PARTE TERZA

CONSIDERAZIONI CRITICHE SULLA STORIA DELLA PROBABILITÀ

Aligia ACCORINTI
Laura BROSIO
Alessandro DE MAGISTRIS
Fiorenzo FERLAINO
Giacinto PLESCIA
Alessandro PORCU

IL CALCOLO DELLE PROBABILITÀ PRIMA DEL XX SECOLO

1.1 - Premessa

Con la fine del XVIII secolo e con l'inizio del XIX si estende lo sforzo per fare chiarezza su alcuni concetti fondamentali, ereditati dalle scuole matematiche dei secoli precedenti, non ancora ben definiti e che tuttavia stavano alla base di importanti teorie, fra cui lo stesso calcolo infinitesimale. L'opera di revisione critica - condotta lungo tutto il XIX secolo da alcuni tra i più grandi matematici come Cantor, Dedekind, Weierstrass - doveva portare all'inizio del XX secolo alla definitiva crisi dei fondamenti della matematica.

L'esigenza di dare basi più solide a teorie scientifiche apparentemente già formalizzate coinvolse rapidamente anche settori più distanti dal corpo centrale delle scienze matematiche, ed in particolare il Calcolo delle probabilità: si tratta di uno sforzo per spezzare il campo di oscurità e di malin

---

(°) Questo capitolo ed i successivi dedicati alla ricostruzione storica dello sviluppo del concetto di probabilità sono una libera sintesi dei primi capitoli del libro: D. COSTANTINI, "Fondamenti del calcolo delle probabilità", Feltrinelli, 1970.

tesi che circondavano il concetto di probabilità, mediante lo apporto di una rigorosa formulazione di ogni parte della teoria: qui però la discussione non investe solo l'interpretazione da attribuire alla probabilità, ma gli scopi stessi che la teoria si prefigge.

Anche in tempi più recenti si incontrano scuole matematiche che collocano la teoria della probabilità tra le scienze sperimentali, altre che la intendono come un ramo della psicologia, ed altre infine che la collegano alla logica formale. Oggi quindi non sembra ancora superata la fase di dibattito intorno ai fondamenti del calcolo delle probabilità; la tendenza resta quella di trovare un terreno comune di confronto tra le varie interpretazioni, riconoscendo porzioni di validità a tutte, in corrispondenza al fatto che tutte rispecchiano tentativi di superare contraddizioni di fondo della teoria e tutte trovano terreni di applicazioni in diversi settori scientifici.

1.2 - La concezione classica

L'interesse verso i fenomeni casuali e quindi verso la probabilità del verificarsi di un evento risale alla più remota antichità, poiché esso nasce con i primi giochi d'azzardo di cui si trova già traccia molti secoli prima di Cristo. Ma i primi studi matematici si sviluppano solo a partire dal XVII secolo, attraverso ricerche di Pascal, Fermat, Bernouilli, ... dirette soprattutto a determinare il numero di combinazioni che conducono allo stesso evento (in quanti modi si può ottenere 7 nel lancio di due dadi?): si tratta dunque di lavori ancora legati strettamente all'interesse per i giochi di azzardo.

La prima opera di sistemazdone teorica viene compiuta da Laplace (1812), il quale cercò anche di fornire una soddisfacente definizione, quella che oggi chiamiamo "definizione classica o di Laplace":

"La probabilità di un evento è il rapporto tra il numero dei casi favorevoli ed il numero dei casi possibili, nell'ipotesi che tutti gli eventi abbiano la stessa possibilità di verificarsi".

La definizione risente ancora evidentemente degli studi sui giochi d'azzardo, ed è utilizzabile solo per spazi campione discreti (vedi Esempio 1 del Cap. 2°, Parte I).

Sottolineiamo l'ultima parte della definizione sulla quale si appuntarono soprattutto le critiche delle scuole che seguirono: supponendo di voler chiarire il concetto di "equipossibilità", ad esempio con riferimento alle 6 facce di un dado, dovremmo ipotizzare una perfetta omogeneità del materiale, mentre l'analisi fisica tende ad escludere tale eventualità. Dunque l'ipotesi di equipossibilità pare inaccettabile, a meno di interpretarla come una ipotesi di "equiprobabilità", ma ciò condurrebbe evidentemente ad un circolo vizioso in cui daremmo una definizione di probabilità che contiene al suo interno il concetto stesso che si cerca di definire.

I limiti di questa definizione erano peraltro già noti fin dal secolo XVIII, anche se il suo uso nelle applicazioni statistiche ha continuato ad estendersi fino ai nostri giorni.

LA CONCEZIONE FREQUENTISTICA

2.1 - La posizione di von Mises

La scuola frequentistica fa soprattutto riferimento agli studi di von Mises e di Reichenbach.

Per von Mises (1919) il calcolo delle probabilità è una branca delle scienze naturali matematizzate, simile per natura alla meccanica razionale ed alla geometria. Come lo scopo della geometria è lo studio dei fenomeni spaziali, così lo scopo della teoria delle probabilità è lo studio dei fenomeni di massa e degli eventi ripetibili.

Non sempre si ha il diritto di parlare scientificamente di probabilità, anche se il linguaggio comune può tranquillamente continuare a farlo: è lecito usare il termine probabilità solo quando esso compare in contesti nei quali sono coinvolti eventi che si presentano ripetute volte. Per comprendere tale distinzione ci serviamo di due esempi: "è probabile che il vertice di Camp David non conduca alla pace in Medio Oriente entro pochi anni"; "è probabile che, lanciando 1000 volte un dado, il numero 2 si presenti 173 volte". Secondo von Mises, la prima affermazione è priva di significato; egli sostiene in fatti: "Il concetto razionale di probabilità si applica solo a problemi in cui lo stesso evento è ripetuto un gran numero di volte, e un gran numero di elementi uniformi sono coinvolti; allo stesso tempo la teoria delle probabilità è la teoria della matematica dei fenomeni di massa".

Nella concezione di von Mises assume un ruolo determinante un nuovo concetto, quello di COLLETTIVO: per ora diciamo solo che esso denota una successione di eventi uniformi che differiscono per certi attributi osservabili. Supponiamo di aver lanciato un dado un gran numero di volte, ad esempio 1000, e di aver constatato che il numero 2 si è presentato 173 volte; il rapporto 173/1000 = .173 è la frequenza relativa dell'attributo 2 nei 1000 lanci. Continuando i lanci, si potranno verificare due casi: la frequenza relativa tenderà a stabilizzarsi attorno ad un certo valore; oppure la frequenza relativa continuerà a presentare notevoli variazioni. Nel primo caso avrà senso parlare di probabilità, mentre sarà del tutto ingiustificato nel secondo caso.

La definizione frequentistica di probabilità è dunque la seguente:

> "La probabilità di un attributo in un dato collettivo è il limite al quale tende la frequenza relativa dell'attributo in questione, al crescere del numero di osservazioni".

Sottolineiamo il fatto che il termine di limite qui usato non coincide con quello abituale nel calcolo infinitesimale; ed infatti è proprio su questo punto che si aprono le critiche alla concezione di von Mises, anche all'interno della scuola frequentistica. Secondo Reichenbach il concetto di limite proposto da von Mises è troppo impreciso per essere posto alla base della teoria della probabilità; egli abbandona dunque qualsiasi riferimento al senso comune ed introduce nella definizione un collegamento con il limite usato nell'Analisi.

Una seconda questione, che attirava le critiche all'impostazione frequentistica di von Mises, riguardava l'esclusione dal la teoria degli "eventi singolari": mentre però le obiezioni concernenti la definizione di limite coinvolgono la logica interna del sistema di von Mises, quelle che investono gli eventi singolari ne minano le stesse fondamenta teoriche.

2.2 - Operazioni sui collettivi

Torniamo al concetto di collettivo e cerchiamo di precisarlo. In primo luogo un collettivo è una successione illimitata nella quale la frequenza relativa di un particolare attributo tende ad un limite determinato. Tale successione deve però soddisfare anche ad un'altra condizione, che potrà essere chiarita dal seguente esempio.

Supponiamo di percorrere una strada dove le distanze sono indicate da cippi: un paracarro ogni 250 metri, una pietra miliare ogni chilometro. Chiediamoci quale sia la probabilità che, fermandoci a caso su questa strada; il cippo più vicino sia una pietra miliare, la risposta è evidentemente 1/4. Infatti, se cominciamo a contare i cippi e indichiamo le pietre miliari con M ed i paracarri con P, otteniamo la successione: P, P, P, M, P, P, P, M, P, ..., ed il limite cui tende la frequenza relativa degli M in questa successione è appunto 1/4.

Supponiamo ora di eliminare dalla successione gli attributi che occupano il 4°, l'8°, il 12°, ... posto, e calcoliamo il limite della frequenza relativa degli M nella successione così ottenuta: questa volta otteniamo 0. In altri termini, eliminando secondo una regola prefissata alcuni cippi, la probabilità in questione cambierebbe. La nuova condizione vuole escludere questa eventualità: affinché una successione illimitata sia un collettivo è necessario che, eliminando secondo una regola prefissata alcuni elementi della successione, il limite della frequenza relativa dell'attributo considerato nella successione residua rimanga invariato (PRINCIPIO DI CASUALITA').

Introduciamo quattro operazioni fondamentali sui collettivi, che permettono di ricostruire le proprietà già note nei secoli precedenti e quindi l'intero calcolo delle probabilità. Useremo spesso il termine DISTRIBUZIONE di un collettivo: esso indica l'insieme di tutti i valori di probabilità relativi ai diversi attributi del collettivo.

La prima delle quattro operazioni sui collettivi è detta SCELTA DI POSTO. Consideriamo il collettivo formato dai lanci di un dado; possiamo formare nuovi collettivi scegliendo solo alcuni tra tutti i lanci: 2°, 4°, 7°, ..., oppure 1°, 4°, 8°, ... Gli attributi del nuovo collettivo restano i medesimi dell'originario e, grazie al principio di casualità, la distribuzione non cambia.

Anche la seconda operazione permette di costruire nuovi collettivi a partire da collettivi dati. Consideriamo ancora il collettivo formato dai lanci di un dado: l'operazione del RAGGRUPPARE consiste nel riunire ad esempio tutti i lanci che danno un risultato pari oppure tutti i lanci che danno un risultato dispari. In questo modo abbiamo raggruppato alcuni attributi (2, 4, 6 oppure 1, 3, 5) per formare un nuovo attributo (pari oppure dispari); se gli attributi originari si escludono a vicenda, la distribuzione del nuovo collettivo si ottiene da quella del collettivo originario sommando le probabilità dei vari attributi originari che concorrono a formare il nuovo attributo.

La terza operazione prende il nome di PARTIZIONE. Domandiamoci quale sia la probabilità che, essendosi verificato un numero pari nel lancio di un dado, esso sia il 4. Per rispondere, lanciamo il dado: se il risultato è dispari, esso non viene trascritto; se invece il risultato è pari, ne teniamo conto. Ad esempio, dal collettivo

661325355 1264453 ...

passiamo al nuovo collettivo

662244 ...

nel quale compaiono solo alcuni degli attributi presenti nel collettivo originario. Evidentemente la distribuzione è cambiata: mentre la probabilità di 4 nel collettivo di partenza vale 1/6; nel nuovo collettivo esso vale 1/3.

Prima di definire l'ultima operazione, introduciamo preliminarmente due nuovi concetti importanti. Il CAMPIONAMENTO CORRELATO è anch'esso un modo per costruire nuovi collettivi mediante altri collettivi. Lanciamo separatamente due dadi: ogni volta che si presenta una determinata faccia nel lancio del primo dado, ad esempio il 3 (e solo allora), trascriviamo il risultato del secondo dado. In questo modo, a partire dai due collettivi formati con i lanci dei due dadi, otteniamo "per campionamento correlato" un nuovo collettivo.

Per quanto riguarda la determinazione della distribuzione di un collettivo ottenuto per campionamento correlato, osserviamo che esso ci porta ad introdurre il concetto di INDIPENDENZA: un collettivo B (secondo dado) è INDIPENDENTE da un altro collettivo A (primo dado) se il campionamento eseguito su B per mezzo di A non ha alcuna influenza sulla distribuzione del collettivo B. Ne segue che la distribuzione del collettivo ottenuto per campionamento è la stessa distribuzione di B, soltanto se A e B sono indipendenti.

Dati due collettivi indipendenti, l'operazione di COMBINAZIONE consiste nel formare un nuovo collettivo trascrivendo simultaneamente gli elementi e gli attributi di entrambi i collettivi. La distribuzione del collettivo nuovo si ottiene moltiplicando tra loro le probabilità dei singoli attributi dei collettivi originari.

Confrontando le considerazioni di questo paragrafo con quelle del paragrafo 2.7 della Parte I, si comprende come le operazioni del raggruppare e di combinazione permettano di giungere ai principi delle probabilità totali e delle probabilità composte.

2.3 - Reichenbach ed il metodo dei posit

Abbiamo già rilevato che la critica alla teoria frequentistica di von Mises si appuntò soprattutto contro l'inadeguatezza del concetto di limite e contro l'esclusione del calcolo delle probabilità per eventi singolari.

Per superare queste critiche Reichenbach (1935) modifica la definizione degli oggetti su cui deve operare il calcolo delle probabilità: in luogo dei collettivi di von Mises sottoposti a condizioni abbastanza restrittive, Reichenbach prende in considerazione successioni di ogni genere, che egli chiama CLASSI DI RIFERIMENTO. Inoltre il concetto di limite usato nel calcolo delle probabilità coincide qui con quello abituale del calcolo infinitesimale.

Mentre lo studio delle successioni costituisce la base indispensabile per la teoria delle probabilità di von Mises, per Reichenbach il concetto di probabilità è slegato da ipotesi sulle successioni prese in esame; ed al contrario è proprio attraverso la definizione di probabilità che si arriva ad una classificazione delle successioni.

Per risolvere l'altro problema cruciale, quello degli eventi singolari, Reichenbach introduce il metodo dei "posit": in esso si continua a negare la possibilità di parlare di probabilità per gli eventi singolari, ma si indica come agire quando si debba prendere una decisione relativa ad un evento singolare.

Supponiamo di effettuare una serie di prelievi da una stessa gettata di cemento e di constatare carichi di rottura superiori alle 8500 lb negli 8/10 dei casi. Come rispondiamo alla domanda: "un nuovo prelievo dalla stessa gettata presenterà un carico di rottura superiore alle 8500 lb?". Non possiamo usare qui il concetto di probabilità, perché si tratta di un evento singolare, e ci si aspetta una sola delle due risposte "sì" o "no". Ebbene Reichenbach propone di rispondere "sì" perché, se ripetessimo la domanda numerose volte e rispondessimo sempre "sì", daremmo una risposta esatta negli 8/10 dei casi.

Riassumendo, un "posit" è una asserzione con la quale ci comportiamo come se fosse vera, sebbene il suo valore di verità sia sconosciuto. Tale comportamento è giustificato solo se è in qualche modo possibile costruire una classe di riferimento per l'evento che vogliamo "porre". Di fronte al fatto che questa presunta possibilità non esiste in numerosi casi (come nell'esempio sul vertice di Camp David del n. 2.1), Reichenbach riconosce la impossibilità di costruire una classe di riferimento in molte situazioni, ma la imputa solo allo stato delle nostre conoscenze, esprimendo quindi la certezza che in futuro sarà possibile superarla.

2.4 - Il problema dell'induzione

Rimane aperto ancora uno dei problemi centrali del calcolo delle probabilità, la cui importanza appare evidente ogni volta che ci si trova di fronte a situazioni specifiche. Infatti la definizione di probabilità come il limite della frequenza relativa non ci dice nulla circa il modo di conoscere le probabilità iniziali, poiché una successione infinita non è realizzabile e non possiamo quindi servirci dell'esperienza per determinare tali probabilità iniziali. In altri termini la definizione frequentistica introduce un'astrazione che porta la teoria fuori dal campo sperimentale e non si ha quindi più nessun diritto di invocare l'esperienza.

Reichenbach affronta il problema, conscio del fatto che, senza la sua soluzione, la teoria frequentistica perderebbe gran parte del suo significato. Poiché solo attraverso una inferenza induttiva è possibile passare dalle frequenze relative alle probabilità, è chiaro che il problema chiama in causa l'induzione.

Reichenbach svolge la sua analisi dei metodi di induzione finora proposti e giunge alla conclusione che l'unica inferen za induttiva non riconducibile al calcolo delle probabilità è l'induzione per enumerazione: tutti gli altri metodi sono esprimibili in termini di teoremi del calcolo delle probabilità. E' proprio l'induzione per enumerazione che permette di passare dalla frequenza relativa (sù una successione finita) alle probabilità, e dunque tutto il calcolo delle probabilità si regge sull'induzione per enumerazione.

Reichenbach propone una regola di induzione che si riallaccia al metodo dei "posit": anche qui si "pone" che una frequenza relativa sia uguale alla probabilità, e non ci si preoccupa di controllare se l'uguaglianza è vera o falsa. Si tratta di una precisazione del normale comportamento induttivo : quando eseguiamo una rilevazione, ci aspettiamo che la precisione nella stima aumenti al crescere del numero delle osservazioni.

Ma ciò presuppone l'esistenza di un limite a cui la frequenza relativa tenda, e questo equivale a supporre che esista la probabilità dell'evento di cui ci stiamo occupando, senza peraltro conoscere esplicitamente il valore limite. Reichenbach accetta questa ipotesi che esista la probabilità di ogni evento ripetibile. La regola consiste allora nel "porre" uguale alla probabilità le frequenze relative via via calcolate aumentando il numero delle osservazioni.

Ci si può domandare quanto della teoria rimanga valido se si lascia cadere l'ipotesi dell'esistenza del limite delle frequenze relative. Evidentemente è necessario sciogliere il dilemma e rispondere alla domanda: esiste o non esiste tale limite? Oppure, poiché l'esistenza di tale limite equivale al principio di uniformità della natura: è o non è uniforme la natura? Reichenbach riconosce che una risposta definitiva non è possibile, ma ribadisce che è ugualmente possibile agire: se un limite esiste, la regola ci condurrà ad esso; se invece il limite non esiste, non potremo raggiungerlo, ma nessuna altra regola potrebbe farci raggiungere il limite inesistente. Dunque l'uso della regola di induzione è razionale. In sostanza la giustificazione dell'induzione da parte di Reichenbach è la seguente: non si sa se la natura sia o meno uniforme, ma è razionale agire come se la natura fosse uniforme.

2.5 - Conclusione

A partire dai primi anni del secolo XX la maggior parte degli studiosi di probabilità e statistica abbandonano la concezione classica di Laplace e accettano, pur con diverse sfumature, le idee dei frequentisti, soprattutto sotto la spinta della scuola inglese di Fischer e Pearson. D'altronde la stessa Parte I del presente testo contiene numerosi esempi in cui prevale in modo evidente una impostazione frequentistica.

Riassumiamo brevemente le maggiori critiche all'interpretazione frequentistica, dividendole in due filoni: quelle interne che, pur riscontrando alcune insufficienze, ne accettano la impostazione di fondo, e quelle esterne che tendono a respingere la definizione stessa di probabilità avanzata dai frequentisti.

Come si è visto, la principale critica interna investiva il concetto di limite su cui von Mises fonda la sua definizione di probabilità: esso faceva un impossibile appello al buon senso comune ed alle capacità dell'osservatore. Reichenbach infatti abbandona ogni riferimento al senso comune e basa la definizione di probabilità sul concetto di limite usato nel calcolo infinitesimale; ma abbiamo visto che, con questo, egli finisce per sconfessare alcuni punti base del frequentismo. In primo luogo egli non introduce alcuna distinzione tra i vari tipi di successione: ora se il concetto di limite dell'analisi ben si applica alle successioni che abitualmente si incontrano in tale settore della matematica, altrettanto non si può dire per le successioni casuali che sono l'oggetto specifico del calcolo delle probabilità. Nella definizione di limite gioca infatti un ruolo determinante la conoscenza della legge di formazione degli elementi della successione, mentre le successioni casuali sono proprio caratterizzate dall'assenza di qualsia si tipo di regolarità.

Una seconda contraddizione nasce dall'assunzione fatta dai frequentisti che il calcolo delle probabilità è una delle scienze naturali matematizzate, la quale studia i fenomeni di massa ed è quindi una teoria empirica. Infatti Waismann ha chiarito che i teoremi della teoria frequentistica delle probabilità, anche se si riferiscono ai fenomeni di massa, sono però del tutto analitici. Mentre le dimostrazioni di questi teoremi discendono solo dalla definizione di probabilità attraverso l'uso di metodi logico-matematici, la determinazione dei valori di probabilità è un fatto empirico: solo l'osservazione sperimentale ci permette di arrivare alla determinazione dei valori di probabilità degli eventi.

Abbiamo visto che proprio questa contraddizione costringe Reichenbach ad operare una netta separazione con l'esperienza, non solo a livello dei teoremi, ma già nel momento stesso della valutazione delle probabilità.

Una terza critica, forse la più grave, avanzata dall'interno contro la teoria di von Mises, è diretta al concetto di collettivo così come era stata formulata originariamente. Non entriamo nel merito di tale critica, per non appesantire la trattazione, ma rileviamo solo che essa dimostrò l'inconsistenza delle condizioni che i collettivi dovevano soddisfare, e cioè dimostrò che il principio di casualità era incompatibile con l'esistenza di collettivi. Va detto che si sono sviluppate ricerche, anche recentemente, che si propongono di formulare un nuovo principio di casualità, il quale non renda vuota la classe dei collettivi.

Per quanto riguarda le critiche esterne, possiamo dire che esse sono riconducibili ad una sola, che sintetizziamo nella seguente affermazione: non è lecito escludere dal calcolo delle probabilità gli eventi singolari.

Von Mises esclude di proposito gli eventi singolari dalla sua teoria; ma è evidente che tale limitazione finisce per escludere una classe troppo vasta di eventi. Infatti non si perdono solamente proposizioni come "il vertice di Camp David non condurrà alla pace in Medio Oriente nel volgere di pochi anni" oppure come "la squadra italiana di pallavolo sarà sconfitta da quella sovietica"; in questo modo si finisce per affermare che nemmeno una proposizione scientifica possa avere un grado di probabilità, e ciò appare una restrizione troppo pesante.

Si è visto come Reichenbach abbia tentato di superare queste critiche con il suo metodo dei "posit"; ma anche questa soluzione va incontro a forti riserve. L'affermazione che per qualsiasi tipo di evento è possibile costruire una classe di riferimento non è convincente: in ogni caso andrebbe sostenuta con validi argomenti e non sulla base di una semplicistica fiducia nel futuro.

Vogliamo infine rilevare come anche la regola di induzione di Reichenbach sia stata sottoposta a critiche: esse hanno condotto Salmon ad una sua giustificazione mediante il principio di invarianza linguistica, sulla cui validità non riteniamo di poterci addentrare.

LA CONCEZIONE LOGISTICA

3.1 - Origini del logicismo

La concezione frequentistica considerava la probabilità come un puro dato sperimentale; la principale obiezione avanzata contro questo indirizzo sottolinea che essa non è in grado di assegnare valori di probabilità a una classe troppo vasta di fenomeni. La concezione logicistica, il cui indirizzo fa capo a Wittgenstein e Keynes, si differenzia dalla frequentistica soprattutto perché sostiene che anche gli eventi singolari debbono essere trattati dal calcolo delle probabilità. Nel contempo i logicisti presentano punti di contatto con i frequentisti, dal momento che essi pure considerano la probabilità un fatto oggettivo.

Posizioni logistiche si riconoscono nell'opera dell'austriaco Wittgenstein, anche se egli si occupa della probabilità solo per poco più di due pagine nel suo "Tractatus logico-filosoficus" edito nel 1921. In esso la probabilità è intesa come una relazione esistente tra certi fatti conosciuti e un fatto non conosciuto; in generale questo è un fatto futuro, ma nulla cambierebbe se si trattasse di un fatto presente o passato.

La probabilità di una certa proposizione è una relazione logica che lega questa a un'altra proposizione e ad una serie di proposizioni che vengono prese come supporto alla prima.

Per usare il linguaggio di Wittgenstein la seconda proposizione dà alla prima proposizione una certa probabilità; sarà dunque possibile costruire la teoria delle probabilità senza alcun riferimento all'esperienza, in modo completamente indipendente da quest'ultima. In questo senso il distacco tra le posizioni frequentistiche e l'atteggiamento di Wittgenstein non poteva essere più marcato.

3.2 - La concezione di Keynes (1921)

Fra le idee di Wittgenstein e di Keynes circa il significato da attribuire al termine di "probabilità" esistono delle notevoli differenze.

Le idee di Keynes costituiscono per certi versi un passo indietro rispetto alla concezione di Wittgenstein, anche se epistemologicamente la trattazione del primo è molto più completa di quella del secondo autore.

D'altra parte si riscontrano nella trattazione di Keynes notevoli lacune, in primo luogo nella definizione di probabilità quale relazione logica. Bisogna riconoscere tuttavia come questo fu il primo tentativo che si proponesse di assiomatizzare le concezioni logicistiche.

Keynes distingue due tipi di conoscenze: una diretta e una indiretta. La prima è quel tipo di conoscenza che otteniamo per esperienza diretta, prevalentemente costituita dalle nostre sensazioni. La conoscenza indiretta è invece quella che ricaviamo dalla prima mediante qualsiasi tipo di inferenza. E' diretta la conoscenza che otteniamo dalla contemplazione degli oggetti della nostra esperienza, è indiretta quella che otteniamo mediante ragionamento.

La teoria delle probabilità tratta di quella parte di conoscenza che otteniamo mediante ragionamento e inoltre dei differenti gradi con cui i risultati ottenuti in questo modo sono conclusivi.

Tutte le proposizioni sono vere o false, ma quello che di esse possiamo conoscere dipende dall'ammontare di conoscenze di cui disponiamo.

Quando nel *linguaggio comune* si parla di proposizioni probabili o improbabili lo si fa sottointendendo delle conoscenze di cui disponiamo; a rigore un'asserzione probabilistica dovrebbe riferirsi a detto insieme, poiché quando si afferma che una proposizione possiede un certo valore di probabilità si esprime una relazione esistente tra la proposizione in questione e un corpus di conoscenze attuali o ipotetiche. Nessuna proposizione è di per sé "probabile" o "improbabile" così come nessun luogo può essere intrinsecamente distante o vicino.

Ciò che Keynes intende dire è che ogni nostra osservazione è relativa ad un insieme di conoscenze ed è proprio per questo che possiamo avere un grado di credenza in una data proposizione. E' necessario distinguere tra credenza razionale e mera credenza, sottolineando che credenza vera e credenza falsa. Può accadere che una credenza razionale sia in effetti falsa.

I termini certezza e probabilità descrivono quindi i vari gradi di *credenza razionale* che, sulla base di differenti insiemi di conoscenze, siamo autorizzati ad avere in una data proposizione. Da ciò deriva la convinzione che la teoria delle probabilità sia una parte della logica. Il calcolo delle probabilità, così inteso, coinvolge unicamente relazioni logiche tra proposizioni (premesse) che assumiamo come conosciute e chiamiamo "*evidenza*" e proposizioni (conclusioni) per cui cerchiamo una conoscenza indiretta che chiamiamo *ipotesi*.

Se un insieme di conoscenze "h" giustifica una credenza razionale in "a" di grado "α", allora dichiamo che c'è una relazione di probabilità di grado "α" tra "a" e "h", e scriviamo

$$\frac{a}{h} = \alpha .$$

L'unica differenza esistente tra la logica tradizionale e la teoria delle probabilità è dovuta al fatto che la prima permette di derivare da premesse vere conclusioni vere, mentre la seconda non può soddisfare una simile condizione. In altri termini il senso delle conclusioni tratte mediante il calcolo delle probabilità è solo formalmente contenuto in quello delle premesse. Le due teorie sono del tutto simili nelle relazioni (oggettive e logiche) che legano le premesse alle conclusioni.

Per la stessa ragione per la quale è senza senso chiamare una proposizione probabile senza specificare l'ammontare di conoscenza a cui essa è correlata, la medesima proposizione può avere più di un grado di credenza razionale o probabilità.

Ammettiamo che una proposizione abbia, sulla base dell'insieme h di conoscenze, una probabilità p; se le nostre conoscenze aumentano, occupando un insieme h_1, la probabilità p_1 della medesima proposizione dovrà essere in generale diverso da p.

Secondo Keynes non è possibile dare una definizione di probabilità, una definizione cioè che lega l'"evidenza" all'"ipotesi". Egli afferma che questa relazione di probabilità entra in gioco quando dalla logica dell'implicazione certa e delle categorie di verità e falsità si passa alla logica della probabilità e delle categorie di conoscenza, ignoranza e credenza razionale.

La convinzione di Keynes è che la probabilità vada valutata mediante una sorta di intuizione logica che permetterebbe di arrivare alla conoscenza diretta di molte relazioni probabilistiche; è chiaro che questa specie di sesto senso probabilistico non può convincere.

Un notevole contributo è stato dato da Keynes allo sviluppo del principio di indifferenza. Come sappiamo le probabilità iniziali giocano un ruolo di fondamentale importanza nel calcolo delle probabilità. Per la loro determinazione ci si rifà appunto al principio di indifferenza: è necessario infat ti disporre di un certo numero di alternative che siano esaustive e a due a due esclusive; se inoltre si riesce a stabilire che queste alternative sono equiparabili, si arriverà mediante una semplice divisione a conoscere i valori di probabilità che ci interessano.

Il principio di indifferenza, che assolve appunto il compito delle equiprobabilità delle alternative, può essere formulato nel modo seguente.

"Se non si conoscono ragioni per predicare di un soggetto una piuttosto che un'altra di parecchie alternative, allora relativamente a tale conoscenza l'osservazione di ciascuna di esse ha uguale probabilità".

Così formulato tuttavia il principio di indifferenza porta a conclusioni paradossali. Si consideri ad esempio una proposizione del cui soggetto si conosca solo il significato e di cui non possediamo alcuna evidenza esterna circa la verità o la falsità. Sia essa "questo libro è rosso", ciò che conosciamo non ci permette di concludere nulla circa la sua verità o la verità di "questo libro non è rosso". In altri termini non possediamo alcuna ragione valida per ritenere che il libro sia rosso oppure no.

Le due proposizioni sono esclusive ed esaustive e per il principio di indifferenza la probabilità della proposizione "questo libro è rosso" è ½, ma possiamo ripetere lo stesso ragionamento per le proposizioni "questo libro è verde, giallo ...". Il risultato è chiaramente paradossale.

Per Keynes le contraddizioni sorgono perché non è sufficientemente chiaro quanto dell'applicazione è puramente meccanico, e quanto è invece lasciato all'intuizione. Le contraddizioni sorgono cioè quando le alternative non sono finali, intendendo per finale un'alternativa che non può essere ulteriormente scomposta. Sulla base di queste considerazioni Keynes suggerisce di aggiungere al principio la seguente condizione: "il principio non è applicabile a una coppia di alternative quando una di esse non è ulteriormente scomponibile in una coppia di alternative possibili e incompatibili della stessa forma dell'alternativa originaria".

3.3 - Conclusione

Ci sembra opportuna qualche considerazione conclusiva. Il primo risultato dell'impostazione logicistica è che lo scoglio degli eventi singolari viene superato, in quanto a qualsiasi evento, ripetibile o no, è possibile assegnare un valore di probabilità relativamente ad una data evidenza. Come questa evidenza sia stata acquisita e se sia valida non è problema che interessi, poiché la teoria logicistica tratta delle relazioni che intercorrono tra l'evidenza e la proposizione di cui cerchiamo la probabilità. La determinazione delle probabilità iniziali può quindi essere fatta senza alcun ricorso all'esperienza e quindi senza alcun appello all'induzione.

Per quanto riguarda Keynes, va sottolineata nella sua opera la chiara e documentata esposizione degli autori precedenti e l'approfondita analisi del principio di indifferenza col conseguente rigetto della sua formulazione classica. Per contro la carenza più marcata del suo sistema è rappresentata dall'intendere la probabilità come "relazione indefinibile che non può essere percepita se non mediante intuizione". D'altra parte l'assoluto logicismo di Keynes trascura nel modo più completo i dati sperimentali, cosicché nella sua opera non vi è mai alcun riferimento all'esperienza, come se nel valutare le probabilità, non avessero alcuna importanza le osservazioni precedentemente effettuate. Con l'impostazione di Keynes si è caduti nell'eccesso opposto a quello dei frequentisti.

Concludiamo evidenziando la convinzione logicistica che il calcolo delle probabilità sia una teoria dell'inferenza induttiva; ne deriva che l'induzione, come la probabilità, è un fatto logico: "la validità e la ragionevolezza di una generalizzazione induttiva è una questione logica e non di esperienza, di leggi formali e non materiali".

LA CONCEZIONE SOGGETTIVISTICA

4.1 - Ramsey e il grado di credenza

Il frequentismo e il logicismo, come abbiamo visto, pur assegnando alla nozione di probabilità due significati diversi mantengono tuttavia un denominatore comune: considerano la probabilità un dato oggettivo. Al contrario, il soggettivismo nega l'oggettività della probabilità.

La probabilità non è un dato di fatto, il limite di una frequenza relativa (frequentismo) o una relazione logica (logicismo), bensì la speranza, o meglio il grado di credenza, che ogni individuo possiede circa il verificarsi di un evento. Per questa ragione le differenze tra logicisti e soggettivisti vanno considerate ben più profonde di quanto è comunemente ritenuto. Anche se nell'impostazione formale i soggettivisti non differiscono notevolmente dai logicisti, usando entrambi lo stesso linguaggio e sostanzialmente la stessa simbologia, le interpretazioni che i due indirizzi danno al termine di "probabilità" non potrebbero essere più divergenti. Tale differenza trae origine dal rifiuto da parte dei soggettivisti del principio di indifferenza, quale momento atto a valutare le probabilità iniziali.

Per i soggettivisti le probabilità iniziali devono essere valutate sulla scorta delle convinzioni individuali e quindi la ricerca di una regola a priori che dovrebbe permettere la loro individuazione è un non-senso. Le ragioni che spingono un individuo ad ammettere un certo grado di credenza circa il verificarsi di un evento potranno essere oggetto di studio della psicologia ma non del calcolo delle probabilità che deve accettarle così come sono. E' invece suo compito precipuo quello di derivare nuovi gradi di credenza a partire dai dati.

Più esplicitamente, per i soggettivisti non esiste alcuna argomentazione che possa impedire a un individuo di valutare la probabilità come crede, nessun ragionamento potrà cioè impedirgli di valutare pari a 1/1.000.000 la probabilità dell'uscita di testa nel lancio di una moneta apparentemente regolare anche se la probabilità in questione è universalmente ritenuta essere 1/2.

Si tratta di un rifiuto perentorio di qualsiasi considerazione inter-soggettiva nella valutazione delle probabilità. Non sembra quindi inesatto raggruppare i frequentisti ed i logicisti sotto la comune denominazione di oggettivisti (alla luce degli sviluppi moderni, le differenze tra i due indirizzi sono meno marcate di quanto comunemente non si creda), e sottolineare che le divergenze interne tra gli oggettivisti sono minori di quelle esistenti tra questi ultimi ed i soggettivisti.

Ramsey (1926) fu il primo a suggerire questa interpretazione in termini moderni, ma ciò non deve indurre a credere che la sua sia una posizione chiusa verso altri indirizzi. Ramsey accetta infatti che la concezione frequentistica abbia una funzione da svolgere nella statistica e nella fisica; inoltre riconosce che: 1°) nel linguaggio comune il termine probabilità è spesso usato come sinonimo di proposizione logica; 2°) se la teoria delle probabilità viene considerata come un ramo della matematica pura e sviluppata conseguentemente, il più semplice e meno controverso modello per il calcolo così costruito è quello che interpreta la probabilità in termini di frequenza relativa; 3°) almeno in via di ipotesi, la probabilità, come vie ne usata nelle teorie statistiche, è in realtà il limite di una frequenza. Non è quindi dalla critica al frequentismo che muove la sua teoria, bensì dalla constatazione che le relazioni di probabilità della teoria logicistica sono così poco chiare da lasciare un lecito dubbio circa la loro esistenza.

L'intuizione logico-probabilistica è ben lungi dall'essere reale. La via che Ramsey intende seguire consiste nel definire la probabilità di un evento come il grado di credenza che un individuo ha nel verificarsi di quell'evento. La teoria della probabilità viene quindi considerata come la logica della credenza parziale; essa è una generalizzazione della logica formale ma in questo processo di generalizzazione si perde uno degli aspetti più importanti della logica simbolica, si perde cioè il suo aspetto oggettivo.

Dicendo che la teoria della credenza parziale è un ramo della logica non si intende dire che la logica ha qualche parte nella determinazione della probabilità degli eventi; la logica della credenza parziale non può suggerire ad un individuo quale deve essere la sua aspettativa nel verificarsi dell'evento "testa" nel lancio di una moneta. Le sue previsioni, e quindi il suo grado di credenza, possono essere quelle che preferisce e il compito della logica a questo riguardo è di rendere il nostro individuo consapevole del fatto che, se nutre certe aspettative, è obbligato, per non contraddirsi, ad averne certe altre, e nulla più. Ne segue che il principio di indifferenza, che abbiamo visto essere così insoddisfacente, diventa del tutto inutile e quindi, con grande vantaggio per il calcolo delle probabilità, eliminabile.

Per parlare sensatamente di grado di credenza parziale, bisogna essere in grado di misurarla. Un sistema soddisfacente deve in primo luogo assegnare ad ogni credenza una grandezza o un grado che abbia una posizione determinata in un ordine di grandezze. Ma questo non è tutto; è necessario anche assegnare dei numeri in modo comprensibile.

Siamo portati quindi a chiederci cos'è un grado di credenza. La risposta che offre il nostro autore è la seguente: il grado di credenza è una proprietà causale di quest'ultima che si può vagamente esprimere come "la misura in cui siamo disposti ad agire in base a questa credenza". Per Ramsey il calcolo delle probabilità deve limitarsi ad escludere che due valori di probabilità siano in contrasto tra di loro. La scelta delle probabilità iniziali è del tutto arbitraria e non può assolutamente essere influenzata dalla teoria delle probabilità che si limita a constatare l'incoerenza se esiste, tra un insieme di probabilità iniziali. (L'incoerenza si riduce, in ultima istanza, al rigetto delle leggi che governano le scelte e pone perciò l'individuo che l'accolga in condizione di subire una perdita certa).

4.2 - B. de Finetti

A de Finetti (1930-31) spetta il merito di aver fatto conoscere l'impostazione soggettivistica. Le pubblicazioni dell'italiano, che operò indipendentemente da Ramsey, sono di pochi anni posteriori al lavoro di quest'ultimo; nonostante ciò, esse suscitano tra i probabilisti un interesse ben maggiore dell'opera dell'inglese.

I due autori, pur arrivando a conclusioni pressoché identiche, muovono tuttavia da due punti di partenza diversi. Mentre per Ramsey è la critica di Keynes che lo spinge alle posizioni soggettivistiche, per de Finetti è la constatazione dell'uso indiscriminato e molte volte errato dello schema bernoulliano (uso che si diffuse in conseguenza dell'accettazione quasi unanime dei probabilisti dell'interpretazione frequentistica) che lo induce a rivedere profondamente la definizione di probabilità.

Il processo di soggettivizzazione è spinto, nell'opera dell'italiano, alle sue estreme conseguenze, nel senso che la possibilità di altre interpretazioni viene negata in modo perentorio. La probabilità non è nulla più dell'espressione dello stato d'animo di un individuo di fronte ad avvenimenti in-certi. Essa è una nozione del tutto soggettiva di stretta pertinenza della psicologia. Il tentativo di oggettivarla è destinato a fallire; non solo, ma generando pericolose confusioni, crea pseudo-problemi che intralciano la corretta formulazione del problema.

Questa posizione nasce dalla convinzione dell'esistenza di due logiche: la logica del certo e la logica del probabile, tra le quali non esiste alcun punto di contatto. Esse sono nettamente separate giacché, la prima si occupa dei fatti che sono veri o falsi mentre la seconda analizza l'insieme dei fatti possibili, assegnando loro un grado di probabilità. La logica del certo è la logica formale, la logica dell'incerto è il calcolo delle probabilità.

Il calcolo delle probabilità non entra nel merito della valutazione delle probabilità dei casi possibili, non suggerisce cioè come debbono essere assegnate le probabilità iniziali. La loro valutazione è del tutto arbitraria e dipende da ciascun individuo, che opererà una scelta sulla scorta delle conoscenze di cui dispone in quel momento. Ciò non implica che la valutazione sia in qualche modo vincolata allo stato di conoscenza che è in quel momento di dominio universale. In altri termini il nostro individuo potrebbe non dare alcuna importanza a ciò che in quel momento costituisce il patrimonio comune della nostra epoca e valutare la probabilità in modo del tutto indipendente da esso. Egli potrebbe quindi valutare 1/1000 la probabilità dell'evento "testa" nel lancio di una moneta, unicamente sulla base di certe sue convinzioni che, ovviamente, sarebbero in netto contrasto con il patrimonio di esperienze acquisito. Tutto ciò che il calcolo può e deve fare è evitare che egli si comporti incoerentemente; nel nostro caso l'individuo sarà obbligato a valutare 999/1000 la probabilità dell'evento "croce". Tutto ciò di cui disponiamo è una frequenza relativa che non deve essere confusa con la probabilità. L'errore di identificare la probabilità con la frequenza relativa, nasce da una parte dal considerare alcuni eventi come ripetibili e dall'altra dalla sopravalutazione dello schema bernoulliano.

de Finetti a questo riguardo è di una chiarezza estrema: eventi ripetibili non esistono, ogni evento è sempre un fatto singolare.

La probabilità è relativa a eventi, e quindi non ha senso parlare della probabilità di un evento ripetuto e identificare la probabilità con la frequenza relativa. Il rovesciamento nei riguardi dell'interpretazione frequentistica non poteva essere più marcato: le conclusioni alle quali arriva de Finetti sono infatti la negazione perentoria e assoluta della concezione oggettivo-frequentistica.

4.3 - Conclusione

Per i soggettivisti, come per i logicisti, è possibile assegnare valori di probabilità a eventi singolari. Accettando che il calcolo delle probabilità si occupi anche degli eventi non ripetibili, si ridimensiona la pretesa di una fondazione esclusivamente frequentistica della teoria delle probabilità e si mettono in luce nuovi e più sottili significati del termine probabilità. (In ciò ravvisiamo l'importanza storica della lotta sostenuta dai soggettivisti contro la pretesa suddetta, ed è doveroso ricordare che il merito di aver contrastato l'esclusivismo frequentista spetta quasi completamente a de Finetti).

Questo non deve tuttavia far dimenticare i limiti ed i pericoli insiti nel punto di vista soggettivistico. Infatti le posizioni radicali assunte da alcuni di loro, rifiutando qualsiasi apertura verso altri indirizzi, rischiano di compromettere i risultati raggiunti.

Per i frequentisti la probabilità è un fatto sperimentale e quindi oggettivo; i soggettivisti controbattono: è un'illusione sostenere che la probabilità sia in qualche modo oggettivabile, essa è un fatto soggettivo irriducibile a fattori non psicologici. I frequentisti affermano che il calcolo delle probabilità trae la sua ragione d'essere dall'evidenza di eventi ripetibili e che solo ad essi è applicabile; con Reichenbach arrivano a sostenere che tutti gli eventi, almeno in linea di principio, possono essere ripetuti. Per i soggettivisti al contrario non esistono eventi ripetibili, ogni evento è un mondo chiuso e finito in se stesso; è quindi priva di senso qualsiasi affermazione probabilistica tendente ad abbracciarne più d'uno.

La posizione dei soggettivisti sembra equivalente alla negazione dei fondamenti stessi della scienza, in ragione del fatto che, se è innegabile l'irrepetibilità di ogni fenomeno, è pure innegabile che gli uomini hanno creato la scienza proprio quando hanno cercato di rendersi conto di quelle regolarità che pure attraverso il continuo mutamento si delineavano.

Il problema non va quindi posto nei termini di una contrapposizione fra i due indirizzi, bensì sottolineando che ciascuna interpretazione sostiene punti di vista solo parzialmente validi e quindi accettando quanto di valido vi è in queste istanze, rifiutando tuttavia le estreme conseguenze.

Detto questo, dall'esame della definizione di probabilità dei soggettivisti e dei suoi presupposti filosofici, risultano palesi i limiti della concezione soggettivistica. Identificare infatti la probabilità con l'espressione dello stato d'animo di un individuo significa sostenere che le sue valutazioni di probabilità non hanno alcun valore per altri individui, poiché questi avranno in generale stati d'animo diversi. Se cioè le valutazioni di probabilità sono relative allo stato d'animo di un individuo allora tutte le conclusioni basate su giudizi probabilistici non hanno alcun significato, indipendentemente dall'individuo che ha valutato la probabilità.

A questo proposito citiamo un passo di Gnedenko (The theory of probability) "... La scienza ha stabilito parecchi risultati positivi sulla base di giudizi probabilistici ... che non differiscono dai risultati ottenuti senza ricorrere alla probabilità. Ad esempio in fisica tutte le proprietà macroscopiche dei gas sono dedotte da assunzioni relative alla natura probabilistica del comportamento delle singole molecole. Se dobbimo attribuire un significato oggettivo indipendente dall'osservatore a queste deduzioni, allora le ipotesi probabilistiche iniziali concernenti il corso dei processi molecolari, devono essere qualcosa di più di un'affermazione dello stato psicologico in cui ci troviamo quando pensiamo al movimento delle molecole."

D'altra parte, sostenere senza riserva la concezione soggettivistica respinge la valutazione della probabilità a un livello prescientifico, poichè, se è vero che agli inizi di ogni disciplina troviamo valutazioni di tipo soggettivo, è innegabile che lo sviluppo scientifico di una disciplina ha inizio quando si accettano criteri di valutazione interrogativi.

Vi è poi un altro pericolo insito nel sostenere incondizionatamente il punto di vista soggettivistico: quello di negare il valore normativo dell'esperienza sulla determinazione della probabilità, cioè, in ultima istanza, di negare il razionalismo scientifico e di accettare involuzioni antirazionalistiche.

L'ASSIOMATIZZAZIONE DEL CALCOLO DELLE PROBABILITÀ

5.1 - Posizione di Kolmogorov

All'inizio degli anni trenta si era concluso il processo di differenziazione tra i sostenitori delle tre principali interpretazioni del termine "probabilità", processo che era iniziato nella seconda metà del XIX secolo. I sostenitori di ciascuna interpretazione erano ormai fermamente convinti che le altre due scuole fossero in errore e conseguentemente ignoravano del tutto le loro argomentazioni (con la parziale eccezione di von Mises).

Proprio in quegli anni il calcolo delle probabilità e le sue applicazioni nella statistica avevano conosciuto un notevole sviluppo e affinamento di metodi; l'attenzione dei probabilisti era quindi incentrata quasi esclusivamente su questi progressi, mentre era riluttante ad occuparsi di un problema così scottante quale la definizione di probabilità.

Teniamo presente che ciascuno dei tre indirizzi in contrasto, oltre a fornire una interpretazione del termine "probabilità", si riprometteva di dimostrare come dalla loro interpretazione derivassero i principi del calcolo delle probabilità e pertanto cercava di far vedere che quest'ultimo era una conseguenza logica di tale interpretazione. Si andava quindi facendo strada la convinzione che si stessero costruendo differenti teorie della probabilità; si pensava cioè che esistesse una teoria frequentistica, una teoria logicistica ed una teoria soggettivistica.

In questo quadro si colloca il lavoro del matematico russo A. Kolmogorov, il quale affermò in modo esplicito (nel 1933) che:

1) il calcolo delle probabilità era, ad onta delle contrastanti interpretazioni, uno e uno solo;

2) era possibile lasciare da parte le diatribe sui fondamenti e, dimenticando ogni tipo di interpretazione, sviluppare in modo puramente formale il calcolo delle probabilità come un ramo della matematica, o meglio come una branca della teoria delle funzioni additive d'insieme.

La convinzione di fondo sulla quale si regge tutto il lavoro di Kolmogorov è che la teoria delle probabilità può e deve essere sviluppata assiomaticamente come la geometria, prescindendo cioè dalle questioni inerenti il significato del termine "probabilità", così come la geometria si disinteressa dei significati dei termini "punto", "retta" e "piano". Dunque Kolmogorov introduce la probabilità come termine primitivo, definito implicitamente dagli assiomi che ne governano l'uso.

5.2 - Gli assiomi del calcolo delle probabilità

Sia E un insieme di elementi E_1, E_2, ..., detti eventi elementari e sia $\mathfrak{F}$ una famiglia di sottoinsiemi di E, i cui elementi sono detti eventi casuali. Introduciamo i seguenti assiomi:

K 1. $\mathfrak{F}$ è un campo di insiemi; ciò significa che, se A e B sono in $\mathfrak{F}$ allora anche $A \cup B$, $A \cap B$ e $A - B$ sono in $\mathfrak{F}$;

K 2. E stesso appartiene a $\mathfrak{F}$;

K 3. ad ogni insieme A di $\mathfrak{F}$ è assegnato un numero reale non negativo $p(A)$; $p(A)$ è detto la probabilità di A;

K 4. $p(E) = 1$;

K 5. se $A \cap B = \emptyset$, allora $p(A \cup B) = p(A) + p(B)$.

Non riteniamo necessario illustrare estesamente questi assiomi: abbiamo visto nella Parte I e nei precedenti capitoli di questa stessa Parte III che tali assiomi sono proprio quei teoremi base del calcolo delle probabilità che frequentisti, logicisti e soggettivisti ottenevano a partire dalle rispettive definizioni di probabilità.

Esponiamo invece la posizione di Kolmogorov nei confronti di un concetto di fondamentale importanza nel calcolo delle probabilità: quello di INDIPENDENZA. Come giustamente Kolmogorov fa notare, questo concetto ha dato alla teoria della probabilità il suo carattere peculiare ed è, con la precisazione della nozione di probabilità, uno dei problemi più importanti della filosofia della probabilità.

Per Kolmogorov l'indipendenza è una proprietà degli esperimenti o delle prove; qui ci limitiamo a presentare la definizione nel caso di due soli esperimenti $^{(1)}$ e $^{(2)}$, cioè di due sole decomposizioni dell'insieme E in sottoinsiemi disgiunti:

$$E = A_1^{(1)} \cup A_2^{(1)}, \qquad E = A_1^{(2)} \cup A_2^{(2)}.$$

In questo caso possiamo dire che i due esperimenti sono indipendenti se:

$$p(A_1^{(1)} \cap A_1^{(2)}) = p(A_1^{(1)}) \cdot p(A_1^{(2)}) \,;$$

$$p(A_2^{(1)} \cap A_1^{(2)}) = p(A_2^{(1)}) \cdot p(A_1^{(2)}) \,,$$

$$p(A_1^{(1)} \cap A_2^{(2)}) = p(A_1^{(1)}) \cdot p(A_2^{(2)}) \,,$$

$$p(A_2^{(1)} \cap A_2^{(2)}) = p(A_2^{(1)}) \cdot p(A_2^{(2)}) \,.$$

Ci sembra opportuno far notare che la definizione ora introdotta di indipendenza è del tutto insoddisfacente. Essa attribuisce infatti la nozione di indipendenza agli esperimenti e non agli eventi, come si è invece soliti fare nei manuali di calcolo delle probabilità. Se si accetta che l'esperimento sia qualcosa di essenzialmente diverso dagli eventi, come sembrano indicare tutte le trattazioni intuitive di questa nozione, la definizione esplicativa di esperimento è del tutto inadeguata poiché riduce l'esperimento ad un evento di tipo particolare.

5.3 - Concetto di variabile aleatoria

Uno dei passaggi più rilevanti nella teoria assiomatica è la trasformazione dello studio di una funzione di probabilità (la funzione p) su eventi (gli elementi di $\mathcal{F}$) nello studio di una funzione di probabilità su numeri reali. Questo passaggio si basa essenzialmente su un nuovo concetto fondamentale : quello di VARIABILE ALEATORIA.

La complessità della definizione formale di variabile aleatoria ci impedisce di approfondire l'argomento: ci limitiamo a ribadire che essa permette di prendere in considerazione, in luogo degli elementi $A \in \mathcal{F}$, numeri reali x o intervalli $I \leq \mathbb{R}$. Quindi, anziché considerare valori $p(A)$, operiamo su valori $P^{(x)}(I)$ o, più precisamente, su valori

$$F^{(x)}(a) = P^{(x)}(-\infty, a),$$

dove la nuova funzione $P^{(x)}(I)$, detta appunto funzione di probabilità, viene definita a partire dal concetto di variabile aleatoria, e corrisponde sostanzialmente con quella che, nella Parte I, abbiamo chiamato distribuzione (cumulativa) di probabilità.

Dalla definizione di variabile aleatoria segue immediatamente che le funzioni di distribuzione godono delle seguenti proprietà:

(i) $\lim_{a \to -\infty} F^{(x)}(a) = 0$;

(ii) $\lim_{a \to +\infty} F^{(x)}(a) = 1$;

(iii) $F^{(x)}(a)$ è non decrescente e continua a sinistra.

Se $F^{(x)}(a)$ è derivabile, la sua funzione derivata

$$f^{(x)}(a) = \frac{d}{da} F^{(x)}(a)$$

è detta densità di probabilità della variabile aleatoria x nel punto a.

Introduciamo ancora la seguente definizione:

sia x una variabile aleatoria; l'integrale

$$M(x) = \int_{-\infty}^{+\infty} t \cdot dF^{(x)}(t)$$

è detto il VALOR MEDIO di x.

Il valor medio gode delle seguenti proprietà:

(i) $|M(x)| \leq M(|x|)$;

(ii) $M(y) \leq M(x)$ se $0 \leq y \leq x$ ovunque;

(iii) $M(kx + ly) = k\, M(x) + l\, M(y)$ in cui k e l sono costanti;

(iv) ogni variabile aleatoria limitata possiede valor medio.

5.4 - Conclusione.

L'opera di Kolmogorov dimostra come si possa costruire tutto il calcolo delle probabilità a partire dagli assiomi, senza fare appello ad alcuna nozione esterna, quindi senza ricorrere ad alcuna interpretazione del termine probabilità.

Questo è senza dubbio il più grande passo avanti compiuto dagli studiosi che hanno sostenuto il punto di vista assiomatico. Ma ciò non deve suonare come una affermazione che nega la possibilità di ulteriori sviluppi. A questo proposito possiamo riportare il pensiero di L. Geymonat: "Il modo stesso con cui abbiamo cercato di delineare - spiegandone le origini storiche - i caratteri e la funzione del metodo assiomatico vale a porre in luce un fatto apparentemente ovvio, ma in realtà tutt'altro che privo di importanza: il fatto che l'assiomatizzazione rappresenta una tappa molto significativa e originale, ma non tale da farci dimenticare l'esistenza di altre tappe che l'hanno preceduta e preparata".

Si possono di fatto individuare delle limitazioni alle quali soggiace l'impostazione assiomatica: nell'intero sviluppo del calcolo non compare alcun numero, se si escludono i casi limite 0 e 1, ma unicamente relazioni di probabilità; il calcolo, basato sugli assiomi di Kolmogorov, ci pone cioè in condizione di lavorare quando si disponga di valori di probabilità, più precisamente quando si siano ricavate le probabilità iniziali. Se non si conosce alcun valore di probabilità, il calcolo non permette, da solo, di costruirne.

Da un lato l'assiomatizzazione del calcolo delle probabilità riafferma l'unicità del calcolo stesso, dall'altro ne sottolinea il carattere di teoria strettamente matematica; nonostante la sostanziale validità di queste tesi, esse contribuirono allo sviluppo di due posizioni di cui solo una è accettabile.

La prima tesi ha avuto un effetto chiarificatore molto importante poiché ha ridimensionato le discussioni sulla definizione della probabilità, riportando quindi entro giusti limiti le ricerche relative alla determinazione dei valori di probabilità.

Non solo: l'esplicita affermazione dell'unicità della teoria della probabilità costituì forse il punto di partenza per quegli autori che una decina d'anni più tardi avrebbero cominciato a porsi il problema di ritrovare una interpretazione in grado di colmare la frattura in atto fra i tre indirizzi storici.

D'altra parte però, la convinzione che il calcolo delle probabilità fosse solamente una branca della matematica spinse molti matematici e probabilisti a ritenere del tutto inutile lo studio dei problemi legati ai fondamenti delle probabilità. Ciò comportò, in primo luogo, la negazione di qualsiasi autonomia alla teoria della probabilità, e inoltre diede adito all'infiltrarsi entro la teoria di nozioni non sufficientemente analizzate, ad esempio l'ipotesi di equiprobabilità, che non si uniformavano al rigore caratteristico dello sviluppo assiomatico della teoria della probabilità.

LA TEORIA DELLA CONFERMA DI R.CARNAP

6.1 - Premessa

Potrebbe sembrare che la pur breve trattazione del punto di vista assiomatico sul calcolo delle probabilità costituisca un punto di arrivo rispetto agli obiettivi delle presenti dispense, quasi avesse posto la parola "fine" alle polemiche tra frequentisti, logicisti e soggettivisti, fornendo una risposta definitiva ai numerosi problemi che erano sollevati nei primi decenni del secolo XX. In effetti questo è l'atteggiamento di gran parte dei manuali di statistica e di probabilità e corrisponde al comportamento "sul campo" della maggior parte di coloro che utilizzano tali strumenti: ciò che interessa (e nemmeno poi sempre!) è constatare come dagli assiomi di Kolmogorov sia possibile ottenere strumenti e teoremi utilizzabili nelle infinite applicazioni concrete del concetto di probabilità.

A noi tutto ciò non pare sufficiente: sottolineiamo qui la profonda scorrettezza della seconda tesi indicata alla fine del capitolo precedente, quella che considera ormai inutile riprendere in esame il problema della determinazione dei valori (iniziali) di probabilità. Il fatto che quasi tutti gli utilizzatori del calcolo delle probabilità condividano pienamente tale tesi non preoccupa tanto sul piano del confronto tra diverse opinioni scientifiche: in realtà ciò non ha affatto impedito - come vedremo - che proseguisse l'indagine teorica sul concetto di probabilità.

Preoccupa invece il rifiuto di prendere in esame il problema di come assegnare le probabilità iniziali perché esso lascia il ricercatore o il tecnico ... libero di assegnare tali probabilità secondo motivazioni dettate da chi lo paga o lo condiziona, e non in base ad un metodo che abbia una sua oggettività.

Proprio per i motivi sopra esposti riteniamo utile fare alcuni brevi cenni all'opera di R. Carnap, uno degli studiosi che più ha contribuito a mantenere vivo il dibattito sul concetto di probabilità, superando le rigide fratture tra frequentisti, logicisti, soggettivisti e riportando il dibattito stesso sul terreno della determinazione dei valori di probabilità.

Precisiamo subito che Carnap non rappresenta un nuovo punto fermo in questo processo di costruzione della teoria della probabilità, ma semmai rappresenta un nuovo e importante punto di partenza.

6.2 - Il grado di conferma

Come si è già accennato, il grande libro "Logical Foundations of probability" (1950) di Carnap supera le posizioni di rigida contrapposizione tra le varie posizioni. Più precisamente egli parte dalla constatazione che nel linguaggio prescientifico esistono due nozioni di probabilità, irriducibili tra loro, indicate rispettivamente con i termini "probabilità 1" (Pr1) e "probabilità 2" (Pr2): la prima corrisponde al concetto di grado di conferma di una ipotesi, introdotto dalla scuola logicistica; la seconda corrisponde alla frequenza relativa della scuola frequentistica.

Carnap sostiene che è compito degli scienziati costruire due "explicata" scientifici distinti a partire da questi due "explicanda" del linguaggio comune; l'errore degli autori precedenti è stato di voler ridurre Pr1 e Pr2 ad un solo concetto scientifico. Mentre Pr1 entra in gioco quando si vuol valutare la probabilità di una certa ipotesi h sulla base di un insieme e di risultati (evidenza), invece Pr2 è usato, all'interno delle teorie scientifiche, per esprimere la frequenza relativa con cui si presenta un certo evento in una successione di prove. La frequenza è dunque un fatto sperimentale, basato sull'osservazione dei fatti; Pr1 esprime invece una relazione puramente logica tra una ipotesi ed una evidenza entrambe formulate in un certo linguaggio.

Nonostante queste profonde differenze tra i due concetti, essi danno origine a sorprendenti analogie: essi sono entrambi funzioni di due argomenti (Pr1 di ipotesi h ed evidenza e, Pr2 di evento e successione); i valori di entrambi sono compresi tra 0 e 1; infine le due teorie si sviluppano in modo quasi identico a partire da Pr1 e Pr2.

Carnap prende anche in esame i contributi della scuola soggettivistica, ma tende (anche in lavori successivi degli anni 60) a far coincidere la loro concezione con quella corrispondente a Pr1. Nonostante il fatto che, dopo i primi paragrafi, Carnap restringa la sua analisi solo a Pr1, è dunque evidente il suo sforzo di riportare il confronto fra le tre posizioni sul piano del dibattito scientifico.

La teoria sviluppata da Carnap ha un carattere strettamente logico in quanto studio dei metodi atti a determinare il grado di conferma di una ipotesi sulla base di una data evidenza. Dati due enunciati h ed e, di cui e deve non essere falso, un'asserzione probabilistica ha quindi la forma

$$c\ (h,\ e) = r\ ,$$

dove r è un numero compreso tra 0 e 1, estremi inclusi, che si legge: il grado di conferma dell'ipotesi h, sulla base dell'evidenza e, è r.

6.3 - Cenni ad alcune questioni di logica

Partiamo da una prima osservazione: gli enunciati probabilistici, cioè gli enunciati della forma $c(h,e) = r$, non appartengono al linguaggio al quale si riferiscono, cioè al linguaggio di cui fanno parte e ed h; tali enunciati appartengono invece ad un metalinguaggio. In altre parole: una cosa è il linguaggio oggetto su cui si opera, altra cosa è il metalinguaggio della logica induttiva.

Una importante conseguenza di questa osservazione è che il valore del grado di conferma non dipende solo dai due enunciati e ed h, ma anche dalla struttura del linguaggio in cui e ed h sono espressi. Un primo decisivo passo di Carnap è dunque quello di una attenta analisi del linguaggio su cui si vuol costruire la logica induttiva, cioè la teoria della probabilità.

Non è qui possibile affrontare una analisi così complessa e articolata; vogliamo solo sottolineare una questione che può assumere maggiore rilievo oggi, man mano che crescono i contributi nella direzione indicata da Carnap: i linguaggi usati da Carnap e da altri recenti studiosi sono estremamente semplici, se li volessimo confrontare con i linguaggi che sarebbero necessari per descrivere la maggior parte delle teorie appartenenti alle scienze sperimentali, le quali peraltro sono proprio il terreno sul quale si vorrebbero applicare i risultati del calcolo delle probabilità. Per essere più chiari, diciamo che il linguaggio più abituale negli studi sui fondamenti della probabilità non è nemmeno in grado di esprimere la teoria dei numeri naturali; si immagini la complessità di un linguaggio che esprima la teoria della relatività einsteiniana oppure certe teorie sociologiche. Orbene, nonostante la semplicità del linguaggio adoperato, la risposta fornita al problema della determinazione dei valori (iniziali) del grado di conferma, cioè della probabilità, rimane di una complessità tale da renderne difficile l'utilizzazione pratica.

Il processo con il quale Carnap giunge a proporre una sua soluzione a questo problema avviene con l'introduzione di suc essive condizioni sulla funzione "grado di conferma", che abbiamo indicato con c(h,e), restringendo sempre più l'insieme delle funzioni c accettabili, fino ad arrivare a cogliere una sola c* (nell'Appendice della sua opera).

6.4 - Considerazioni conclusive

I lavori di Carnap successivi al 1950, insieme con quelli di numerosi altri autori, portano a nuovi significativi progressi. Carnap stesso (1963), Hintikka (1965) ed altri studiosi giungono a definizioni assiomatiche della probabilità che tentano di fornire una soluzione unitaria alle vecchie polemiche: da un lato gli assiomi di Kolmogorov costituiscono il primo blocco dei sistemi di Carnap e di Hintikka; dall'altro lato i successivi assiomi permettono di costruire una funzione "grado di conferma" che tiene conto (attraverso l'introduzione di "pesi" opportunamente determinati nelle varie circostanze) sia del punto di vista frequentistico-sperimentale, sia del punto di vista logicistico-induttivo. E proprio la possibilità di scegliere i "pesi", per chi deve applicare il calcolo delle probabilità nei casi specifici, sembra assicurare un riconoscimento anche alle posizioni soggettivistiche.

Si è già detto che le soluzioni finora proposte per la determinazione dei valori iniziali di probabilità sono alquanto macchinose; tuttavia la strada percorsa negli ultimi decenni appare promettente ed ha comunque posto fine alle contrapposizioni sterili tra frequentisti, logicisti e soggettivisti. Tuttavia a noi sembra interessante concludere questa parte delle dispense sottolineando i limiti del processo di assiomatizzazione; a questo proposito condividiamo il punto di vista di L. Geymonat esposto alla fine del capitolo precedente e che ribadiamo con le seguenti parole di E. Casari: "... non v'è ragione di ritenere che la tappa costituita dal metodo assiomatico debba essere considerata come una tappa definitiva".

Carnap R., Logical foundations of probability, The University of Chicago Press, Chicago, 1950.

Replies and sistematica expositions, The philosophy of R. Carnap, ed. P. Schlipp, La Salle Illinois, 1963.

Casari E., Questioni di filosofia della matematica, Feltrinelli, Milano, 1964.

Costantini D., Fondamenti del calcolo delle probabilità, Feltrinelli, Milano, 1970.

Finetti de B., Fondamenti logici del ragionamento probabilistico, Boll. UMI, vol. IX-5, dic. 1930.

Sul significato soggettivo delle probabilità, Fund. Mat., vol. 17, Warsaw 1931.

Teoria delle probabilità, Einaudi, Torino, 1970.

Geymonat L., Filosofia e filosofia della scienza, Feltrinelli, Milano, 1960.

Gnedenko B.V., The theory of probability, New York, 1968.

Hintikka J., Towards a theory of inductive generalization, Logic Meth. and Phil. of Sc. Proc., Int. Congr., ed. Bar-Hillel, Amsterdam, 1965.

Keynes J.M., A treatise of probability, London, 1921.

Kolmogorov A., Grundbegriffe der Wahrscheinlichkeitsrechnung, Berlin, 1933.

Laplace P.S., Theorie analytique des probabilites, Paris, 1812.

Essai philosophique sur les probabilites, 1814.

Mises von R., Grundlagen der Wahrscheinlichkeitrechnung, Math. Zeit., 1919, 279-294.

Wahrscheinlichkeit, Statistik und Wahrheit, Wien, 1928.

Ramsey F.P., Truth and probability, 1926.

Reichenbach H., Wahrscheinlichkeitslehre, Berlin, 1935.

Wittgenstein L., Tractatus logico-philosophicus, 1921.

* * *