Università
Ca' Foscari
Venezia

Corso di Laurea magistrale
in Scienze filosofiche

Tesi di Laurea

Calcolo delle Probabilità. Interpretazioni e significati tra scienza e scommessa su Dio

Relatore
Ch. Prof. Paolo Pagani

Correlatore
Ch. Prof. Enrico Jabara

Laureando
Andrea Piccolo
Matricola 871909

Anno Accademico
2019/2020

I numeri non sono significativi che all'interno di un quadro teorico. In mancanza di questo quadro, il loro accumulo illumina il pensiero meno di quanto non lo schiacci.

Olivier Rey

Non può mancare un pensiero e un ringraziamento a mia moglie Lidia, che ha condiviso la fatica e le soddisfazioni di questi anni di studio.

Elisabetta, Matteo, Francesca e Gabriele hanno vivacemente partecipato all'impresa, spronando e incoraggiando il papà-studente.

Infine, ma all'inizio di tutto, i miei genitori.

1. Sommario

1. Teorie della probabilità

La probabilità, il calcolo delle probabilità e un ampio impiego di metodi statistici sono all'ordine del giorno in ogni ambito della nostra vita. Dalle attività quotidiane più consuete alle procedure industriali più tecnologicamente avanzate, dalle dinamiche relazionali sociali alle sofisticate strategie di politica internazionale, che si tratti di un ambito tecnico scientifico come l'obsolescenza programmata degli apparati tecnologici, o creativo come il marketing pubblicitario, i numeri del calcolo statistico-probabilistico emergono ovunque; non si tratta però del numero che del mondo rivela l'essenza. Il numero che è in questione nel calcolo della probabilità non svela i caratteri nascosti della realtà che indaga.

Nella storia del pensiero occidentale, da Pitagora a Galileo, il numero e la geometria rappresentano a volte il fondamento, a volte l'archetipo o l'alfabeto usato per costituire il cosmo nel suo ordinamento, ma sempre con una corrispondenza reale, anche se da scoprire, tra numero e mondo. Quel pensiero che ha portato alla nascita e al fiorire delle scienze positive, indicando a lungo la matematica come il linguaggio capace di cogliere la realtà per quello che è, si trova oggi alle prese con un mondo più imprevedibile e indescrivibile di quel che sarebbe stato disposto ad ammettere. La *scienza* usa ancora il linguaggio dei numeri, ma questi non pretendono di descrivere più puntualmente il mondo fisico e le sue dinamiche[1], quanto piuttosto la confidenza che le cose stiano in un certo modo, che la descrizione che diamo del mondo sia approssimativamente corretta, si mantenga entro certi margini di errore e si muova verso una direzione individuata in modo ragionevolmente verosimile.

[1] Questo è generalmente vero anche per le scienze positive, o scienze esatte, con la sola eccezione della matematica, che nella sua veste *non applicata* continua ad avere sempre lo stesso rigore di forma e precisione puntuale, senza più alcuna pretesa di parlare di realtà mondane. La diagonale del quadrato unitario e il rapporto tra circonferenza e diametro continuano ad essere numeri irrazionali, calcolabili con precisione assoluta fino a un qualunque numero di cifre decimali finito, se non si hanno problemi di tempo, ma non sanno dire più nulla di cosa sia la realtà in se stessa, sensibile o soprasensibile che sia.

Gli oggetti matematici, sempre più complessi e articolati nelle loro strutture formali, sono semplicemente modelli, abilmente selezionati o costruiti, per descrivere nel modo migliore possibile il mondo e interagire con esso per mezzo degli strumenti che la matematica mette a disposizione: quelli del calcolo. Se il modello non si mostra più adeguato o si vuole instaurare una diversa interazione con il mondo, si cerca un nuovo modello e si abbandona il precedente. Questo progresso che una volta rappresentava il procedere nella comprensione del mondo così com'è, oggi non rappresenta necessariamente un perfezionamento della conoscenza: con l'avvento delle scienze statistiche e la loro penetrazione in ogni disciplina e tecnica, si sancisce la rinuncia a una conoscenza puntuale in favore di un'approssimazione più o meno grossolana per la quale ci si accontenta strumentalmente di una certa capacità predittiva dei fenomeni. Non si tratta qui di recuperare la distinzione e separazione kantiana tra fenomeno e noumeno, accessibile il primo e inconoscibile il secondo. L'approccio statistico alla conoscenza ha sfumato nei contorni e progressivamente eroso proprio la conoscibilità del fenomeno che Kant indicava come unico possibile oggetto di conoscenza.

Non è comunque Kant il riferimento cui pensare, occupandosi di scienze statistiche, quanto piuttosto il determinismo di matrice positivista; e mi sembra particolarmente indicativo di questa cesura tra conoscenza e realtà conosciuta il fatto che Domenico Costantini e Ludovico Geymonat aprano il loro saggio sulla probabilità proprio facendo chiarezza su quello che gli autori chiamano il dogmatismo laplaciano:

> La pretesa di definire esplicitamente la probabilità segue da un approccio dogmatico ai fondamenti delle discipline statistico-probabilistiche – il dogmatismo laplaciano – consistente nel ritenere che la probabilità abbia un solo e ben definito significato, mente tutte le altre interpretazioni che possono essere e di fatto sono state attribuite a questo termine siano false. Come vedremo, la prima massiccia comparsa di questo dogmatismo può essere individuata nel tentativo, per altro pienamente riuscito, di derivare la teoria classica delle probabilità dalla definizione di probabilità. Avvertiamo però che un atteggiamento simile si riscontra anche in autori precedenti e che la posizione di Laplace non è ancora pienamente dogmatica, mentre sarà solo nel nostro secolo che il dogmatismo laplaciano diventerà, con pochissime eccezioni, il modo privilegiato di guardare

> ai fondamenti delle discipline statistico-probabilistiche.[2]

In particolare qui ci interessa osservare che, indipendentemente dal fatto che si tratti di una consapevolezza riflessa o implicita, in tutti i casi in cui si ricorre al calcolo delle probabilità e alla scienza statistica, facilmente si presuppone una lettura del mondo sostanzialmente deterministica, e Laplace nel suo *Essai philosophique sur les probablités* sposa chiaramente questa posizione:

> Dobbiamo dunque considerare lo stato presente dell'universo, come l'effetto del suo stato anteriore, e come la causa di quello che subito segue. Una intelligenza che, a un dato istante, conoscesse tutte le forze da cui la natura è animata, e la collocazione rispettiva degli esseri che la compongono, se d'altra parte fosse abbastanza vasta per sottoporre questi dati all'analisi, comprenderebbe nella stessa formula i movimenti dei più grandi corpi dell'universo e quelli del più leggero atomo: nulla sarebbe incerto per essa, e l'avvenire come il passato, sarebbe presente ai suoi occhi.[3]

Ma questa intelligenza in grado di analizzare la natura in modo onnicomprensivo non c'è, o per lo meno non è appannaggio dell'uomo, così si ricorre alla probabilità per maneggiare questa natura altrimenti ingestibile: «La curva descritta da una semplice molecola d'aria o di vapore, è regolata in modo altrettanto certo, quanto le orbite dei pianeti: non ci sono differenze tra esse, se non quel che vi pone la nostra ignoranza. La probabilità è relativa in parte a questa ignoranza, in parte alle nostre conoscenze»[4]

Questa lettura deterministica non è però l'unica possibile, e lo stesso Laplace

---

[2] D. Costantini e L. Geymonat, *Filosofia della probabilità*, Feltrinelli, Milano 1982, p. 19
[3] P.S. Laplace, *Essai Philosophique Sur Les Probabilités*, Bachelier, Imprimeur-Libraire, Paris 1840, pp. 3-4. Traduzione mia.
[4] *Ivi*, pp. 6-7. Traduzione mia.

quando definisce i principi che stanno alla base del calcolo delle probabilità, sembra costretto a far rientrare dalla finestra il problema che pensava di aver messo alla porta, e si possono sollevare molte fondate obiezioni sul fatto che egli sia riuscito a definire la probabilità senza fare ricorso al concetto stesso di probabilità, come si vedrà più avanti.

In particolare, quando analizzeremo il *pari* di pascal sarà indispensabile distinguere tra la concezione usuale contemporanea, prettamente deterministica, che fa ricorso alla probabilità come ad uno strattagemma per gestire la parte inafferrabile delle regole che governano il mondo, ed altre concezioni possibili.

Va detto che in ambito accademico esistono riflessioni e confronti anche vivaci sul significato teoretico della probabilità e le sue implicazioni sulla filosofia della scienza, ma oggi come oggi sono generalmente circoscritti a una nicchia piuttosto ristretta della ricerca filosofica, e l'uso ovunque diffuso della statistica fa riferimento, per lo più in maniera implicita, a una concezione meccanicistica e determinista della probabilità.

Per questo mi pare importante premettere all'analisi del *pari* una rapida carrellata storica sul calcolo delle probabilità e sulla statistica, che non vuole essere una ricerca storica, e non ne avrà pertanto la corrispondente dettagliata profondità e ampiezza, ma ci serve per fare una sorta di inventario della terminologia, dei concetti fondamentali e di come si sono sviluppati, per assicurarci un bagaglio minimo che ci consenta di esaminare il *pari* senza incasellarlo in modo preconcetto all'interno di categorie che non rappresentano l'intero orizzonte possibile quando si parla di probabilità.

2. Origini storiche

Storicamente il calcolo delle probabilità viene fatto nascere a metà del secolo XVII, quando lo scrittore francese Antoine Gombaud, noto come Cavaliere De Méré, accanito scommettitore nei salotti, chiese aiuto all'amico Blaise Pascal per trovare strategie favorevoli nel gioco dei dadi. Dalle richieste e problemi posti da De Méré, nel 1654 nacque tra questi e Pierre de Fermat un carteggio dove di fatto si trova il primo abbozzo di calcolo delle probabilità.

L’argomento affrontato da Pascal è così originale nella formulazione e nel modo in cui viene affrontato, che Ian Hacking intitola il saggio in cui espone la storia di questa disciplina *L’emergenza della probabilità*[5], a sottolineare proprio come in tutto il panorama scientifico fino a quel momento lo studio delle probabilità fosse assente, anche come problema irrisolto. A riprova di ciò, Hacking cita in apertura del suo saggio l’opera di Isaac Todhunter, *A History of the Mathematical Theory or Possibility from the time of Pascal to that of Laplace*, del 1865, che raccoglie meticolosamente i lavori significativi sulla probabilità che erano stati pubblicati tra il 1654 e il 1822, da Pascal a Laplace. Il lavoro di Todhunter è effettivamente dettagliato e cita di ogni autore le opere e in ogni opera i passi che affrontano argomenti probabilistici, la descrizione dei problemi esposti e come viene impostata la loro soluzione con dovizia di esposizione anche di formule e calcoli. Con una osservazione molto pertinente Hacking fa notare che delle oltre 600 pagine dell’opera di Todhunter, quelle dedicate agli autori che precedono Pascal sono solo 6, per sottolineare il fatto che non esisteva un calcolo delle probabilità prima di Pascal. Curioso il fatto che il primo autore menzionato da Todhunter sia Dante Alighieri, i cui versi iniziali del sesto Canto del *Purgatorio* rappresentano «la più antica indicazione di differenti probabilità dei vari lanci che possono essere fatti con tre dadi»[6].

Rispetto all’osservazione di Hacking, è il caso di dire che se Todhunter non poteva trovare materiale per la sua raccolta quasi bibliografica prima di Pascal, ciò vale in primo luogo perché della probabilità il solo punto di vista che gli interessasse era quello meramente matematico. Anche la menzione di Dante, con una chiave di lettura dell’*incipit* del sesto canto del *Purgatorio*[7] che difficilmente il sommo poeta

---

[5] Hacking, *L'emergenza della probabilità*, trad.it. M. Piccone, Il Saggiatore, Milano 1987.

[6] Todhunter, *A History of the Mathematical Theory or Possibility from the time of Pascal to that of Laplace*, Macmillan and CO., Cambridge 1865, p 1. traduzione mia. Todhunter fa riferimento al lavoro di Guillaume Libri che, nel suo *Histoire des Sciences Mathématiques en Italie* interpreta in questa chiave un commentario alla *Divina commedia* pubblicato a Venezia nel 1477.

[7] I primi tre versi del Canto VI del *Purgatorio* recitano: «Quando si parte il gioco de la zara,/ colui che perde si riman dolente,/ repetendo le volte, e tristo impara.

avrebbe condiviso, mostra come a metà del XIX secolo la probabilità fosse ormai vista secondo una prospettiva fortemente improntata al calcolo.

Il fatto che Pascal abbia inventato il calcolo delle probabilità, almeno nella sua prima formulazione, non deve però fare dimenticare che lo stesso Pascal nel *pari* impiega una metodologia di calcolo per affacciarsi su un orizzonte che non è affatto, o non solamente matematico. Questa ambivalenza della probabilità, senza nulla togliere al genio di Pascal e alla sua capacità di escogitare applicazioni originali di metodi noti, pare indicare una natura della probabilità che travalica il mondo della matematica.

2.1. Il gioco dei dadi

Il calcolo delle probabilità nasce indagando problemi legati al gioco dei dadi; nessuno mette in discussione ciò. Quel che suscita curiosità e interrogativi è come mai questo gioco d'azzardo, antichissimo nelle sue varianti, del quale esistono testimonianze addirittura nei resti archeologici e nelle pitture egizie, non sia mai stato analizzato da un punto di vista matematico-probabilistico prima del XVII secolo. Hacking[8] passa in rassegna le ipotesi che sono state formulate per spiegare questo ritardo: una concezione deterministica del mondo che avrebbe ostacolato la comparsa dell'idea di casualità, il prevalente impiego divinatorio dei dadi nell'antichità che avrebbe reso irrispettoso, se non empio, un calcolo che ne prevedesse le risposte, la mancanza di dadi abbastanza precisi da suggerire il concetto di equiprobabilità, un mercato assicurativo sviluppatosi solo in epoca moderna che con la sua necessità di metodi per stimare condizioni contrattuali appetibili e redditizie fa da innesco, una conoscenza della matematica non ancora sufficientemente progredita. Se le cause ipotizzate possono avere avuto una qualche influenza nel determinare il momento in cui il calcolo delle probabilità ha fatto la sua comparsa nella storia della matematica, Hacking mostra come nessuna sia veramente risolutiva della questione né sia in grado di reggere di fronte a una

[8] Hacking, *L'emergenza della probabilità*, cit., pp. 12-20.

verifica critica seria.

Una cosa resta incontestabile: l'occasione che ha avviato la riflessione formale sulla probabilità è la curiosità di una mente geniale per una scommessa su un lancio di dadi, indagata con l'interesse accanito di un giocatore d'azzardo.

2.2. Un'idea ambivalente e una definizione difficile

Posto che si voglia collocare nel 1654 la nascita del moderno concetto di probabilità, l'idea che vi sta dietro non è né univoca né, almeno sul momento, esplicitamente riflessa:

> È importante osservare come la probabilità che comparve così all'improvviso abbia carattere bifronte. Da un lato, infatti, essa è statistica e riguarda le leggi stocastiche dei processi casuali, dall'altro è epistemologica, volta a valutare gradi ragionevoli di credenza in proposizioni assolutamente prive di contenuto statistico.[9]

La ambivalenza in parte è dovuta alla matematica che compare come elemento nuovo nella teoria della probabilità, con tutto quel che comporta di espansione del campo semantico del termine, aprendo la strada alla statistica. Il tentativo poi di catturare nella terminologia quello che emerge come novità e distinzione nell'ambivalenza, è anche riflesso del rigore che la matematica chiede al formalismo tecnico e al linguaggio che impieghiamo per descriverlo, e al contempo indice di una resistente inconciliabilità tra i diversi modi di intendere la probabilità. Hacking ne dà una esemplificazione significativa:

> Carnap sosteneva che bisogna distinguere. una. «probabilità$_1$» da una «probabilità$_2$» e più tardi parlò di probabilità induttiva e statistica. Poisson e

[9] *Ivi*, p. 22.

> Cournot proposero di usare le parole francesi *chance* e *probabilité* per sottolineare questa stessa distinzione. Prima di loro, Condorcet propose il termine *facilité* per indicare i concetti aleatori e *motif de croire* per indicare quelli epistemici. Bertrand Russell per questi ultimi usa la parola «credibilità». Molti altri sono i termini cui si è fatto ricorso. Troviamo *Zuverlässigkeit*, «propensità», «proclività», così come una moltitudine di modifiche aggettivali della parola «probabilità», tutte usate per indicare tipi diversi di probabilità.[10]

Tra la nascita del calcolo delle probabilità, che ha preso il via con Pascal, e l'assiomatizzazione della teoria, ad opera di Andrej Nikolaevič Kolmogorov nel 1933, passano quasi tre secoli. Durante questo lungo lasso di tempo la discussione sulla natura della probabilità non cessa e il suo carattere ambivalente non trova soluzione, tanto che la teoria assiomatica della probabilità fonda un sistema di calcolo che si può impiegare indipendentemente dalla concezione della probabilità che si ritenga valida. Hacking osserva con una certa ironia che «I filosofi sembrano curiosamente incapaci di considerare come due cose distinte l'aspetto aleatorio e quello epistemologico della probabilità. Questo fa pensare che ci troviamo sotto l'influenza di forze più oscure di quelle ammesse nell'ontologia positivista»[11]. Quello che Hacking chiama aspetto epistemologico della probabilità va riferito alla concezione "frequentista", come appare dalla correlazione che fa con la *Probabilità*$_2$ di Carnap nella citazione precedente. La corrispondenza risulta anche dall'associazione che Geymonat fa parlando di probabilità a priori e probabilità a posteriori rispetto a quella distinzione[12].

Hacking in realtà fa un passo in più e va oltre la mera constatazione di una divergenza di fatto:

---

[10] *Ivi*, p. 23.
[11] *Ivi*, p. 25.
[12] L. Geymonat, *Riflessioni sulle ricerche di Carnap intorno alla probabilità e all'induzione*, in «Rivista Critica di Storia della Filosofia», 10, 1955, pp. 452-453.

> consideriamo le diverse interpretazioni di cui è stata oggetto l'opera di Jacques Bernoulli, il cui libro sulla probabilità venne pubblicato postumo nel 1713. Siccome introdusse il termine «soggettivo» nella filosofia della probabilità, Bernoulli è stato definito soggettivista. Altri sostengono che egli anticipò la teoria «logicista» della probabilità di Carnap. Altri ancora lo definiscono il precursore dell'estrema versione frequentista della statistica di cui Jerzy Neyman è il più celebre esponente. Ricercatori di scuole diverse possono quindi, non a torto, far risalire all'opera di Bernoulli posizioni come quelle di Neyman, di Carnap e dei soggettivisti, benché esse siano giustamente considerate incompatibili.
>
> Tutto questo sembra confermare, a un livello ancora rozzo e indefinito, che, nonostante tutti i nostri progressi nella tecnica matematica, diversi aspetti del dualismo nella nozione di probabilità erano presenti fin dall'inizio.[13]

Detto in altre parole, la possibilità di fare riferimento a uno stesso autorevole pensatore muovendo da posizioni concettuali radicalmente diverse, pone il fondato dubbio che si dia effettivamente una separazione radicale tra le posizioni di partenza. La dualità della probabilità non è dicotomia mascherata da un uso impreciso del linguaggio, che assegna lo stesso vocabolo a cose distinte, ma indizio della sua effettiva natura.

2.3. Prima di Pascal

Naturalmente il termine *probabilità* non nasce con Pascal. Con lui e i numerosi pensatori che ne hanno seguito le intuizioni, approfondendo e arricchendo la teoria delle probabilità, alla novità dei concetti non corrisponde necessariamente una novità di vocaboli. La novità evidente e incontestabile sta nel fatto che prima della seconda metà del XVII secolo la probabilità è un concetto privo di qualsiasi riferimento al caso e alla casualità. Ancora Hacking[14], esaminando il significato del termine nel corso dei secoli, fa notare come sia sempre stato impiegato per

[13] Hacking, *L'emergenza della probabilità*, cit., p. 26.
[14] *Ivi*, pp. 29-segg.

esprimere un grado dell'opinione, la credibilità di ciò rispetto a cui viene predicata la probabilità, la sua possibilità e plausibilità o l'accreditamento fornito da una figura autorevole; e questo uso comune del termine si è mantenuto fino a noi. In realtà in tempi recenti l'impiego dell'ultima accezione, ovvero l'autorità che, in virtù del suo stesso prestigio o del ruolo che ricopre, con il suo pronunciamento rende credibile ciò che afferma, tende a essere concettualmente negata o disconosciuta, vuoi per influssi di pensiero illuminista e positivista, vuoi per vicinanza al pensiero debole, ma puntualmente in ogni discussione accalorata viene un momento in cui le tesi vengono sostenute ricorrendo a titoli ed esperienza di chi le condivide o all'incompetenza di chi le avversa.

Al di là delle cause che possono essere ipotizzate per spiegare la comparsa del calcolo delle probabilità, passate in rassegna nel paragrafo 2.1, Hacking individua nel *segno* e nel modo in cui si è trasformato il suo riferirsi alle cose dandone evidenza, l'ambito epistemologico che mutando ha favorito il radicarsi della nuova idea. Non si vuole intendere qui che sia mutata la struttura conoscitiva antropologica, ma che parallelamente al progredire della scienza e del modo in cui l'uomo si avvicina allo studio del mondo, si è trasformato il modo in cui leggiamo i segni che guidano la nostra ragione alla conoscenza. Nel mondo antico il cosmo parla di se stesso per il fatto stesso di manifestarsi, e i segni della presenza degli enti indicano la loro natura: la luce delle stelle rivela il fuoco che le costituisce e la solidità della materia rivela qualcosa della sua intima natura:

> Infatti nascono, dai triangoli che abbiamo scelto, quattro generi, tre dei quali da quel solo triangolo che ha i lati diseguali, mentre il quarto, e questo solo, è costituito dal triangolo isoscele. […] Distribuiamo i generi che abbiamo ottenuto con il ragionamento, in fuoco, terra, acqua e aria. […] Dunque i corpi semplici e primi si sono prodotti in virtù di queste cause.
>
> Per quanto riguarda, poi, le diverse specie che si sono prodotte in ciascuno di quei generi, si deve attribuire la causa alla costituzione di ciascuno dei due triangoli elementari, perché mediante ciascuna composizione non si è prodotto da principio il triangolo di una grandezza sola, ma più grandi e più piccoli, e tanti di numero quante sono le sottospecie dei generi.

> Perciò, mescolati questi triangoli con se medesimi e tra di loro, ne risultano infiniti per varietà.
>
> E di questa varietà bisogna che diventino osservatori coloro che intorno alla natura vorranno fare uso di un ragionamento verosimile.[15]

Questa concezione del mondo e di come è strutturato giustifica le sensazioni impresse dai corpi fisici su di noi.

> Infatti, che l'impressione del fuoco sia qualcosa di acuto, lo sentiamo tutti. E bisogna rilevare la sottigliezza degli spigoli e l'acutezza degli angoli e la piccolezza delle parti e la rapidità del movimento: per tutte queste cose essendo il fuoco veemente e tagliente in modo acuto, taglia sempre ciò in cui si imbatte. E bisogna ricordare altresì la generazione della forma di esso, in quanto è soprattutto questa natura e non un'altra che, dividendo i nostri corpi e sminuzzandoli in piccole parti, produce questa impressione che ora chiamiamo calore, e insieme il suo nome.[16]

La struttura del mondo era concepita essenzialmente come un tutto ordinato che va al di là dell'aspetto semplicemente fisico, e che i greci indicavano con il termine "cosmo":

> E i sapienti dicono, o Callicle, che cielo, terra, dèi e uomini sono tenuti insieme dalla comunanza, dall'amicizia, dalla temperanza e dalla giustizia: ed è proprio per tale ragione, o amico, che essi chiamano questo intero universo «cosmo», ordine, e non, invece, disordine o dissolutezza.[17]

---

[15] Platone, *Timeo*, 54B-57D. Trad it di G. Reale su testo greco di J. Burnet.
[16] Platone, *Timeo*, 61E-62A. Trad it di G. Reale su testo greco di J. Burnet.
[17] Platone, *Gorgia*, 507E-508A. Trad it di G. Reale su testo greco di J. Burnet.

Questa idea di cosmo ordinato e intelligibile nella sua struttura, anche in una prospettiva trascendente e metafisica, supera i confini del mondo greco per conservarsi in occidente durante tutto il Medioevo.

Nel Rinascimento, Paracelso ribalta la concezione galenica della medicina e sostiene il principio di somiglianza. I segni che la natura pone nel libro che squaderna di fronte ai nostri occhi mostrano quello che possiamo apprendere; così, con un procedimento analogico, dal fatto che le ramificazioni nelle corna del cervo indicano la sua età e che i nomi delle stelle indicano il loro influsso, impariamo a curare gli organi del corpo con piante le cui foglie hanno forma di reni o fegato, mentre le lacerazioni dei tessuti interni vanno curate con le pungenti foglie di cardo.

In questa rapida carrellata di concezioni fisiche del mondo, che con terminologia contemporanea diremmo evoluzione della corrispondenza tra segno significante e significato, occupa certamente una posizione di rilievo Galileo Galilei con *Il saggiatore*:

> Parmi oltre a ciò di scorgere nel Sarsi ferma credenza che nel filosofare sia necessario appoggiarsi all'opinioni di qualche celebre autore, sicché la mente nostra, quando non si maritasse col discorso di un altro, ne dovesse in tutto rimanere sterile ed infeconda; e forse stima che la filosofia sia un libro e una fantasia di un uomo come l'*Iliade* e l'*Orlando Furioso*, libri ne' quali la meno importante cosa è, che quello che vi è scritto sia vero. Signor Sarsi, la cosa non istà così. La filosofia è scritta in questo grandissimo libro che continuamente ci sta aperto innanzi agli occhi (io dico l'universo), ma non si può intendere, se prima non s'impara a intender la lingua, e conoscer i caratteri ne' quali è scritto. Egli è scritto in lingua matematica, e i caratteri son triangoli, cerchi ed altre figure geometriche, senza i quali mezzi è impossibile intenderne umanamente parola; senza questi è un aggirarsi vanamente per un oscuro laberinto[18].

[18] G. Galilei, *Il saggiatore*, G. Barbera, Firenze 1864, pp. 59-60.

Questo scritto famoso di Galileo del 1623, assieme alle sue *Lettere copernicane* pubblicate tra il 1613 e il 1616, rappresenta il momento fondante della scienza moderna, in cui si espone e si dà dimostrazione di come applicare il metodo scientifico, che consiste nella continua verifica delle ipotesi nate dall'osservazione. È curioso il fatto che Galileo fornisca il metodo di lavoro che servirà a confutare sue tesi contenute nello stesso volume in cui lo espone: infatti, *Il saggiatore* è dedicato allo studio delle comete, che Galileo riteneva essere meri simulacri, effetto di giochi di riflessione e rifrazione della luce, anziché corpi celesti materiali.

Il passo citato di Galileo è però importante e ci interessa in particolar modo per due ragioni. La prima, di ordine generale rispetto all'argomento che trattiamo, è che con Galileo la scienza si avvia alla modernità definitivamente ancorata alla concezione platonica, che vede nella matematica, e più nello specifico la geometria, il linguaggio con cui è scritto l'universo. Se per Galileo il mondo è già universo e non più cosmo, ci si può chiedere però fino a che punto egli rompa consapevolmente rispetto a Platone, considerando questo alfabeto composto da triangoli, cerchi e altre figure geometriche, non sostanziale rispetto alla natura che descrive. Ai nostri giorni, un ingegnere considera la matematica altrettanto irrinunciabile di quanto non faccia Galileo, ma in quanto si tratta del solo strumento utilizzabile dall'ingegno umano per costruire un modello concettuale che consenta di rappresentare il mondo; chiaramente anche un fisico usa la matematica per costruire modelli rappresentativi del reale, ma è all'ingegneria che demandiamo la realizzazione di modelli matematici capaci di scaricare effetti riscontrabili nel mondo reale[19]. L'immagine del mondo come libro aperto, disponibile per chiunque sia intenzionato a leggerlo per comprenderlo, è ampiamente usata anche prima di Galileo, ma in lui

[19] Un fisico costruisce modelli matematici precisi, in grado di descrivere fenomeni la cui verifica a volte può avvenire in modo molto indiretto e anche ricorrendo alla coerenza con altri modelli teorici: ad esempio Albert Einstein con le onde gravitazionali o Stephen Hawking con la perdita di massa dei buchi neri. Un ingegnere costruisce modelli matematici approssimati (eventualmente semplificando modelli fisici precisi) la cui validità definitiva viene verificata pragmaticamente: l'aereo progettato effettivamente si alza dal suolo, nel punto di trivellazione effettivamente si estrae il petrolio, ecc. Non è un caso che l'ingegneria in quanto prassi della scienza abbia preso forma e si sia istituzionalizzata, strutturandosi in maniera sempre più articolata, via via che le scienze positive sganciavano la loro attività teorico speculativa dal mondo fenomenico.

si avvia in modo irrevocabile la separazione tra strumento di indagine e oggetto di conoscenza.

La seconda cosa da notare del testo citato è il termine *filosofia* impiegato per indicare la fisica, come la intenderemmo noi, ma più ancora come l'avrebbe intesa Aristotele. Pur non avendo qui l'interesse a fare una indagine filologica del testo galileiano, è chiaro che non si tratta né di una svista né di un uso caratteristico suo: celebre il verso del coevo Shakespeare che fa dire ad Amleto «Vi sono più cose in cielo e in terra, Orazio, di quante ne sogni nella tua filosofia», per apostrofare l'amico sgomento durante l'apparizione del fantasma[20]. La scienza, che da sempre esisteva nell'alveo della riflessione filosofica, trova ancora eco di ciò nel vocabolo impiegato per riferirvisi, ma ha ormai imboccato un percorso proprio dove la conoscenza riduce la metafisica alla fisica, e la teoretica si accontenta di essere teoria[21]. È in questo nuovo orizzonte del conoscere che *il caso* perde ogni riferimento a fato e destino, e può essere indagato dal calcolo delle probabilità.

3. Concezioni e definizioni della probabilità – Oggettivismo e soggettivismo

Entrando nel merito del significato della probabilità e dovendone dare una definizione, cosa necessaria anche solo per poterne parlare intendendosi su ciò di cui si discute, si individuano due filoni di pensiero riferiti alla relazione che sussiste tra la probabilità e chi la valuta. Uno è l'oggettivismo, che consiste nel ritenere la probabilità insita nelle cose; ad esempio, lanciando un dado le sue facce escono con una probabilità che non dipende in nessun modo dalla persona che ha effettuato il lancio. L'oggettivismo di fatto non richiede immediata evidenza di una

---

[20] W. Shakespeare, *Amleto*, atto I scena V. Traduzione mia. La prima pubblicazione è del 1603

[21] È evidente, come la storia ha poi anche mostrato, che non è possibile una teoria scientifica che prescinda assolutamente da una visione teoretica come orizzonte di comprensione e condizione della sua stessa possibilità. Il percorso della scienza, infatti, è stato poi inevitabilmente recuperato in ambito filosofico con la filosofia della scienza, che vede importanti contributi di autorevoli esponenti delle scienze positive. I teoremi di incompletezza di Goedel sono forse l'esempio contemporaneo più emblematico di un pensiero squisitamente matematico e allo stesso tempo squisitamente filosofico.
La stessa fisica nella sua riflessione più ampia sul mondo ha recuperato un'idea di odine nell'universo, chiamando quella branca di studi "cosmologia".

res che manifesti un comportamento probabilistico: la semplice formulazione di un modello matematico delle probabilità, che impiegato in un calcolo fornisca risultati invarianti rispetto a chi lo applica, esprime l'idea di una probabilità oggettiva, che può essere indagata e verificata allo stesso modo da chiunque. L'altro importante filone è il soggettivismo, che sostiene che in ogni calcolo probabilistico entri in gioco una valutazione dipendente dal soggetto che effettua il calcolo, e questa valutazione soggettiva concorre in qualche modo a formare il valore della probabilità stimata. Domenico Costantini e Ludovico Geymonat in *Filosofia della probabilità*[22] schematizzano le interpretazioni possibili della probabilità in "logicista", "frequentista", e "soggettivista". Le prime due fanno capo alla corrente oggettivista. Categorizzare gli approcci alla teoria delle probabilità come interpretazioni, significa di fatto dare alla probabilità un significato che non è solo applicazione di un formalismo, ma ne indica in qualche modo una natura che potrebbe essere indagata più approfonditamente. A questo proposito, Costantini e Geymonat riconoscono a Carnap, che ammette l'esistenza di più definizioni di probabilità, il merito di volersi sottrarre all'atteggiamento di dogmatismo, predominante nella riflessione sulla probabilità; siamo infatti all'inizio del XX secolo e Kolmogorov sta formalizzandone l'assiomatizzazione. I nostri autori rimproverano a Carnap di scivolare in una sostanziale «indifferenza filosofica circa la natura della probabilità»[23], coerentemente con il neopositivismo di cui lui è autorevole esponente. Per Costantini e Geymonat però la natura della probabilità è una questione che non può essere accantonata: deve essere recuperato il rapporto teoria-prassi che era presente all'inizio e ha determinato le diverse interpretazioni della probabilità, prima di sfumare man mano che la teorizzazione diventava dominante; la tesi di questi autori è infatti che la probabilità non sia un assoluto immutabile ma una nozione variabile nel tempo e in funzione degli ambiti di applicazione.

[22] D. Costantini e L. Geymonat, *Filosofia della probabilità*, cit.
[23] *Ivi*, p. 14.

3.1. Interpretazione logicista e definizione classica

L'interpretazione logicista è «il grado di credenza che è ragionevole assegnare al verificarsi di certi fenomeni»[24] e ad essa fa capo la definizione classica della probabilità che sinteticamente viene enunciata come «il rapporto tra il numero dei casi favorevoli ed il numero dei casi egualmente probabili»[25]. Perciò per determinare il valore della probabilità, dopo che si è avuto l'accortezza di individuare correttamente i casi favorevoli, è sufficiente procedere a una semplice divisione tra due numeri.

I casi che possono verificarsi prenderanno il nome di *eventi* nella formalizzazione teorica delle probabilità. La distinzione tra *eventi* e *casi* può confondere in quanto di fatto nella enumerazione delle possibilità coincidono. Concettualmente un evento è la possibilità che effettivamente si dà, o si ipotizza data, in risposta a un problema la cui probabilità di essere soddisfatto può essere misurata con il rapporto tra casi favorevoli e casi possibili. I casi favorevoli sono tutti quegli eventi semplici che, se si verificassero, rappresenterebbero una risposta ugualmente soddisfacente al problema. Mi pare abbastanza indicativo della costante difficoltà a focalizzare la definizione di probabilità il fatto che, con la formalizzazione della teoria, sono stati definiti concetti come "l'insieme dello spazio campionario" associato a un certo problema, gli "eventi" definiti come sottoinsiemi dello spazio campionario, gli "eventi semplici" che sono i sottoinsiemi con un solo elemento, ma la definizione di probabilità in sé continua a fare riferimento a "casi" (nel senso di eventualità), favorevoli e possibili, che poco hanno a che fare col caso (cioè con la casualità) e illustrati a mezzo di esempi più che definiti[26].

La più tradizionale tra le esemplificazioni della definizione classica è quella che illustra le possibilità che possono darsi nel caso del lancio di un dado o di una

---

[24] *Ivi*, p. 21.
[25] B. de Finetti, *Filosofia della probabilità*, a cura di A. Mura, il Saggiatore, Milano 1995, p. 208.
[26] Esiste una corrispondenza tra i casi possibili e gli eventi semplici. Resta il fatto che tra tutti i testi sulla probabilità da me esaminati il rapporto tra casi favorevoli e possibili viene illustrato con una esemplificazione discorsiva e senza fornire una definizione di cosa sia un *caso*.

moneta: quale *grado di credenza* posso assegnare al fatto che in seguito al lancio una particolare faccia del dado si troverà in alto? La faccia particolare che sto considerando è una, le possibili facce che possono presentarsi sono sei, il grado di credenza o valore della probabilità che sto cercando è 1/6; analogamente per il lancio della moneta si determina che la probabilità che si presenti una data faccia è 1/2. Correntemente, le facce si considerano contrassegnate (ad esempio i numeri da 1 a 6 per quelle del dado o testa e croce per la moneta), e con terminologia storicamente mutuata dalle bische, il presentarsi di una data faccia si chiama *uscire*: si dice pertanto che in seguito a un lancio di dado esce, ad esempio, il 3, quando tra tutti gli eventi possibili si verifica quello per cui la faccia contrassegnata dal numero 3 si presenta in alto.

Poiché l'insieme degli eventi favorevoli è sempre un sottoinsieme, al più improprio, di quello dei casi possibili e la cardinalità minore è al numeratore, la probabilità viene espressa da un numero che è sempre compreso nell'intervallo tra 0 e 1. Gli estremi dell'intervallo rappresentano rispettivamente l'evento impossibile e l'evento certo; ad esempio la probabilità che al lancio di un dado esca la faccia numero 7 vale 0 (non può capitare), mentre la probabilità che si presenti una faccia con numero compreso tra 1 e 6 vale 1 (avviene sicuramente).

Il presupposto insito in questa interpretazione della probabilità è che gli eventi possibili siano tutti equiprobabili, ovvero ogni evento possa presentarsi con la stessa facilità di ogni altro. Presupposto talmente imprescindibile da venire inserito da Laplace al secondo posto tra i sette principi che egli individua a fondamento del calcolo delle probabilità:

> Principi generali del calcolo delle Probabilità.
>
> I principio
>
> Il primo di questi principi è la definizione stessa della probabilità che, come abbiamo visto, è il rapporto del numero dei casi favorevoli con quello di tutti i casi possibili.
>
> II principio

> Ma questo presuppone i diversi casi ugualmente possibili [*également possibles*]. Se non lo sono si determinerà in primo luogo le loro possibilità rispettive la cui corretta valutazione è uno dei punti più delicati della teoria del caso.[27]

I primi due principi di Laplace, di fatto, rappresentano la prima definizione dell'interpretazione classica della probabilità; definizione impiegata correntemente fino ad oggi in modo sostanzialmente invariato, salvo far collassare i due principi in un unico periodo per sintesi e brevità.

Come si faceva osservare nell'introduzione, esiste il problema di una autoreferenzialità non facilmente superabile nella definizione di probabilità: se da una parte parlare di casi ugualmente "possibili" evidenzia l'accortezza di impiegare termini differenti per non rendere tautologica la definizione, dall'altra riesce difficile definire o comunque spiegare cosa sia la "possibilità" di un caso senza dire che si tratta della sua probabilità. Di fatto, l'imbarazzo per la evidente circolarità di una definizione che demarca un ambito della matematica, costituisce una forte spinta alla riflessione successiva a Laplace per cercare di identificare un sistema assiomatico per il calcolo delle probabilità, e a mano a mano che i tentativi di assiomatizzazione si rivelavano sempre più efficaci, l'imbarazzo per la circolarità si è trasformato nello scrupolo per una nota a margine (almeno sul fronte della matematica). Infatti se prendiamo un testo universitario recente di teoria del calcolo delle probabilità, la definizione classica viene riportata con questa formulazione:

> Nell'interpretazione *classica* la probabilità di un evento A è il rapporto tra i possibili risultati favorevoli all'evento $A, n(A)$, e il numero N dei possibili risultati cioè
>
> $$P(A) \triangleq \frac{n(A)}{N}$$
>
> Si deve notare che $n(A)$ ed N non sono i <u>risultati effettivi</u> di un esperimento, ma i <u>possibili risultati</u> di esso, cioè sono la *cardinalità* di A (numero di elementi di A) e

[27] P.S. Laplace, *Essai Philosophique Sur Les Probabilités*, cit., p. 12.

di N rispettivamente.

La definizione classica presenta delle ambiguità e conduce a risultati non corretti nel caso in cui i vari risultati possibili non siano *equiprobabili*. Si deve quindi inserire nella definizione la condizione che i diversi risultati possibili sono equiprobabili (*principio della ragione insufficiente)*[28]

Oramai, una volta stabiliti gli assiomi e verificato che la teoria che ne deriva consente di recuperare tutti i progressi nel calcolo probabilistico fatti fino a quel momento, è evidente come la matematica accantoni ogni remora nel lasciare emergere l'aporia della circolarità nella definizione classica di probabilità, declassando la questione a una condizione al contorno che bisogna ricordarsi di porre.

Il "principio di ragione insufficiente" di Jacob Bernoulli, menzionato da Galati e Pavan nel testo citato sopra, anche detto "principio di indifferenza", si contrappone al "principio di ragione sufficiente di Leibnitz". La contrapposizione tra i due principi è più concettuale che intenzionale, dato che Bernoulli non fa riferimento a Leibnitz e al suo principio, anche perché benché i due fossero contemporanei Bernoulli non poté leggere le pubblicazioni dove Leibnitz lo esponeva né risultano contatti tra loro.

Bernoulli argomenta in questo modo:

I numeri di casi possibili sono noti, ad esempio, nei giochi con i dadi, dal momento che conosciamo il numero di facce di ciascuno di essi e che tutte hanno la stessa propensione a presentarsi, dal momento che non vi `e alcuna ragione, vista la somiglianza delle facce e la distribuzione uniforme del peso del dado, perché una faccia debba essere più propensa a presentarsi rispetto ad un'altra, come

[28] G. Galati e G. Pavan, *Teoria dei fenomeni aleatori*, TeXmat, Roma 2006, p. 21.

potrebbe succedere se le facce avessero una forma diversa ovvero se il dado fosse di due materie diverse. Similmente sono noti i numeri di casi favorevoli all'estrazione di una scheda bianca o una nera da un'urna ed è noto che sono tutte egualmente possibili, dal momento che sono determinati e noti i numeri di schede di entrambi i tipi e non si vede alcun motivo per cui debba essere estratta una piuttosto che un'altra.[29]

In pratica se non riusciamo a individuare nessuna ragione che possa giustificare il verificarsi di un accadimento piuttosto che un altro, o a favorirne uno rispetto a un altro, si deve ritenere che tutti i casi possibili abbiano la stessa possibilità di presentarsi.

Ci si può chiedere se questa conclusione, ben sintetizzata da Laplace quando afferma che «la probabilità è relativa in parte a questa ignoranza, in parte alle nostre conoscenze»[30], sia un modo accorto di procedere rispetto a quel che non conosciamo. La scienza che studia i fenomeni aleatori non risponde a questa domanda, come abbiamo visto, paga del fatto che l'assiomatizzazione della teoria è riuscita a fornire un sistema formale che appiana le difficoltà. Dopotutto un assioma, per definizione, non deve essere giustificato.

La cosa non si chiude con altrettanta disinvoltura in ambito filosofico, dato che resta il problema di capire se quando calcoliamo una probabilità stiamo misurando un aspetto della realtà e se stiamo facendo riferimento alla natura di un qualcosa. De Finetti commenta così la definizione classica:

È illusorio pensare che definendo così la probabilità se ne chiarisca davvero il significato. In realtà non si va oltre la consueta caratterizzazione assiomatica, alla quale si aggiunge solo l'ulteriore vincolo formale della equiprobabilità delle alternative elementari. Per fare di questa pretesa definizione una definizione vera

[29] J. Bernoulli, *Ars Conjectandi*, Tournes, Basilea 1713, p. 224. Traduzione tratta dalla dispensa del prof Riccardo Rosso, *Storia della probabilità*, pro manuscripto, Pavia 2020

[30] P.S. Laplace, *Essai Philosophique Sur Les Probabilités*, cit., p. 7. Traduzione mia.

> e propria, bisognerebbe *in primis* chiarire che cosa si intende per «uguale probabilità». E poiché la definizione di uguale probabilità è una parte essenziale della definizione di probabilità, il risultato è una definizione circolare. A prova di ciò basti considerare che una persona che non sapesse che cosa significa «probabilità», non potrebbe apprenderlo mediante la definizione classica, dal momento che si richiederebbe la comprensione del termine «uguale probabilità».
>
> Si potrebbe tentare di eliminare la circolarità sostituendo alla condizione che le alternative elementari siano equiprobabili, una condizione che giustifichi l'eguale probabilità attribuita ad esse. Ma quand'è che due eventi hanno uguale probabilità? Se domandate a una persona che adottasse l'imposizione classica: «perché quelle alternative elementari sono equiprobabili?», tale persona potrebbe rispondere in molti modi diversi, nessuno dei quali sarebbe sufficiente per *definire* ciò che si intende con il termine «ugualmente probabili».
>
> Perciò tutti coloro che adottano la definizione classica incorrono nella circolarità, indipendentemente dal modo con cui giustificano il giudizio di equiprobabilità che la sua applicazione richiede. Vi incorrono sia coloro che basano il giudizio di equiprobabilità su considerazioni di simmetria fisica, sia coloro che lo basano su considerazioni di frequenza.[31]

La conclusione di de Finetti è inappellabile: la definizione classica in realtà non definisce nulla, in quanto non è in grado di sfuggire alla circolarità autoreferenziale che la caratterizza; nemmeno le *considerazioni di simmetria*, richiamate in conclusione della citazione, sono risolutive. Richiamare la regolarità simmetrica delle dimensioni, della disposizione e delle forme delle facce che compongono un dado, potrebbe indurre alla convinzione che un dado ideale goda della equiprobabilità delle facce in un lancio ideale per le sue caratteristiche di progettazione, avvalorando il principio di ragione insufficiente richiamato in precedenza (è usuale in matematica riconoscere la verità di certe proposizioni rispetto a un dato problema *per costruzione*, quando il modo in cui sono state poste le condizioni del problema rende automaticamente vere quelle proposizioni). Il fatto

---

[31] B. de Finetti, *Filosofia della probabilità*, cit., pp. 208-209.

di non disporre di nessuna informazione che possa privilegiare l’uscita di una faccia di un dado rispetto a un’altra, e le facce di un dado ideale in un lancio ideale sono in questa condizione, non comporta necessariamente e senza ulteriori condizioni che le probabilità siano uguali.

Alla voce «Interpretations of Probability» della *Stanford Encyclopedia of Philosophy*[32] c’è un interessante e paradossale esempio che mostra il limite del principio di ragione insufficiente e che ripropongo qui adattato: Si immagini una fabbrica che produce cubi di dimensione variabile in modo casuale, con lo spigolo che ha una lunghezza compresa tra 0 e 1 metro, qual è la probabilità che lo spigolo di un cubo scelto a caso misuri meno di 1/2 metro? Se assumiamo una distribuzione uniforme delle dimensioni dello spigolo dei cubi prodotti, la probabilità vale 1/2. Se la fabbrica produce cubi di dimensione variabile in modo casuale con la faccia che ha una superficie compresa tra 0 e 1 metro quadrato, qual è la probabilità che la faccia di un cubo scelto a caso misuri meno di 1/4 di metro quadro? Se assumiamo una distribuzione uniforme delle dimensioni delle facce dei cubi prodotti, la probabilità vale 1/4. Se la fabbrica produce cubi di dimensione variabile in modo casuale con il volume compreso tra 0 e 1 metro cubo, qual è la probabilità che il volume di un cubo scelto a caso misuri meno di 1/8 di metro cubo? Se assumiamo una distribuzione uniforme dei volumi dei cubi prodotti, la probabilità vale 1/8. Le assunzioni sull’omogeneità della distribuzione di lunghezza, area e volume le possiamo fare in base al principio di ragione insufficiente, ma se osserviamo che si tratta di una sola fabbrica che produce cubi con dimensione casuale, con gli estremi indicati sopra che sono simultaneamente soddisfatti per la dimensione lineare, di superficie e di volume, lo stesso principio ci dice che è indifferente su quante dimensioni scegliamo la distribuzione uniforme, perciò la probabilità di scegliere un cubo con lo spigolo minore di 1/2 metro vale, a piacimento, 1/2, 1/4 o 1/8.

Questo risultato, oltre ad essere palesemente paradossale, mostra come

[32] Hájek, Alan, "Interpretations of Probability", The Stanford Encyclopedia of Philosophy (Fall 2019 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/fall2019/entries/probability-interpret/>

l'interpretazione classica, pur essendo l'idea più antica della probabilità, e ancora adesso il percorso di apprendimento privilegiato per il neofita che voglia apprendere il calcolo delle probabilità e le scienze statistiche, non abbia mai trovato una formulazione che ne rimuova le ambiguità.

3.2. Interpretazione frequentista

L'interpretazione frequentista nasce dalla considerazione che, se un evento ha una certa probabilità di verificarsi, in una serie di esperimenti ripetuti tenderà a presentarsi con una frequenza che dipende dalla sua probabilità, per cui la probabilità dell'evento corrisponde alla frequenza di distribuzione con cui si presenta in un numero molto grande di esperimenti ripetuti:

> [La frequenza relativa] fa riferimento alla possibilità di effettuare per un numero N di volte e sempre nelle medesime condizioni, un esperimento in cui si può osservare se l'evento di interesse si è verificato oppure no.
>
> La frequenza relativa $f(A)$ di un evento A è il rapporto tra il numero di volte $n(A)$ in cui si ha come risultato un elemento dell'insieme A ed il numero N di prove dell'esperimento.
>
> $$f(A) \equiv \frac{n(A)}{N}$$
>
> La frequenza relativa approssima un valore "stabile" quando il numero delle prove è "molto grande". Si dice che la probabilità dell'evento A è il limite di $f(A)$ per N tendente all'infinito.
>
> La precedente considerazione è empirica e non corrisponde ad alcun limite in senso matematico, cioè per N comunque grande sono possibili valori del rapporto $\frac{n(A)}{N}$ che si scostino da un qualsiasi valore prefissato per quantità finite.[33]

[33] G. Galati e G. Pavan, *Teoria dei fenomeni aleatori*, cit., pp. 20-21.

Diamo anche la definizione di “statistica”, termine che abbiamo già usato e incontreremo spesso, dal momento che si tratta di una scienza strettamente legata al calcolo delle probabilità, in particolare se inteso nella sua interpretazione frequentista:

> La statistica è una disciplina di tipo e di valore applicativo che collega i concetti teorici della teoria della probabilità con la realtà. Il dominio delle affermazioni statistiche non è più un arbitrario spazio campione S, ma piuttosto lo spazio S_n generato da n prove ripetute. Queste possono essere pensate come ripetizioni di un medesimo esperimento (“*osservazioni*” del medesimo fenomeno). Lo scopo è quello di stabilire delle *inferenze* sulla legge probabilistica che governa il fenomeno, cioè dedurre, dalle osservazioni, proprietà di tale legge.[34]

L’interpretazione frequentista ricevette un grande impulso a partire dal XIX secolo, sia per il crescente interesse verso i fenomeni demografici e la possibilità che finalmente si presentava di manipolare grandi quantità di dati con criteri statistici, per gestire collettività molto numerose in modo analitico e predittivo; sia per l’intuizione di Maxwell, che cercando un modo per studiare il comportamento dei fluidi nella termodinamica, pensò di ricorrere all’uso di metodi statistici. La difficoltà di gestire la meccanica dei fluidi, con cui si trovò alle prese Maxwell, consiste nel fatto che un fluido in ogni situazione reale significativa è composto da un numero enorme di molecole, ognuna delle quali è soggetta alle leggi della meccanica classica; se si considera una singola particella, le leggi del moto sono note ed è facile calcolarne traiettoria e comportamento anche in seguito ad urti con altre particelle, ma se si considera che un volume di gas pari a 22,5 litri, alla temperatura di 0 gradi centigradi e alla pressione atmosferica, contiene circa 6×10^{23} molecole (un numero di molecole rappresentabile in numerazione decimale con il numero 6 seguito da 23 zeri) continuamente in movimento e in continua collisione tra loro,

[34] *Ivi*, p. 316.

l’impresa è titanica. La possibilità di calcolare le traiettorie e le interazioni di così tante molecole in movimento, non solo era fuori della portata di Maxwell ma lo sarebbe anche dei più moderni computer, per non parlare della impossibilità di determinare le condizioni iniziali del sistema in termini di velocità e direzione di ogni particella a un dato momento. L’idea di Maxwell fu di trascurare completamente il comportamento delle singole particelle e considerare solo l’effetto della loro azione complessiva come risultato di un comportamento valutato statisticamente: non possiamo determinare l’effetto di una singola particella che si muove con velocità ignota lungo una traiettoria ignota, ma un numero enorme di particelle di cui si ignora il movimento producono un effetto complessivo statisticamente prevedibile e misurabile in termini di pressione e temperatura.

Senza l’approccio statistico allo studio del comportamento dei fluidi, oggi non esisterebbe, ad esempio, la scienza della meteorologia, che studia masse di gas enormi, composte da un numero di particelle che possiamo provare a scrivere ma non riusciamo certo a rappresentarci. È vero che la statistica non coincide con il calcolo delle probabilità, ma questo fonda quella, e l’idea sottesa è proprio che particelle contenute in uno stesso volume di spazio hanno uguali probabilità di muoversi e urtarsi. «Per gli statistici della fine dell’Ottocento il termine “probabilità” era quindi diventato sinonimo di frequenza. Questo è il dato di fatto che colpisce l’attenzione di colui che legga le opere dei grandi ricercatori di quel periodo, in particolare del più illustre di loro, cioè di Pearson»[35]

La definizione di probabilità come frequenza relativa fa riferimento a prove che devono essere ripetute in grande numero per poter avere una distribuzione che si stabilizzi attorno a un certo valore. Per spiegare meglio con un esempio, se vogliamo ricorrere all’interpretazione frequentista per sapere qual è la probabilità che esca testa lanciando una moneta, usando la stessa notazione della definizione riportata sopra: il lancio della moneta è l’esperimento ripetibile, N il numero di lanci effettuati, A il fatto che esce testa e $n(A)$ il numero totale di volte in cui è uscita

[35] D. Costantini e L. Geymonat, *Filosofia della probabilità*, cit., p. 34.

testa e la frequenza relativa $f(A)$ corrisponde alla probabilità che stiamo cercando. Al primo lancio ovviamente si danno due casi: esce testa oppure non esce testa; applicando la formula $f(A) = n(A)/N$, se è uscita testa la probabilità vale 1, se non è uscita testa vale 0. Con un solo esperimento dovremmo concludere che la probabilità che esca testa nel lancio di una moneta sia 1 o 0, a seconda dell'esito dell'esperimento, e poiché l'interpretazione frequentista serve a determinare sulla base di fatti accaduti i fatti possibili, assegnare questi valori di probabilità significa dire rispettivamente che è certo che esca testa ogni volta che lancio una moneta e che è impossibile che esca testa ogni volta che lancio una moneta. L'evidente risultato paradossale, oltretutto contraddittorio con l'esperienza di chiunque abbia provato a lanciare una moneta più di 2 o 3 volte, mette bene in evidenza il fatto che secondo l'interpretazione frequentista la necessità di fare un gran numero di esperimenti non è solo una raccomandazione ma caratterizza strutturalmente quella che è l'idea di una probabilità come frequenza di distribuzione. La ragione di questa caratterizzazione, è dovuta al fatto che la misura della probabilità in senso frequentista è una conoscenza acquisita tramite inferenza induttiva: sulla base di osservazioni particolari si giunge a una conoscenza di valore universale.

A mano a mano che aggiungiamo un esperimento alla catena di quelli già fatti, il valore della distribuzione cambia, e progressivamente può discostarsi anche in modo significativo dal valore che aveva prima: se al primo lancio dove è uscito testa ne segue uno in cui non esce testa, la frequenza relativa degli eventi cercati passa da 1 a 0,5; se al lancio successivo ancora non esce testa, la frequenza relativa passa a 0,33 per poi risalire nuovamente quando uscirà testa.

Nella definizione riportata sopra, i due periodi «la frequenza relativa approssima un valore "stabile" quando il numero delle prove è "molto grande"» e «si dice che la probabilità dell'evento A è il limite di $f(A)$ per N tendente all'infinito»[36] sono cruciali per comprendere il senso di questa interpretazione della probabilità: stanno ad indicare il fatto che all'aumentare dei lanci di moneta, per restare sull'esempio di

---

[36] G. Galati e G. Pavan, *Teoria dei fenomeni aleatori*, cit., p. 21.

prima, la frequenza relativa tenderà ad attestarsi su un valore da cui non dovrebbe poi discostarsi molto, e se i lanci diventano numerosi al punto da tendere all'infinito, questo valore limite è la probabilità dell'evento.

Per comprendere meglio cosa questa definizione comporti, e per rafforzare il filo logico del ragionamento, riportiamo nella *Tabella1* i casi di lancio di una moneta ripetuto sei volte: vengono rappresentati i casi che si possono dare con le sequenze di esiti su tentativi ripetuti; il simbolo 1 viene usato per indicare che il tentativo ha avuto successo (nel nostro caso, al lancio della moneta è uscito testa), 0 per indicare che non ha avuto successo (è uscito croce, in quanto unica alternativa possibile).

Per registrare l'esito di un esperimento, non va trascurata la tecnica che si sceglie per rappresentare il risultato, perché a seconda di come si procede le informazioni raccolte sono in grado di darci una conoscenza più o meno ampia, soprattutto in funzione della direzione in cui si sta muovendo l'indagine. Nel caso che stiamo immaginando adesso, poiché si tratta del lancio di una moneta, non c'è differenza tra il tenere traccia del successo o dell'esito del lancio. Se invece stessimo lavorando con un dado a sei facce e considerassimo un successo l'uscita della faccia col numero 1 (perché ad esempio vogliamo verificare con che frequenza si presenta), potremmo ugualmente tenerne traccia segnando un 1 quando esce la faccia col numero 1 e uno 0 quando esce una qualunque delle altre facce; ma in questo caso, dato che il fallimento si può verificare con diversi risultati, raccoglieremmo un'informazione ugualmente corretta ma più "povera" rispetto alla capacità di alimentare la riflessione. Quanto ai dati nella Tabella 1, si può osservare che, per coprire completamente la casistica dei risultati, i valori presenti nella colonna "Esiti" rappresentano la sequenza di numeri da 1 a 64 (la sesta potenza di 2) in notazione binaria.

Esiti	Totale esiti positivi	Frequenza relativa
000000	0	0
000001	1	0,166666667
000010	1	0,166666667
000011	2	0,333333333
000100	1	0,166666667
000101	2	0,333333333
000110	2	0,333333333
000111	3	0,5
001000	1	0,166666667
001001	2	0,333333333
001010	2	0,333333333
001011	3	0,5
001100	2	0,333333333
001101	3	0,5
001110	3	0,5
001111	4	0,666666667
010000	1	0,166666667
010001	2	0,333333333
010010	2	0,333333333
010011	3	0,5
010100	2	0,333333333
010101	3	0,5
010110	3	0,5
010111	4	0,666666667
011000	2	0,333333333
011001	3	0,5
011010	3	0,5
011011	4	0,666666667
011100	3	0,5
011101	4	0,666666667
011110	4	0,666666667
011111	5	0,833333333
100000	1	0,166666667
100001	2	0,333333333
100010	2	0,333333333
100011	3	0,5
100100	2	0,333333333
100101	3	0,5
100110	3	0,5
100111	4	0,666666667
101000	2	0,333333333
101001	3	0,5
101010	3	0,5
101011	4	0,666666667
101100	3	0,5
101101	4	0,666666667
101110	4	0,666666667
101111	5	0,833333333
110000	2	0,333333333
110001	3	0,5
110010	3	0,5
110011	4	0,666666667
110100	3	0,5
110101	4	0,666666667
110110	4	0,666666667
110111	5	0,833333333
111000	3	0,5
111001	4	0,666666667
111010	4	0,666666667
111011	5	0,833333333
111100	4	0,666666667
111101	5	0,833333333
111110	5	0,833333333
111111	6	1

Tabella 1

Frequenza relativa	Casi	Probabilità totale casi
0	1	0,015625
0,17	6	0,09375
0,33	15	0,234375
0,5	20	0,3125
0,67	15	0,234375
0,83	6	0,09375
1	1	0,015625

Tabella 2

Ogni riga della tabella corrisponde alla simulazione di sei lanci di moneta. Nella colonna "Esiti" è riportata la sequenza di risultati dei sei lanci da sinistra verso destra, dove compare il simbolo 1 quando in quel lancio è uscita testa e il simbolo 0 quando non è uscita testa (cioè non si è verificato l'evento che ci interessava); nella colonna "Totale esiti positivi" viene riportato il numero di successi (è uscita testa) della sequenza di lanci alla riga corrispondente; infine la colonna "Frequenza relativa" riporta la frequenza di successi in quella serie di esperimenti ovvero il numero di successi (seconda colonna) diviso 6 (gli esperimenti fatti). La tabella contiene 64 righe che corrispondono a tutte le sequenze di risultati differenti che si possono avere lanciando sei volte una moneta.

La prima cosa cui fare attenzione è che la tabella non va considerata nel suo complesso, altrimenti ricadiamo nella definizione classica della interpretazione logicista con casi favorevoli (192 volte esce testa) e casi possibili (384 lanci) il cui rapporto è banalmente 1/2; dato che per la definizione classica il moltiplicarsi degli esperimenti ha la sola conseguenza di moltiplicare numeratore e denominatore per il numero di esperimenti, con effetto nullo sul rapporto. Letta in maniera frequentista, invece, la tabella rappresenta sei soli esperimenti che dobbiamo immaginare effettivamente eseguiti, e le 64 righe rappresentano le possibili storie di questi 6 esperimenti: 64 possibilità di cui se ne verifica una sola. Se lanciamo sei volte una moneta, quindi, una sola riga della tabella viene selezionata e ci viene restituita la frequenza relativa che approssima il valore della probabilità che stiamo cercando. Cercando di misurare la probabilità di ottenere testa lanciando una moneta possiamo scoprire che corrisponde a uno tra i valori appartenenti all'insieme:

$$\{0; 0{,}17; 0{,}33; 0{,}5; 0{,}67; 0{,}83; 1\}$$

Potrebbe stupire che con l'approccio frequentista la valutazione della probabilità di un evento altamente dubbio, come l'uscire testa al lancio di una moneta, non escluda i valori 0 e 1 che restano una possibilità anche se facciamo una serie di esperimenti molto più numerosi, ma si deve ricordare che in questo caso i valori ottenuti sono solo approssimazioni della probabilità cercata: 0 non significa impossibile, ma "estremamente difficile che avvenga"; e 1 non significa certo, ma "estremamente difficile che non avvenga". La ragione pratica di ciò è che lanci successivi di una moneta sono eventi indipendenti, ovvero quando si lancia una moneta non ha importanza se in precedenza sia già uscita testa una, dieci o cento volte consecutive: la possibilità che esca di nuovo è reale. Anche la probabilità pari a 0,5 (che secondo la definizione classica è quella effettiva) che compare nella tabella, è una approssimazione della probabilità; infatti se facciamo un settimo lancio, la frequenza relativa pari a 0,5 non è più possibile, come in ogni caso in cui, nell'esempio in questione, il numero di esperimenti fatti sia dispari.

Nella *Tabella 1* si può vedere che la sequenza di esiti che generano la probabilità (approssimata) di 0 e 1 sono rispettivamente uno solo per entrambi, mentre per il valore di probabilità 0,5 sono 20 sequenze. Se proviamo a iterare l'analisi della distribuzione, ragionando sulla distribuzione della frequenza relativa rispetto alle sequenze di risultati, otteniamo il risultato della *Tabella 2*, dove nella prima colonna compaiono le frequenze relative presenti nella *Tabella 1*, riportate una volta sola e arrotondate al secondo decimale; nella seconda colonna c'è il numero totale di volte in cui il valore della prima colonna compare nella *Tabella 1* (vale a dire come è distribuito); e nella terza colonna la probabilità di quella frequenza relativa, ottenuta moltiplicando 1/64 (la probabilità della sequenza di 6 lanci) per il numero di presenze nella *Tabella 1* (seconda colonna della *Tabella 2*).

Questo secondo passaggio, con cui si compongono i dati della *Tabella 2*, è stato fatto applicando la definizione classica della probabilità: non abbiamo infatti esperimenti ripetibili e valori approssimati che si modifichino in conseguenza di quelli, ma ci serve per fare notare come i dati di distribuzione e probabilità graficamente possano essere rappresentati in questo modo:

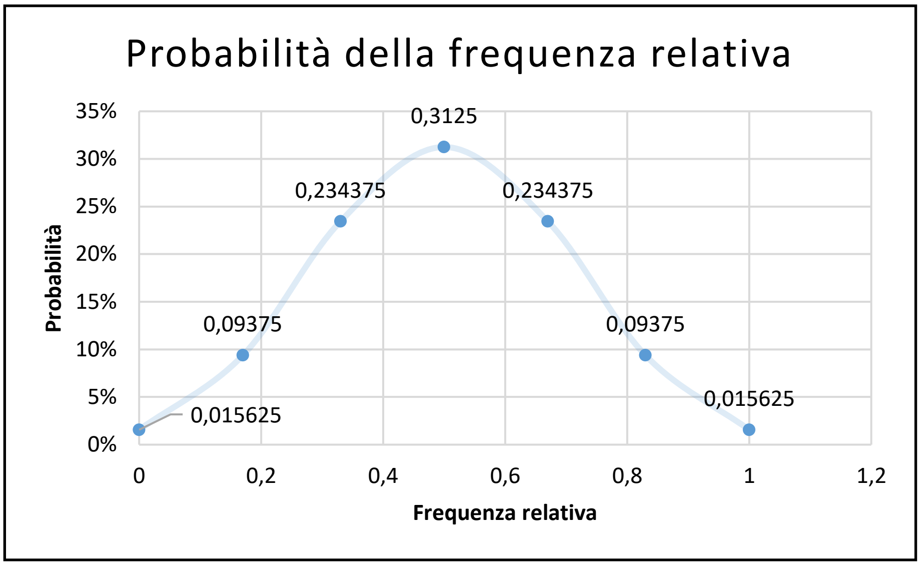

Figura 1

L'analisi di come è distribuita la frequenza relativa della *Tabella 1* mostra che lanciando una moneta 6 volte c'è il 78% di probabilità (nel senso classico) di scoprire, con un approccio frequentista, che la probabilità di avere testa lanciando una moneta è compresa tra 0,33 e 0,67. La somiglianza del grafico in *Figura 1* con la curva degli errori di Gauss, nota come "gaussiana" o "distribuzione normale", non è casuale e rappresenta il fatto che ogni volta che misuriamo qualcosa, e con l'esperimento ripetibile di lanciare una moneta stiamo cercando di misurare la probabilità di avere testa, incorriamo in un errore che può essere più o meno grande, ma se ripetiamo la misura più volte, è facile avere spesso errori piccoli (la zona centrale della curva a campana) e più raramente errori grandi (i rami della curva a campana che tendono asintoticamente a zero verso sinistra e verso destra). Si può mostrare[37], che l'area compresa tra l'asse x e la curva gaussiana vale 1, ovvero il valore dell'evento certo; in termini di teoria degli errori si può esprimere dicendo che ogni volta che io faccio una misurazione ho la certezza di ottenere un valore dalla misura, ma la correttezza di questo valore dipende da come è distribuito rispetto agli altri valori trovati, cioè da dove cade sulla curva gaussiana.

---

[37] Non ci interessano qui i passaggi matematici per dimostrare che l'area sottesa alla gaussiana vale 1, per i quali si rimanda a un qualunque testo di analisi matematica o di teoria dei fenomeni aleatori.

Questo breve sconfinamento nella teoria degli errori, che non è propriamente oggetto della nostra indagine, era utile per spiegare come mai l'approccio frequentista alla probabilità introduca nella definizione un numero di esperimenti infinito, anche se come limite. La definizione «probabilità dell'evento A è il limite di $f(A)$ per N tendente all'infinito» non ha molto senso, soprattutto alla luce della precisazione immediatamente successiva per cui «la precedente considerazione è empirica e non corrisponde ad alcun limite in senso matematico»[38]. Possiamo immaginare un limite che non sia matematico ma sia la risultanza di verifiche empiriche; riesce però difficile capire come calcolarlo operativamente, se "empiricamente" vogliamo spingerci verso l'infinito. Il fatto che calcolare la probabilità, secondo questa interpretazione, consista nell'operare una inferenza induttiva, rende la generalizzazione propria dell'induzione qualcosa di molto poco empirico.

Mi pare che la confidenza che la frequenza relativa si assesti attorno a un certo valore al crescere degli esperimenti, come postula la definizione dell'interpretazione frequentista, esprima una convinzione che nasce da considerazioni legate alla teoria degli errori: teoria certamente feconda di sviluppi teorici e applicativi, ma che sembra fondata ultimamente su una visione classica della probabilità.

Una certa difficoltà nella definizione di probabilità come frequenza di distribuzione, e più in generale nell'assumere l'interpretazione frequentista, emerge anche più evidente se prendiamo in esame l'altro grande ambito di applicazione della teoria delle probabilità, che non può essere ridotto all'interpretazione classica come il lancio di una moneta o il gioco dei dadi: le valutazioni statistiche riferite a contesti umani e sociali.

La pervasività della statistica ci ha abituato all'idea che sia normale ragionare in termini di probabilità per parlare di natalità, aspettativa di vita e flussi migratori,

[38] G. Galati e G. Pavan, *Teoria dei fenomeni aleatori*, cit., p. 21.

come non è pensabile fare a meno del calcolo delle probabilità per realizzare prodotti assicurativi, pianificare reti di trasporti o redigere la legge finanziaria di uno stato, solo per fare alcuni esempi tra i più comuni e macroscopici. Ma cosa significa chiedere quale sia la probabilità che una persona raggiunga gli 80 anni di età o la probabilità che la prossima estate si rechi in vacanza all'estero? Il tecnico addetto ai lavori risponderebbe subito che questo tipo di probabilità si calcola facilmente, ricorrendo a tavole di mortalità piuttosto che archivi storici delle prenotazioni fatte presso operatori turistici, eventualmente integrate con altri dati utili per circoscrivere la popolazione di riferimento su base geografica, sociale o patrimoniale, così da ricavare distribuzioni di eventi che ci permettono di calcolare la probabilità che ci interessa. La prassi effettiva, ci permette di ricondurre senza dubbio questo tipo di calcolo delle probabilità alla interpretazione frequentista.

Ma cosa significa effettivamente domandarsi quale sia la probabilità che una certa persona possa morire entro l'anno o la probabilità che uno specifico bambino nato oggi riesca a conseguire un titolo universitario entro il compimento del 25° anno di età? È chiaro che in questo caso il concetto di "esperimento ripetibile" cui abbiamo fatto riferimento finora deve essere inteso in modo diverso. L'individuo vivente, e in particolare l'uomo, si pone in contrapposizione alla ripetibilità sperimentale, a cominciare proprio dalla sua individualità: siamo tutti fermamente convinti di non avere "repliche", al punto di trovare quantomeno ridicola l'idea di chi si proponesse di girare il mondo per verificarlo direttamente, mettendosi a confronto con gli altri abitanti del pianeta.

Se parliamo della probabilità di avere un incidente utilizzando un determinato mezzo di trasporto, è chiaro che stiamo facendo riferimento alla frequenza di un certo tipo di evento, ma è altrettanto evidente che nel momento in cui l'impiego del mezzo diventa un viaggio effettivamente compiuto da un soggetto, questo avvenimento assume una connotazione che non è più generalizzabile: il viaggio può andar bene o male, ma non è chiaro cosa significhi affermare, durante il viaggio, che c'è una certa probabilità di fallimento; anche perché il valore di questa probabilità non si riduce mano a mano che il viaggio si approssima alla meta, e il viaggio è un'esperienza irripetibile per chi vi partecipa, benché costui possa poi fare altri viaggi sullo stesso percorso. E tuttavia, è frequente che il viaggiatore attribuisca una probabilità di successo alla sua impresa.

La critica di de Finetti in merito all'interpretazione frequentista va a cogliere proprio questo salto di prospettiva, tra l'esperienza puntuale del soggetto e la elaborazione che ne viene fatta in termini di frequenza:

> Ci sono impostazioni nelle quali il ruolo della frequenza diventa ancor più fondamentale, dal momento che essa viene inserita nella stessa definizione di probabilità. Secondo tali impostazioni, un evento sarebbe tanto più probabile quanto più fosse frequente. Per rendere plausibile questa definizione si deve intendere la parola «evento» in un senso generico, come termine generale di cui ogni evento singolo sarebbe una «prova». Ma codesto non è che un trucco verbale per parlare di «prove di uno stesso evento» anziché di eventi che essendo (sotto qualche punto di vista) analoghi, vengono ritenuti *a priori* ugualmente probabili. Parlare di «prove di un evento» anziché più semplicemente di eventi, è una mera convenzione terminologica, che non può, da sola, essere sufficiente per eliminare il carattere singolo e irripetibile di ciascuna «prova»[39]

De Finetti porta l'attenzione sul fatto che parlare in termini di frequenza della probabilità di eventi, che possono accadere o non accadere, lascia aperto il problema di come si possa misurare la probabilità di eventi che possono manifestarsi in "prove" (gli "esperimenti" della definizione frequentista) che mancano di caratteri di generalità, e non possono quindi essere interpretati in chiave frequentista senza snaturarli, o per lo meno leggerli in maniera riduttiva; ma a questo punto perde di senso il valore della probabilità ottenuta come frequenza relativa, se si vuole intendere che è il valore di probabilità dei singoli fatti per i quali ci sì è interrogati se l'evento è verificato o no.

Nondimeno, le compagnie assicurative vendono prodotti che sono indubbiamente

---

[39] B. de Finetti, *Filosofia della probabilità*, cit., p. 209.

redditizi, altrimenti smetterebbero di offrirli sul mercato, e per definire premi e garanzie da offrire nelle polizze, in modo da bilanciare la possibilità di venderle con l'eventualità di dover pagare dei rimborsi mantenendole redditizie, studiano l'incidenza dei sinistri sulla popolazione dei clienti potenziali, ricorrendo proprio a probabilità calcolate in base alla frequenza di distribuzione degli eventi oggetto di garanzia. Questa considerazione sembra indebolire l'obiezione di de Finetti, dal momento che la probabilità calcolata su una distribuzione di eventi viene poi impiegata con successo per prevedere l'incidenza dei sinistri sulle polizze vendute.

Un esempio portato da Cohen aiuta a chiarire la questione che potrebbe sembrare solo di lana caprina: Si ipotizzi che in un istituto accademico sia attivato un corso di storia del fagotto cui sono iscritti un certo numero di studenti. Se la frequenza relativa tra gli studenti che studiano storia del fagotto e il numero totale di studenti è uno a cinquecento, diciamo che la probabilità che uno studente scelto a caso studi storia del fagotto è 0,002 (1/500). Ma, una volta scelto un determinato studente, poniamo George,

> l'analisi frequentista non può autorizzare un valore, diciamo, 1/500 per la probabilità diadica che George sia uno storico del fagotto, dato che è uno studente. Una ragione perché non possa farlo è che un simile giudizio di probabilità non ha letteralmente senso in tale analisi. Sono le proprietà o le caratteristiche, e non le proposizioni – siano attorno a individui o a qualsiasi altra cosa – che accadono con certe frequenze relativamente alle classi di riferimento. Ed un'altra ragione è che la probabilità che George sia uno storico del fagotto, dato che è uno studente, può benissimo non essere eguale alla probabilità che, in generale, un individuo qualsiasi sia uno storico del fagotto dato che è uno studente. George può essere un caso speciale, perché suo padre è un noto solista di fagotto. Perciò la *sua* probabilità di essere uno storico del fagotto, se proprio deve essere identificata con una frequenza relativa, può essere identificata solamente con la frequenza relativa degli studenti di storia del fagotto fra gli studenti i cui padri sono famosi solisti di fagotto. E allora, poiché George è l'unico studente di quel tipo, sembrerebbe che noi non possiamo conoscere la probabilità che egli studi storia del fagotto senza conoscere se egli sia o non sia, di fatto, uno storico del fagotto. Ma ciò sarebbe paradossale e inaccettabile, perché noi otterremmo una probabilità affidabile se e solamente se fossimo al

> 100% certi che egli è o non è uno storico del fagotto. In breve, un'analisi frequentista non solo non permette che giudizi individuali di probabilità siano provvisti di significato, ma comporta delle difficoltà anche per ogni tentativo sistematico di considerarli come dei giudizi implicitamente generali.[40]

Secondo Cohen dunque, per tornare alla considerazione che si faceva subito prima, possiamo realizzare un prodotto assicurativo basato sull'analisi delle frequenze di accadimento di una certa proprietà (veicoli incidentati sulle strade pubbliche in una determinata regione), ma interrogarsi sulla probabilità (frequentista) di una proposizione come "Giovanni guida la sua automobile da Milano a Venezia senza avere incidenti", non ha semplicemente senso.

A questo punto mi pare che assuma nuova luce un'assunzione, tacita e indiscussa, nell'impostazione frequentista del calcolo delle probabilità: la frequenza relativa è tanto più accurata, nel senso che approssima meglio la probabilità che si vuole misurare, quanto più i "fatti" dove si misura l'accadimento che interessa sono omogenei. Questa omogeneità può essere assicurata dal modo in cui realizzo l'esperimento ripetibile, ad esempio il lancio di un dado, o dal modo in cui seleziono il campione (in senso statistico) di misura, ad esempio per cercare le probabilità di avere un incidente in auto prenderò in considerazione solo le auto sulle strade di una certa regione, e poi di queste quelle circolanti in un certo orario, escludendo i giorni festivi, e così via; ma la ricerca di omogeneità in ciò che si misura non è forse chiedere una uguale probabilità? La probabilità frequentista è tanto più precisa quanto più misurata su fatti equiprobabili, ma questo significa fare rientrare dalla finestra il principio di ragione insufficiente che ci pareva di aver già messo alla porta.

Nel procedimento che si segue per analizzare la frequenza di un evento, il principio in questione non pare avere un ruolo, e in effetti non costituisce una premessa metodologica, come invece nel caso della probabilità misurata secondo la

[40] L. J. Cohen, *Introduzione alla filosofia dell'induzione e della probabilità*, Giuffrè, Milano 1998, pp. 57-58.

definizione classica. Infatti se noi vogliamo misurare la distribuzione degli incidenti automobilistici, sulla base di quanto detto finora, non considereremo indifferentemente tutti i viaggi in automobile, ma di questi saremo noi a selezionare solo quelli che risultano omogenei, senza assumere nulla a priori. Per cui, se vogliamo essere particolarmente precisi, tornando all'esempio appena fatto, oltre a regione, orario e giorno della settimana, introdurremo criteri di selezione del campione in base al tipo di strada, condizioni atmosferiche, età del guidatore e sua esperienza di guida, livello di manutenzione del veicolo, ecc.

Quando avremo terminato di introdurre tutti i criteri che ragionevolmente possono produrre delle differenze nell'esito di un viaggio, e avremo misurato la frequenza relativa degli incidenti, otterremo un numero compreso tra 0 e 1. Fino a questo punto non abbiamo introdotto nessuna assunzione implicita particolare, ma nel momento in cui diciamo che il valore della frequenza rappresenta anche la probabilità di avere un incidente, facendo un viaggio con tutte le caratteristiche che corrispondono al campione selezionato, stiamo applicando il principio di ragione insufficiente: infatti, con un salto logico che non trova giustificazione nella metodologia applicata, affermiamo che tutti i viaggi con quelle particolari caratteristiche (e in particolare quelli che abbiamo inserito nel campione da misurare) hanno uguale probabilità, proprio perché non vediamo nulla che ragionevolmente ci permetta di distinguerli, allo stesso modo in cui, ai fini delle probabilità, non consideriamo differenti le sei facce di un dado anche se riportano una numerazione o sono colorate in modo diverso, benché la cosa ci permetta di distinguerle.

Il principio di ragione insufficiente fa nuovamente la sua comparsa, quindi, non mentre calcoliamo la frequenza relativa di un evento, ma nel momento in cui consideriamo il valore della frequenza come una probabilità, e in tale senso vogliamo utilizzarlo.

C'è un altro elemento che viene alla luce con questo esempio, e che vale la pena mettere in evidenza, in quanto mostra un carattere paradossale e in qualche modo contraddittorio dell'interpretazione frequentista. La definizione dice che "la frequenza relativa approssima un valore stabile al crescere del numero di esperimenti”, o di elementi presenti nel campione se non si tratta di un esperimento ripetibile. Questo principio posto dalla definizione, è stato pacificamente accolto dalla statistica, che per lavorare fa ricorso al calcolo delle probabilità, e proprio dalla statistica probabilmente lo abbiamo appreso la prima volta: il dato statistico è tanto più accurato e capace di prevedere un comportamento, proprio nella misura in cui viene calcolato sulla base di una popolazione numerosa.

Ma mentre affermiamo di aver bisogno di una base di calcolo il più ampia possibile, nella prassi operativa rimuoviamo la maggior parte dei casi (nell’esempio fatto: tipo di strada, orario, condizioni meteo, e tutto ciò che non corrisponde ai criteri che ci siamo dati), per assicurare omogeneità alla base di calcolo, al punto da dare l’impressione che si stia cercando di ricostruire le condizioni di un esperimento ripetibile anche quando ciò non sia possibile, col risultato di inseguire l’ideale del limite per un numero di esperimenti che tende all’infinito, operando una progressiva selezione il cui limite, all’aumentare della nostra capacità di rilevare differenze, restituisce il caso singolo: esattamente l’opposto di quel che richiede la misura di una distribuzione per restituire un valore che abbia senso. Infatti la probabilità frequentista, abbiamo visto, è definita come limite della distribuzione relativa quando il numero di esperimenti ripetibili tende a infinito; se lanciamo un dado indagando la probabilità che, ad esempio, esca un numero pari, compiamo, almeno idealmente, un esperimento ripetibile; e dato che il valore della distribuzione tende a stabilizzarsi intorno al valore della probabilità al crescere dei tentativi (come posto esplicitamente dalla definizione), possiamo migliorare l’accuratezza della misura semplicemente ripetendo più volte il lancio, con il solo limite del tempo che siamo disposti a dedicare al reiterarsi delle prove. La scienza statistica, fondata su una concezione essenzialmente frequentista del calcolo delle probabilità, ha da subito recepito l’indicazione e l’importanza di disporre di dati in grande quantità, ma il più delle volte si è trovata a dover misurare frequenze di eventi che ben difficilmente si inquadrano in un contesto di “esperimento ripetibile”: se si vuole misurare la probabilità di vivere fino a 80 anni, non si può certo far vivere una persona molte volte per avere una distribuzione dell’età della morte, anche se non volessimo dare un limite al tempo che siamo disposti a dedicare alle ripetizioni; pure nel caso dei viaggi in automobile, abbiamo visto che fare un secondo viaggio sullo stesso percorso non si può immediatamente identificare come la ripetizione di uno stesso esperimento. Per ovviare a quello che è evidentemente un limite pratico nell'applicazione della teoria, si prendono in considerazione, per restare sugli esempi fatti, vite di persone diverse e diversi viaggi percorsi in auto; sulla base di elementari considerazioni, però, è immediatamente evidente che la vita di una persona nata e cresciuta in una metropoli dell'occidente industrializzato e quella di una persona nata e cresciuta in un paese del terzo mondo, non possono in alcun modo essere considerate alla stregua di una ripetizione: sono troppo profonde le differenze di condizioni igieniche, alimentazione, accesso all'istruzione, lavoro, reddito, sistema sanitario, e così via. Insomma, o rinunciamo all'idea di un esperimento ripetibile, anche solo per analogia, e quindi abbandoniamo l'idea di una probabilità frequentista, oppure teniamo i due casi in serie di misure diverse. Ma essere attenti a una certa omogeneità nei casi, per rispettare il dettato di esperimento ripetibile, porta ad escludere dalla misura un certo numero di quei casi che si hanno a disposizione; numero tanto più grande quanto più siamo accurati nell'esaminare le differenze significative tra un caso e l'altro, con la conseguenza di avere un numero sempre più piccolo di "esperimenti ripetibili" da misurare. Insomma, per rispettare la definizione dovremmo raccogliere il maggior numero di casi, ma per farlo in modo da rispettare la definizione ci troviamo a ridurne continuamente il numero.

Ci si può chiedere, a questo punto, se l'interpretazione frequentista della probabilità non rappresenti effettivamente una valida teoria degli errori – nello specifico errori di misura – e solo in via secondaria, e in un modo che deve essere precisato meglio, una teoria delle probabilità. Abbiamo già accennato al fatto che la misura della frequenza si lega alla accuratezza della misura e alla teoria degli errori, parlando della curva gaussiana. Questo legame non è occasionale e accompagna costantemente lo sviluppo della teoria delle probabilità in chiave frequentista:

> Ma proprio alla fine dell'Ottocento von Bortkiewicz scoprì che un'altra legge di probabilità, quella che Poisson aveva dedotto come limite delle distribuzioni di esperimenti binomiali caratterizzati da probabilità molto piccole, trovava riscontro nella distribuzione di un fenomeno casuale, quello dei morti per calcio di cavallo

> nell'esercito prussiano rispetto ai corpi d'armata ed agli anni. Come se non bastasse, Pearson si era accorto che l'adattamento della curva di Gauss ai risultati sperimentali era molto minore di quanto avevano entusiasticamente supposto Quetelet e Galton.
>
> Ripetute rilevazioni sperimentali lo avevano convinto che nelle distribuzioni empiriche erano molto frequenti lievi asimmetrie, cosa che rendeva necessaria l'individuazione di nuove curve atte a rappresentare siffatte distribuzioni. Uno dei grandi meriti di Pearson è proprio di aver elaborato dodici tipi di curve di frequenze – termine che Pearson con pochissime eccezioni usò al posto di curva di probabilità, sostituzione questa che diventerà usuale fra gli statistici così da rendere palese anche nella terminologia il mutamento in atto nei fondamenti della statistica – con l'intento di fornire nuovi modi di descrivere le frequenze con cui si presentavano le varie modalità delle grandezze osservate, fossero esse altezze barometriche, misure fenotipiche di granchi oppure durate di matrimoni.[41]

Si noti, nella citazione precedente, l'equivalenza terminologica tra curva di frequenza e curva di probabilità. Questa meticolosità nel ricercare una misura sempre più accurata, un dettaglio sempre più puntuale di come si manifestano i fenomeni, mi pare indice del fatto che l'approccio frequentista sia effettivamente e principalmente una tecnica di misura; in particolare di misura degli errori. L'idea di probabilità associata all'interpretazione frequentista, è innanzitutto la probabilità di misurare correttamente. La meticolosa ripetitività della misura per identificare il modello che rappresenti al meglio i dati raccolti, produce delle curve che Pearson giustamente chiamava "di frequenza", in quanto descrivono ciò che è già stato, e che chiamiamo "di probabilità" nel momento in cui ci rappresentiamo un futuro che per lo più, ne siamo fiduciosi, replicherà quello che è già stato.

Che questa traslazione temporale sia almeno problematica, lo mostra il fatto che la statistica non è mai soddisfatta delle informazioni raccolte, e le proiezioni di probabilità vengono sempre sottoposte nuovamente alla verifica frequentista. Un

---

[41] D. Costantini e L. Geymonat, *Filosofia della probabilità*, cit., pp. 34-35.

esempio per tutti: nel 2019 l'esperto statistico più bravo del mondo, cui fossero stati messi a disposizione tutti i dati relativi ai flussi di traffico della rete stradale, con tutte le serie storiche, non sarebbe riuscito a fare una previsione minimamente attendibile sul traffico nel 2020. Nessuno avrebbe ritenuto veramente possibile che la tangenziale di Milano potesse essere deserta durante l'orario del cosiddetto "coprifuoco" o che sarebbe stato possibile vedere il fondo dei canali nella laguna di Venezia. A chi obbiettasse che la pandemia del 2020 è stato un evento dalla unicità imprevedibile, non si potrebbe che dare ragione e ringraziarlo per averci confermato la problematicità della questione: probabilità implica casualità, vale a dire una reale possibilità di trovarci spiazzati ed essere smentiti nelle nostre previsioni più prudenti; proprio quello che manca, e addirittura si vuole escludere, in una analisi della frequenza di ciò che è stato per fare ipotesi su ciò che sarà.

Il continuo aggiustamento delle informazioni da raccogliere e dei criteri impiegati per selezionarle, con la conseguente variazione della frequenza calcolata, per alcuni rappresenta l'inevitabile imponderabilità che il caso, o ciò che è imprevedibile nella sua intenzionalità, introduce nella ripetizione della misura; per chi ha una concezione neopositivista, o comunque strettamente determinista, questa variabilità è indice del progredire delle nostre conoscenze e del nostro modo di gestire la parte che ignoriamo delle leggi che regolano il mondo. Se davvero vogliamo fare valere nel suo significato deterministico la frase di Laplace – «la probabilità è relativa in parte a questa ignoranza, in parte alle nostre conoscenze»[42]–, le diverse curve di probabilità individuate da Pearson nel suo lavoro, come tutti i modelli probabilistici che descrivono il mondo in modo sempre meno generico e più particolareggiato, sono segnale della nostra conoscenza che conquista nuovo terreno prima dominato dall'ignoranza. Questo progresso si sviluppa un passo dopo l'altro, senza alcun limite o termine apparente, allo stesso modo della successione dei numeri naturali verso l'infinito. Resta comunque aperta la questione se questa potenziale infinità della conoscenza scientifica possa dare

[42] P.S. Laplace, *Essai Philosophique Sur Les Probabilités*, cit., p. 7. Traduzione mia.

ragione della sconfinata vastità di tutto il nostro mondo brulicante di vita, anche volendo dedicarvi l'infinita pazienza che una tale progressione richiede, quando sappiamo già in partenza che l'infinito dei numeri naturali non esaurisce tutto l'infinito della sola matematica[43].

Una delle difficoltà dell'interpretazione frequentista, poi, è quella di suggerire che quel che è altamente improbabile sia impossibile:

> Corrado Gini (il fondatore dell'Istituto centrale di statistica), uno statistico che faceva spesso osservazioni acute, per lo più rivolte contro la tendenza ad attribuire ai risultati della teoria delle probabilità una certezza che non hanno, soleva portare il seguente esempio. C'è una probabilità molto piccola – diceva – che comprando un biglietto di lotteria, si acquisti proprio il biglietto vincente: se, per esempio, v'è un milione di biglietti, tale probabilità è un milionesimo. Se si ammette che un milionesimo è una probabilità così piccola che un evento con quella probabilità debba essere considerato come praticamente impossibile, allora se ne dovrebbe concludere che, con pratica certezza, colui che ha vinto alla lotteria lo ha fatto in base a un trucco e dovrebbe pertanto venire arrestato! Ragionamenti come questo, non si limitano a mostrare quanto siano strampalate certe vedute purtroppo assai diffuse, ma mostrano anche (sebbene questa conclusione che traggo non coincida con il punto di vista degli statistici, Gini incluso) che la probabilità non può identificarsi con la frequenza.[44]

Insomma, da un punto di vista frequentista, non soltanto eventi altamente improbabili possono verificarsi facilmente, ma in casi come quello della lotteria si verificano certamente.

---

[43] Cantor ha dimostrato che esistono diverse "grandezze" infinite. La numerosità degli elementi (cardinalità) di un insieme infinito non può essere ricondotta a un generico "infinito" che li denoti indistintamente, in quanto si dimostra matematicamente che ci sono infiniti "più grandi" di altri. In questo ordinamento di insiemi infiniti maggiori di altri, quello dei numeri naturali è il più piccolo.

[44] B. de Finetti, *Filosofia della probabilità*, cit., pp. 144-145.

3.3. Assiomatizzazione

La formalizzazione degli assiomi della teoria della probabilità rappresenta il punto di arrivo di un lungo sforzo che aveva l'obbiettivo di unificare le differenti interpretazioni e dare alla prassi del calcolo delle probabilità, che aveva da subito iniziato a produrre metodi di calcolo e teoremi, il fondamento più rigoroso che si possa dare in ambito matematico, ovvero quello assiomatico. Questo sforzo per ricondurre tutto sotto una comune radice, come si era già fatto notare, ha richiesto quasi tre secoli di riflessioni e tentativi, e nel 1933, finalmente, Kolmogorov propose un insieme di assiomi che incontrarono rapidamente il favore dei matematici e rimasero sostanzialmente invariati fino a oggi.

Ne riprendiamo la definizione dal già citato manuale di Galati e Pavan:

> La teoria assiomatica, introdotta da *A. N. Kolmogorov* nel 1933 […] si basa sulla definizione di alcune caratteristiche (assiomi) che deve possedere la probabilità $P(A)$ dell'evento A.
>
> Essa deve soddisfare i seguenti tre assiomi:
>
> **I** $P(A)$ è un numero non negativo: $P(A) \geq 0$.
>
> **II** L'evento certo S ha probabilità unitaria: $P(S) = 1$.
>
> **III** Se due eventi A e B non hanno elementi comuni (sono *incompatibili* o *disgiunti*) la probabilità dell'evento unione è uguale alla somma delle probabilità dei singoli eventi: $P(A \cup B) = P(A) + P(B)$.
>
> L'assioma III comporta che, se $A_1, A_2, \dots, A_n$ sono disgiunti, la probabilità dell'evento somma è uguale alla somma delle probabilità degli eventi singoli:
>
> $$P(A_1 \cup A_2 \cup \dots \cup A_n = P(A_1) + P(A_2) + \cdots + P(A_n)$$
>
> Ciò va esteso al caso di un numero infinito di eventi:
>
> $P(A_1 \cup A_2 \cup \dots = P(A_1) + P(A_2) + \cdots$ (assioma **III** bis dell'additività infinita).[45]

---

[45] G. Galati e G. Pavan, *Teoria dei fenomeni aleatori*, cit., p. 20.

La prima cosa da precisare è che gli assiomi della teoria della probabilità non hanno lo scopo di strutturare l'intero sistema di calcolo, infatti non sostituiscono gli assiomi della matematica, che continua ad esistere legata ai soliti principi (assiomi); questo è particolarmente evidente se si considera che Bayes e Poisson formularono e dimostrarono teoremi, tuttora impiegati nel calcolo delle probabilità, ben prima che di questo ne venissero formulati gli assiomi. Questi assiomi non definiscono un sistema che si ponga accanto alla matematica, come invece accade nel caso della geometria. Essa gode di una relativa autonomia non triviale, infatti basata su assiomi che non presuppongono immediatamente quelli della matematica: possiamo ragionare di punti, linee e piani senza fare necessariamente calcoli, così come possiamo fare considerazioni sulla congruenza di segmenti e figure o sulle tipologie di angoli e le loro caratteristiche senza dover ricorrere, fino a un certo punto, alla matematica, anche se nessuno si accontenterebbe di una geometria priva della potenza del calcolo.

Non così per gli assiomi della probabilità. Già dal primo si fa esplicito riferimento ai numeri e ad un ordinamento tra di essi: la probabilità di un evento è un numero maggiore o uguale a 0. La funzione degli assiomi non è quindi quella di definire un sistema parallelo alla matematica, che a questa si affiancherebbe, ma di "ritagliare", se così possiamo dire, all'interno della matematica stessa un territorio chiaramente delimitato, che costituisca l'ambito di riferimento in cui spazia il calcolo delle probabilità.

L'altro elemento importante da notare, è che non viene data alcuna definizione di cosa sia un evento né cosa sia la probabilità che di quell'evento si predica.

Già parlando di "casi favorevoli" e "possibili", e ancor più parlando di "esperimenti ripetibili", non era stata data una definizione precisa di cosa dovesse intendersi per "caso" o per "esperimento". Non deve però sfuggire un aspetto fondamentale di quella che abbiamo visto essere un'ambivalenza finora ineludibile, parlando di probabilità: quando parliamo di casi che possono verificarsi lanciando un dado, ci pare di afferrare facilmente il concetto, anche se poi approfondendo la questione non è più tutto così pacifico; anche parlando di esperimenti ripetibili ci pare che, da Galileo in poi, l'idea di cosa sia una verifica sperimentale sia stata acquisita chiaramente, benché, a un esame più attento, le cose si complichino.

Ma "casi" ed "esperimenti" sono concetti cui facciamo ricorso per definire nella maniera più precisa possibile cosa sia la probabilità. Se poi questa non ne esce delineata in modo netto, ciò può essere dovuto alla natura del problema oppure al fatto che ricorriamo ad un linguaggio che è per natura sua ricco di sfaccettature; è però evidente, tanto nella definizione classica quanto in quella frequentista, l'intenzione di afferrare, delimitare e comunicare cosa si debba intendere per probabilità.

Nella teoria assiomatica delle probabilità, invece, si abbandona intenzionalmente ogni tentativo di definire cosa sia un evento e cosa sia la probabilità; infatti questi termini si suppongono dati nel loro significato, al punto che vengono impiegati per definire gli assiomi, ma quale sia questo significato è assolutamente irrilevante ai fini della coerenza del sistema assiomatico.

> La grande innovazione che Kolmogorov introdusse nelle ricerche sui fondamenti della probabilità e, indirettamente, della statistica fu di abbandonare ogni pretesa volta a sviluppare la teoria delle probabilità a partire da una delle definizioni di probabilità: il dogmatismo laplaciano. Questa teoria doveva essere affrontata assiomaticamente, sviluppata cioè a partire da un gruppo di condizioni la cui validità non venne da lui né posta in discussione né, tantomeno, giustificata.
>
> Da allora, questa separazione fra teoria delle probabilità e ricerche sulla definizione di probabilità divenne categorica. Naturalmente ciò non significa che le condizioni scelte quali assiomi lo furono in modo arbitrario; infatti esse erano quelle che da ormai tre secoli guidavano ogni ricerca sulla probabilità e sulla statistica. Lo stesso Kolmogorov si impegna filosoficamente, almeno in modo implicito, nel senso che l'unica definizione che prende in considerazione nella sua monografia è la frequentista. Tuttavia questa definizione è presentata come un esempio di applicazione dei suoi assiomi. Ogni giustificazione di questi, e quindi anche quella che potrebbe venirgli dall'interpretare la probabilità come una

frequenza relativa, è esplicitamente respinta.[46]

Si potrebbe discutere sul fatto che fare riferimento all'interpretazione frequentista, per mostrarne la conciliabilità con gli assiomi, costituisca effettivamente un decidersi, ancorché implicitamente, per una posizione specifica. Per il resto, tutti concordano: lo scopo dell'assiomatizzazione era proprio quello di fornire un fondamento rigoroso alle tecniche di calcolo della probabilità che, indipendentemente dalla teoria cui facevano riferimento, erano ormai diffuse e ampiamente utilizzate; e farlo in modo da non doversi schierare per nessuno dei possibili significati di "probabilità".

Non si deve pensare, però, che stabilire principi e regole per operare sull'entità concettuale "probabilità" rinunciando a definirla costituisca un *unicum* nel campo dell'assiomatizzazione. L'intera matematica oggi è fondata sulla teoria assiomatica degli insiemi, che fa riferimento al sistema di assiomi ZFC: Zermelo-Fraenkel più l'assioma della scelta (la C di ZFC sta infatti per *Choice*). I testi di algebra che espongono la teoria ZFC precisano che gli oggetti della matematica sono definiti partendo da alcune nozioni *primitive*, a volte per ciò stesso considerate *semplici* e *intuitive* quanto a significato, e tra queste nozioni vi è quella di "insieme", che non viene perciò definita e per il cui significato si rimanda alla conoscenza che dovrebbero averne tutti in quanto nozione primitiva. Da notare che la teoria assiomatica degli insiemi ZFC non solo non dà una definizione di insieme, ma pur costituendo il fondamento di tutta la matematica non menziona in nessun modo i numeri; e sì che storicamente la matematica senza gli insiemi è esistita, mentre senza i numeri non si dà. L'elemento fondante è rappresentato dal fatto che i numeri possono essere considerati gli elementi di un insieme, e quindi definiti in questi termini; non è tuttavia una questione intuitiva nel senso di semplice, come mostra la riflessione filosofica sul *numero* fin da Pitagora.

[46] D. Costantini e L. Geymonat, *Filosofia della probabilità*, cit., pp. 46-47.

Anche la geometria è un sistema assiomatico che prescinde dalla definizione degli oggetti propri: il punto, la retta, il piano.

Naturalmente, se le scienze positive possono ignorare il significato e la natura ultima di ciò che indagano, la riflessione filosofica non può cessare di interrogarsi anche in merito alle questioni cui altri rinunciano intenzionalmente; capita infatti che l'astensione non sia un'opzione, e manifestare l'intenzione di non pronunciarsi sia già una presa di posizione:

> A scanso di equivoci nel libro di Kolmogorov non si trova nessuna definizione di probabilità, ma solo come usarla in modo coerente. Per alcuni questo suona non soddisfacente, lascerebbe infatti senza risposta la domanda: «cos'è la probabilità?». Non è affatto ovvio che sia veramente necessario rispondere a tale questione; qualcuno è forse in grado di dare una definizione precisa di lunghezza? Questa (necessaria) difficoltà nel definire in modo preciso i concetti fondamentali della scienza è stato espresso in modo chiaro dal fisico John A. Wheeler (1911-2008): «La probabilità, come il tempo, è un concetto inventato dagli esseri umani, e gli esseri umani devono assumersi la responsabilità delle oscurità che lo circondano.»[47]

Se davvero probabilità e tempo fossero solo invenzione umana, potremmo aspettarci di riuscire a congegnarli in modo da rendere possibili i viaggi nel tempo e i sistemi di gioco al lotto con vincita garantita. Si tratta di obiettivi per i quali c'è sempre qualcuno che si industria, ma per il momento l'ingegno umano è riuscito solo a progredire dagli orologi ad acqua o a sabbia agli orologi atomici, e la responsabilità delle oscurità che circondano l'essere umano non paiono tutte in capo a lui, come invece ritiene Wheeler.

Quanto alla teoria assiomatica delle probabilità in sé, non vi è molto altro da dire,

---

[47] A. Vulpiani, *Caso, probabilità e complessità*, Ediesse, Roma 2014, p. 199.

se non interessa specificamente i teoremi che da questa si deducono e la maniera in cui si sviluppano le applicazioni pratiche e i metodi di calcolo che ne derivano. Essa è vuota di significato e, come abbiamo visto, lo è intenzionalmente.

Questa intenzionalità, assente in altre teorie assiomatiche, benché anch'esse trascurino di definire le nozioni primitive, mi pare che possa rappresentare un potenziale fattore di indebolimento della capacità speculativa della scienza positiva. C'è il rischio che il vuoto "forzato" di significato degli assiomi, svuoti anche il significato della definizione classica di probabilità, la cui circolarità nella definizione viene di fatto liquidata come una conseguenza della limitazione linguistica di una formulazione inadeguata. Allo stesso modo ci si libera delle difficoltà che reca con sé il principio di indifferenza piuttosto che le altre associate all'interpretazione frequentista viste prima: la nozione di probabilità e di evento sono intuitive come quelle di numero, punto e retta, e capiamo tutti di cosa stiamo parlando anche se non ne troviamo una definizione soddisfacente. Il risultato potrebbe essere un'indifferenza rispetto alla possibilità della teoria di descrivere effettivamente il mondo reale.

Non va poi trascurata quella che mi pare una differenza significativa tra la teoria delle probabilità e le altre teorie assiomatiche: guardando la storia del calcolo delle probabilità, chi ha dato una definizione di probabilità non si è mai mostrato insoddisfatto di come ne esponeva il concetto, salvo poi trovarsi a discutere con altri che avanzavano diverse formulazioni, per difendere la propria posizione; chi prova a spiegare le nozioni primitive della matematica o della geometria, invece, procede esitante, ritoccando continuamente la definizione, perché è lui, non altri, ad essere insoddisfatto della formulazione data. Questo potrebbe essere indice del fatto che "probabilità" ed "evento" non sono tanto *nozioni primitive* rispetto agli assiomi, quanto piuttosto termini sui quali non si è ancora trovato un accordo. In altre parole, l'ambivalenza della *probabilità* potrebbe essere dovuta, almeno in parte, al fatto di impiegare lo stesso termine per dire cose diverse.

3.4. Interpretazione soggettivista

Può sorprendere che un argomento così matematizzato come la teoria della probabilità annoveri tra le proprie interpretazioni una che fa esplicito riferimento al soggettivismo. Se ci si riflette, una componente soggettiva è sempre esistita nella valutazione delle probabilità, almeno nella cerchia delle bische: quando il Cavaliere De Méré si rivolgeva a Pascal, per capire come mai un gioco ai dadi che aveva ideato per attirare scommettitori non gli era favorevole come uno escogitato in precedenza, si era fatto certamente un'idea di cosa poteva ottenere in un caso e nell'altro, in termini di scommesse e di vincite. È meglio scommettere puntando sull'uscita di un sei lanciando un dado per quattro volte, o sull'uscita di un doppio sei lanciando due dadi per 24 volte? Il Cavaliere De Méré si era fatto l'idea che fossero entrambe a lui vantaggiose, e una certa possibilità di vantaggio avranno pensato di averla anche gli scommettitori che sfidavano al gioco il Cavaliere. Il ricorso al genio di Pascal mostra che le convinzioni di De Méré non avevano riscontro alla prova dei fatti.

Costantini e Geymonat condividono la presenza dell'elemento soggettivo e la sua presenza per lo più latente all'inizio:

> È certamente molto difficile, ed in qualche misura arbitrario, individuare con precisione il sorgere di questo modo di intendere la probabilità [l'interpretazione soggettivista] il cui discostarsi dal logicismo può, forse, essere individuato nella differenza fra il "grado delle possibilità" di Leibniz e il "grado di certezza" di Bernoulli. Certamente la differenza, che all'inizio è quasi inesistente, diventerà via via più marcata fino a rappresentare una sorta di spartiacque negli studi moderni di filosofia della probabilità.[48]

Gli esponenti più importanti di questa interpretazione sono Frank P. Ramsey e Bruno de Finetti, entrambi nati all'inizio del XX secolo; la probabilità soggettivistica fiorisce dunque in concomitanza con l'assiomatizzazione. Potrebbe quasi sembrare

---

[48] D. Costantini e L. Geymonat, *Filosofia della probabilità*, cit., pp. 21-22.

che, nel momento della formalizzazione e del rigore scientifico al massimo grado, per una qualche reazione di compensazione si sia manifestata al massimo grado anche l'interpretazione meno scientifica.

Il termine "soggettivo" evoca sicuramente una opinabilità, una prevalenza della *doxa* che potrebbe fare intendere la probabilità soggettivista come non era nelle intenzioni dei suoi più autorevoli teorizzatori; lo stesso Alberto Mura, che scrive l'introduzione a *Filosofia della Probabilità* di de Finetti, espone il punto di vista dei suoi massimi teorizzatori in questi termini: «Sia Ramsey che de Finetti (riprendendo un'idea che risale alle origini stesse del concetto di probabilità) ritennero che i gradi di credenza, in quanto grandezze psicologiche, possano essere desunti dal comportamento in condizioni di incertezza»[49]. Indicare i gradi di credenza come "grandezze psicologiche" potrebbe non essere una buona scelta dei termini, dato che lo psicologismo rafforza l'idea di una soggettività capricciosa e volubile.

De Finetti esprime in questi termini la questione:

> Per conto mio […] la probabilità ha soltanto un senso soggettivo. Ritengo cioè che non abbia senso chiedere quale probabilità abbia di per sé, in astratto, un evento. Per «evento», d'altro canto, intendo un caso singolo ben determinato. È un evento, per esempio, il fatto che il rapido proveniente da Milano oggi sia arrivato con un ritardo compreso fra trenta e trentacinque minuti. E se parlo della probabilità di questo evento, mi riferisco al caso singolo indicato e non, per esempio, alla «probabilità di ritardi in generale per tutte le linee» o «per tutti i giorni». Per «evento», quindi, intendo una circostanza di cui si può sapere se si è verificata o non si è verificata.[50]

Salta subito all'occhio che "soggettivo" è usato in contrapposizione al procedimento di astrazione generalizzante, presente nelle altre concezioni di probabilità (almeno

[49] B. de Finetti, *Filosofia della probabilità*, cit., p. 13.
[50] *Ivi*, p. 63

secondo de Finetti); Inoltre ci si premura di definire nel modo più chiaro possibile cosa sia un "evento", invece di demandarne la comprensione a una intuibilità di validità universale, come nel caso dell'assiomatizzazione. La probabilità è riferita esclusivamente a situazioni concrete e particolari, che non sono valutate da un dispositivo che svolga un calcolo matematico, ma da una persona – concreta come la situazione valutata – , quindi da un "soggetto". Il soggettivismo della probabilità non è dunque una reazione di pancia allo sforzo teso ad assiomatizzare; si tratta invece di una concezione che, collocando la probabilità in una situazione fattuale, recupera la figura del soggetto, ristabilendo un confine alla progressiva oggettivazione illimitata del mondo, operata da una certa idea della scienza.

La probabilità soggettivista è una posizione che si può non condividere, ma che non dovrebbe essere liquidata frettolosamente come banale o ingenua, anche dal punto di vista filosofico, soprattutto se si considera che Ramsey e de Finetti erano in primo luogo dei matematici; matematici che non hanno mai avuto ripensamenti sulla matematica e i suoi metodi, ma che nel caso della probabilità si sono sentiti in dovere di riportare in primo piano il soggetto.

Lasciamo ancora la parola a de Finetti per chiarire meglio come vada intesa la probabilità:

> Il punto di vista che sostengo si basa sulla tesi che non ha senso parlare della probabilità di un evento se non in relazione all'insieme di conoscenze di cui una persona dispone. E se proprio volessimo parlare di una probabilità oggettiva, allora dovremmo dire che ogni evento ha probabilità 0 se non si verifica e probabilità 1 se si verifica. Pertanto, dato che noi non sappiamo se la probabilità oggettiva sarà 0 oppure 1, facciamo una stima di essa, a seconda che propendiamo più per l'uno o per l'altro valore. […] Come si può definire la probabilità? Come si può dare un significato alla parola «probabilità», se non si accetta né l'uno né l'altro dei due modi più comuni di definirla [ovvero la definizione classica e quella frequentista]? C'è invece un modo, che ritengo sia l'unico, il quale permette di dire esattamente che cosa si intende per «probabilità». Si tratta del metodo delle *regole di penalizzazione appropriate (proper scoring rules)*, le quali consistono nel domandare ad un individuo (chiamiamolo «X») qual è la probabilità che egli attribuisce ad un evento E, con

> l'avvertenza che riceverà una penalità in dipendenza e della risposta che darà e del valore della «probabilità oggettiva» di E (nel senso più sopra precisato: 0% se E è falso, 100% se E è vero). Supponiamo che X risponda dicendo, per esempio, «40%». Se non ci fosse alcuna penalizzazione, A potrebbe dire quel che gli pare, senza che per lui ne derivi alcuna conseguenza negativa. Invece, in presenza di una regola di penalizzazione appropriata, A sarebbe indotto dal suo stesso interesse – che è quello di rendere minima la previsione della penalizzazione – a dare la risposta esatta.[51]

Con la precisazione che:

> Parlare della probabilità di un evento *tout court*, senza alcuna qualificazione, non ha alcun significato concreto. Si deve invece aver presente che la probabilità è sempre relativa allo stato di conoscenza in cui si trova chi giudica.[52]

La probabilità soggettiva non è dunque tale perché chiunque può dire la prima cosa che gli passa per la mente, ma perché il soggetto propende per una certa possibilità che un certo evento si verifichi basandosi sulle conoscenze di cui dispone, ed è trattenuto dal pronunciarsi a caso dal meccanismo di penalizzazione. Le conoscenze impiegate per valutare la probabilità soggettiva, sono del soggetto anche quando questi ricorre ad informazioni che provengono da altri, dato che la valutazione dell'attendibilità dell'informazione e il suo impiego in misura maggiore o minore stimando la probabilità sono ancora in capo al soggetto.

La regola di penalizzazione appropriata, secondo de Finetti, «è esattamente pari al quadrato della differenza tra la probabilità dichiarata e quella oggettiva»; si tratta della regola di Brier, usata spesso nella valutazione di accuratezza delle previsioni meteorologiche. Altri fanno riferimento al giusto prezzo che si è disposti a pagare

[51] *Ivi*, pp. 64-65]
[52] *Ivi*, p. 101

per avere un premio se si verifica un certo evento, e la penalizzazione è rappresentata da quel che si è pagato se l'evento non si verifica; si tratta di richiami alla definizione di probabilità data da Bayes:

> DEFINIZIONE 5. La probabilità di qualsiasi evento è il rapporto tra il valore al quale dovrebbe essere calcolata un'aspettativa che dipende dall'accadimento dell'evento, e il valore della cosa attesa nel momento in cui sta accadendo.[53]

Ad esempio, se voglio andare in ferie negli Stati Uniti, devo presentare una domanda di visto turistico che è soggetta ad approvazione. Per partire sono disposto a pagare un biglietto che costa 100 (nell'esempio possiamo trascurare l'unità di misura) ma devo acquistarlo subito per trovare posto, e ho dei dubbi sul fatto che mi venga rilasciato il visto; proprio per l'incertezza della situazione potrei quindi non esser disposto a pagare più di 50 il biglietto. L'evento è l'approvazione della domanda di visto, la probabilità che lo approvino è 50/100=1/2.

Mura, nel suo saggio introduttivo a *Filosofia della probabilità,* riformula in questi termini la definizione di Bayes:

> DEFINIZIONE 1. Siano E un evento, B un oggetto di valore v e sia v'il valore dell'offerta «B se si verifica E, nulla altrimenti». Dicesi «probabilità di E» il rapporto $\frac{v'}{v}$

La probabilità come prezzo che si è disposti a pagare, richiama l'argomento delle scommesse sportive o da bisca, da cui eravamo partiti parlando di interpretazione soggettivista, e anche de Finetti nei suoi lavori fa riferimento alla probabilità sia in

[53] T. Bayes, *An Essay towards Solving a Problem in the Doctrine of Chances*, in «Philosophical Transactions», 53, 1763, pp. 376.

termini di penalizzazioni appropriate sia in termini di prezzo:

> Il modo più semplice – sebbene qualcuno possa ritenerlo non abbastanza elevato – di caratterizzare sia la probabilità sia la previsione è quello di pensarli come *prezzo* di una scommessa.[54]

Ma c'è un punto da chiarire per evitare possibili malintesi. Parlando di componente soggettiva nella valutazione di una probabilità è facile che vengano alla mente le puntate al casinò o le scommesse sui cavalli, forse per la caratteristica impulsività e gli insondabili processi mentali di chi scommette, come è vero che l'interpretazione soggettivista chiama inevitabilmente in causa le scommesse a motivo del cosiddetto "allibramento olandese", che vedremo tra poco. Però, il Cavaliere de Méré non stava affidandosi a una valutazione soggettiva delle probabilità quando chiamava in causa Pascal, anzi, stava cercando di trovare il valore corretto di una probabilità concepita in modo essenzialmente classico, anche se per impostarne correttamente il calcolo ha avuto bisogno dell'amico matematico. De Finetti riterrebbe comunque soggettiva la probabilità stimata da de Méré, non per come viene calcolata, a motivo della circolarità della definizione classica, ma per la valutazione che lo porta a disporre di un dado che si comporterà in un certo modo (non truccato) nel momento in cui lo sta lanciando (si rammenti la definizione di "evento"), come non è ritenuta oggettiva la probabilità ricavata da lanci ripetuti:

> bisognerebbe distinguere il concetto di previsione di frequenza da quello di probabilità, distinzione che di solito le teorie che si basano sulla definizione statistica (le quali confondono la probabilità del caso singolo con una media) non fanno. Non si può dire che quella media è la probabilità, perché se si facesse la media delle probabilità di vittoria su tutte le partite, si otterrebbe un valore che

[54] B. de Finetti, *Filosofia della probabilità*, cit., p. 253.

> sarebbe privo di valore pratico rispetto al caso singolo. Naturalmente, si può ben dire che si tratta di una media. Ma se essa venisse intesa come probabilità, servirebbe soltanto ove si avesse l'obbligo di attribuire a tutte le partite la stessa probabilità (ma allora non avrebbe significato neanche giocare).[55]

Per il resto, de Finetti ritiene che la capacità dei giocatori abituali di valutare correttamente la probabilità sia facilmente viziata da ragioni (queste sì squisitamente psicologiche) decisamente verosimili:

> I giocatori, per esempio (quando non sono degli illusi, in quanto sopravvalutano le possibilità che hanno di vincere - specialmente se, come può capitare benissimo, hanno per caso vinto la prima volta che hanno giocato), giocano perché provano piacere a rischiare e sono indotti spesso a fare scelte che essi stessi giudicano non convenienti sotto il profilo del guadagno. Il comportamento di questi giocatori non è necessariamente incoerente. E tuttavia le regole di penalizzazione, applicate a queste persone, non fornirebbero la risposta corretta, dato che tali persone non avrebbero come esclusivo interesse quello di minimizzare la previsione della penalizzazione.[56]

Per de Finetti, comunque, la probabilità soggettiva non è priva di qualsivoglia vincolo, ed è richiesto che non violi il calcolo delle probabilità. La stima della probabilità che viene fatta dal soggetto è in ogni caso libera, ed è indotta ad essere onesta dalle regole di penalizzazione; ma in aggiunta è richiesto che la probabilità rispetti le regole del calcolo delle probabilità, quelle stesse che sono imposte dalla teoria assiomatica, perché violandole si possono presentare delle situazioni in cui operare sulla base delle probabilità diventa un vantaggio o uno svantaggio certo.

Questa condizione spiacevole o slealmente vantaggiosa, a seconda del lato da cui

---

[55] *Ivi*, p. 68.
[56] *Ivi*, p. 78.

la si guarda, è nota come “allibramento olandese” o “scommessa olandese”, e si tratta di una scommessa in cui l’evento certo non vale 1. Si tenga presente che in una scommessa dove si danno più esiti possibili, l’evento certo è dato dalla somma di tutte le possibilità che possono darsi, dato che almeno una di queste dovrà sicuramente presentarsi. Ad esempio, è certo che lanciando un dado si presenterà una delle 6 facce, per cui, anche se fosse truccato e alcune di queste avessero probabilità più grande di 1/6, mentre altre avessero probabilità minore, la somma delle probabilità di tutte le facce deve essere 1. Questo si ha perché il dado rispetta le regole del calcolo delle probabilità o, per essere più precisi, noi abbiamo assegnato le probabilità alle facce rispettando queste regole, in modo che la somma risultasse uguale a 1. Le conseguenze del violare queste regole e scommettere con un allibramento olandese sono state illustrate da vari autori, ricorrendo anche al formalismo richiesto da una dimostrazione matematica che deve valere per una casistica generale, in termini di numero di scommesse e di scommettitori. L’esempio che ne dà de Finetti non è una dimostrazione esaustiva, ha però il merito di illustrare il problema in modo semplice, e chiarire il motivo per cui si richiede che anche la probabilità soggettiva rispetti le regole del calcolo delle probabilità:

> Il soddisfacimento degli assiomi del calcolo delle probabilità costituisce la condizione necessaria e sufficiente affinché un insieme di valutazioni di probabilità non dia luogo a conseguenze manifestamente dannose. Per esempio, il calcolo impone a una persona *X* che attribuisca a un certo evento *E* il valore di probabilità *p* di attribuire all’evento $\tilde{E}$ la probabilità 100%-p. Ebbene, se *X* attribuisse a $\tilde{E}$ una probabilità q diversa da 100%-p, egli offrirebbe a un competitore l’opportunità di guadagnare a colpo sicuro. Infatti, se *q* < 100%-*p* (se, per esempio, *p* fosse 40% e *q* fosse 50%) il competitore potrebbe ragionare così: «faccio contemporaneamente due scommesse, una su *E*, l’altra su $\tilde{E}$; pago 40 lire per la prima e 50 per la seconda (con una spesa complessiva di 90 lire), in cambio della promessa di ricevere (per ciascuna delle scommesse) 100 lire in caso di vincita. Poiché è certo che vincerò una (e una sola) delle due scommesse, riceverò 100 lire con certezza a fronte delle 90 lire spese, realizzando così a colpo sicuro un guadagno netto di 10 lire». Se invece *q* > 100%-*p* (se per esempio *p* fosse pari a 40% e *q* fosse pari a 70%), allora il competitore potrebbe ragionare

così: «accetto due scommesse, una su *E* l'altra su $\tilde{E}$, facendomi pagare 40 lire per la prima e 70 lire per la seconda (e quindi in tutto 110 lire) ed impegnandomi in cambio a pagare 100 lire (per ciascuna delle scommesse) in caso di vincita. Poiché è certo che vincerò una e una sola di queste due scommesse, pagherò 100 lire con certezza a fronte delle 110 lire incassate, realizzando così a colpo sicuro un guadagno netto di 10 lire». Dal punto di vista delle scommesse la *coerenza* si riduce a questo.[57]

Va subito notato che se l'allibramento olandese nel calcolo delle probabilità è indice di un errore, nel mondo reale delle scommesse rappresenta invece una condizione ricercata intenzionalmente nella maggior parte dei casi: non esiste casa da gioco che non abbia stabilito regole che garantiscano al banco una vincita sicura dal punto probabilistico. Ad esempio, al gioco della roulette francese la vincita per il banco, dal punto di vista pratico, è data dal fatto che escano numeri per i quali non è stata fatta nessuna puntata, ma il gioco è fatto in modo che anche nell'ipotesi peggiore per il banco, ovvero se i giocatori fossero in numero tale e con disponibilità tali da poter coprire il tavolo di puntate a ogni mano, il banco avrebbe comunque un vantaggio; infatti la vincita sul numero pieno è di 36 volte la posta, mentre i numeri che possono uscire sono 37 per la presenza dello 0, e la puntata in più costituisce il vantaggio certo del banco. Il fatto che, nel caso di puntate equivalenti per valore, su tutto il tavolo si puntino 37/37 mentre la vincita sia 36/37 è proprio l'allibramento olandese; l'effettiva differenza tra le puntate viene poi assorbita dalla distribuzione statistica di puntate e numeri usciti. Discorso analogo per gli altri tipi di puntate: il rosso e il nero vincono 2 volte la posta, ma lo 0 è verde e determina il vantaggio del banco rispetto alle scommesse sul colore, e così via.

La critica alla interpretazione soggettivista mi pare che nasca soprattutto dalla difficoltà di conciliare il "soggetto" con l'oggettività scientifica, dato che il soggetto è

[57] *Ivi*, p. 189.

in ultimo sempre indisponibile, mentre l'oggetto della scienza è presente a piacere nella virtualmente infinita disponibilità ripetibile del laboratorio. Il calcolo delle probabilità con i suoi metodi e le sue applicazioni statistiche, di ragione o di fatto, è la più fruttuosa risorsa della ricerca scientifica del XX secolo, e al momento non sembra debba esser diversamente nel XXI; si capisce che l'idea di introdurre un forte elemento di soggettività in ogni valutazione probabilistica possa causare un certo sconcerto, se lo si intende come una forma di arbitrio.

Galati e Pavan, nel loro manuale di *Teoria dei fenomeni aleatori,* menzionano l'interpretazione soggettiva della probabilità, relegandone però l'effettiva utilità a quei casi in cui siamo alle prese con eventi singoli o che, pur essendo ripetibili, non hanno ancora fornito sufficienti dati da poter essere letti in termini frequentisti. Inoltre la probabilità soggettiva viene legata non solo allo stato di informazione ma anche alla propensione al rischio dell'individuo. L'abbinamento stretto tra i concetti "informazione disponibile" e "propensione al rischio", oltre al fatto che gli autori siano professori di ingegneria (Galati emerito), il ramo accademico che più ha contribuito alla strutturazione delle imprese industriali, fa pensare che stiano facendo riferimento alla branca di teoria delle decisioni, disciplina che si è sviluppata fortemente in ambito di gestione di impresa nel XX secolo, oscillando tra impostazioni statistico-matematiche e psicologistiche, documentata più nella "letteratura di management" che in quella accademica. Nei fatti, Galati e Pavan fanno una debole critica della probabilità soggettiva, relegandola a un ambito marginale, di scarso interesse dal punto di vista matematico e scientifico, infatti non riprendono più l'argomento.

Decisamente più categorica la critica di Costantini e Geymonat, che rifiutano senza mezzi termini l'interpretazione soggettivista in quanto inconsistente. In particolare, rispetto alla definizione classica, contestano la critica fatta da de Finetti al criterio di equiprobabilità, in quanto se uguali conoscenze non consentono a persone diverse di pervenire a un uguale giudizio di probabilità, allora, per coerenza, le stesse informazioni in capo allo stesso soggetto producono valutazioni di probabilità differenti in tempi diversi; un soggettivismo coerente, insomma, sarebbe qualcosa di decisamente strampalato:

> Una simile concezione dell'equiprobabilità, generando valutazioni istantanee della probabilità, oltre ad andare contro le nostre convinzioni più radicate, non servirebbe proprio a nulla. Questa constatazione sembra essere sufficiente a concludere che, con riferimento alle valutazioni di equiprobabilità, l'unica concezione coerente del soggettivismo è del tutto implausibile.[58]

De Finetti dice chiaramente che la valutazione soggettiva della probabilità, proprio perché dipende dalle conoscenze a disposizione, può cambiare nel tempo, e sulla base di questa premessa è anche verosimile che la stessa persona, disponendo delle stesse informazioni, le analizzi in tempi diversi con una diversa comprensione, giungendo a una diversa valutazione.

Mutabilità e provvisorietà sono costanti del sapere scientifico, ed evidenza stessa del suo progredire; sorprende perciò il tono della critica di Costantini e Geymonat all'interpretazione soggettivista della probabilità. Lo stesso quando rispondono alle obiezioni al frequentismo:

> Passando ora alle critiche dei soggettivisti contro la concezione frequentista, è necessario premettere che non è affatto semplice riferirle dal momento che nei lavori dei soggettivisti non si trovano, a questo proposito, affermazioni della stessa chiarezza di quelle che abbiamo visto rivolte contro la concezione classica.[59]

Ma abbiamo visto che la definizione di "evento" data da de Finetti è di una chiarezza inequivocabile, ed è una definizione che esclude ogni applicazione frequentista della probabilità. Lo stesso de Finetti, altrove, dice altrettanto chiaramente che la

---

[58] D. Costantini e L. Geymonat, *Filosofia della probabilità*, cit., p. 66.
[59] *Ivi*, p. 66.

misura di una frequenza non è una misura di una probabilità. Il sospetto è che Costantini e Geymonat, verificando che l'interpretazione soggettivista è una concezione profondamente diversa dalle altre, del termine "soggettivo" abbiano colto soprattutto l'accezione più affine ad "arbitrario".

3.5. Dalla probabilità alla statistica soggettivista – Il teorema di Bayes

Abbiamo visto che, indipendentemente dalle convinzioni personali, se si guarda l'interpretazione soggettivista della probabilità senza preconcetti, non siamo di fronte a una teoria fumosa e imprecisa, a cominciare dalla richiesta di rispettare le regole del calcolo delle probabilità così come consegue dalla teoria assiomatica.

Nei fatti, il soggettivismo ha implicazioni profonde, anche se raramente note al grande pubblico, proprio nel campo dominato dall'interpretazione frequentista: la statistica. Il collegamento tra soggettivismo e statistica è dato dal concetto di "probabilità condizionata" e dal teorema di Bayes, tanto che spesso la "probabilità soggettiva" viene anche detta "probabilità bayesiana"; ho preferito evitare questa seconda terminologia, però, dato che il teorema di Bayes è un teorema a tutti gli effetti e pertanto, ammessi gli assiomi, vale in ogni caso, a prescindere dal personale punto di vista sulla probabilità.

Con "probabilità condizionata" si intende la probabilità che, dati due eventi non incompatibili, si verifichi uno dei due, posto che l'altro si è verificato.

Definizione di probabilità condizionata

Se A e B sono due eventi di uno spazio campione S, con $P(B) \neq 0$, si definisce probabilità condizionata di A rispetto a B (o "probabilità di A dato B"), e si indica con $P(A|B)$, il rapporto:

$$P(A|B) = \frac{P(AB)}{P(B)}$$

(si ricordi che AB è l'evento intersezione di A con B).[60]

Se gli eventi sono incompatibili, la probabilità condizionata vale 0, ad esempio se lanciamo due dadi possiamo chiederci quali siano le probabilità che la somma dei valori usciti sia 8. Ma se sappiamo che per uno dei dadi è uscito 1, qualunque sia l'esito dell'altro dado è impossibile che la somma dia 8. L'evento "il primo dado dà esito 1" e l'evento "la somma dei dadi dà esito 8" sono due eventi possibili singolarmente ma incompatibili tra loro.

Se invece sappiamo che l'esito di uno dei due dadi è 3, ha senso chiedersi quali siano le probabilità che la somma dia 8, perché gli eventi non sono più incompatibili. Posto $B = "il\ primo\ dado\ dà\ esito\ 3"$ e $A = "la\ somma\ dei\ dadi\ è\ 8"$, quando è verificato B ci sono 6 possibili esiti del secondo dado: $\{(3,1); (3,2); (3,3); (3,4); (3,5); (3,6)\}$ e una sola coppia di risultati dà 8 come somma; quindi, se il primo dado dà esito 3, la probabilità che la somma sia 8 è 1/6.

Il significato della formula della probabilità condizionata si può vedere anche ragionando in termini di definizione classica: se B si verifica, A si verifica se e soltanto se si verifica AB (l'evento intersezione), per cui $P(AB)$ rappresenta i casi favorevoli, mentre $P(B)$ sono quelli possibili.

Il teorema di Bayes fa un passo in più, ed esprime una probabilità condizionata in funzione della probabilità non soggetta a condizione.

Tra i vari modi di illustrare questo teorema, mi pare che l'esposizione che ne dà de Finetti sia tra le più interessanti. La riporto quasi integralmente, nonostante la lunghezza, perché mostra bene il legame tra teorema di Bayes e statistica soggettivistica:

---

[60] G. Galati e G. Pavan, *Teoria dei fenomeni aleatori*, cit., p. 26.

Per semplicità prescinderò dal far riferimento alle regole di penalizzazione appropriate. Farò riferimento, invece, all'interpretazione della probabilità e della previsione come prezzo. La probabilità di un evento E, infatti, può essere considerata come il prezzo equo (secondo la valutazione di chi la fa) del diritto di ricevere 1 ove E si verificasse e niente ove E non si verificasse.

1) $\boldsymbol{P}(E) = \begin{cases} 1\ se\ E \\ 0\ se\ \tilde{E} \end{cases}$

Considerando E come un numero aleatorio che assume valore 1 o 0 a seconda che E sia vero o falso, la probabilità di E può essere considerata come il prezzo di un'offerta aleatoria pari a E. Parlando di prezzo intendo una grandezza lineare. D'altro canto $\tilde{E}$ può essere identificato con il numero aleatorio $1 - E$. La linearità implica:

$$\boldsymbol{P}(\tilde{E}) = \boldsymbol{P}(1 - E) = 1 - \boldsymbol{P}(E)$$

Detto a parole: la probabilità del complemento $1 - E$ di E (cioè di $\tilde{E}$) è uguale al complemento della probabilità di E, cioè a $1 - \mathbf{P}(\mathrm{E})$.

Consideriamo ora l'operazione aritmetica $A + B$. Essa non dà luogo a un evento. Dà invece luogo a un numero aleatorio che vale 0 se tanto A che B sono falsi, 1 se uno dei due eventi è vero e l'altro è falso, 2 se tanto A che B sono veri:

$$A + B = \begin{cases} 0\ se\ \tilde{A}\tilde{B} \\ 1\ se\ \tilde{A}B\ oppure\ A\tilde{B} \\ 2\ se\ AB \end{cases}$$

La somma aritmetica deve essere distinta dall'unione di due eventi la quale è sempre – diversamente dalla somma – un evento.

[…si tralascia un esempio esplicativo su corrispondenze e differenze tra operazioni logiche e aritmetiche…]

Naturalmente, tutti questi eventi devono essere considerati rispetto a uno stato di informazione H_0. Cioè si dovrebbe scrivere:

$$(A|H_0), (B|H_0), (C|H_0), (D|H_0)\ ecc.$$

Il riferimento a H_0 è importante, anche se nella pratica può essere sottinteso per semplificare la notazione.

Vorrei adesso mostrare che rapporto sussiste tra la probabilità di un evento E, $\boldsymbol{P}(E|H_0)$ e la probabilità di E subordinatamente al realizzarsi di una condizione ipotetica H, probabilità che indichiamo con $\boldsymbol{P}(E|HH_0)$. Che cosa vuol dire $\boldsymbol{P}(E|HH_0)$ o (sottintendendo lo stato di informazione H_0) $\boldsymbol{P}(E|H)$? È il prezzo che si giudica equo pagare per ricevere 1 nel caso che siano veri sia H che E (di H_0 si sa già che è vero) e nulla altrimenti, con la clausola che ove H non si verificasse (indipendentemente dal verificarsi o non verificarsi di E) allora non si farebbe alcuna operazione, cioè non si pagherebbe nulla e non si riceverebbe nulla: se si fosse già pagato, si verrebbe rimborsati. Quindi abbiamo:

2) $$\mathbf{P}(\mathrm{E}|\mathrm{H}) \equiv \begin{cases} 1 \text{ se EH} \\ 0 \text{ se } \tilde{\mathrm{E}}\mathrm{H} \\ \mathbf{P}(\mathrm{E}|\mathrm{H}) \text{ se } \tilde{\mathrm{H}} \end{cases}$$

Vediamo quale relazione sussiste tra $\boldsymbol{P}(E)$ e $\boldsymbol{P}(E|H)$. In base a 1) vale:

3) $$\boldsymbol{P}(E) \equiv \begin{cases} 1\ se\ EH \\ 1\ se\ \tilde{E}H \\ 0\ se\ \tilde{H} \end{cases}$$

In virtù della supposta linearità dei prezzi possiamo moltiplicare membro a membro 3) e 2). Si ottiene in tal modo:

4) $$\mathbf{P}(\mathrm{H})\mathbf{P}(\mathrm{E}|\mathrm{H}) \equiv \begin{cases} 1 \text{ se EH} \\ 0 \text{ se } \tilde{\mathrm{E}}\mathrm{H} \\ 0 \text{ se } \tilde{\mathrm{H}} \end{cases}$$

Da 1) segue (considerato che $EH = HE$):

5) $$\boldsymbol{P}(HE) \equiv \begin{cases} 1\ se\ EH \\ 0\ se\ \tilde{E}H \\ 0\ se\ \tilde{H} \end{cases}$$

Confrontando 5) con 4) si ottiene:

6) $$\boldsymbol{P}(HE) = \boldsymbol{P}(H)\boldsymbol{P}(E|H)$$

Analogamente varrà:

7) $$\boldsymbol{P}(EH) = \boldsymbol{P}(E)\boldsymbol{P}(H|E)$$

e poiché $EH = HE$, dalle relazioni 6) e 7) si ricava subito:

8) $$\boldsymbol{P}(H)\boldsymbol{P}(E|H) = \boldsymbol{P}(E)\boldsymbol{P}(H|E)$$

e quindi, dividendo entrambi i membri dell'identità 8) per $\boldsymbol{P}(H)$, si ottiene infine:

9) $$\boldsymbol{P}(E|H) = \boldsymbol{P}(E)\frac{\boldsymbol{P}(H|E)}{\boldsymbol{P}(H)}$$

L'equazione 9) fornisce la formulazione più semplice del teorema di Bayes. Qui $\boldsymbol{P}(E)$ è la probabilità che si attribuisce a E nello stato di conoscenza iniziale H_0. Se invece supponiamo che H si verifichi (oppure si viene a sapere che H si è verificato), allora $\boldsymbol{P}(E)$ - la probabilità iniziale di E - si altera nella stessa proporzione in cui si modifica $\boldsymbol{P}(H)$ subordinatamente a E, dando luogo alla nuova probabilità $\boldsymbol{P}(E|H)$.

Tutta la statistica soggettivistica si basa su questo semplice teorema del calcolo delle probabilità. Ciò fa sì che la statistica soggettivistica abbia un fondamento molto semplice e generale. Inoltre, fondandosi sui soli assiomi di base della probabilità, la statistica soggettivistica non dipende da quelle definizioni della

> probabilità che ne restringerebbero il campo di applicazione (come per es. quelle che si basano sull'idea di eventi ugualmente probabili).[61]

La concezione soggettivista fonda la statistica soggettivista (o statistica bayesiana) sul teorema di Bayes, in contrapposizione alla statistica frequentista. Si noti che il teorema di Bayes, come si accennava prima, vale anche per la statistica classica, è però diverso il significato che si attribuisce al teorema. La trasformazione algebrica e le inferenze logiche che conducono alla dimostrazione del teorema – o a ricavarlo come fa de Finetti nella citazione riportata sopra – sono valide per tutti. Per un soggettivista però, che non è vincolato alla tesi di una probabilità esistente nell'"oggetto" e perciò immutabile anche se magari ignota, $\boldsymbol{P}(E)$ è la probabilità di una ipotesi E in una condizione iniziale H_0; una plausibile verosimiglianza che è suscettibile di variazione quando le conoscenze disponibili mutano. In particolare, la probabilità iniziale $\boldsymbol{P}(E)$ cambia quando si verifica l'evento H, o quando si ha una certa confidenza sul verificarsi di H, e varia proprio in funzione della quantità di informazione introdotta dall'evento H che determina il nuovo complesso di conoscenze disponibili.

Questa impostazione bayesiana vale per l'evento E, sia che si tratti di un accadimento, una circostanza particolare di cui si esprime, tramite la probabilità $\boldsymbol{P}(E)$, la confidenza nel suo verificarsi, sia nel caso che E rappresenti una teoria scientifica, magari affiancata ad altre teorie alternative F, G, ecc. che inizialmente, per via delle condizioni H_0, hanno ognuna una particolare probabilità di essere valide, e che all'accadere dell'evento H, o in considerazione di una certa probabilità che questo si verifichi, acquistano una nuova probabilità che rende la teoria E una migliore (o peggiore) spiegazione scientifica rispetto alle altre teorie, proprio per il livello di credibilità dato dalla probabilità $\boldsymbol{P}(E|H)$.

Lavorando sule ipotesi scientifiche, il teorema di Bayes afferma proprio che la

[61] B. de Finetti, *Filosofia della probabilità*, cit., pp. 95-99.

validità della teoria E viene accresciuta dall'evento H, nella misura in cui H è spiegabile se si presuppone E

Questo modo di procedere ricorsivamente sulle ipotesi, selezionando quelle più promettenti e introducendone di nuove, è indubbiamente familiare nel modo di procedere dell'indagine scientifica, e detto in questi termini non rappresenta certo una novità. Purtuttavia, esiste un discrimine tra strumenti operativi al servizio della ricerca, e il punto nodale della separazione tra la statistica classica e quella soggettivistica è rappresentato proprio dall'ammettere una probabilità iniziale da cui muovere induttivamente:

> Questa rivoluzione nei fondamenti [rappresentata dall'affermarsi dei metodi frequentisti nella statistica nella seconda metà del XIX secolo] ebbe una conseguenza di enorme importanza sulla statistica induttiva: il teorema di Bayes, che fino ad allora, per il tramite delle probabilità inverse, aveva rappresentato il principio informatore di ogni inferenza statistica, venne abbandonato. E ciò non fu dovuto al rifiuto del suo ruolo teorico, bensì al fatto che, in accordo col nuovo modo di determinare le probabilità, non era più possibile pervenire, se non in casi particolari, alla valutazione delle probabilità a priori.[62]

Nonostante sia poco nota al di fuori dell'ambito specialistico, la statistica bayesiana o soggettivista non è un mero esercizio concettuale, e su impulso del lavoro di Bruno de Finetti e Harold Jeffreys – anche quest'ultimo matematico ma senza la forte connotazione filosofica di de Finetti – nel 1992 è stata fondata la International Society for Bayesian Analysis (ISBA), che ha sede negli Stati Uniti presso la Duke University, a Durham nel North Carolina[63].

[62] D. Costantini e L. Geymonat, *Filosofia della probabilità*, cit., p. 31.

[63] Chi vuole prendersi del tempo per andare a curiosare nel sito dell'associazione (https://bayesian.org/), può verificare direttamente come oltre alla sede, anche tutta l'amministrazione e la gestione delle attività associative sia gestita da accademici. ISBA pubblica dal 2006 la rivista trimestrale Bayesian Analysis (online e integralmente a libero accesso: https://projecteuclid.org/journals/bayesian-analysis), cui contribuiscono numerosi ricercatori da tutto il mondo; mette anche a disposizione conferenze registrate e altra documentazione relativa all'analisi statistica di matrice bayesiana in vari settori della scienza, rimandando anche a siti web dedicati. Singolare il fatto che, nata dal

Questo per dare la misura di come la statistica, che negli ultimi 2 secoli si era assestata stabilmente nell'alveo della prospettiva frequentista, negli ultimi decenni sia in fermento, per il ripensamento radicale dei suoi fondamenti in un'ottica soggettivista.

La differenza tra i due sistemi di analisi non è solo formale, e determina una significativa differenza nella valutazione dei risultati delle prove scientifiche. Particolarmente indicativo, anche per l'interesse suscitato in letteratura, il dibattito che si è sviluppato in seguito all'annuncio, fatto il 4 luglio 2012 al CERN (Conseil Européen pour la Recherche Nucléaire) di Ginevra, che il *bosone di Higgs* era stato scoperto.

Il bosone di Higgs è una particella del Modello Standard della fisica moderna che era stata ipotizzata nel 1964 per spiegare – semplificando molto – il passaggio dallo stato dell'universo immediatamente successivo al Big Bang, dove esisteva una sola forza elettrodebole e particelle prive di massa, allo stato a più bassa energia, dove i campi di forze si differenziano e le particelle acquistano massa (con l'eccezione del fotone, responsabile del campo elettromagnetico). Le ipotesi sulle condizioni iniziali dell'universo sono giustificate dal fatto che il Modello Standard spiega sufficientemente bene la fisica delle alte energie, quella che osserviamo negli acceleratori di particelle. Il bosone di Higgs, secondo la teoria, è responsabile della rottura della *simmetria elettrodebole* che segna la transizione di fase dall'universo iniziale a quello che conosciamo noi – o almeno la transizione dalla fisica delle alte energie a quella delle basse energie –; si capisce perciò che la ricerca dell'evidenza scientifica di questa particella abbia rappresentato, per i fisici e per gli appassionati, un'avvincente caccia al tesoro che nella seconda metà del XX secolo ha visto fiorire la produzione sia di letteratura accademica sia di quella divulgativa.

La scoperta delle particelle elementari viene fatta per via indiretta, cercando le

lavoro di un matematico italiano e di uno britannico, ISBA abbia capitoli in tutti i continenti tranne l'Europa, anche se non mancano comitati locali, anche italiani, attivi per opera di studenti e ricercatori.

tracce e le interferenze che esse producono sull'ambiente circostante, dato che, per evidenti motivi di dimensioni, non possono essere osservate direttamente. Questo significa anche che l'indagine non può essere che statistica, dato che le possibili prove devono essere verificate rispetto alla compatibilità con la teoria, con la presenza della particella ricercata e con la possibilità di un errore di rilevamento.

L'esperimento era stato condotto con una impostazione frequentista: se ipotizziamo che il bosone di Higgs non esista, ci si attende un certo risultato dall'esperimento; alla prova dei fatti però i risultati dell'esperimento ideato per mostrare che "il bosone di Higgs non esiste" si discostavano dal valore atteso per più di cinque deviazioni standard[64]. Dato che uno scostamento così estremo può capitare accidentalmente solo in un caso su due milioni, gli statistici del CERN conclusero che il bosone di Higgs era stato effettivamente scoperto e l'ufficio stampa ne diede l'annuncio.

Poco dopo l'annuncio, Tony O'Hagan, professore di statistica all'università di Sheffield, inviò una-mail alla newsletter dell'ISBA ponendo degli interrogativi in merito alla scoperta:

> Abbiamo sentito parlare molto del bosone di Higgs. Le notizie dicono che l'LHC [Large Hadron Collider, acceleratore di particelle del CERN,] aveva bisogno di prove convincenti prima che potessero annunciare che era stata trovata una particella che assomigliava a (nel senso di avere alcune delle giuste caratteristiche di) l'inafferrabile bosone di Higgs. Nello specifico, la notizia si riferiva a un intervallo di confidenza con limiti di 5 sigma.
>
> Ora, questo sembra corrispondere a un test di significatività frequentista con un livello di significatività estremo. Cinque deviazioni standard, assumendo la normalità [cioè la distribuzione dei dati secondo una gaussiana], significano un

---

[64] Nella distribuzione *normale*, quella descritta dalla curva gaussiana con la caratteristica forma a campana, il valore atteso è rappresentato dall'asse di simmetria della curva. Senza addentrarci nel formalismo matematico, la deviazione standard (indicata anche con la lettera greca σ) è la distanza a destra e sinistra del valore atteso, tale per cui il 68% dei valori della distribuzione cadono all'interno di quell'intervallo; all'interno di due deviazioni standard (o 2σ) cade il 95% dei dati e all'interno di tre deviazioni standard il 99%. La probabilità di trovare dati della distribuzione che cadono a più di tre deviazioni standard tende rapidamente a zero, anche se solo asintoticamente.

valore di probabilità di circa 0,0000005. Mi sorgono spontanee dalla mente una serie di domande.

1. Perché un requisito di prova così estremo? Sappiamo da una prospettiva bayesiana che questo ha senso solo se (a) l'esistenza del bosone di Higgs (o qualche altra particella che condivide alcune delle sue proprietà) ha una probabilità a priori estremamente ridotta e / o (b) le conseguenze di annunciare erroneamente la sua scoperta sono estremamente gravi. Non sembra che si dia nessuno dei due casi, quindi perché 5-sigma?

2. Piuttosto che una giustificazione *ad hoc* per un valore di probabilità, è ovviamente meglio fare una corretta analisi bayesiana. La comunità della fisica delle particelle ha sposato del tutto l'analisi frequentista? In caso affermativo, qualcuno ha cercato di spiegare quale cattiva scienza sia questa?

3. Sappiamo che, forniti sufficienti dati, è quasi sempre possibile che un test di significatività rigetti l'ipotesi nulla [quella per cui si allestisce il test ipotizzandola vera, ma che si ritiene falsa] con valori di probabilità arbitrariamente bassi, semplicemente perché il parametro non sarà mai esattamente uguale al suo valore nullo. E a quanto pare l'LHC ha accumulato una grande quantità di dati. Quindi anche questo estremo valore di probabilità potrebbe essere illusorio?

Se qualcuno ha delle risposte a queste o ad altri argomenti correlati, sarei interessato a saperlo.[65]

Come si può immaginare, questo intervento nella newsletter ISBA suscitò una ridda di commenti e articoli in risposta, da parte di esponenti di entrambe le scuole di statistica. In un suo successivo intervento, O'Hagan ammise l'intenzione volutamente provocatoria del suo intervento, in particolare per quel che riguarda il secondo punto, pur ribadendo la necessità di studiare più a fondo la validità di una

[65] Questo intervento nel forum di ISBA viene riportato da vari autori, sia facendo riferimento al sito www.isba.org (ad es. Jan Sprenger), vecchio dominio di ISBA, ora registrato a nome di una associazione di giuristi, sia con riferimento alla URL https://bayesian.org/forums/news/3830 (ad es. Kent W. Staley), che corrisponde al corretto dominio del sito istituzionale di ISBA ma a nessuna pagina visibile. Tramite web.archive.org, una organizzazione che si occupa di archiviare pagine web non più disponibili online, è possibile recuperare il post sul forum ISBA dove O'Hagan, nel 2015, ha postato un commento dove annunciava di aver finalmente messo mano a un documento di riepilogo di tutti i commenti fatti e per il quale rimandava al suo sito personale. Il testo è stato dunque ripreso da http://tonyohagan.co.uk/academic/pdf/HiggsBoson.pdf. Traduzione mia.

scoperta dimostrata ricorrendo all'ipotesi nulla e la significatività dei dati statistici raccolti:

> L'analisi statistica che alla fine conduce ad un risultato di 5 sigma si basa sul confronto di due ipotesi. L'ipotesi nulla è il MS [Modello Standard] senza H [bosone di Higgs], mentre l'ipotesi alternativa è il MS con H. Un commentatore [del forum] ha chiesto cosa potrebbe significare l'ipotesi nulla se il MS prevede H. Presumo si voglia dire che il MS è vero eccetto che nella sua previsione di H, ma rimane un potenziale problema al riguardo. Se il MS si sbaglia su H, quali parti di esso devono essere cambiate in modo che H non sia più una conseguenza? Probabilmente ci sono molti candidati per questo genere di cambiamenti. Senza indicare esplicitamente quali modifiche devono essere applicate, l'ipotesi nulla non è definita adeguatamente e non è possibile calcolare la probabilità dei dati sotto condizione nulla.[66]

L'osservazione di O'Hagan sottolinea un punto estremamente interessante, anche in termini di filosofia della scienza: l'allestimento di alcuni esperimenti richiederebbe una cura delle condizioni al contorno non solo per quel che riguarda l'ambiente fisico in cui sono svolti, ma anche per la teoria che implicano e che, necessariamente, ha ripercussioni sull'esito della prova. Il ruolo centrale occupato dal bosone di Higgs in una teoria complessa come il Modello Standard, impone, se si vuole inserire come ipotesi nulla la sua mancanza, di descrivere le possibili varianti del modello, quale si intende testare e perché l'esperimento non tiene conto delle altre; in caso contrario l'esito dell'esperimento rischia di essere chiaro quanto a misura, ma estremamente vago quanto a significato.

Il punto cruciale, mi pare, è che la scienza, indagando il mondo fisico, se vuole procedere con dimostrazioni per assurdo non può limitarsi a replicare lo schema operativo impiegato in matematica, ma deve introdurre condizioni specifiche e

[66] Come nota precedente.

chiarire le premesse in base alle quali i risultati hanno effettivamente significato rispetto all'ipotesi nulla. Il motivo di questi vincoli aggiuntivi sta nel fatto che in matematica l'esito di un'inferenza logica che parte da una premessa assurda conduce a un risultato autocontraddittorio, mentre un esperimento fisico che parte da un'ipotesi falsa produce dati che possono essere inaspettati, sbagliati, eventualmente incomprensibili, ma per il fatto stesso di essere "dati" non saranno mai autocontraddittori.

Qui naturalmente non ci interessa né determinare se il bosone di Higgs sia stato effettivamente scoperto, né se si tratti di una ricerca per la quale sarebbe preferibile un'analisi bayesiana piuttosto che frequentista. La cosa importante, ai fini della nostra riflessione sulle teorie della probabilità, è che la discussione sui metodi del calcolo delle probabilità, che sembrava definitivamente chiusa con l'affermarsi della statistica frequentista, si è riaccesa con una vivacità che lascia ben sperare in merito allo stato di salute della ricerca scientifica.

4. Il problema dell'induzione

La nostra conoscenza, almeno per la parte acquisita razionalmente, si accresce secondo un procedimento che inferisce da elementi noti altri elementi che erano ignoti ed entrano così, da quel momento, nel novero di ciò che si conosce e che può fare da premessa per ulteriori possibili conoscenze.

I processi inferenziali sono principalmente di due tipi: deduttivi e induttivi. Dall'inizio del XX secolo si risveglia un significativo interesse anche per l'abduzione, che classicamente rappresentava un tipo di sillogismo dalle premesse certe e dalla conclusione solo probabile; ma che con il lavoro di Peirce diventa l'inferenza che permette di formulare ipotesi che rendono conto di fatti osservati non coerenti con altre teorie note[67]:

[67] Charles Sanders Peirce (1839-1914) è all'origine della teoria epistemologica nota come pragmatismo. La sua produzione scientifica pubblicata in vita fu più che modesta rispetto a quella postuma e in parte tuttora inedita. Alla sua morte lasciò circa 100.000 pagine manoscritte che vennero acquistate da Harvard e pubblicate in buona parte

Essendosi ormai conclusa la stagione delle ricerche formalizzanti d'impronta neopositivistica, si fa strada [tra le ipotesi che descrivono il processo di ricerca scientifico] una nuova concezione indiziaria, congetturale, "sfumata", ma non per questo irrazionale. Anzi, tutto ciò ha permesso di portare maggiore attenzione a un procedimento inferenziale che non si può ricondurre né alla deduzione né all'induzione. È in questa prospettiva che si è indagata la struttura del ragionamento abduttivo.

Dobbiamo a Ch.S. Peirce la prima e tuttora più completa analisi dell'abduzione. Tra le molte definizioni che egli propone, forse la più interessante, per il nostro percorso, è la seguente:

«L'abduzione è il processo di formazione d'ipotesi esplicative. È l'unica operazione logica che introduce una nuova idea, in quanto l'induzione non fa che determinare un valore e la deduzione sviluppa semplicemente le conseguenze necessarie di una pura ipotesi. La deduzione trova che qualcosa *dev'essere*; l'induzione mostra che qualcosa *è realmente* operativa; l'abduzione meramente suggerisce che qualcosa *può essere*.»

L'abduzione è uno strumento per generare ipotesi, evidentemente quando ne siamo sprovvisti, o quando quelle consuete sembrano non funzionare. Di fronte ad alcuni fatti che non rientrano nello schema esplicativo consueto per quel tipo di fenomeni, occorre inventare delle ipotesi che ne diano ragione. L'abduzione è l'inferenza che permette d'ipotizzare tale spiegazione, riconducendo l'eccentricità a una – diversa – normalità. In che modo?

«La forma dell'inferenza è la seguente: si osserva un fatto sorprendente C; ma se A fosse vero, C sarebbe spiegato come fatto naturale; dunque c'è ragione di sospettare che A sia vero».[68]

---

tra il 1931 e il 1958, negli 8 volumi dei *Collected Papers*. Questo scostamento temporale tra la sua attività speculativa e la pubblicazione della maggior parte della sua produzione scientifica, spiega come le sue idee, in particolare per l'aspetto che ci interessa, si siano diffuse parecchi anni dopo la sua morte. Anche la sua teoria di maggior successo, pubblicata nel 1878 nel saggio *How to make our ideas clear*, fu diffusa e resa nota al grande pubblico col nome di pragmatismo solo a partire dal 1898 ad opera di William James.

[68] G. Boniolo e P. Vidali, *Filosofia della scienza*, Bruno Mondadori, Milano 1999, p. 282.

Questa pagina di Boniolo e Vidali, che a loro volta citano Peirce, mi pare interessante non solo perché dà una visione d'insieme sulle tre famiglie di inferenze logiche, ma anche perché le colloca nella prospettiva contemporanea in cui sono per lo più studiate e interessano, ovvero quella della conoscenza scientifica e delle sue giustificazioni nell'ambito della filosofia della scienza.

Le inferenze deduttive costituiscono il cuore e il metodo di lavoro della logica classica, e sono state studiate a cominciare da Aristotele con il sillogismo: in estrema sintesi, data una o più premesse vere, le conclusioni cui si giunge applicando correttamente le regole della logica sono necessariamente vere anch'esse (dimostrazione apodittica); per cui, in un certo senso ed effettivamente, «il metodo deduttivo propriamente detto non "scopre" nulla, poiché ciò che si raggiunge nella conclusione di un'inferenza deduttiva è tutto contenuto nelle sue premesse»[69]. Va precisato che questa "incapacità di scoprire" dell'inferenza deduttiva deve essere intesa correttamente, poiché se è vero che non si produce nessuna conoscenza che non sia implicitamente contenuta nelle premesse, in quanto le conclusioni sono ottenute da una opportuna applicazione quasi algoritmica di regole logiche, questo non vuol dire che non si tratti di conoscenza in senso assoluto: il fatto che la somma degli angoli interni di un triangolo corrisponda a 180° può evidentemente essere dedotto dalla definizione di triangolo e dagli assiomi della geometria come premesse, ma sarebbe almeno azzardato sostenere che per questo motivo non aggiunga nessuna conoscenza, a meno di voler sostenere che un corso di laurea in matematica non fornisca nessuna conoscenza che vada al di là degli assiomi e delle regole della logica, trattandosi essenzialmente di un lungo tirocinio di pratica deduttiva. Anche in questo caso, lo scarso apporto di conoscenza fornito dall'inferenza deduttiva è una sottolineatura caratteristica di una concezione moderna, se non contemporanea, delle teorie della conoscenza, dove è una costante di ogni disciplina scientifica la tensione per identificare una struttura formale di assiomi (o principi generali), che fondino a priori

[69] *Ivi*, p. 221.

le prassi operative di quella disciplina, ma è trascurabile il fatto che a posteriori questa struttura venga poi dotata di un significato, come abbiamo visto nel caso degli assiomi della teoria delle probabilità nel paragrafo 3.3. È evidente che Aristotele non sarebbe stato disposto in nessun caso a considerare il sillogismo un “girare in tondo intorno agli assiomi”.

L’induzione è l’inferenza che più ci interessa, in quanto la totalità degli autori contemporanei in letteratura, almeno quelli che sono riuscito a verificare, considerano il calcolo delle probabilità sostanzialmente una questione di induzione. La cosa potrebbe sorprendere, in quanto tutti i teoremi relativi al calcolo delle probabilità, come abbiamo visto, sono dimostrati deduttivamente a partire dagli assiomi. Certamente, se pensiamo alla frequenza e all’impiego statistico che se ne fa, possiamo dire che in molti casi si tratta di “indurre” avvenimenti futuri sulla base delle conoscenze pregresse, ma non rende veramente ragione del fatto che Costantini e Geymonat parlino di «metodo di induzione basato sulla frequenza relativa»[70] o che L. J. Cohen scriva un testo dal titolo *Introduzione alla filosofia dell’induzione e della probabilità*. Per capire meglio come induzione e probabilità siano legate e si richiamino, e per evitare malintesi e mettere in luce il rischio di possibili fraintendimenti, è opportuno vedere brevemente come il significato di “induzione” sia cambiato nel tempo:

Dal lessico che sta a corredo della *Storia della filosofia antica* di Giovanni Reale leggiamo alla voce “induzione”:

> INDUZIONE (ἐπαγωγή)
>
> *È il procedimento attraverso cui dal particolare si ricava l’universale*. Aristotele sottolinea che l’induzione non è propriamente un ragionamento, bensì un «essere condotto» dal particolare all’universale. *Il processo astrattivo coincide sostanzialmente con l’induzione*. In questo processo che va dal particolare all’universale, la induzione si oppone alla deduzione sillogistica che è invece un

[70] D. Costantini e L. Geymonat, *Filosofia della probabilità*, cit., p. 33.

> procedimento dall'universale al particolare. *La deduzione* (sillogistica), però, *presuppone l'induzione*, come dice Aristotele in questo passo esemplare: «[…] Apprendiamo o per induzione o per dimostrazione, La dimostrazione procede dagli universali, mentre l'induzione procede dai particolari. Ma è impossibile considerare gli universali se non per induzione, poiché è possibile rendere noto per induzione anche i risultati della cosiddetta astrazione» (*Analitici Secondi*, A 18, 81 a 40 segg.)[71]

Per Aristotele, non esiste dubbio sul fatto che entrambe le forme di inferenza siano fonte di conoscenza, anche se non ogni conoscenza deriva da una inferenza. Inoltre induzione e deduzione si richiamano vicendevolmente, anticipando la conclusione contemporanea del tormentato percorso di riflessione sui metodi della conoscenza scientifica, proprio della filosofia della scienza. La differenza fondamentale è che mentre per Aristotele la conoscenza (di necessità o di fatto) è sempre riferita a una qualche realtà sostanziale, le teorie epistemologiche contemporanee non paiono interessarsi a molto più della coerenza del sistema formale (non si possono dimostrare contraddizioni), tralasciando anche la possibilità di trovare tutte le risposte da quando nel 1930 Kurt Gödel dimostrò i famosi teoremi di incompletezza[72].

Dopo Aristotele, gli stoici negano la validità dell'induzione, in quanto si può predicare necessariamente qualcosa di un ente solo se si tratta di un attributo proprio dell'ente, e quel che verifichiamo essere un carattere costante delle nostre osservazioni particolari non è detto che sia necessario, mentre gli epicurei sostengono che se nulla viene a contraddire l'inferenza (oggi diremmo: in mancanza di controesempi) la conclusione è vera. Sesto Empirico distingue tra

---

71 G. Reale, *Storia della filosofia antica*, Vita e pensiero, Milano 1995, vol. V, pp. 142-143.

72 Gödel dimostrò nel 1930 che in un sistema formale coerente, sufficientemente espressivo da poter assiomatizzare i numeri naturali con le operazioni di somma e prodotto, esistono proposizioni formalmente corrette ma per le quali non si può dimostrare né la verità né la falsità della proposizione (si dice che sono indecidibili). Una interessante esposizione della scoperta di Gödel, ampia sia per la ricchezza di esempi che per la esposizione delle implicazioni filosofiche, si trova nella *Enciclopedia Treccani* online, alla voce *La seconda rivoluzione scientifica: matematica e logica. I teoremi di incompletezza di Gödel*, di Carlo Cellucci (URL https://www.treccani.it/enciclopedia/la-seconda-rivoluzione-scientifica-matematica-e-logica-i-teoremi-di-incompletezza-di-godel_%28Storia-della-Scienza%29/)

induzione completa e induzione incompleta, a seconda che per generalizzare si siano presi in esame tutti i particolari o solo una parte, e fino all'epoca moderna la discussione, ove presente, procede nel solco di questa distinzione e, naturalmente, chi sostiene la necessità di una induzione completa di fatto nega la validità del processo induttivo per la evidente impossibilità pratica di applicarlo.

Nel XVI secolo, Francesco Bacone propone una vera e propria metodologia per procedere induttivamente nella ricerca scientifica, e mentre la scienza comincia a impiegare ampiamente l'induzione, si sviluppa la riflessione su quello che viene sempre più percepito come "il problema dell'induzione", vale a dire l'effettiva possibilità di basarsi sull'induzione in ambito scientifico e la giustificazione dell'impiego di questo tipo di inferenza. Dice a questo proposito Hume:

> Tutte le inferenze tratte dall'esperienza suppongono, come loro fondamento, che il futuro rassomiglierà al passato e che poteri simili saranno uniti a simili qualità sensibili. Se ci fosse qualche sospetto che il corso della natura potesse cambiare e che il passato non servisse di regola per il futuro, ogni esperienza diverrebbe inutile e non potrebbe dare origine ad alcuna inferenza o conclusione. È impossibile perciò che argomenti tratti dall'esperienza possano provare la rassomiglianza del passato con il futuro: giacché tutti gli argomenti siffatti sono fondati sulla supposizione di quella rassomiglianza. Sia pure ammesso che il corso delle cose è stato sempre regolare: questo solo, senza alcun argomento o inferenza nuova, non prova che per il futuro continuerà così.[73]

Hume pone l'attenzione su una circolarità concettuale che impegnerà la riflessione filosofica fino ai giorni nostri: il progredire della conoscenza scientifica non è giustificazione sufficiente della validità dei metodi induttivi cui la scienza fa ampiamente ricorso, perché questi poggiano sull'assunzione che la realtà sia strutturata in modo tale da essere adeguatamente descritta in senso induttivo. È

[73] N. Abbagnano, *Dizionario di filosofia*, UTET, Torino 1992., voce *Induzione*, p. 483; cita Hume, *Inquiry Concerning Human Understanding*, 1748, IV, 2.

evidente che se l’indagine scientifica, ricorrendo a metodi induttivi, scopre le leggi dell’aerodinamica, che mi consentono di progettare un dispositivo in grado di volare e questo, alla prova pratica, vola davvero, ciò deve pur dire qualcosa della nostra capacità di indagare e comprendere il mondo, e questo qualcosa non può essere liquidato come una coincidenza occasionale e fausta. L’obiezione di Hume svuota però il metodo induttivo della pretesa di raggiungere risultati necessari, e questo, per una scienza che si prefigge di individuare leggi universali e necessarie, è un problema.

Abbagnano, nel *Dizionario filosofico* alla voce *Induzione*, individua tre soluzioni che sono state date per affrontare il problema: quella oggettivistica, quella soggettivistica, e quella pragmatica.

La soluzione oggettivistica sostiene che nella natura sono presenti regolarità e uniformità, la realtà ha una trama omogenea che consente di generalizzare le esperienze simili; è questa costante di fondo che permette di stabilire un principio di causalità. La risposta oggettivistica è presente già nel pensiero filosofico antico e medievale, e non rappresenta una risposta in senso stretto all’obiezione di Hume, non tanto per la sequenzialità cronologica, ma principalmente perché toglie forza alla sua argomentazione ampliando l’orizzonte di riferimento, ma sposta solo il problema in un modo che innesca un circolo vizioso.

La soluzione soggettivistica è proposta da Kant che vede nell’intelletto, e in particolare nella sua struttura categoriale, quell’uniformità che l’oggettivismo pone nella natura. Le leggi particolari devono essere desunte dall’esperienza, perché di fatto avviene così, e comunque le sole categorie non potrebbero produrre direttamente un così gran numero di leggi particolari. Questo significa che nella natura è richiesta comunque una certa regolarità, che in questo caso non emerge dalla natura in sé; sono invece le categorie dell’intelletto a strutturare con una regolarità di spazio e tempo la natura, giustificando così l’induzione come metodo di indagine scientifica.

La soluzione pragmatica nasce nel pensiero contemporaneo. Essa rinuncia a cercare una giustificazione teoretica dell’induzione e ne fornisce invece una probabilistica come riportiamo direttamente dal *Dizionario filosofico* di Abbagnano:

> Quando un determinato carattere ricorre in una certa *proporzione* dei campioni esaminati, si può assumere che questa proporzione valga per tutti gli altri esempi del caso, salvo prova in contrario. Quando la proporzione è uguale al cento per cento dei campioni esaminati, cioè quando il carattere in questione ricorre in *tutti*, si ha il caso della generalizzazione *uniforme* o *completa*. È questo il caso quando si afferma che «tutti gli uomini sono mortali» per il fatto che l'essere mortale si è sempre trovato costantemente congiunto con l'essere uomo. Dall'altro lato quando il valore numerico di quella proporzione si assume come misura della *possibilità* che il carattere in questione ricorra in un nuovo esempio, si ha un giudizio di *probabilità*. Ovviamente la generalizzazione completa o il giudizio di probabilità sono aspetti della generalizzazione statistica. Stando ciò, la giustificazione dell'induzione da un punto di vista pragmatico può essere fatta asserendo: *a)* che l'induzione è il solo mezzo di ottenere previsioni; *b)* che essa è il solo metodo suscettibile di autocorrezione.[74]

Abbagnano riporta anche le obiezioni alle asserzioni *a)* e *b)* fatte da Max Black nel suo testo *Problems of analysis – Philosophical essays*: se l'induzione è il solo metodo per fare previsioni, e lo si si afferma in questo modo sostanzialmente tautologico, poco importa poi ai fini di validare il metodo induttivo che le previsioni si avverino o meno. Quanto all'autocorrezione, essa suppone una direzione in cui procedere e la consapevolezza che si tratti della direzione buona, ma non si capisce cosa dovrebbe garantire che si proceda effettivamente così.

Le definizioni riportate mostrano come l'idea di induzione si sia modificata profondamente nel corso del tempo; tanto profondamente che la recente impostazione pragmatica non ha più nulla dei contenuti dell'induzione come concepita da Aristotele.

---

[74] N. Abbagnano, *Dizionario di filosofia*, cit., p. 485.

A questo punto appare chiaro anche come mai nei testi di calcolo delle probabilità, soprattutto in quelli con un taglio filosofico, si insista tanto sulla probabilità come risultato di inferenze induttive: la concezione pragmatica dell'induzione è, né più né meno, l'interpretazione frequentista della probabilità.

Vediamo ancora una definizione di induzione dal testo già citato di Boniolo e Vidali sulla filosofia della scienza che, pur non aggiungendo altri punti di vista all'idea di induzione la contestualizza bene all'interno della riflessione contemporanea e ci offre la possibilità di chiarire meglio i termini della questione:

> Perché dev'essere difficile rendere conto dell'induzione, cioè del nostro estendere al presente e al futuro la conoscenza ottenuta nel passato? Perché è così difficile dar ragione di una fra le procedure più ordinarie del nostro agire razionale, al punto da ritenerla essenziale per la nostra stessa sopravvivenza? Perché, insomma, per Hume, come oggi per noi, l'induzione è un problema?
>
> Cominciamo a rispondere illustrando che cosa s'intenda per "induzione". Una lunga tradizione, che inizia con Aristotele e arriva almeno fino all'800, sostiene che un ragionamento induttivo inferisce dal particolare al generale, a differenza di quello deduttivo, che procede dal generale al particolare. È una definizione impropria, se non errata. Vi sono infatti induzioni con premesse generali, o con conclusioni particolari, e deduzioni con premesse particolari, o con conclusioni generali. Meglio allora fondare la specificità dell'induzione non sulla quantificazione degli enunciati ma sul tipo di nesso inferenziale.
>
> Per questa via possiamo dire che:
>
> «[…] un argomento deduttivo è quello secondo cui la conclusione segue dalle premesse con necessità assoluta, questa necessità non essendo questione di grado, né dipendendo in alcun modo da qualunque altra cosa possa verificarsi; in netto contrasto, un argomento induttivo è quello secondo cui la conclusione segue dalle premesse solo con un certo grado di probabilità, questa probabilità essendo questione di grado e dipendendo da quant'altro possa verificarsi» (I.M. Copi, C. Cohen, *Introduzione alla logica*, Il Mulino, Bologna 1997, p. 75).
>
> […] L'avverbio "probabilmente" è cruciale: mentre è impossibile che la conclusione di un ragionamento deduttivo corretto sia falsa se le sue premesse sono vere, la conclusione di un argomento induttivo non è mai certa, ma solo

> probabile, con probabilità compresa tra 0 e 1.[75]

Non solo l'avverbio "probabilmente" è cruciale, ma anche l'escursione di questa probabilità che cade nell'intervallo numerico tra 0 e 1, indica una volta di più la sostanziale identità di concezione tra induzione e probabilità. Gli autori poi articolano il problema dell'induzione, sviluppando temi che abbiamo già esposto con un dettaglio che qui non ci interessa analizzare; quello che invece ci preme notare è, in primo luogo, come la concezione classica di induzione e deduzione (e quindi di via alla conoscenza) sia considerata «impropria, se non errata».

Nel cammino della ragione verso la conoscenza, Aristotele lega l'induzione alla forma sostanziale di ciò che si conosce, oggi leghiamo l'induzione alla probabilità. Forse più che di una concezione impropria si tratta di due concezioni completamente diverse. Possiamo pensare che Aristotele sbagliasse perché non giungeva alle nostre stesse conclusioni, solo se stiamo cercando di dire la stessa cosa, ma è talmente cambiata la prospettiva su questo argomento, che bisogna avere cautela prima di emettere un giudizio. Boniolo e Vidali giustificano il superamento della concezione classica di induzione e deduzione riportando in nota una serie di esempi di inferenze, corrette, dove i quantificatori non rispettano la regola per cui l'induzione va dal particolare all'universale, mentre la deduzione andrebbe dall'universale al particolare. Prendiamo il caso dell'induzione con conclusione particolare, che è quello che meglio si presta a sottolineare il salto di paradigma:

> Per esempio: "Mario giocava a rugby e si è infortunato alla gamba, Luigi giocava a rugby e si è infortunato al ginocchio, Pietro giocava a rugby e si è infortunato alla schiena. Quindi, anch'io, che gioco a rugby, probabilmente incorrerò in un

[75] G. Boniolo e P. Vidali, *Filosofia della scienza*, cit., pp. 222-223.

infortunio”.[76]

La cosa interessante di questo esempio è che si tratta di una inferenza allo stesso tempo giusta e sbagliata, ma fortunatamente non sotto il medesimo rispetto. Aristotele direbbe che, benché la conclusione sia corretta, il procedimento non è stato applicato correttamente, in quanto si sono fatti collassare in un’inferenza sola un’induzione e una deduzione.

Un esperto di statistica nostro contemporaneo, invece, direbbe che l’inferenza induttiva dell’esempio rappresenta, in modo semplificato ma ineccepibile nella forma, la corretta impostazione per determinare una probabilità, mediante il calcolo della frequenza relativa.

Per illustrare meglio il punto di vista classico, l’inferenza corretta, con tutti i passaggi, avrebbe dovuto essere:

Momento induttivo (premesse particolari e conclusione universale)

> Mario giocava a rugby e si è infortunato alla gamba.
>
> Luigi giocava a rugby e si è infortunato al ginocchio.
>
> Pietro giocava a rugby e si è infortunato alla schiena.
>
> Chi gioca a rugby si infortuna nel corso della sua carriera sportiva

Sillogismo deduttivo

> Chi gioca a rugby si infortuna nel corso della carriera sportiva.
>
> Io gioco a rugby.
>
> Io mi infortunerò nel corso della carriera sportiva.

Ho sostituito “probabilmente” con “nel corso della carriera sportiva” per eliminare il richiamo probabilistico, che non avrebbe senso in un contesto classico e comunque sarebbe inteso in un modo diverso; rimane comunque l’idea di aspettativa e

[76] *Ivi*, p. 222, nota 3 a piè di pagina.

possibilità. Ho anche tralasciato le possibili considerazioni sul fatto che gli infortuni capitati a Mario, Lugi e Pietro costituiscano le sole osservazioni disponibili per inferire deduttivamente la conclusione, in particolare ignorando la totale assenza di sportivi non infortunati, dato che qui ci interessa solo il meccanismo inferenziale dell'esempio.

Che si facciano collassare due tipi di inferenze in una, non porta necessariamente ad un errore, non è pertanto detto che sia un'operazione scorretta in ogni caso, anche se deve comunque essere fatta all'interno di un quadro di regole logiche consentite; a questo proposito, la logica induttiva, formulata da Rudolf Carnap nel 1950, è un esempio di logica non standard, studiata per fornire delle regole di inferenza che rispondono a nuove esigenze[77].

Considerare questa inferenza dal particolare al particolare un esempio di induzione non è una distrazione, e neppure una interpretazione caratteristica di Boniolo e Vidali. Hacking in *L'emergenza della probabilità* fa notare che c'è

> una certa ambiguità presente nel termine ≪evidenza induttiva≫ usato dai filosofi. Esso significa due cose. Da un lato è l'evidenza a favore di una generalizzazione o anche di una legge di natura, ottenuta mediante osservazioni ed esperimenti particolari. Dall'altro, è l'induzione dal particolare al particolare. Hume, in realtà, considera soprattutto quest'ultimo significato, come quando si pone il problema se *questo* pezzo di pane che ho davanti a me sia nutriente. J. S. Mill arrivò fino al punto di sostenere che ogni inferenza è dai particolari ai particolari, in quanto

---

[77] Le logiche non standard sono una famiglia di logiche letteralmente esplose, in varietà numerica e in capacità di attrarre l'interesse dei ricercatori, dopo il 1930 e la pubblicazione dei teoremi di incompletezza di Gödel: l'esistenza di proposizioni indecidibili ha costituito un fastidioso elemento di discontinuità nella concezione della matematica come fonte di certezze senza limiti; fastidio che ha fatto da pungolo all'innovazione. L'idea che sta alla base del fiorire di queste logiche è che, se si estende o si indebolisce il sistema assiomatico della logica standard, si possono generare sistemi formali rigorosi, capaci di superare i limiti degli strumenti concettuali tradizionali e fornire semantiche in grado di lavorare con problemi considerati, fino ad allora, ingestibili. Oltre al problema di coerenza e completezza della matematica, Logiche non standard sono state ideate per affrontare tematiche legate, ad esempio, al calcolo delle probabilità o alla meccanica quantistica, anche spaziando in ambiti puramente teorici, sperimentando sistemi formali che ambiscono a valutare la verità di una proposizione in tutti i mondi possibili di leibniziana memoria, o a lavorare rimuovendo il vincolo di non contraddizione. In molti casi queste logiche si sono trovate alle prese con versioni più complesse dei problemi che si proponevano di superare.

> le generalizzazioni sono semplicemente lo schema dell'inferenza particolare.[78]

Questa ambiguità del termine "induzione", però, mi sembra che non derivi tanto da una ricchezza semantica in sé – quasi una difficoltà dei filosofi a trovare un punto di incontro – quanto dall'aver cominciato a usare un termine esistente con un significato completamente nuovo. Parlando di induzione, perciò, diventa necessario precisare a quale tipo di inferenza ci si vuole riferire e con quale significato si impieghi il termine. Ad esempio, Costantini e Geymonat, nel loro testo sulla *Filosofia della probabilità*, eliminano ogni dubbio già dall'introduzione:

> Un'inferenza induttiva è un ragionamento di carattere non dimostrativo che, servendosi di un metodo di induzione, cioè di una o più assunzioni relative alla funzione di probabilità usata nel corso dell'inferenza, mira ad assegnare un valore di probabilità ad un'ipotesi sulla base di una data evidenza.[79]

L'interpretazione frequentista e i metodi di calcolo delle frequenze relative, paiono dunque uscire vincitori dall'arena dove si sono confrontate le diverse concezioni della probabilità, al punto da far valere una nuova idea di induzione nelle teorie della conoscenza che si prefiggono di giustificare la scienza. Si tratta di una vittoria che deve molto al successo pratico di quello che è il "braccio operativo" del calcolo delle probabilità, la statistica, in quanto le obiezioni e le criticità messe in evidenza trattando della concezione frequentista rimangono per il momento dei punti aperti.

Tra i punti aperti, uno che la riflessione filosofica non dovrebbe trascurare di indagare e approfondire, è se la concezione moderna di induzione e inferenza induttiva, al di là degli indiscutibili successi operativi della statistica, possa davvero fare a meno dell'idea classica di generalizzazione e astrazione. La difficoltà mi

---

[78] Hacking, *L'emergenza della probabilità*, cit., pp. 49-50.
[79] D. Costantini e L. Geymonat, *Filosofia della probabilità*, cit., p. 17.

sembra risiedere in particolare nella generalizzazione completa che, abbiamo visto, si ha quando un certo evento si verifica per tutti i campioni presi in esame. Generalizzare dicendo «che "tutti gli uomini sono mortali" per il fatto che l'essere mortale si è sempre trovato costantemente congiunto con l'essere uomo»[80], accontenta l'inferenza induttiva di matrice frequentista perché fornisce un modo di gestire una frequenza relativa che vale 1 (l'evento certo). Sta di fatto, però, che noi, e ciò vale anche per tutti quelli di cui abbiamo notizia, abbiamo veduto uomini viventi in numero molto più grande di quanti ne abbiamo visti viventi e poi morti, per cui chiunque volesse verificare direttamente la probabilità di morire in chiave frequentista, a un primo rapido esame dovrebbe concludere che gli uomini mortali siano una specie piuttosto rara. Ma anche volendo valutare in maniera più approfondita, prendendo per buone le informazioni che altri riferiscono in merito a persone effettivamente morte, quindi senza pretendere di verificare direttamente la distribuzione, le cose non vanno tanto meglio: in base al "Population Reference Bureau", che fornisce dati statistici sulla popolazione mondiale (e poiché stiamo ragionando in termini di frequenze e probabilità, i dati statistici sono compatibili e in linea di principio ammissibili), dal 50.000 a.C. ad oggi si stima che siano nati circa 109 miliardi di uomini[81] e oggi quasi 8 miliardi di questi sono viventi per cui, approssimativamente, dato che il 93% di tutti gli uomini esistiti è morto, c'è un 7% di probabilità di vivere per sempre; se poi invece degli ultimi 50.000 anni prendiamo in considerazione solo gli ultimi 50, la percentuale degli uomini morti crolla vertiginosamente, ne sono nati infatti 6,64 miliardi e sono attualmente viventi 5,91 miliardi di uomini sotto i 50 anni[82]; questo vuol dire che l'89% degli uomini nati dal 1970 sono vivi, e se consideriamo quanto si è allungata la vita media nell'ultimo secolo, potremmo concludere che ci siano buone possibilità che qualcuno dei nati dopo le guerre mondiali possa finalmente rientrare nel novero degli uomini che non sono mortali. Questa conclusione paradossale suona falsa al nostro orecchio, e non sono possibili vizi di campionamento della popolazione, dato che stiamo

[80] N. Abbagnano, *Dizionario di filosofia*, cit., p. 485.
[81] Fonte: Population Reference Bureau; https://www.prb.org/howmanypeoplehaveeverlivedonearth/
[82] Fonte: United Nations, Department of Economic and Social Affairs; https://population.un.org/wpp/

prendendo in esame l'intera popolazione mondiale di tutta la storia dell'umanità, e comunque ogni partizione della superficie terrestre abitata condurrebbe a una analoga frequenza relativa, che suggerisce una "elevatissima probabilità" che tutti gli uomini siano mortali; eppure noi siamo certi che proprio tutti siano mortali.

Sta di fatto che "tutti gli uomini sono mortali" non è in alcun modo una generalizzazione completa di una inferenza induttiva probabilistica. In questo caso la conoscenza certa, se deve avere origine da una induzione basata su evidenze, può nascere solo affermando che ciò che si è verificato per tutte le generazioni negli ultimi 50.000 anni non può che avvenire anche per la nostra. Ma ciò significa recuperare nei fatti la risposta oggettivista (classica) o soggettivista (kantiana) al problema dell'induzione, che la risposta pragmatica (probabilistica) considerava inconcludenti.

Non si vuole suggerire che l'impostazione pragmatica – probabilistica nei confronti del problema dell'induzione non funzioni. Anzi, funziona talmente bene che si può essere portati a pensare che sia in grado di fornire tutte le risposte.

Anche senza toccare temi che evocano facilmente risvolti trascendentali e metafisici, come la mortalità dell'uomo, l'induzione meramente probabilistica manifesta comunque dei limiti: "tutti i corvi sono neri" è un altro esempio di conoscenza ottenuta induttivamente, almeno nell'accezione contemporanea; anche in questo caso però, non si tratta di una generalizzazione completa, infatti basta fare una ricerca di immagini su internet, impostando come criterio "corvo albino", per verificare che esistono esemplari di corvi che non sono neri. Ciononostante un corvo continua ad essere un animale di colore nero, infatti mentre per un cane o un bovino non sapremmo pronunciarci sul colore che deve avere (a meno di fare riferimento a varietà specificamente selezionate), i corvi sono animali dal corpo interamente nero anche se in alcuni rari casi possono essere bianchi.

Un frequentista meticoloso può richiamare l'attenzione sul fatto che l'albinismo è una anomalia genetica possibile in tutti gli animali, e di conseguenza "tutti" riferito ai corvi deve essere sostituito da un aggettivo che non intenda la totalità. È vero. Come è vero che tutti i corvi hanno due ali e due zampe dotate di quattro dita, ma anche in questo caso anomalie genetiche possono produrre (e di fatto si verificano) malformazioni che impediscono la corrispondenza coi caratteri morfologici della specie. Oltretutto ali, zampe e dita possono mancare anche in seguito ad eventi traumatici.

È difficile mantenere una impostazione puramente probabilistica per giustificare questa variabilità marginale rispetto alla specie, perché in questa prospettiva il colore dei corvi può solo derivare o da un'osservazione degli esemplari della specie, o da una definizione a priori che si vuole dare di cosa rientra nella categoria "corvo"; ma questo è da escludere, altrimenti un corvo albino non apparterrebbe alla specie dei corvi; perciò il colore del corvo deve derivare per forza dall'osservazione. Un bambino cresciuto in campagna è in grado di distinguere un corvo da altre specie di uccelli neri, e se gli capita di vederne uno albino non lo confonde con una colomba, perché sa che si tratta di un corvo e sa, contrariamente all'evidenza, che quel corvo dovrebbe essere nero, come dovrebbe avere due zampe se fosse mutilato. In questo senso è corretto dire che *tutti* i corvi sono neri, anche se non si tratta di una generalizzazione completa; è proprio del corvo, oltre ad avere tutti i caratteri comuni degli uccelli, avere penne e becco nero, ma questo sapere di ciò-che-è-proprio emerge senza che si possa ricondurlo al prodotto di un'induzione pragmatica-probabilistica.

5. Il *pari* di Pascal

Non si può parlare di probabilità senza toccare l'argomento del *pari* (scommessa) che Pascal affronta nei *Pensieri*. Si tratta di un'opera postuma, contenente le riflessioni che Pascal annotava per un testo apologetico che stava preparando. In parte per la natura apologetica, in parte per la redazione rimasta largamente incompiuta, non si tratta di materiale organizzato in modo sistematico e neppure strutturato secondo una disposizione simile tra le diverse edizioni critiche. Qui farò riferimento al testo dell'edizione di Jacques Chevalier.

Pascal espone i termini della scommessa nel frammento 451 dei *Pensieri*, immaginando un dialogo nel quale si confronta con un interlocutore immaginario, di cultura libertina, sostenendo il proprio punto di vista sulla fede e le ragioni per cui il suo interlocutore dovrebbe decidersi in favore di essa:

> 451. *Infinito-nulla. La nostra anima è gettata nel corpo, dove essa trova numero, tempo, dimensioni. Essa vi ragiona sopra e chiama tutto questo natura, necessità, e non può credere ad altra cosa.*
>
> […] Se vi è un Dio, egli è infinitamente incomprensibile, perché, non avendo né parti né limiti, non ha nessun rapporto con noi. Noi siamo dunque incapaci di conoscere ciò che Egli è, né se è.
>
> - Esaminiamo dunque questo punto, e diciamo: «Dio esiste, o non esiste». Ma da qual parte inclineremo? La ragione qui non vi può determinare nulla; vi è un caos infinito che ci separa. Si gioca un gioco, all'estremità di quella distanza infinita, in cui uscirà testa o croce. Su cosa scommetterete? Con la ragione, voi non potete fare né l'una né l'altra scelta; con la ragione, non potete sostenere nessuna delle due.
>
> Non accusate dunque di errore quelli che hanno fatto una scelta, perché non ne sapete nulla.
>
> - No, ma io li biasimo di aver fatto, non questa scelta, ma una scelta; perché, sebbene sia chi sceglie croce sia l’altro commettano lo stesso errore, sono tutti e due in errore: giusto è non scommettere.
>
> - Sì, bisogna ma scommettere. Questo non è lasciato al libero volere, voi siete imbarcati. Ma cosa sceglierete dunque? Vediamo. Dal momento che bisogna scegliere, guardiamo ciò che vi interessa di meno. Avete due cose da perdere: il vero e il bene, e due cose da imparare: la vostra ragione e la vostra volontà, la vostra conoscenza e la vostra beatitudine; e la vostra natura ha due cose da fuggire: l'errore e la miseria.
>
> La vostra ragione non patisce maggior offesa se sceglie l’una o l’altra, dal momento che bisogna necessariamente scegliere. Ecco un punto risolto. Ma la vostra beatitudine? Pesiamo il guadagno e la perdita, se viene croce, che Dio esiste Valutiamo questi due casi: se vincete, vincete tutto, se perdete, non perdete nulla. Scommettete, dunque, che Dio esiste, senza esitare.
>
> - Ammirevole! Sì, bisogna scommettere, ma forse scommetto troppo.
>
> - Vediamo. Poiché vi è uguale possibilità di guadagno e di perdita, se aveste da

> guadagnare due vite contro una, potreste scommettere ancora. […] Ma vi è un'eternità di vita e di felicità. Stando così le cose, quand'anche ci fosse una infinità di casi, di cui uno solo favorevole, voi avreste ancora ragione di scommettere uno per avere due. […] Ma qui vi è una infinità di vita infinitamente felice da guadagnare, una probabilità di vincita contro un numero finito di probabilità di perdita, e ciò che rischiate è finito. Ciò toglie ogni incertezza: dovunque ci sia l'infinito, e non vi sia una infinità di probabilità di perdere contro quella di vincere, non vi è motivo di esitare: bisogna dar tutto.
>
> […] Ogni giocatore rischia in modo cero per guadagnare in modo incerto; e nondimeno arrischia con certezza il finito per vincere il finito con incertezza, senza peccare contro la ragione. […] E così, la nostra proposizione ha una forza infinita, quando c'è da arrischiare il finito in un gioco in cui vi è uguale probabilità di vincita come di perdita, e c'è l'infinito da vincere.[83]

La struttura dell'argomentazione è chiaramente quella di una scommessa, con una premessa in cui si prende in esame quel che si conosce (i dati noti del problema), quel che si potrebbe arrivare a sapere e quel che resta comunque ignoto. Sulla base di queste valutazioni si deve stabilire la scelta più conveniente. Una cosa immediatamente evidente del *pari*, che ne costituisce anche un po' l'assoluta originalità, è il fatto di avventurarsi con una valutazione di carattere probabilistico, là dove normalmente il calcolo delle probabilità non si spinge coi suoi pronunciamenti: la metafisica, la riflessione sull'esistenza di un Dio trascendente.

Il pari di Pascal è stato visto da tanti studiosi delle teorie della probabilità come il primo tentativo di fondare una teoria della decisione:

> La scommessa non è di facile comprensione. I logici l'hanno messa da parte, e hanno avuto torto. Nelle pagine di Pascal troviamo tre argomenti distinti, tutti validi. Ciascuno è un tipico argomento di teoria della decisione, la cui struttura è

[83] B. Pascal, *Oeuvres complètes - Pensées*, a cura di J. Chevalier, Bibliothèque de la Pléiade, Parigi 1954, traduzione di A. Bausola e R. Tapella. Frammento 451

> stata classificata e caratterizzata solo in questo secolo. Anche se Pascal non formulò esplicitamente dei principi, è chiaro che sapeva quello che stava facendo. Il ragionamento era nuovo, ma, data la popolarità che ebbero i Pensieri, divenne comunemente noto-che i giochi di sorte potevano servire come modello per altri problemi di decisione in condizioni di incertezza.[84]

A mio avviso, non era nelle intenzioni di Pascal dare delle linee guida per una teoria della decisione, almeno se non si vogliono ampliare i confini di questa fino a comprendere qualunque cosa comporti una scelta. In primo luogo la differenza sta nel fatto che la teoria della decisione si occupa di come regolarsi di fronte a scelte operative particolari, mentre il *pari* chiede di prendere posizione di fronte a una scelta esistenziale, una scelta che dà orientamento a un'esistenza personale e dalla quale, di conseguenza, dipenderanno tante altre scelte. Poi, la teoria della decisione serve a scegliere tra alternative dalle conseguenze incerte ma il cui impatto è difficilmente nullo e certamente di portata finita, mentre nel *pari* il premio della scommessa non ha vie di mezzo tra zero e infinito. Infine, benché il calcolo delle probabilità non si esima a priori dal maneggiare l'infinito, abbiamo infatti visto che il terzo assioma delle probabilità ha la variante "dell'additività infinita", tuttavia quando capita si tratta sempre dell'infinito numerabile della matematica, un infinito potenziale e incrementabile; Pascal invece fa qui riferimento a un infinito di perfezione, in atto e non incrementabile.

Sicuramente una chiave di lettura del *pari* è rappresentata dal modo in cui Pascal si rapporta all'infinito che è Dio; e a seconda di come connotiamo questo Dio, la scommessa di Pascal espone il fianco a una critica di parzialità e limitatezza per come viene impostato il problema:

> La premessa dell'argomento è che o Dio non esiste, oppure esiste un Dio i cui attributi sono quelli correttamente descritti dalla Chiesa. Il Dio dei musulmani, per

[84] I. Hacking, *L'emergenza della probabilità*, cit., p. 75.

> esempio, non è considerato una possibilità. La conseguenza è che l'argomento di Pascal vale per qualunque problema decisionale che abbia la stessa struttura formale. «Un imano potrebbe ragionare esattamente nella stessa maniera», come osservava Diderot. Può darsi che la partizione degli stati di cose scelta da Pascal sia oggi fuori luogo, ma questo non è che un pensiero in un libro di pensieri. Gli altri pensieri portano altre ragioni a sostegno di tale partizione. Ci sono anche argomenti rivolti ad altre classi particolari di persone; per esempio, quelli che si rivolgono agli ebrei ortodossi, la cui partizione, ovviamente, non ha nulla a che fare con quella, poniamo, di un parigino del tempo. «Dio è o non è»: così Pascal formula la sua partizione.[85]

È chiaro che se in considerazione della natura apologetica dei *Pensieri* si pone l'alternativa «Dio esiste o non esiste», si rischia di interpretarla come "i cristiani hanno ragione o non hanno ragione". In questo caso l'alternativa, ai fini della scommessa, diventa una partizione effettivamente arbitraria e limitata dell'universo degli eventi, e non si capisce per quale motivo, nel momento in cui si parla di probabilità, non si dovrebbero prendere in considerazione altre divinità, assegnando un valore di probabilità anche alla loro esistenza; naturalmente il tutto complicherebbe non poco l'argomentazione relativa al premio in caso di vincita:

> L'argomento è valido, ma le premesse sono discutibili, se non decisamente false. Sono pochi i non credenti disposti oggi ad ammettere che la partizione di Pascal esaurisce tutte le possibilità. Se ammettiamo almeno un'altra alternativa, per esempio la dottrina di alcune sette fondamentaliste, secondo cui Jehovah danna tutti coloro che si trastullano con «l'acqua benedetta e i sacramenti», allora la strategia cattolica non è più dominante. Esiste cioè uno stato di cose possibile in cui tale strategia non dà la vincita più alta.[86]

[85] *Ivi*, p. 80.
[86] *Ivi*, p. 81.

Oltretutto, sempre nell'ipotesi che la scommessa debba giocarsi sull'esistenza di Dio come viene professato nelle varie religioni, Pascal avrebbe tenuto conto invano dell'universo religioso: i Testimoni di Geova e Scientology, ad esempio, non esistevano ancora, e noi avremmo dovuto riformulare e ricalcolare probabilità e utilità dei premi alla comparsa di ogni nuova religione. Cosa dire poi della equiprobabilità che Pascal postula, dal momento che «vi è uguale possibilità di guadagno e di perdita», nel momento in cui dovessimo trovarci alle prese con il "Prodigioso Spaghetto Volante", dio della chiesa Pastafariana[87]? Ampliando così la platea delle religioni sembrerebbe più ragionevole una valutazione soggettiva della probabilità che un certo dio esista, ma intendere il *pari* in chiave soggettivista significa svuotarlo di ogni validità e rendere inconsistente l'idea che sta dietro: convincere l'interlocutore a decidersi per Dio.

Se consideriamo equivalente scommettere su Dio e scommettere sulla religione, diventa valida l'obiezione portata da Hacking:

> Fin dall'inizio, tuttavia, è stata avanzata un'obiezione fallace, esemplificata da questa annotazione del contemporaneo di Pascal, Daniel Huet: «il ragionamento di Pascal si adatta a tutte le religioni; un argomento che dimostra tutto non dimostra niente, eccetto la necessità di abbracciare una qualche religione, ma non per forza quella cristiana». Ci troviamo di fronte a un errore, a meno che non si abbiano motivi indipendenti per escludere la tesi di una religione un po' eccentrica, secondo cui tutte le persone religiose ed esse soltanto possono essere dannate.[88]

La conclusione dell'ultima obiezione suggerisce un motivo in più per ritenere che

[87] Il Pastafarianesimo è una religione fondata nel 2005 da Bobby Henderson nel Kansas (USA) per protestare contro l'introduzione nelle scuole dell'insegnamento dell'Intelligent Design quale alternativa alla teoria evoluzionista. Nonostante il carattere fortemente parodistico dei principi, dei rituali e della gerarchia, la Chiesa Pastafariana reclama per sé il riconoscimento quale religione e in alcuni stati, come l'Olanda, è annoverata tra le religioni ufficialmente riconosciute.
[88] I. Hacking, *L'emergenza della probabilità*, cit., p. 81.

Pascal non facesse questione di religioni, pur non nascondendo in nessun modo la sua identità cristiana: il *pari* prende in considerazione la puntata (la vita che si mette in gioco) il premio (una infinità di vita infinitamente felice), e in caso di perdita è solo il bene finito che si è puntato ad essere perduto. Nessun accenno a castighi o pene dopo la morte, che pure sono una componente importante di molte credenze religiose e che non ci aspetteremmo che uno scommettitore esperto come Pascal dimentichi, valutando vantaggi e svantaggi del gioco.

Il *pari* va dunque letto non secondo la prospettiva di una pratica religiosa per cui optare, ma di un'esperienza che Pascal ha fatto e che vuole comunicare: l'esperienza di Dio che si manifesta come infinito assoluto e fonte di infinita felicità. L'esperienza dell'infinito è filo conduttore dei *Pensieri* e continuamente richiamata in modo esplicito in tutto il frammento 451. “Infinito nulla dello spirito umano di fronte a Dio”, così si apre il discorso che introduce la scommessa; infinita l'aspirazione di beata felicità dell'uomo, che non trova soddisfazione né appagamento al suo desiderio di felicità in nessuno dei beni di cui può godere in questa vita e su questa terra. Solo l'infinito di Dio è in grado di soddisfare l'infinito desiderio dell'uomo[89].

Pascal in questo riprende Agostino, anche se per alcuni versi la sete di infinito sembra dovuta a una privazione dovuta alla caduta originaria più che al limite intrinseco della finitezza umana. L'aspettativa di felicità, innegabilmente presente in ogni uomo, non può essere appagata da nessun bene finito se non in modo transitorio e mai adeguato. Questo perché l'uomo è ontologicamente aperto su un orizzonte infinito che è domanda, attesa, cammino. L'apertura infinita è costantemente in atto – in quanto apertura – e per questo niente di finito può appagare la speranza dell'uomo. Solo un infinito di perfezione può corrispondere all'apertura infinita dell'uomo, presentandosi come risposta, compimento, meta, e perciò felicità beata. Un Dio infinito, fonte di infinita felicità e beatitudine, esaudisce l'infinita inquietudine dell'uomo prima di ogni determinazione che nasce dalla pratica religiosa. Non si vuole affermare che le religioni siano sostanzialmente

[89] B. Pascal, *Oeuvres complètes - Pensées*, cit. Cfr. frammento 370

invarianti rispetto al contenuto dottrinale, e neppure mettere in discussione l'utilità di una pratica religiosa, ma che se l'infinito di perfezione che chiamiamo Dio si dà all'uomo, allora il suo darsi è, per l'orizzonte infinito su cui si apre l'uomo, risposta che viene prima di ogni dottrina e ogni pratica. Per questo Pascal, pur restando cristiano, invita a scommettere per Dio senza dover articolare la proposta sulle religioni particolari.

Il contributo portato da Pascal alla scienza è enorme. La paternità del calcolo delle probabilità, assieme a Fermat, è certamente il tema più noto al grande pubblico, ma a lui si devono anche principi fisici e teoremi matematici, un importante trattato sulle coniche e invenzioni meccaniche, come la pressa idraulica e strumenti di calcolo (la pascalina). Credo che vada considerato questo aspetto della personalità di Pascal, esaminando il *pari*, e vada tenuto presente nella sua connotazione difettiva: in tutto il frammento 451 dei *Pensieri* non è riportato nemmeno un calcolo delle probabilità, neppure abbozzato; eppure non si parla che di scommessa, posta, vincita. Sulla base del testo si possono costruire matrici decisionali e prospetti che riepilogano gli argomenti portati da Pascal e come questi gravano uno sull'altro, ma se leggiamo il testo senza filtri viene difficile pensare che a parlare di scommesse sia lo stesso autore dei carteggi appassionati scambiati col Cavaliere De Méré o con Fermat, e non per la mancanza di passione, ma per l'assenza di calcoli.

Alla luce di questa constatazione, mi pare che non si possa interpretare correttamente il significato del *pari* se non si tiene presente la distinzione che fa Pascal tra *esprit de finesse* ed *esprit de géométrie*, esemplarmente riassunta nella celebre frase «Il cuore ha le sue ragioni, che la ragione non conosce»[90]

Se lo guardiamo in questa prospettiva, il *pari* non rappresenta un aspetto particolare del calcolo delle probabilità; invece dovremmo dire che il calcolo delle probabilità viene impiegato come metodo che aiuta ad indicare e riferirsi a contenuti che non si esauriscono in un'analisi dello spirito geometrico. Che la ragione non

---

[90] B. Pascal, *Oeuvres complètes - Pensées*,cit. Cfr. frammento 477.

conosca le ragioni del cuore significa che non è capace di padroneggiarle e percorrerle autonomamente, non che il cuore paia assurdo alla ragione. Se la fede ha una ragionevolezza che non annienta la razionalità di chi crede, il solo modo per raccontare la fede è farlo calcando i sentieri della ragione in modo da suggerire la meta opportuna, anche là dove la ragione non arrivi a individuarla. L'esperienza che fa Pascal dell'infinito, che gli si presenta come quel Dio che riempie di beatitudine l'infinito nulla umano, o viene conservata come un segreto del cuore o, se la si vuole comunicare, deve essere in qualche modo tradotta in un discorso che ricorra ad analogie e simboli. Pascal vuole testimoniare la sua esperienza, e quale cosa più naturale che recuperare dalla propria storia, dalle frequentazioni libertine e dal proprio talento le parole per spiegare ciò che per natura sua è ineffabile? Altri avrebbero fatto ricorso alla musica, all'arte pittorica, o alla poesia, o ancora all'architettura e alla tecnica e abilità artigiana.

È normale che per parlare di ciò che più ci appassiona e raccontiamo con trasporto si faccia ricorso ai temi che più hanno segnato – nel bene o nel male – la nostra vita. È normale che Pascal per parlare della sua esperienza di Dio come bene infinito faccia ricorso ai temi delle scommesse e del calcolo delle probabilità.

Se il *pari* è questo, se cioè rappresenta la riflessione ragionevole su "come giocarsela" dal momento che «siamo imbarcati», non possiamo considerarlo né un aspetto formale della teoria delle probabilità, né la fondazione di una teoria della decisione. Pascal è molto chiaro sul fatto che il *pari* non è la dimostrazione dell'esistenza di Dio ma della ragionevolezza del nostro investire la vita sulla promessa di felicità e vita infinite. Valutare l'opportunità di una scelta di questo tipo, comporta soppesare pro e contro di natura tale che non si possono incasellare all'interno di un calcolo numerico.

6. Conclusioni

Un obiettivo di questo lavoro era cercare di mostrare che il calcolo delle probabilità, nonostante il successo e il credito di cui gode, non è quella disciplina certa e dai principi assodati che si ritiene comunemente.

In parte perché la comunità scientifica sta ancora riflettendo sul corretto impiego dei suoi metodi operativi, come abbiamo visto nel confronto tra statistica bayesiana e statistica frequentista; e in questo caso il termine “corretto” non sta a indicare che una delle due scuole sia sbagliata, ma che a seconda del tipo di problema che ci si trova ad affrontare un approccio potrebbe essere preferibile rispetto all’altro.

D’altra parte, l’ambivalenza che caratterizza la probabilità, e che buona parte dei suoi studiosi sono disposti a riconoscere, è forse dovuta a una mancanza di chiarezza sul significato della probabilità. Dopotutto l’assiomatizzazione ha rimosso autoritativamente ogni discorso sul senso e la natura della probabilità, ma questo non vuol dire che si possa davvero trattarla come se fosse una questione irrilevante.

La mancanza di una definizione di probabilità soddisfacente per tutti potrebbe essere anche l’indizio di una sovrapposizione di significati che vengono veicolati dallo stesso termine. Carnap distingueva tra *probabilità*$_1$, il grado di credenza razionale, e *probabilità*$_2$, la frequenza. Forse questa distinzione dovrebbe essere ripresa, approfondita, meglio articolata e arricchita.

Ultimo, ma non certo di importanza minore, il calcolo delle probabilità si è sviluppato in un contesto di determinismo assoluto con Laplace, per recuperare recentemente il punto di vista del soggetto con de Finetti, e tra questi due estremi non c’è un’evoluzione. Si tratta anche in questo caso dell’impiego di uno stesso termine per riferirsi a significati diversi, ma con un elemento nuovo: se il soggetto non si riduce a una nuova ipotesi deterministica, deve essere riconosciuta al soggetto la possibilità di agire in un modo non preventivamente calcolabile. Un interessante percorso di indagine filosofica potrebbe essere lo studio di come il soggetto, nel momento in cui valuta la probabilità di un evento, possa essere dotato di libertà, anche nello specifico della filosofia della scienza; de Finetti non mi pare, per quel che ho potuto verificare, che prenda posizione in merito a una espressione libera del soggetto, ma l’espressione libera del soggetto che valuta una probabilità è un argomento che a mio avviso meriterebbe di essere indagato più a fondo da chi si occupa di antropologia filosofica.

Questo elenco di problemi aperti e percorsi da approfondire rispetto alla natura della probabilità e il suo significato, può essere poi declinato con un'infinità di sfumature e punti di attenzione particolari, a cominciare dalla domanda “cosa misuriamo quando misuriamo una probabilità”, che abbiamo visto essere quella cui ci si sottrae più volentieri ma che inevitabilmente torna di continuo.

Sebbene in tanti casi si faccia conto sul calcolo delle probabilità per descrivere il mondo reale, in alcuni casi questa rappresentazione approssimativa parte da una idealizzazione del mondo che non è nemmeno un'approssimazione.

Laplace identificava nella probabilità il mezzo di cui disponiamo per padroneggiare il mondo per quella parte che sfugge alle nostre conoscenze, e i progressi nella termodinamica e nella fisica delle particelle gli hanno dato ragione, ma non deve sfuggire che anche qui troviamo delle limitazioni che non si conciliano con una certa idea di probabilità.

James C. Maxwell, fisico noto per il lavoro sull'elettromagnetismo ma che diede anche un grande contributo alla teoria cinetica dei gas elaborandone un modello probabilistico, ideò un esperimento mentale noto come “diavoletto di Maxwell”. Immaginiamo di avere due contenitori, che chiamiamo A e B, affiancati e comunicanti tramite una piccola botola, pieni di un fluido qualunque in equilibrio termico, vale a dire che il fluido è tutto alla stessa temperatura in entrambi i contenitori. La temperatura, come insegna la termodinamica, dipende dal continuo movimento delle molecole che, in un moto caotico, possono avere singolarmente velocità anche molto diverse, ma che statisticamente hanno una velocità media strettamente correlata alla temperatura del fluido. Se in vicinanza della botola attraverso cui sono messi in comunicazione il contenitore A e il contenitore B si pone un *diavoletto* dalla vista così acuta da poter vedere le molecole in movimento, e così abile nei movimenti da poter aprire e chiudere la botola in modo da far passare nel contenitore B le molecole che nel contenitore A si muovono lentamente quando si trovano nei pressi della botola, e nel contenitore A quelle che si muovono velocemente in B, vedremmo che senza bisogno di fornire energia dall'esterno il fluido nel contenitore A si scalderebbe mentre quello nel contenitore B si raffredderebbe.

Nella realtà non può succedere che un fluido si riscaldi senza che venga fornita energia dall'esterno, e questo divieto è stato sancito nel secondo principio della termodinamica che afferma che esiste una funzione chiamata entropia, che descrive lo stato del sistema termodinamico, e questa funzione può essere soltanto crescente.

La cosa che è importante notare, è che la teoria delle probabilità non impedisce affatto che particelle in movimento si dispongano in modo da avere concentrazioni di particelle lente e altre concentrazioni di particelle veloci. Ma se noi vedessimo una caraffa d'acqua che congela per metà mentre l'altra metà evapora bollendo non penseremmo di aver assistito a un evento dalle probabilità remote, ma che le nostre più basilari concezioni sull'universo sono sbagliate. Vincere alla lotteria e assistere alla violazione del secondo principio della termodinamica sono eventi, da un punto di vista strettamente probabilistico, altrettanto rari, ma uno lo concepiamo possibile e l'altro no, per cui nell'universo oltre a ciò che ci vedeva Laplace (ciò che sappiamo e la probabilità) dobbiamo *probabilmente* aggiungere qualcosa.

In riferimento ai modelli che non hanno una vera corrispondenza con la realtà, il lancio di una moneta non truccata, ad esempio, è qualcosa che esiste solo in letteratura e negli esperimenti mentali, e non per le impercettibili imperfezioni della moneta che la sbilanciano. Si pensi a un giocoliere che facendo roteare le clave le recupera sempre dal manico, il movimento ritmico e sempre uguale delle mani imprime alle clave un numero completo e sempre uguale di rotazioni; per quale motivo allora siamo propensi a ritenere che le probabilità che il giocoliere recuperi la clava sono molto alte, mentre quelle che lanciando una moneta esca sempre testa o sempre croce sono molto basse? Chiunque può verificare che se lancia ripetutamente una moneta, il movimento della mano tenderà a regolarizzarsi e alla fine il risultato del lancio dipenderà in realtà dalla faccia che la moneta mostra prima di essere lanciata. Allo stesso modo, la regolarità geometrica delle facce di un dado non le rende equiprobabili in un lancio, ma facilmente prevedibili nella misura in cui normiamo le condizioni in cui viene fatto.

Ciononostante, anche la prossima generazione di statistici si formerà cominciando a riflettere sul comportamento di un dado "onesto", e ciò non è sbagliato, perché ricorrendo a quegli esempi di casualità riusciamo a comunicare e capirci. Vale però la pena riflettere sul fatto che questa capacità di intendersi su concetti che non sono compiutamente definiti, potrebbe indicare che tra i significati di “probabilità” se ne conserva ancora uno primigenio, che è stato messo in ombra dal fenomeno de *L’emergenza della probabilità*, come la chiama Hacking dando il titolo al suo testo.

Prima di essere una teoria ricondotta a un linguaggio formale e metodi di calcolo, la probabilità è un confrontarsi con ciò che calcolabile non è. Non tutto quel che non sappiamo prevedere ci tormenta richiedendo la nostra attenzione, ma a volte è così. Non tutto ciò che è imponderabile si presenta rivendicando una centralità nella nostra esistenza, ma per qualcosa avviene. Ciò che non si conosce a volte è solo ignoranza dei fatti, altre volte è mistero che interpella e richiede che si prenda una posizione. Per questo, forse, il *pari* di Pascal, che da un punto di vista formale è quanto di più lontano possa esserci dal calcolo delle probabilità, ha sempre esercitato un’attrazione invincibile in chi delle probabilità si occupa; perché seppure spoglio di formalismi è forse l’argomento che più si avvicina al significato originario di probabilità: se «siamo imbarcati» e non possiamo sottrarci, se «con la ragione non possiamo fare né l’una né l’altra scelta», in assenza di garanzie «da quale parte inclineremo?»

Il *pari* merita uno studio specifico, la cui profondità e la cui ampiezza vanno decisamente al di là di quel che si poteva fare in questa sede; approfondimento cui mi auguro di poter fornire un contributo prossimamente. Benché non si tratti di una teoria di calcolo, forse la scommessa di Pascal indica una interpretazione della probabilità finora assente nell’elenco ufficiale.

7. Bibliografia

N. Abbagnano, *Dizionario di filosofia*, UTET, Torino 1992.

M. Allais, *Le Comportment de l'Homme Rationnel Devant la Risque: Critique des Postulats et Axiomes de l'École Américaine*, in «Econometrica», 21, 1953, pp. 503-546.

P. Bartha, *Many Gods, Many Wagers: Pascal's Wager Meets the Replicator Dynamics*, in «Probability in the Philosophy of Religion», A cura di J. Chandler e V. S. Harrison, Oxford University Press, Oxford 2012, pp. 187-206.

T. Bayes, *An Essay towards Solving a Problem in the Doctrine of Chances*, in «Philosophical Transactions», 53, 1763, pp. 370-418.

J. Bernoulli, *Ars Conjectandi*, Tournes, Basilea 1713.

M. Black, *Problems of analysis - Philosophical essays*, Cornell university press, Ithaca 1954.

G. Boniolo e P. Vidali, *Filosofia della scienza*, Bruno Mondadori, Milano 1999.

J. Broome, *The Two-Envelope Paradox*, «Analysis», 55, 1995, pp. 6-11.

G. Brown, *A Defence of Pascal's Wager*, in «Religious Studies», 20, 1984, pp. 465-479.

J. Cain, *Infinite Utility*, in «Australasian Journal of Philosophy», 73, 1995, pp. 401-404.

J. Cargile, *Pascal's Wager*, in «Philosophy», 35, 1966, pp. 250-257.

R. Carnap, *The Continuum of Inductive Methods*, The University of Chicago Press, Chicago 1952.

R. Carnap, *The Logical Foundations of Probability*, in «Journal of Philosophy», 60, 1963, pp. 362-364.

L. J. Cohen, *Introduzione alla filosofia dell'induzione e della probabilità*, Giuffrè, Milano 1998.

M. Colyvan, *Relative Expectation Theory*, in «Journal of Philosophy», 105, 2008, pp. 37-54.

M. Colyvan e A. Hàjek, *Making Ado Without Expectations*, in «Mind», 125, 2016, pp. 829-857.

J. Conway, *On Numbers and Games*, Academic Press, Londra 1976.

D. Costantini, *Ma chi lo ha detto che bisogna decidere?*, in «Statistica Applicata - Italian Journal of Applied statistics», 4, 1992, pp. 65-72.

D. Costantini e L. Geymonat, *Filosofia della probabilità*, Feltrinelli, Milano 1982.

B. de Finetti, *Filosofia della probabilità*, a cura di A. Mura, il Saggiatore, Milano 1995.

K. J. Devlin, *La lettera di Pascal*, Rizzoli, Milano 2008.

A. Duff, *Pascal's Wager and Infinite Utilities*, in «Analysis», 46, 1986, pp. 107-109.

K. Easwaran, *Conditional Probability*, in AA.VV.*The Oxford handbook of probability and philosophy*, a cura di A. Hàjek e C. Hitchcock , Oxford, Oxford University press, 2016, pp. 167-182.

W. Feller, *An Introduction to Probability Theory and its Applications*, vol. II, Wiley, Londra 1971.

H. Fischer, *A Brief History of Probability Theory from 1810 to 1940*, in AA.VV.*The Oxford handbook of probability and philosophy*, a cura di A. Hàjek e C. Hitchcock , Oxford, Oxford University press, 2016, pp. 87-111.

R. Foley, *Pragmatic Reasons for Belief*, in «Jordan» , 1994b, 1994, pp. 31-46.

G. Galati e G. Pavan, *Teoria dei fenomeni aleatori*, TeXmat, Roma 2006.

M. C. Galavotti, *Anti-realism in the Philosophy of Probability: Bruno De Finetti's Subjectivism*, in «Erkenntnis», 31, 1989, pp. 239-261.

M. C. Galavotti, *The Origins of Probabilistic Epistemology: Some Leading 20th-century Philosophers of Probability*, in AA.VV.*The Oxford handbook of probability and philosophy*, a cura di A. Hàjek e C. Hitchcock , Oxford, Oxford University press, 2016, pp. 130-154.

G. Galilei, *Il saggiatore*, G. Barbera, Firenze 1864.

L. Geymonat, *Riflessioni sulle ricerche di Carnap intorno alla probabilità e all'induzione*, in «Rivista Critica di Storia della Filosofia», 10, 1955, pp. 450-461.

C. Glymour, *Probability and the Explanatory*, in «The British Journal for the Philosophy of Science», 66, 2015, pp. 591-604.

I. Hacking, *L'emergenza della probabilità*, trad.it. M. Piccone, Il Saggiatore, Milano 1987.

I. Hacking, *The Logic of Pascal's Wager*, in «American Philosophical Quarterly», 9, 1972, pp. 186-192.

A. Hàjek, *Interpretations of Probability*, in «The Stanford Encyclopedia of Philosophy», 2019.

A. Hájek, *Objecting Vaguely to Pascal's Wager*, in «Philosophical Studies», 98, 2000, pp. 1-16.

A. Hàjek, *Pascal's Wager*, in «The Stanford Encyclopedia of Philosophy», 2018.

O. Hermer e N. Rescher, *On the Epistemology of the Inexact Sciences*, in «Management Science», 6, 1959, pp. 25-52.

F. Huber, *What Should I Believe About What Would Have Been the Case?*, in «Journal of Philosophical Logic», 44, 2015, pp. 81-110.

W. James, *The Will to Believe and Other Essays*, Longmans, Green,and Co, Londra 1897.

J. Jordan, *Pascal's Wager Revisited*, in «Religious Studies», 34, 1998, pp. 419-431.

J. M. Joyce, *How Probabilities Reflect Evidence*, in «Philosophical Perspectives», 19, 2005, pp. 153-179.

A. N. Kolmogorov, *Concetti fondamentali di teoria della probabilità*, TEKNOS, Roma 1995.

A. La Caze, *Frequentism*, in AA.VV.*The Oxford handbook of probability and philosophy*, a cura di A. Hàjek e C. Hitchcock , Oxford, Oxford University press, 2016, pp. 341-359.

P.S. Laplace, *Essai Philosophique Sur Les Probabilités*, Bachelier, Imprimeur-Libraire, Paris 1840.

D. Lewis, *A Subjectivist's Guide to Objective Chance*, AA.VV. *IFS. The University of Western Ontario Series in Philosophy of Science*, vol XV, Springer, Dordrecht 1980.

D. Lewis, *Philosophical Papers*, vol. II, Oxford University Press, Oxford 1986.

A. Lyon, *Kolmogorov's Axiomatization and Its Discontents*, in AA.VV.*The Oxford handbook of probability and philosophy*, a cura di A. Hàjek e C. Hitchcock , Oxford, Oxford University press, 2016, pp. 155-166.

G. I. Mavrodes, *David Hume and the Probability of Miracles*, in «International Journal for Philosophy of Religion», 43, 1998, pp. 167-182.

P. B. Medawar, *Induzione e intuizione nel pensiero scientifico*, Armando, Roma 1970.

D. H. Mellor, *Probability in the Philosophy of Religion*, in «Analysis», 73, 2013, pp. 548-554.

E. Nelson, *Radically Elementary Probability Theory*, in «Annals of Mathematics Studies», 117, 1987.

B. Pascal, *Oeuvres complètes - Pensées*, a cura di J. Chevalier, Bibliothèque de la Pléiade, Parigi 1954.

C. S. Peirce, *Collected Papers*, Harward University Press, Cambridge 1935-1966.

G. Reale, *Storia della filosofia antica*, Vita e pensiero, Milano 1995, vol. V.

N. Rescher, *Probabilities as Potentially Problematic*, in «Mind & Society», 15, 2016, pp. 27-32.

N. Rescher, *Leibniz, Keynes, and the Rabbis on a Problem of Distributive Justice*, in «The Journal of Philosophy», 86, 1989, pp. 337-352.

J. Sprenger, *Bayesianism vs. Frequentism in Statistical Inference*, in AA.VV.*The Oxford handbook of probability and philosophy*, a cura di A. Hàjek e C. Hitchcock , Oxford, Oxford University press, 2016, pp. 382-405.

I. Todhunter, *A History of the Mathematical Theory or Possibility from the time of Pascal to that of Laplace*, Macmillan and CO., Cambridge 1865.

N. Trotignon, *Pascal, Fermat et la géométrie du hasard,* arXiv:1309.2824v1 , 2013.

C. van Fraassen, *Belief and the Will*, in «The Journal of Philosophy», 81, 1984, pp. 235-256.

A. Vulpiani, *Caso, probabilità e complessità*, Ediesse, Roma 2014.

D. Williams, *The Challenging Situation in the Philosophy of Probability*, in «Philosophy and Phenomenological Research», 6, 1945, pp. 67-86.

J. R. G. Williams, *Probability and Nonclassical Logic*, in AA.VV.*The Oxford handbook of probability and philosophy*, a cura di A. Hàjek e C. Hitchcock , Oxford, Oxford University press, 2016, pp. 248-276.

L. Zynda, Subjectivism, in AA.VV.*The Oxford handbook of probability and philosophy*, a cura di A. Hàjek e C. Hitchcock , Oxford, Oxford University press, 2016, pp. 360-381.