrario), con una qualsiasi unità di misura (una lira, un dollaro, un milione, ecc.).

Si pensi ad esempio a una lotteria tra amici, fatta mediante i numeri della tombola. Ognuno paga una certa quota, diciamo 1.000 lire, e chi vince (avendo il numero estratto) prende le 90.000 lire raccolte. Allora la probabilità secondo l'impostazione soggettiva risulta $1/90$, cioè il prezzo pagato, riferito al piatto come unità di misura: è proprio il valore che si ottiene dalla definizione classica, che in questo caso appare appropriata.

Si è parlato di 'prezzo equo', e questo richiede una precisazione che viene data dalla condizione di equità (si dice anche di *coerenza*): non si devono valutare le probabilità in modo tale che sia possibile ottenere una vincita certa o una perdita certa.

Chiariamo subito questo punto. Supponiamo di presentare una moneta dicendo che è truccata in modo tale che la probabilità $P(T)$ di ottenere 'testa' in un lancio è $1/2$ mentre quella di ottenere 'croce', $P(C)$, è $1/4$. Molti obietteranno subito che non può andare, perché la somma di $P(T)$ e $P(C)$ deve essere 1. È un'esigenza intuitiva, alla quale la condizione di coerenza dà una giustificazione convincente. Infatti facendo contemporaneamente due scommesse, una su T e una su C , si pagherebbe $1/2$ per la prima più $1/4$ per la seconda, ricevendo comunque 1, con un guadagno netto, non aleatorio, di $1/4$, in contraddizione con la condizione di coerenza. Questo perché la somma $P(T) + P(C)$ è minore di uno; se fosse maggiore si avrebbe una perdita certa: la condizione di coerenza impone che sia $P(T) + P(C) = 1$.

Questa relazione richiama la legge delle probabilità totali vista sopra, e in effetti sono ricavabili matematicamente l'una dall'altra. Insomma anche la definizione soggettiva porta alle stesse leggi (I) già derivate dalle altre definizioni.

Un breve inciso per osservare che in generale chi organizza il gioco (si pensi alle lotterie, al totocalcio, ecc.) trattiene una parte della somma raccolta, ed entro certi limiti è giusto: la condizione matematica di equità in queste condizioni sarebbe in realtà iniqua (su questo torneremo nel cap. 10).

Naturalmente anche la definizione soggettiva presenta dei problemi: uno in particolare è legato all'atteggiamento delle persone nei confronti del rischio, che a volte si cerca — come già si è osservato — a volte invece si tenta di evitare (con le assicurazioni). Questo porta delle complicazioni per il concetto di prezzo equo e in generale per il ricorso alle scommesse.

La critica più diffusa a questa impostazione è proprio di essere soggettiva, cioè di fondare la probabilità sull'opinione dei singoli, ovviamente variabile da persona a persona, e anche per la stessa persona a seconda delle circostanze in cui si trova e delle informazioni in suo possesso. Si è sviluppato un lungo confronto, spesso polemico, tra gli 'oggettivisti' che accusano l'impostazione soggettiva di rendere impossibile la comunicazione tra persone diverse, e i 'soggettivisti' che denunciano l'illusorietà della pretesa oggettività delle altre impostazioni.

6. Un confronto tra le diverse concezioni.

L'impostazione soggettiva è a mio avviso quella più convincente. La definizione classica fornisce un valore che non presenta alcun contenuto al di là della sua definizione numerica, e da qui deriva l'accusa di tautologia che le viene mossa. La definizione frequentista ha un significato sostanziale convincente, ma è troppo legata a una situazione di prove ripetute che spesso non si verifica. L'impostazione soggettiva ha un significato chiaro, e la giustificazione fornita dalla condizione di coerenza è convincente: non accettare le regole che ne derivano significa ammettere la possibilità di una scommessa certamente vincente, che annulla il concetto stesso di aleatorietà. Ovviamente la traduzione in termini monetari spesso non è possibile, ma in ogni campo, per esempio in quello scientifico, resta il concetto di fondo della condizione di coerenza: si deve escludere la possibilità di trasformare in certezza, con una opportuna combinazione, quello che certamente risente di una variabilità aleatoria.

Va sottolineato il fatto che un soggettivista per una seria valutazione delle probabilità deve tener conto delle informazioni in suo possesso: in particolare, se è il caso, delle frequenze del passato o delle condizioni di simmetria delle alternative possibili, ma senza accettarle acriticamente. D'altra parte anche l'oggettivista sa bene che l'uguale possibilità dei casi, o l'uguaglianza delle condizioni nella ripetizione delle prove, sono ipotesi che non possono sfuggire a una sua valutazione critica, necessariamente — almeno entro certi limiti — soggettiva. E, come abbiamo osservato, le leggi matematiche della probabilità sono comuni alle tre impostazioni.

Può sembrare a questo punto che la differenza tra i diversi punti di vista sfumi fino a diventare insensuale, ma in realtà rimane la profonda distanza concettuale. Essa si manifesta con notevoli differenze operative nell'*induzione statistica* e nella *teoria statistica delle decisioni*. Ma anche nelle valutazioni pratiche si coglie spesso la divergenza delle interpretazioni.

Per un giocatore regolare del lotto la concezione frequentista è soddisfacente: quello che gli interessa sono i risultati a lungo andare, e la frequenza in un gran numero di prove risponde alle sue esigenze. Ben diversa è la situazione di chi deve sottoporsi a un intervento chirurgico. Se egli sa che nel passato il 95% delle operazioni ha avuto pieno successo, può ragionevolmente aspettarsi che anche tra le prossime cento operazioni 95, più o meno, andranno bene; ma a lui interessa il suo intervento, non gli altri 99. Certamente la frequenza registrata nel passato gli dà anche per la sua — unica — operazione una certa fiducia, con una chiara interpretazione di tipo soggettivo. E questo grado di fiducia gioca un ruolo determinante nella decisione, quando questa è consentita.

7. L'impostazione assiomatica.

L'impostazione assiomatica, o ipotetico-deduttiva, è largamente preferita in ogni campo della matematica perché permette di ottenere un livello accettabile di rigore logico. Si parte da certe ipotesi, o assiomi, o postulati, che fissano le relazioni tra certi enti

astratti, non altrimenti definiti, e si sviluppano le conseguenze di tali postulati. La teoria così costruita può essere applicata ogni volta che nella situazione allo studio (sia essa astratta o concreta) si presume si siano verificate le relazioni espresse dai postulati.

Il calcolo delle probabilità non poteva sfuggire a questa esigenza di sistemazione. E il fatto già notato che le diverse impostazioni arrivano tutte alle stesse leggi matematiche espresse dalle formule (1) rende naturale prendere tali leggi come base per una costruzione assiomatica.

Molti studiosi in realtà hanno ritenuto opportuno dare alla teoria matematica una struttura più completa, integrando le leggi (1). Il sistema di postulati che ha avuto più fortuna è stato quello proposto nel 1933 da A. N. Kolmogorov, che struttura la probabilità in modo da inserirla nella *teoria della misura* dell'analisi matematica, permettendo di sfruttarne i metodi ben sviluppati e collaudati.

Alcuni, soprattutto tra i soggettivisti, contestano il sistema assiomatico di Kolmogorov, perché le ipotesi aggiunte non sono 'naturali' e restringono il campo di validità della teoria. C'è da dire però che le situazioni escluse con questa restrizione sono piuttosto artificiose e non rientrano nei problemi usualmente affrontati nelle applicazioni.

L'assiomatizzazione di Kolmogorov è stata perciò largamente, se non unanimemente, adottata, e ha permesso un poderoso sviluppo del calcolo delle probabilità, che veniva incontro ai problemi sempre più complicati posti dalla scienza, dalla tecnologia e dalle scienze sociali.

8. La probabilità come logica dell'incerto.

Il calcolo delle probabilità nelle sue diverse concezioni, e in particolare in quella soggettiva, può essere interpretato come *logica dell'incerto*.

Considerando equivalenti, in queste considerazioni sommarie, proposizioni (o enunciati) ed eventi (o fatti), la logica (classica) assegna a essi dei valori di verità che possono essere *vero* o *falso*; la probabilità invece introduce una graduazione tra i due estremi, con un valore p compreso tra 0 (= falso) e 1 (= vero), che misura, potremmo dire, la *quantità di verità* attribuita all'evento o, nell'impostazione soggettiva, il grado di fiducia nel verificarsi dell'evento. In tale impostazione appunto questo aspetto della probabilità viene meglio evidenziato, facendo derivare la probabilità dalla nostra ignoranza: quando non siamo in grado di sapere se un fatto è vero o falso, possiamo almeno cercare di precisare in che misura lo riteniamo vero.

Allora il calcolo delle probabilità si presenta ancora come lo studio delle relazioni tra valori di verità, ma, con un ampliamento rispetto alla logica, considera in generale i valori p invece dei soli valori 0 e 1.

Questo aspetto viene studiato dall'*impostazione logica* della probabilità che ne dà anche diverse formalizzazioni (v. Costantini, 1970).

Sul piano pratico il calcolo delle probabilità si propone come guida al ragionamento, e conseguentemente all'azione. È una pretesa ambiziosa, che però certamente ha un fondamento. Quando vediamo un

cielo nuvoloso e dobbiamo decidere se prendere l'ombrello, è certamente una valutazione di probabilità che, implicitamente e senza calcoli, ci muove nella scelta; la formalizzazione e le elaborazioni si propongono naturalmente in circostanze di diverso livello.

9. La media.

Alla probabilità è strettamente legato il concetto di media, che anzi, come si è già osservato, è emerso prima nella consapevolezza degli studiosi. Esaminiamolo anzitutto dal punto di vista soggettivo.

Come abbiamo visto, la probabilità $P(A)$ dell'evento A è il prezzo equo per ricevere 1 se l'evento A si verifica, e $x \cdot P(A)$ è il prezzo equo per ricevere un premio di valore x . Se vi è anche un altro premio di valore y , in corrispondenza a un evento B , e gli eventi A e B sono incompatibili, cioè non si possono presentare entrambi, sarà giusto pagare $x \cdot P(A) + y \cdot P(B)$ per aspirare a entrambi i premi.

Più in generale consideriamo una lotteria in cui si possono vincere più premi, di valore $x_1, x_2, \dots, x_n$, rispettivamente con probabilità $p_1, p_2, \dots, p_n$. La vincita è allora un *numero aleatorio*, o *variabile aleatoria*, che indicheremo con X , cioè un numero che per il momento è ignoto e si conoscerà solo dopo una prova dal risultato aleatorio, ma che comunque è ben definito, perché è descritto in modo univoco il procedimento per arrivarci. Il discorso fatto sopra equivale a dire che il prezzo equo da pagare per ricevere la variabile aleatoria X è $x_1 \cdot p_1 + x_2 \cdot p_2 + \dots + x_n \cdot p_n$; e questa è per definizione la *media* di X , che viene indicata con $M(X)$:

$$M(X) = x_1 p_1 + x_2 p_2 + \dots + x_n p_n. \quad (2)$$

La media di una variabile aleatoria X è quindi, nell'impostazione soggettiva, il prezzo equo da pagare per ricevere in cambio la variabile aleatoria X . Si dice anche che la media è l'equivalente certo del valore aleatorio X ; ma l'equivalenza va presa con molta cautela, perché un valore aleatorio non può mai essere del tutto assimilato a un valore certo (è questo un aspetto già notato, sul quale ritorneremo nel cap. 10).

Vale la pena di osservare che la probabilità si presenta ora come un caso particolare della media. Una scommessa sull'evento A equivale a pagare per ricevere in cambio una variabile aleatoria che può assumere i valori 1 e 0 rispettivamente con probabilità $P(A)$ e $1 - P(A)$; e la sua media risulta $1 \cdot P(A) + 0 \cdot [1 - P(A)] = P(A)$.

Anche nell'impostazione classica (e, in modo analogo, in quella frequentista) si può dare un significato alla media generalizzando il concetto di probabilità di un evento A , che è il rapporto tra il numero n_A dei casi favorevoli e il numero totale dei casi n . A ciascuno dei casi assegniamo il numero 1 se il caso è favorevole ad A , e il numero 0 in caso contrario: abbiamo così n numeri $x_1, \dots, x_n$ che caratterizzano i diversi casi possibili; allora esattamente n_A degli n numeri x sono uguali a 1 (e gli altri a 0), cosicché $n_A = x_1 + \dots + x_n$. Di conseguenza

$$P(A) = \frac{n_A}{n} = \frac{x_1 + x_2 + \dots + x_n}{n}. \quad (3)$$

Questa formula è quella, ben nota, della media aritmetica, e può sorprendere vederla ora presentata per la probabilità. In realtà ciò mette in evidenza il fatto già osservato che la probabilità è un caso particolare della media. Nella formula sopra i numeri x possono essere 1 o 0 (caso favorevole e caso contrario); se invece attribuiamo agli x dei valori qualsiasi otteniamo la media, cioè:

$$M(X) = \frac{x_1 + x_2 + \dots + x_n}{n} \quad (4)$$

Se ad esempio i valori $x_1, \dots, x_n$ sono le altezze di n persone, la formula (4) dà l'altezza media. Quando invece interessa sapere solo se l'altezza è superiore, diciamo, a 170 cm, possiamo assegnare a ciascuna persona il numero $x = 1$ se la sua altezza è superiore a tale valore, e il numero $x = 0$ altrimenti; la stessa formula fornisce allora il rapporto tra il numero dei casi favorevoli e quello dei casi possibili, cioè la probabilità che una persona scelta a caso nel gruppo abbia l'altezza desiderata.

10. Giochi d'azzardo e assicurazioni.

Non sembri strano l'accostamento nel titolo: i trattati sulle assicurazioni specificano che vengono prese in esame le attività organizzate e offerte al pubblico, mentre i contratti isolati, in uso fin dai tempi antichi, vanno piuttosto considerati come scommesse.

Dal punto di vista matematico un gruppo di polizze differisce ben poco da una serie di puntate alla roulette: c'è nelle assicurazioni una maggiore complessità degli eventi e una maggiore difficoltà nella valutazione delle probabilità. Anche la terminologia viene a volte mutuata dai giochi: è ampiamente sviluppata, ad esempio, la *teoria della rovina del giocatore* che diventa per le assicurazioni lo studio del rischio di fallimento.

Un'altra importante caratteristica comune è che, come già notato più volte, il giocatore nel primo caso e l'assicurato nel secondo pagano, per partecipare, più del 'prezzo equo' matematico. Nella teoria matematica delle assicurazioni ciò viene messo in evidenza calcolando anzitutto un *premio puro* che corrisponde al 'prezzo equo' matematico e aggiungendo un *caricamento* per tener conto dei costi accessori. Un'analogia distinzione può essere fatta per i giochi d'azzardo. Così nel gioco del lotto la vincita per l'ambo è 250 volte la posta giocata, mentre dovrebbe essere 400,5 volte, essendo 1/400,5 la probabilità di un ambo; si riceve perciò il 62% di quanto sarebbe 'equo'. Questo è, possiamo dire, il 'rendimento medio' della somma giocata: a lungo andare il giocatore si deve aspettare di ricavare dalle vincite il 62% di quanto ha speso (perdendo quindi il 38%).

La distinzione tra i due casi non sta negli aspetti quantitativi, ma nella motivazione. L'assicurato paga per avere la garanzia di poter far fronte a possibili imprevisti che rischierebbero di squilibrare le sue finanze. Nel gioco il costo aggiuntivo è motivato, oltre che dalla passione, dalla speranza di una notevole vincita, che permetterebbe di realizzare aspirazioni altrimenti inattuabili. Questo aspetto viene sfruttato nella gestione del lotto con una forte progressività: il 'rendi-

mento medio', pari al 62% per l'ambo, scende al 36% per il terno, e addirittura al 2% per la cinquina, che paga un milione di volte la posta invece di 44 milioni circa (v. Costantini e Monari, 1996).

Nel gioco interviene pure l'illusione di sconfiggere il caso mediante 'sistemi' sicuramente vincenti. È un'illusione che ha degli appigli reali, come nella strategia del raddoppio: alla roulette puntando successivamente 1, 2, 4, 8 lire ecc. sul rosso, fino a che il rosso non esce, si arriverà con probabilità uguale a 1 a vincere, cioè, tenendo conto delle puntate perse fino a quel momento, a guadagnare una lira. Ma questo risultato è solo teorico, e si raggiungerebbe giocando senza limitazioni di tempo e di denaro: tenendo conto di queste limitazioni, che possono intervenire a livelli bassi o altissimi ma sono comunque inevitabili, un gioco inizialmente 'equo' rimane equo e un gioco 'sfavorevole' rimane sfavorevole.

11. Il cammino della probabilità.

Non si vuole qui esporre l'evoluzione matematica dello studio della probabilità, che dopo la storia delle origini diverrebbe rapidamente inaccessibile ai non specialisti, ma soltanto sottolineare la diffusione che esso ha raggiunto nelle applicazioni e accennare ai passi più significativi compiuti nelle varie direzioni. Si può dire che in ogni campo del sapere e dell'agire la probabilità è entrata a pieno titolo o se ne esplora con successo l'utilizzazione.

Come si è già osservato, il calcolo delle probabilità si è sviluppato con riferimento ai giochi d'azzardo, per la semplicità e naturalezza dei loro aspetti quantitativi. Gli strumenti iniziali del calcolo delle probabilità erano troppo rudimentali per affrontare i complessi problemi posti dalle applicazioni, e i dati disponibili erano in generale insufficienti, soprattutto mancavano i concetti e i procedimenti della statistica, passaggio obbligato per collegare la teoria matematica alle osservazioni: si può dire che essa è lo 'strumento di misura' per una grandezza atipica come la probabilità.

La statistica ha come suo primo scopo la descrizione dei fenomeni quantitativi; ha poi una seconda parte, l'*induzione o inferenza statistica*, che costituisce lo strumento induttivo più efficace. Tipicamente essa cerca di ottenere informazioni su una *popolazione* (intesa nel senso più ampio di insieme di unità) attraverso un *campione* estratto dalla popolazione stessa; più in generale tende a risalire dal particolare al generale, dalle osservazioni (sperimentali o sul campo) alle leggi. Il complesso dei procedimenti si basa in modo essenziale sul calcolo delle probabilità, e ha raggiunto considerevoli sviluppi e una notevole raffinatezza, arrivando di conseguenza a un'enorme diffusione nelle varie scienze, e in particolare nelle scienze sociali. Lo sfruttamento della probabilità permette così in ogni campo valutazioni e stime più attendibili, possiamo dire 'più certe'.

Veniamo ad alcuni settori di applicazione specifici. I primi interventi significativi si ebbero già alla fine del XVII secolo nel campo assicurativo, dove la situazione era più favorevole: cominciavano a essere disponibili in Inghilterra dati sulla mortalità e, so-

prattutto, le frequenze osservate permettevano di valutare direttamente, senza complicati procedimenti di là da venire, le probabilità di sopravvivenza alle diverse età, mediante le quali calcolare il valore medio di una rendita vitalizia.

È interessante osservare che a questi studi contribuirono diversi astronomi, tra i primi Halley, che tuttavia erano ben lontani dal collegare la probabilità ai dati astronomici già disponibili.

In effetti tra lo studio nascente della probabilità e la statistica che muoveva anch'essa i primi passi non si scorgevano ancora i legami che oggi appaiono inscindibili. Avendo per la stessa grandezza diverse misurazioni, Halley prendeva come misura più valida la media aritmetica, con vaghe giustificazioni. Oggi si prende in considerazione un ipotetico valore 'vero', da cui ogni misurazione si discosta per variazioni accidentali, a carattere aleatorio; i procedimenti dell'induzione statistica indicano la media aritmetica come la 'migliore' stima del valore 'vero'.

Il calcolo delle probabilità fece il suo ingresso nella fisica più tardi, ma con maggiore incisività. Infatti col tempo ci si rese conto che l'aleatorietà entrava non solo nelle osservazioni ma nello stesso processo fisico, e che le leggi classiche, che enunciavano risultati deterministici, erano inadeguate in molti casi a rappresentare i fenomeni studiati. Si arrivò così a formulare dei modelli aleatori, nei quali le leggi prevedono non già un valore fisso, ma un numero aleatorio con la sua distribuzione di probabilità. Nella seconda metà del secolo scorso Ludwig Boltzmann rivoluzionava lo studio dei gas con la teoria cinetica basata sulla probabilità e arrivava ad affermare che il linguaggio della fisica è il calcolo delle probabilità.

Più o meno allo stesso periodo appartengono le ben note ricerche di G. Mendel, che con lo studio probabilistico della trasmissione ereditaria dei caratteri dava inizio alla genetica moderna. Una situazione in cui risulta più chiaro, ed evidentemente inevitabile, l'uso di una legge aleatoria: il numero di discendenti con determinate caratteristiche previsto dalla teoria non è deterministico, ma è un numero aleatorio, di cui la teoria fornisce la distribuzione di probabilità.

I progressi nelle scienze ebbero naturalmente una profonda influenza sulla tecnologia. Le operazioni più complesse, come il lancio di un satellite, vengono ora programmate e controllate con l'aiuto del calcolo delle probabilità. È ampiamente sviluppata la teoria dell'affidabilità che valuta quanto si può contare sul corretto funzionamento di un macchinario, tenendo conto dell'aleatorietà che incide sul funzionamento dei vari pezzi.

La teoria delle file d'attesa studia, oltre alle comuni code agli sportelli, l'arrivo di 'clienti' dei tipi più diversi: chiamate a un centralino telefonico (è qui che la teoria è stata più sviluppata), aerei a una pista di atterraggio, pazienti a un posto di pronto soccorso, ecc. Vengono studiati la lunghezza della coda e il tempo d'attesa per i clienti (ovviamente aleatori), dando la possibilità di dimensionare le strutture in relazione ai costi e all'urgenza del servizio.

Ricordiamo anche i metodi statistici del controllo di qualità che fanno ormai parte integrante di ogni processo produttivo di una certa dimensione.

Vale la pena di citare anche i procedimenti di simulazione. Quando un processo a carattere aleatorio, naturale o organizzato, è troppo complesso per la sua trattazione matematica, lo si può studiare cercando di riprodurlo con meccanismi aleatori artificiali (mediante computer) e stimando poi le probabilità desiderate attraverso la frequenza in un gran numero di prove. Un tale strumento di lavoro si rivela prezioso in molte circostanze, per esempio per prevedere le conseguenze di interventi sul traffico; un esempio più insolito è uno studio sul restauro condotto agli Uffizi.

Ma l'utilizzazione della simulazione si è estesa anche alla soluzione di problemi deterministici. Il primo esempio sembra dovuto al naturalista George Louis Buffon, il quale calcolò che, lanciando a caso un ago su un foglio su cui sono tracciate delle rette parallele a distanza uguale alla lunghezza dell'ago, la probabilità che l'ago intersechi una retta è uguale a $2/\pi$. Allora se si lancia l'ago un gran numero di volte la frequenza f dei successi sarà vicina alla probabilità $2/\pi$, cosicché $2/f$ fornisce un valore approssimato di π . Questo procedimento, originale ma senza pratica utilità in questo esempio, attualmente viene ampiamente utilizzato nel calcolo numerico per ottenere il valore di complicate espressioni matematiche difficilmente raggiungibile per altra via. Il calcolo delle probabilità appare così utile anche per risolvere problemi che in sé non hanno nulla di aleatorio.

La probabilità è entrata anche in settori più lontani. Diversi musicisti del XVIII secolo, fra cui Bach e Mozart, si sono divertiti a ottenere composizioni associando a caso, con un dado, diversi frammenti preparati in precedenza. Il procedimento si è sviluppato oggi nell'ambito della musica elettronica.

Il matematico A. A. Marcov all'inizio di questo secolo, applicando alcuni suoi originali risultati di calcolo delle probabilità, analizzò statisticamente il romanzo *Eugenij Onegin* di Puškin. I suoi procedimenti furono il punto di partenza per la statistica linguistica, ma non mancano esempi precedenti, già nel XIX secolo. Metodi statistici e probabilistici sono alla base dei procedimenti per l'analisi della voce che compaiono, a fini polizieschi, nel romanzo *Il primo cerchio* di A. Solženicyn e che oggi vengono sviluppati per la trascrizione automatica dei testi parlati.

12. La probabilità e le scienze sociali.

Le scienze sociali sono un campo in cui si può seguire bene il processo che ha portato faticosamente all'odierno massiccio impiego delle probabilità e della statistica partendo dai primi tentativi di applicazione che si limitavano a esprimere un'esigenza (v. Hacking, 1990; v. Stigler, 1986).

A fatti sociali si riferivano già, come si è detto, gli studi di Port Royal. Seguono numerosi riferimenti sporadici: tra i più noti il tentativo di John Arbuthnot, all'inizio del XVIII secolo, di dimostrare l'intervento della divina provvidenza attraverso la regolarità del rapporto tra i sessi alla nascita. Ma dopo gli inizi concreti nel campo demografico, già citati a proposito delle assicurazioni, bisogna aspettare più di un secolo per progressi sostanziali.

Alla fine del XVIII secolo veniva affermata l'esigenza dello studio quantitativo dei fatti sociali, e in particolare da Condorcet, che teorizzò una *matematica sociale* (preferendo questo termine a *morale* e *politico* allora più in uso) che includeva la probabilità.

Il XIX secolo, specialmente nella prima metà, fu determinante. Si sviluppò anzitutto un'ampia discussione sul concetto e sulla possibilità stessa di 'leggi' per i fatti sociali. Esse erano viste in analogia con le scienze fisiche: "I fatti di ordine sociale sono soggetti, come quelli di ordine fisico, a leggi invariabili", affermava André-Michel Guerry, studioso di criminologia, e Adolphe Quételet (nella seconda, più ampia, edizione) dava il titolo di *Physique sociale* al suo lavoro fondamentale.

Questo accostamento inseriva le scienze sociali nel generale processo per cui nelle scienze fisiche e naturali si passava dai modelli deterministici ai modelli probabilistici: nella fisica e nella nascente genetica, come abbiamo visto, ma anche in campi vicini, come la psicologia.

L'interesse per i fatti sociali non si limitava alle discussioni teoriche. I dati che cominciavano a essere disponibili con una certa larghezza davano luogo a numerosi studi, anzitutto sulle cause di morte e sulle malattie: l'impulso veniva ancora una volta dalle assicurazioni, all'inizio a carattere cooperativo e poi con i primi interventi pubblici. Particolare attenzione fu rivolta ai suicidi, per i quali Guerry propose, e in parte ottenne, la rilevazione di tutti i pensabili fattori sociali, fisici e ambientali. Naturalmente fu investito anche il vicino campo della criminologia, con l'intento, che rimase un'aspirazione, che questi studi dovessero servire a orientare la legislazione.

Il protagonista di questo processo fu Adolphe Quételet, lo studioso che pose le basi, a un tempo, della statistica e della sua applicazione, in particolare alle scienze sociali. Dall'astronomia, suo primo e perdurante interesse, egli mutuò il concetto di *curva normale o degli errori*, per descrivere, e in qualche modo spiegare attraverso le motivazioni fornite dalla teoria delle probabilità, la variabilità dei fenomeni. Fu sempre Quételet a dare l'avvio all'analisi statistica dei dati, al di là delle primissime elaborazioni già in uso, arrivando anche ai primi passi dell'induzione statistica.

Egli si occupò anche concretamente dei vari settori menzionati sopra, specialmente della demografia. In particolare diede inizio allo studio dell'antropometria, nella quale introdusse il concetto di *uomo medio*, da allora oggetto di studi e controversie.

A questo punto l'avvio dello studio quantitativo delle scienze sociali può dirsi consolidato. I passi successivi, di enorme portata, furono legati allo sviluppo della metodologia statistico-probabilistica, e non è il caso di seguirli.

Un discorso a parte merita il campo giuridico, che fu oggetto di attenzione fin dall'inizio. Nel XIX secolo si sviluppò una discussione sulla probabilità di errore nel verdetto di una giuria, legato alla specifica maggioranza richiesta. Ma restavano discorsi astratti. Il passaggio all'effettiva applicazione richiedeva che il concetto di probabilità e il relativo calcolo fossero accettati dalla sensibilità comune, e ciò è avvenuto molto più recentemente.

Oggi molti giornali presentano, alla vigilia degli incontri di calcio, le probabilità valutate da esperti per i tre risultati della schedina. I sondaggi e le 'proiezioni' sono sempre più effettuati e seguiti, e c'è sostanzialmente la consapevolezza di dover rinunciare spesso a una irraggiungibile certezza per accettare un valore di probabilità. Quando, a proposito del disastro della navetta spaziale *Challenger* del 1986, si lesse sui giornali che era venuto meno un pezzo la cui probabilità di tenuta era del 90%, nessuno si scandalizzò, come sarebbe accaduto in altri tempi, che la vita umana fosse affidata alla probabilità; fu solo criticato, giustamente, il livello troppo basso di cui ci si era accontentati.

Bibliografia

- COSTANTINI, D., *Fondamenti del calcolo delle probabilità*, Milano 1970.
- COSTANTINI, D., MONARI, P. (a cura di), *Probabilità e giochi d'azzardo*, Padova 1996.
- DABONI, L., *Calcolo delle probabilità e statistica matematica*, Torino 1970.
- DALL'AGLIO, G., *Calcolo delle probabilità*, Bologna 1987.
- DALL'AGLIO, G., *La probabilità nei tribunali*, in "Archimede", 1996, II, n. 3.
- DAVID, F., *Games, gods, gambling*, Huntington, W. Va., 1962.
- FINETTI, B. DE, *Teoria della probabilità*, Torino 1970.
- HACKING, I., *The emergence of probability*, Cambridge 1975 (tr. it.: *L'emergenza della probabilità*, Milano 1987).
- HACKING, I., *The taming of chance*, Cambridge 1990 (tr. it.: *Il caso domato*, Milano 1995).
- STIGLER, S.M., *The history of statistics*, Cambridge, Mass., 1986.
- TARUFFO, M., *La prova dei fatti giuridici. Nozioni generali*, Milano 1992, pp. 166-215.