

Incertezza: il pensiero sistemico di Bruno de Finetti

Giordano Bruno¹

Desideriamo la verità e troviamo in noi solo incertezza.

Blaise Pascal, Pensieri, BUR Rizzoli, 1983

Sembrerebbe pertanto naturale che i modi abituali di pensare, di ragionare, di decidere, dovessero esplicitamente e sistematicamente impennarsi sul fattore *incertezza* come sull'elemento concettualmente preminente e determinante.

Bruno de Finetti, Teoria delle probabilità, Einaudi, 1970

Sunto

L'approccio scientifico *meccanicistico*, basato su principi quali quello per cui il mondo microscopico è più semplice di quello macroscopico, e che quello macroscopico può essere espresso attraverso la conoscenza precisa di infiniti dettagli, è stato messo in crisi da molte questioni, una delle quali è rappresentata dal cosiddetto "Problema dei Tre Corpi" (Barrow-Green, J., 1996).

Nuovi approcci conoscitivi sono stati introdotti basati, solo per citarne qualcuno, sul ruolo teorico dell'osservatore, sulla non-

¹ ISIA (Istituto Superiore Industrie Artistiche)
bigi@dmmm.uniroma1.it

linearità, su principi di incertezza, sul costruttivismo, sul punto di vista sistemico e sull'emergenza.

Parole chiave: Probabilità, soggettivo, oggettivo, coerenza di eventi, logica dell'incerto

Introduzione

In questo lavoro, voglio mettere in rilievo come il pensiero di Bruno de Finetti si inserisca perfettamente in questi tipi di approccio, anzi li anticipi, soffermandomi nell'evidenziare quale sia il ruolo essenziale di colui che effettua valutazioni di probabilità, che chiameremo *l'osservatore*.

Al contrario degli orientamenti di tipo oggettivistico, in particolare quelli che si richiamano ad una visione prettamente statistica della realtà, Bruno de Finetti ha avuto, nel secolo scorso, l'intuizione fondamentale per cui le valutazioni di probabilità espresse relativamente ad un evento non rappresentano altro se non il *grado di fiducia* che un *individuo coerente* ha nel realizzarsi di quell'evento. Tale grado di fiducia viene determinato sulla base delle *informazioni* che l'individuo possiede in merito a quell'evento e in tal senso l'individuo svolge il ruolo di osservatore, come accade nella moderna *teoria della complessità*.

de Finetti ci ha insegnato che la logica dell'incerto esplora il contesto del possibile, accettando la condizione di essere inadatta a fare *predizioni*, fornendo al contrario gli strumenti per assegnare probabilità agli eventi, in modo da poter effettuare *previsioni* e prendere decisioni.

Tradizionalmente il concetto di probabilità è stato considerato come una conseguenza della nostra *ignoranza*, delle nostre limitazioni. Nello stesso modo i principi dell'incertezza erano

considerati dal pensiero meccanicistico come un limite della nostra conoscenza di infiniti dettagli.

Ora possiamo, invece, considerare la probabilità come il nostro *strumento* per descrivere la natura.

La metodologia seguita da de Finetti si basa sulla *coerenza* nelle assegnazioni di probabilità a singoli eventi e a famiglie di singoli eventi.

La coerenza è lo strumento essenziale che permette di effettuare valutazioni di probabilità e, in particolare, di aggiornare la valutazione di un evento quando una nuova (o supposta) informazione divenga utilizzabile (d'interesse).

Mi propongo di far vedere come il problema di un'assegnazione *coerente* di probabilità ad una famiglia di eventi, da parte di un *osservatore*, è correlata all'*emergenza*, nel modo in cui viene considerata nella scienza moderna.

Ricordo che l'*emergenza* (Corning 2002; Minati and Pessa, 2002; Pessa 1998; 2002), in questo ambito, può essere considerata come un processo di formazione di nuove *entità collettive* (che richiedono un livello di descrizione differente da quella usata per i singoli elementi), quali sciami, stormi, traffico, distretti industriali, mercati; *effetti collettivi*, quali superconduttività, ferromagnetismo, effetto laser. Tali enti sono determinati dal comportamento *coerente* (così percepito da un osservatore) di componenti interagenti.

In riferimento al modello cognitivo di (Anderson 1983) usato dall'osservatore, l'*emergenza* può essere vista come un processo dipendente solamente dall'osservatore, considerando cioè che le proprietà collettive *emergono* ad un livello più alto (vale a dire più astratto) rispetto a quello da lui utilizzato per trattare i componenti. Le proprietà collettive sono rilevate dall'osservatore come nuove, in dipendenza dal modello cognitivo assunto, adatto a stabilire la presenza di coerenza.

Ma di più la concezione soggettiva della probabilità di de Finetti è tutta costruita come un *sistema cognitivo logicamente aperto*. Infatti l'osservatore (il valutatore) svolge un ruolo *attivo*, sensibile al *contesto*, si muove in un *ambito non oggettivistico*, è *orientato al processo*, *impara dinamicamente* dall'esperienza in maniera *induttiva*, è *parte integrante del sistema* e *genera la sua esistenza*, infine *può scegliere le regole* di cui servirsi (Minati 2004).

Partirò da quest'ultimo punto per illustrare quanto detto. Nell'assegnare una misura dell'incertezza ad una famiglia di eventi, di cui non conosciamo il valore di verità, l'osservatore può seguire due strade, a seconda del contesto in cui si trova (o a cui sceglie di far riferimento).

La prima consiste nell'introdurre tra gli eventi un *ordinamento naturale*: se infatti supponiamo, come è lecito, che un evento confrontato con un altro possa essere o *più possibile* o *meno possibile* o *ugualmente possibile* di questo, egli può facilmente costruire una *relazione d'ordine* di *non meno possibile* nell'ambito della famiglia di eventi considerati.

Tale relazione gode delle seguenti proprietà:

Un evento incerto E_h della famiglia è sempre più possibile dell'evento impossibile e meno possibile dell'evento certo ed è non meno possibile di se stesso.

Se un evento E_h è non meno possibile di E_k allora non può essere E_k non meno possibile di E_h , a meno che E_h ed E_k non siano ugualmente possibili.

Se un evento E_h è non meno possibile di E_k e E_k di E_j allora E_h è non meno possibile di E_j .

Se un evento E è incompatibile con E_k e con E_j e E_k è non meno possibile di E_j allora l'unione di E con E_k risulta non meno possibile dell'unione di E con E_j .

In questo modo l'osservatore introduce una *misura qualitativa* dell'incertezza di un evento.

Si osservi, di sfuggita, che nel caso particolare in cui si abbia a che fare con una partizione dell'evento certo in casi giudicati ugualmente possibili, ne segue che l'ordinamento qualitativo introdotto si traduce immediatamente in una misura quantitativa della loro incertezza come rapporto tra il numero dei casi favorevoli e il numero dei casi possibili.

Se, poi, si considera un'ulteriore proprietà relativa agli eventi condizionati:

gli eventi E_h e E_k implicino E allora $E_h|E$ è non meno possibile di $E_k|E$ se E_h è non meno possibile di E_k ;

la sua introduzione insieme alle precedenti permette di sviluppare *qualitativamente* tutta la teoria della probabilità (de Finetti 1937).

La seconda strada è quella per cui l'osservatore esprime il proprio *grado di fiducia* nel verificarsi di un evento per mezzo di una valutazione di tipo quantitativo.

Per operare in tal modo occorre uno *strumento* di misura adatto allo scopo. Bruno de Finetti ha proposto due criteri di misura equivalenti: quello della *scommessa* e quello della *penalizzazione* (de Finetti 1974).

Richiamo qui solo quello relativo alla scommessa, che appare più naturale poiché rispecchia anche storicamente quanto avvenuto nella costruzione della teoria della probabilità.

Supponiamo di dover puntare una certa somma per vincerne un'altra nel caso che un determinato evento si verifichi. Questo è quanto comunemente accade in tutti i campi riguardanti delle scommesse.

Allora effettuare una *scommessa* su di un evento E significa che si è disposti a pagare una *parte* di una certa somma S , che possiamo indicare con pS , per ricevere S se E si verifica e 0 nel caso contrario.

Se introduciamo la funzione *guadagno* G_E relativamente a quella data scommessa, otterremo la seguente schematizzazione:

$$G_E = \begin{cases} S-pS, & \text{se } E \text{ si verifica} \\ -pS, & \text{se } E \text{ non si verifica.} \end{cases}$$

Si osservi che è del tutto evidente che in una scommessa ciascuno di noi ha un obiettivo quello di massimizzare il proprio guadagno, quindi questo fatto potrebbe portarci a distorcere la nostra valutazione; come difenderci da ciò per evitare che lo strumento scommessa divenga arbitrario e quindi inefficace allo scopo a cui deve rispondere?

Intanto bisogna garantirsi che chi scommette pS per ricevere S (dall'altro scommettitore) se E si verifica, deve analogamente essere disposto a pagare S per ricevere pS se E si verifica, cioè a scambiare i termini della scommessa con l'altro scommettitore, questo garantirà che la valutazione effettuata dall'individuo rispecchi il proprio grado di fiducia senza essere influenzato dal desiderio di realizzare un maggiore guadagno, ché altrimenti questo potrebbe essere realizzato dall'altro. Ad esempio, supponiamo che in una scommessa su E io valuti di poter pagare 70 per ricevere 100 nel caso in cui E si verifichi, potrei pensare di aumentare il mio ipotetico guadagno (in questo caso pari a 30), dichiarando di essere disposto a pagare 40 ; ma se sono disposto a scambiare le quote di scommessa con l'altro competitore allora il mio ipotetico guadagno potrebbe diminuire sensibilmente (-60)!

Ma ciò non basta, bisogna garantirsi che i possibili valori del guadagno G_E non siano entrambi dello stesso segno, perché solo in tal caso non si avrebbe una vincita o una perdita certa, indipendentemente dal verificarsi o meno di E . Con la conseguenza che solo in tal caso un individuo accetterebbe di scommettere. Tale condizione fu opportunamente chiamata da de Finetti *coerenza*.

La coerenza in una scommessa su di un evento E porta come è ben noto a stabilire che, posto S pari a 1 , (ma vale anche per $S \neq 1$) in ogni caso la quota p che un individuo è disposto a pagare per

ricevere I se E si verifica è sempre compresa tra 0 e 1 . Inoltre la stessa coerenza impone che se E è *certo* allora p deve essere necessariamente 1 e se E è *impossibile* allora p deve essere necessariamente 0 .

Si osservi subito però che $p=1$ non implica E *certo*, né $p=0$ implica E *impossibile* (per un approfondimento in tal senso si veda de Finetti 1974).

Seguendo de Finetti quindi la probabilità di un evento E (ovvero la *misura numerica* dell'incertezza su E) è la quota p che un individuo è disposto a pagare in una *scommessa coerente* per ricevere I se E si verifica e 0 altrimenti.

Ma appunto una scommessa sarà coerente solo se l'*osservatore* (l'individuo) che valuta l'incertezza di E la renderà tale.

Esisteranno quindi, relativamente ad E infinite valutazioni coerenti, purché comprese tra 0 e 1 !

Come farà allora l'osservatore a sceglierne una? Dovrà basarsi sull'*informazione* che possiede in merito ad E ed esprimere tale informazione attraverso un numero. Naturalmente quanto più l'informazione sarà *ricca* tanto più si sarà meno dubbiosi nello scegliere uno tra infiniti numeri!

Ora nel misurare l'informazione un individuo, come sempre accade, non possiede un criterio *oggettivo* di valutazione: in essa infatti interverranno le convinzioni personali, gli stati d'animo e tutte quelle caratteristiche che contribuiscono alla formazione di un giudizio. Quindi egli dovrà soltanto formulare, potremmo dire così, oggettivamente quanto valuta soggettivamente!

Così si esprime de Finetti ne *L'invenzione della verità* (2006):

“Qualunque comportamento...nel senso di ritenere plausibile il verificarsi di questo o di quell'evento non dipende quindi che da un sentimento, da quello stesso sentimento che bisogna lealmente presentare come il vero punto di partenza, e che taluno preferisce invece ignorare e pudicamente nascondere dietro una barricata di congegni logici altrettanto macchinosi quanto vuoti”.

Ovviamente in alcuni casi sarà tutto più semplice: ad esempio se un individuo giudicasse ugualmente possibile il verificarsi di E e quello della sua negazione E^c , ad entrambi assegnerà il valore di probabilità $\frac{1}{2}$ (ma siamo sicuri che il giudicare ugualmente possibili due eventi *reali* sia così semplice e naturale?). Mentre se giudicasse cinque volte più probabile E rispetto a E^c allora assegnerà ad E probabilità $\frac{5}{6}$ e ad E^c $\frac{1}{6}$.

In altri casi potrà fare ricorso a valutazioni che si basano sulla frequenza osservata, ma solo, come ha ben precisato de Finetti, quando l'evento E faccia parte di una famiglia di eventi *scambiabili*, cioè quando la valutazione della probabilità di una qualunque n -pla di eventi della famiglia considerata dipende solo dal numero degli eventi fissati e non dai particolari eventi fissati; in breve: dipende solo da *quanti* eventi e non da *quali* si considerano (de Finetti 1974)!

Così ad esempio, nell'estrazione con restituzione di palline da un'urna di composizione incognita, cioè per la quale si conosca il numero totale di palline ma non la percentuale di quelle *rosse*, un individuo che volesse valutare la probabilità di ottenere pallina *rossa* alla n -esima estrazione, essendone state estratte $(n-1)/3$ rosse su $(n-1)$, potrebbe stimare e pari a $(n-1)/3$ la probabilità che esca pallina *rossa* alla n -esima estrazione, perché gli eventi della famiglia considerati sono scambiabili (ci interessa solo *quante* palline rosse vengono estratte). Ma se volesse valutare la probabilità di vittoria di un determinato pugile al 101 -esimo incontro della sua carriera, sapendo che ne ha vinti 85 sui 100 precedenti, basterebbe ciò per stimare tale probabilità pari a $85/100$? Ovviamente no, perché potrebbe, nel caso peggiore, aver perso tutti gli ultimi 15 incontri, e il grado di fiducia nei suoi confronti... potrebbe non essere così elevato!

Ho svolto finora considerazioni relative a singoli eventi o a famiglie di eventi analoghi. Spesso però ci imbattiamo in fenomeni aleatori, descrivibili in alcuni casi attraverso numeri aleatori. Come

ad esempio il numero X di incidenti d'auto mortali in un anno relativamente a persone che non indossavano la cintura di sicurezza, per i quali si possono considerare eventi del tipo ($X = n$). Oppure in fenomeni più complessi, rispetto ai quali possiamo essere interessati ad eventi del tipo: la quantità di pioggia caduta in una determinata città italiana nel prossimo anno sarà maggiore di quella dell'anno trascorso, oppure il grado di umidità medio sarà minore, o il livello di smog misurato sarà maggiore, o il consumo di carburante sarà uguale.

Nel primo caso, quello dei numeri aleatori, è possibile formulare diversi modelli probabilistici che ci permettono di valutare la probabilità di qualsivoglia evento relativo ad essi, ma non dobbiamo dimenticare che queste valutazioni non sono oggettive, come apparentemente possono sembrare, perché dato il modello basta applicare delle formule per ricavarle; viceversa è sempre l'osservatore che sceglie, in base alle proprie informazioni, il modello che ritiene più adatto a descrivere il fenomeno considerato. Cosa accade, invece, nel secondo caso?

Nuovamente gli aspetti sistemici dell'impostazione adottata da de Finetti forniscono una risposta esemplare.

Ricordo che nell'approccio *classico* (comunemente utilizzato nelle applicazioni) relativo ad un fenomeno aleatorio si procede così: prima si definisce uno spazio Ω dei risultati o casi elementari possibili, ovvero si costruisce una partizione dell'evento certo, successivamente si attribuisce una probabilità a ciascuno di questi casi (o *costituenti*) e poiché qualunque evento riferito a quel fenomeno si può ottenere come unione di costituenti, a ciascuno di essi viene attribuita in maniera *lineare* una probabilità.

Faccio osservare che naturalmente rimane aperto il problema di come assegnare le probabilità ai costituenti!

Al contrario de Finetti basa la sua concezione sul fatto che ogni evento è singolo e per ciascuno di essi possiamo esprimere attraverso una valutazione qualitativa o quantitativa il nostro grado

di fiducia. Se poi oltre a trovarci davanti ad un unico evento ed averne valutato la sua probabilità, abbiamo la necessità di assegnare delle probabilità ad ulteriori eventi come dobbiamo comportarci?

Si possono presentare diversi casi. Se si ha a che fare con una famiglia di eventi E_i che forma una *partizione* dell'evento certo, allora per la coerenza la somma delle probabilità dei singoli E_i deve essere uguale ad 1 e la probabilità dell'unione di n eventi *incompatibili* deve essere uguale alla somma delle singole probabilità.

Se, ancora sono dati n eventi ed una loro assegnazione coerente di probabilità, allora risulta determinata la probabilità di un evento che dipenda *linearmente* dai primi n .

Nei casi, invece, in cui tra gli eventi E_i considerati ed un nuovo evento E vi sia un legame di tipo *logico*, allora sempre per la coerenza si deve procedere in questo modo: si costruiscono i costituenti relativamente alla famiglia data (cioè tutte le intersezioni possibili tra gli eventi in modo che in ciascuna compaia uno di essi o la sua negazione, ad es. $E_1 \wedge E_2 \wedge \dots \wedge E_h \wedge E_{h+1} \wedge \dots \wedge E_{n-1} \wedge E_n$), poi si individuano due eventi E' ed E'' , che sono rispettivamente il massimo evento unione di tutti i costituenti implicanti E ed il minimo evento unione di tutti i costituenti implicati da E : necessariamente la probabilità di E dovrà essere compresa nell'intervallo chiuso $[P(E'), P(E'')]$.

Anche in questa situazione quindi la probabilità di E non risulta univocamente determinata, può semplicemente restringersi l'intervallo $[0,1]$ all'interno del quale l'osservatore può valutarla per essere coerente.

Ma la situazione più interessante si presenta quando si ha a che fare con singoli eventi che vengono esaminati a partire dal primo e poi gli altri successivamente. A ciascuno di questi l'osservatore in base alle proprie informazioni attribuisce un valore di probabilità coerente, cioè compreso tra 0 e 1 . In tal caso però tra gli eventi

possono sussistere dei legami logici di cui non si è tenuto conto o di cui se ne è avuta informazione solo quando sono stati introdotti tutti gli eventi, come fare allora per verificare se la valutazione è complessivamente coerente?

Occorrerà costruire i costituenti a partire dagli eventi considerati, e poiché ciascuno di questi ultimi risulterà unione di alcuni dei costituenti ottenuti, la sua probabilità deve essere uguale alla somma delle probabilità (non ancora determinate, quindi diciamo così *incognite*) di tali costituenti. Si otterrà così un sistema di n equazioni in s ($\leq 2^n$) incognite x_i , con i vincoli $x_1 + x_2 + \dots + x_s = 1$ e $x_i \geq 0$. Se esiste una s -pla soluzione del sistema allora la valutazione data potrà dirsi coerente.

E' interessante osservare che il sistema potrebbe non ammettere soluzioni, in tal caso la valutazione sarebbe incoerente; che può ammettere un'unica soluzione, ma che possono esistere anche più soluzioni, cioè un insieme di valutazioni diverse tutte coerenti.

Chiarirò quest'ultimo interessante aspetto attraverso degli esempi.

Siano dati tre eventi A, B, C ed un osservatore abbia valutato le loro probabilità al modo seguente $P(A) = 1/2$, $P(B) = 2/5$, $P(C) = 1/5$

(ovviamente ciascuna di esse rappresenta una valutazione coerente!).

Sia noto poi che $ABC = \phi$ (dove $ABC \equiv A \wedge B \wedge C$).

Allora i costituenti possibili sono

$$C_1 = A^c B C, C_2 = A B^c C, C_3 = A B C^c, C_4 = A B^c C^c, C_5 = A^c B C^c, C_6 = A^c B^c C, C_7 = A^c B^c C^c.$$

Per stabilire allora se, nelle condizioni date, la valutazione complessiva espressa dalla terna $P(A) = 1/2$, $P(B) = 2/5$, $P(C) = 1/5$ risulti coerente bisogna stabilire se il seguente sistema ammette almeno una soluzione:

$$\left\{ \begin{array}{l} x_2 + x_3 + x_4 = 1/2 \\ x_1 + x_3 + x_5 = 2/5 \\ x_1 + x_2 + x_6 = 1/5 \\ x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 = 1 \\ x_i \geq 0, i = 1, \dots, 7 \end{array} \right.$$

Se si pone, come è lecito fare, $x_1 = 0, x_3 = 0, x_6 = 0$, con semplici passaggi si ottiene la seguente soluzione del sistema

$$x_1 = 0, x_2 = 1/5, x_3 = 0, x_4 = 3/10, x_5 = 2/5, x_6 = 0, x_7 = 1/10,$$

quindi le attribuzioni di probabilità assegnate agli eventi A, B, C determinano una valutazione complessiva coerente!

Si osservi, che se avessimo posto $x_2 = 0, x_3 = 0, x_6 = 0$, avremmo ottenuto una diversa soluzione del sistema

$$x_1 = 1/5, x_2 = 0, x_3 = 0, x_4 = 1/2, x_5 = 1/5, x_6 = 0, x_7 = 1/10$$

e anche in tal caso la valutazione complessiva sarebbe risultata coerente!

Non solo, ma se inizialmente l'osservatore avesse valutato $P(A)=\square\square\square P(B)=\square\square\square P(C)=\square\square\square$ con la condizione $ABC = \phi$, avremmo ottenuto il seguente sistema

$$\left\{ \begin{array}{l} x_2 + x_3 + x_4 = \alpha \\ x_1 + x_3 + x_5 = \beta \\ x_1 + x_2 + x_6 = \gamma \\ x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 = 1 \\ x_i \geq 0, i = 1, \dots, 7 \end{array} \right.$$

$$0 \leq \alpha \leq 1, 0 \leq \beta \leq 1, 0 \leq \gamma \leq 1 \square \square \square$$

Con le scelte degli $x_i = 0$, fatte nei due casi esaminati in precedenza, avremmo avuto rispettivamente le seguenti soluzioni:

$x_1 = 0, x_2 = \gamma, x_3 = 0, x_4 = \alpha - \gamma, x_5 = \beta, x_6 = 0, x_7 = 1 - (\alpha + \beta)$,
che darebbe luogo ad una valutazione complessiva coerente, purché sia $\alpha \geq \gamma$ \square e $\alpha + \beta \leq 1$ \square

$x_1 = \gamma, x_2 = 0, x_3 = 0, x_4 = \alpha, x_5 = \beta - \gamma, x_6 = 0, x_7 = 1 - (\alpha + \beta)$,
che darebbe luogo ad una valutazione complessiva coerente, purché sia $\beta \geq \gamma$ \square e $\alpha + \beta \leq 1$ \square

Passerò ora ad esaminare gli ulteriori aspetti dell'assunto iniziale, in particolare per quanto riguarda *l'imparare induttivamente dall'esperienza*, riferendomi agli eventi condizionati.

Ricordo che, dati due eventi E ed H qualsiasi (purché $H \neq \Phi$), si può considerare un nuovo ente E/H e , a cui viene attribuito il nome di *evento condizionato*, che ha il seguente significato

E/H =	VERO,	se H vero ed E vero
	FALSO,	se H vero ed E falso
	INDETERMINATO,	se H falso

Allora se vogliamo effettuare una scommessa (condizionata) relativamente ad E/H dovremo, secondo de Finetti², comportarci al seguente modo:

PAGARE p	per RICEVERE	1 se H vero ed E vero
		0 se H vero ed E falso
		p se H falso

² de Finetti Bruno, *Teoria delle probabilità*, Einaudi, 1970

Da questa definizione di scommessa condizionata segue che possiamo misurare l'incertezza dell'evento E/H attraverso la quota p che un individuo *coerente* è disposto a pagare con gli esiti precedenti e questa misura viene chiamata *probabilità condizionata* e indicata con $P(E/H)$.

de Finetti prova, inoltre, come la probabilità condizionata, così introdotta, verifichi tutte le proprietà di una probabilità e, infine, come l'unica *lettura* possibile di $P(E/H)$ sia quella della probabilità di E *supposto vero* H .

In particolare, egli fa vedere come la semplice (e naturale) condizione di coerenza: il guadagno aleatorio non uniformemente negativo in un insieme di scommesse, conduca al *teorema delle probabilità composte*:

considerati due eventi qualsiasi E ed H ($H \neq \Phi$) si ha

$$P(E \cap H) = P(H) P(E/H);$$

ed al suo corollario, il *teorema di Bayes* (con anche $E \neq \Phi$):

$$P(H/E) = K P(H) P(E/H), \quad \text{con } K = 1/P(E), P(E) \neq 0,$$

il quale ci dice che la probabilità di H condizionatamente ad E è uguale al prodotto della probabilità di H per la probabilità di E condizionatamente ad H , a meno di un fattore di proporzionalità K non nullo.

Soffermiamoci, ora, sul *significato* del teorema di Bayes. In realtà, esso ci dice molto di più se ne diamo una *lettura* di questo tipo: sia E un evento che rappresenti l'esito di un esperimento di un dato fenomeno aleatorio e sia H una ipotesi relativa allo stesso fenomeno, allora il teorema afferma che la probabilità dell'ipotesi H condizionatamente ad E è proporzionale alla probabilità dell'ipotesi H non condizionata moltiplicata per la probabilità di E condizionata ad H .

Per rendere più chiaro quanto espresso ricorriamo al classico esempio dell'urna di *composizione incognita*. Consideriamo cioè un'urna di cui si sappia che contiene N palline, ma delle quali non è

nota la percentuale di palline rosse (potendo, quindi, essere presenti nell'urna da 0 a N palline rosse). L'evento E rappresenti un possibile risultato relativo all'estrazione di n palline dall'urna (ad es., senza reimpulso): *h palline rosse su n*, e l'evento H sia l'ipotesi: *sono presenti nell'urna r palline rosse su N*. Allora il teorema di Bayes ci permette di valutare la probabilità dell'ipotesi H condizionatamente all'evento E (che possiamo chiamare *probabilità finale*), attraverso la probabilità di H (che possiamo chiamare *probabilità iniziale*) e la probabilità di E/H, cioè di E supposta vera H (che può essere facilmente calcolata e che possiamo chiamare *verosimiglianza*), sempre a meno del fattore di proporzionalità K.

In definitiva, il teorema di Bayes ci dice come dobbiamo *aggiornare* una nostra valutazione di probabilità in presenza di una nuova *informazione* (o meglio supposto che ci divenga nota una nuova informazione):

probabilità finale = K x *probabilità iniziale* x *verosimiglianza*.

Si osservi che *iniziale* e *finale* hanno, in questo contesto, il solo significato rispettivamente di prima e dopo che divenga noto (si supponga di conoscere) E.

Naturalmente, allo stesso modo può valutarsi la probabilità finale dell'ipotesi contraria H^c , risultando

$P(H^c/E) = K P(H^c) P(E/H^c)$, con $K = 1/P(E)$, $P(E) \neq 0$.

E, più in generale, volendo valutare la probabilità finale di m diverse ipotesi H_j , formanti una partizione dell'evento certo, si ha la seguente espressione del teorema di Bayes:

$P(H_j/E) = K P(H_j) P(E/H_j)$, $j=1,2,\dots,m$; con $K = 1/P(E)$, $P(E) = \sum_{j=1,2,\dots,m} P(H_j) P(E/H_j)$.

Tornando al nostro esempio, le H_j rappresenterebbero le possibili ipotesi sulla composizione dell'urna e noi dopo averne dato una valutazione iniziale, potremo attraverso il teorema di Bayes, esprimerne una finale.

In generale nelle impostazioni oggettivistiche della probabilità (quella classica e quella frequentista), poiché le valutazioni di probabilità sono essenzialmente dei rapporti numerici, l'osservatore (l'individuo che valuta) ha il solo compito di eseguire bene dei calcoli, basati su situazioni di simmetrie o su ripetizioni di un dato fenomeno (apparentemente, però, perché comunque in ogni caso sono state fatte delle precise e soggettive scelte in precedenza: i casi possibili considerati equiprobabili e le prove ripetute considerate equiprobabili ed indipendenti!). In queste concezioni il teorema di Bayes perde il suo significato ed il suo valore intrinseco e assume la veste di un risultato puramente matematico.

Nella concezione soggettivistica di Bruno de Finetti, al contrario, esprime al meglio il significato di come si può (e si deve) apprendere dall'esperienza!

E' lo stesso processo logico che porta un medico, per fare un altro esempio, ad effettuare una diagnosi. Egli nel valutare la presenza o meno di una malattia in un paziente, prima lo visita ed esprime una prima idea riguardo alla malattia, poi gli fa compiere una serie di accertamenti diagnostici-strumentali, i cui risultati utilizza per arrivare a stabilire la presenza o meno della malattia.

Si osservi, per inciso, che questa procedura sembra far parte della logica del certo, mentre viceversa più in generale riguarda l'ambito della logica dell'incerto: sempre e solo di valutazioni di probabilità si tratta, anche se fortunatamente spesso le probabilità finali di accertamento della malattia sono prossime a 1 o a 0 (si pensi solo al fatto che oltre alla valutazione personale anche i test clinici possono essere affetti da errore)! In ogni caso quando ricorriamo ad un medico, l'evento: *sono affetto da una data malattia*, è sempre *possibile*, né certo, né impossibile!

Spesso nelle applicazioni per stimare il valore di qualche parametro relativo ad un dato fenomeno aleatorio, si ricorre al cosiddetto *metodo della massima verosimiglianza*: in breve e schematizzando si tratta di valutare i diversi valori delle probabilità di E (o la

densità di probabilità, nel caso di un parametro continuo) condizionatamente alle ipotesi H_j (le *verosimiglianze*) e di ricavarne il valore massimo, che viene scelto come *stima*.

Ritornando, ancora una volta, al nostro esempio dell'urna di composizione incognita si calcolano le

$P(E/H_j)$, per ogni j , e poi se ne trova quella più grande di tutte; se questa, ad esempio è la $P(E/H_3)$ si dice che l'ipotesi H_3 è una stima della *vera* composizione dell'urna.

Esemplificando ulteriormente, in un'urna ci siano 10 palline, di cui non è noto quante sono quelle rosse. Si effettua un campionamento, ossia un'estrazione ripetuta (con restituzione) per un totale di 5 palline e di queste se ne trovano 3 rosse e sia E questo evento. Si valutano le $P(E/H_j)$, $j=0,1,\dots,10$, ricavando

$$P(E/H_j) = \frac{5!}{3!2!} \left(\frac{j}{10}\right)^3 \left(\frac{10-j}{10}\right)^2;$$

e distintamente

$$P(E/H_0) = 0$$

$$P(E/H_1) = 0.0081$$

$$P(E/H_2) = 0.0512$$

$$P(E/H_3) = 0.1323$$

$$P(E/H_4) = 0.2304$$

$$P(E/H_5) = 0.3125$$

$$P(E/H_6) = 0.3456$$

$$P(E/H_7) = 0.3087$$

$$P(E/H_8) = 0.2048$$

$$P(E/H_9) = 0.0729$$

$$P(E/H_{10}) = 0$$

Quindi poiché il valore massimo è 0.3456 , che si ha relativamente all'ipotesi H_6 , se ne deduce che la *stima* (di massima verosimiglianza) per la composizione incognita dell'urna è *6 palline rosse su 10*.

Questa metodologia ha due controindicazioni, entrambe di tipo logico!

La prima è quella di effettuare una stima relativamente ad ipotesi, utilizzando le probabilità condizionate inverse $P(E/H_j)$ rispetto a quelle che più correttamente andrebbero confrontate: le $P(H_j/E)$; solo attraverso queste, in maniera logicamente inequivocabile, è possibile stabilire quale delle ipotesi di composizione dell'urna è quella più probabile, osservato (supposto di osservare) E , e prenderla come stima della composizione!

La seconda è che in tale procedura non intervengono le probabilità delle ipotesi $P(H_j)$, che possono essere sempre valutate e la cui non considerazione può essere assolutamente dannosa per le conclusioni raggiunte.

Infatti, anche nel precedente esempio, se si utilizzasse (come è necessario) il teorema di Bayes, nella valutazione delle $P(H_j/E)$ entrerebbero le $P(H_j)$, e una qualunque distribuzione di queste diversa dall'equiprobabilità, potrebbe rendere massima la probabilità finale di una delle H_j diversa da H_6 . Se, per un qualche motivo, si sapesse che è molto più probabile che siano state messe nell'urna tante palline rosse quante di un altro colore rispetto alle altre ipotesi, e cioè nel nostro caso si avesse, ad esempio $P(H_5)=0.6$, $P(H_4)=P(H_6)=0.15$, e per semplicità $P(H_0)=P(H_1)=P(H_2)=P(H_3)=P(H_7)=P(H_8)=P(H_9)=P(H_{10})=0.0125$, si ricaverebbe

$$P(H_0/E) = 0$$

$$P(H_1/E) = 0.00035$$

$$P(H_2/E) = 0.00225$$

$$P(H_3/E) = 0.00581$$

$$P(H_4/E) = 0.12185$$

$$P(H_5/E) = 0.66111$$

$$P(H_6/E) = 0.18278$$

$$P(H_7/E) = 0.01357$$

$$P(H_8/E) = 0.00902$$

$$P(H_9/E) = 0.00320$$

$$P(H_{10}/E) = 0$$

E quindi l'ipotesi largamente più probabile risulterebbe H_5 .

Desidero presentare un esempio ancor più significativo, per i suoi aspetti paradossali.

Paolo non si reca ad un colloquio di lavoro presso la ditta Smile, indichiamo con E questo evento. Il dirigente del reparto assunzioni vuole capire perché e formula le seguenti ipotesi

H_1 = Paolo ha trovato un altro lavoro

H_2 = Paolo è finito in prigione

H_3 = Paolo ha vinto la lotteria o qualunque altra ragione diversa dalle altre.

Se egli dovesse decidere sulla base delle verosimiglianze, in ogni caso dovrebbe scegliere l'ipotesi H_2 , in quanto H_2 implica E e quindi $P(E/H_2)=1$.

Mentre se si utilizza il teorema di Bayes, l'ipotesi più probabile potrebbe non essere H_2 . Infatti basta che siano $P(E/H_1)=0.6$, $P(E/H_3)=0.2$ e $P(H_1)=0.7$, $P(H_2)=0.25$, $P(H_3)=0.05$ per ottenere

$$P(H_1/E)=0.618$$

$$P(H_2/E)=0.368$$

$$P(H_3/E)=0.014,$$

da cui si ricaverebbe che l'ipotesi più probabile è la *più ragionevole!*

Il teorema di Bayes, in definitiva, è l'architrave della *coerenza* su cui poggia l'*aggiornamento* delle valutazioni di probabilità. Attraverso di esso si può solo ridurre l'*incertezza*, mai eliminarla giungendo a conclusioni certe. Da questo punto di vista rappresenta un efficace forma di *pensiero non lineare*.

L'*aggiornamento* delle valutazioni di probabilità deve quindi seguire un solo principio: quello della *coerenza*; questo garantisce all'*osservatore* di *non violare le regole* nell'attribuzione di probabilità: tra ciò che aveva valutato prima (le *probabilità iniziali*) e ciò che valuta dopo (le *probabilità finali*). Di ciò si tratta quando ci si riferisce *all'imparare induttivamente dall'esperienza!*

In ogni caso, da tutti gli esempi illustrati, si ricava come un insieme di eventi si può trasformare in un *sistema di eventi* quando un osservatore faccia *emergere* per essi una *valutazione coerente* di probabilità. Come abbiamo visto, in generale, di queste valutazioni coerenti ce ne possono essere più di una, rimane quindi alla *responsabilità* dell'osservatore scegliere quella che ritiene rappresentare al meglio il suo *stato di informazione*, rispetto all'insieme di eventi considerati!

Quindi l'apertura sistemica della concezione di de Finetti esplica in tal modo tutta la sua ricchezza metodologica ed esalta il ruolo dell'osservatore, come portatore ed elaboratore di scelte consapevoli e appunto responsabili.

Da un punto di vista sistemico si può dire che in questo tipo di logica induttiva le interazioni tra agenti (eventi in questo caso) devono essere *coerenti*, invece che lineari o deducibili le une dalle altre. In tale circostanza la *coerenza* non è qualcosa di relativo alle regole della logica formale come nella deduzione, per il calcolo deterministico, ma è relativa all'*emergenza*.

La coerenza, pertanto, non è qualcosa di deterministicamente calcolato e ricavato: è progettato, appreso, sperimentato e *quindi* formalizzato nella più generale costruzione di modelli e nella loro

simulazione, come viene anche fatto nell'ambito della *logica fuzzy* (Zadeh *et al.*, 1996).

Da questo punto di vista gli elementi di un sistema sono eventi.

L'osservatore modella il sistema emergente, che è una configurazione di eventi considerata interagente con le probabilità assegnate dall'osservatore, oltre che dalle interazioni fisiche.

Tale approccio sembra essere necessario per il cruciale ruolo teorico dell'osservatore nei processi di emergenza e nella loro modellizzazione.

Avendo a che fare con sistemi (quali quelli fisici, biologici e sociali) considerati emergenti a causa delle interazioni dei componenti e assumendo in modo oggettivistico che l'osservatore non faccia parte del sistema, oppure ne faccia parte ma adotti una logica incompatibile con il sistema considerato (assuma cioè una logica lineare, pensando di agire in uno spazio deterministico) si usano delle strategie, che sebbene non siano errate, sono quanto meno inefficaci.

L'approccio introdotto da de Finetti porta a considerare sistemi di probabilità, partendo da singoli eventi ed arrivando a sistemi di eventi.

Conclusione

In conclusione, è importante sottolineare come il linguaggio e la logica dell'incerto, sviluppati da Bruno de Finetti, abbiano un ruolo cruciale nella vita di tutti i giorni. Dispositivi di pensiero che ci permettono di misurarci con la *complessità* e produrre discipline che interagiscano sistemicamente, in modo che vengano usati coerentemente concetti, analogie, corrispondenze e invarianti.

Strumenti che fino ad oggi ci hanno consentito di arricchire enormemente le nostre conoscenze e che se usati, come Bruno de Finetti ha fatto non solo nella costruzione della sua teoria delle

probabilità ma anche riguardo alla sua concezione *fusionista* della scienza e non solo, ancor di più ci consentiranno un incremento di tale arricchimento che si fondi sul superamento di quegli steccati disciplinari, cui egli ha sempre cercato di abituarci a diffidare.

Bibliografia

- [1] Anderson, J. R., 1983, *The Architecture of Cognition*. Harvard University Press, Cambridge, MA.
- [2] Barrow-Green, J., 1996, *Poincaré and the Three Body Problem*. Amer. Math. Soc.
- [3] Bruno, G., 2002, *Logica dell'incerto e didattica della matematica*, Atti Convegno Nazionale UMI, Loano.
- [4] Bruno, G., 2004, *Logica dell'incerto: un'esperienza didattica*, Atti Convegno, Insegnare la matematica nella scuola di tutti e di ciascuno, Bari.
- [5] Bruno, G., 2006, *Logica dell'incerto e didattica*, Periodico di matematiche, serie VIII, vol. 6, n. 3.
- [6] Cerasoli M., Eugeni F., Rizzi B., 1983, Sulla probabilità del k-MCD di m naturali scelti a caso, Rendiconti di Matematica, Roma, pp. 367-379 (Generalizza un teorema di de Finetti caso $k=1$).
- [7] Corning, P., 2002, *The Re-emergence of "Emergence": A Venerable Concept in Search of a Theory*, Complexity **7(6)**:18-30.
- [8] Davydov, A.S., 1979, *Bioenergetics and the mechanism of muscle contraction*. International Journal of Quantum Chemistry , **16**: 5-17.
- [9] de Finetti, B., 1937, *La prevision: ses lois logiques, ses sours subjectives*, Annales de l'Institut Poincaré, vol.7, fasc.1, pp.1-68.

- [10] de Finetti, B., 1974, *Theory of Probability: A Critical Introductory Treatment*, 2 volumes (Translated by A. Machi and A. Smith). Wiley, London.
- [11] de Finetti, B., 2006, *L'invenzione della verità*, Cortina Editore.
- [12] Di Gennaro F., Eugeni F., Maturo A., 1999, Sulla probabilità di indovinare la chiave ed alterare i messaggi in sistemi di autenticazione, atti Congresso nazionale Mathesis Vol. 2, Teramo, pp. 45-54.
- [13] Heisenberg, W., 1971, *Physics and Beyond*. Harper & Row, New York.
- [14] Klir, G. J., and Bo, Y., 1995, *Fuzzy sets and Fuzzy Logic: Theory and applications*. Prentice Hall, Englewood Cliffs, NJ.
- [15] Minati, G., and Pessa, E., (eds.), 2002, *Emergence in Complex Cognitive, Social and Biological Systems.*, Proceedings of the Second Conference of the Italian Systems Society, Kluwer Academic/Plenum Publishers, London.
- [16] Minati, G., and Pessa, E., (eds.), 2002, *Emergence in Complex Cognitive, Social and Biological Systems*. Kluwer Academic/Plenum Publishers, New York.
- [17] Pelusi D. (2002), Un modo di ragionare fuzzy a livello elementare, Periodico di Matematiche, Serie VIII, Vol. 2, Nr. 4, pp. 23-30.
- [18] Peskin, E. M., and Schroeder, D. V., 1995, *An Introduction to Quantum Field Theory*, Westview Press.
- [19] Pessa, E., 1998, *Emergence, Self-Organization, and Quantum Theory*. In Proceedings of the First Italian Conference on Systemics (G. Minati, ed.), Apogeo scientifica, Milano, Italy.
- [20] Pessa, E., 2002, *What is emergence?* In *Emergence in Complex Cognitive, Social and Biological Systems* (G. Minati and E. Pessa, eds.), Kluwer Academic/Plenum Publishers, New York.
- [21] Rosen, R. 1985 *Anticipatory Systems*, Pergamon, New York.
- [22] Scozzafava R., 2002, *Incertezza e Probabilità*, Zanichelli.

[23] von Bertalanffy, L., 1968, *General Systems Theory*, George Braziller, New York.

[24] von Foerster, H., 2003, *Understanding Understanding: Essays on Cybernetics and Cognition*. Springer-Verlag, New York.

[25] Zadeh, L. A., Klir, G. J., (eds.), Yuan, B., (ed.), 1996, *Fuzzy Sets, Fuzzy Logic, and Fuzzy Systems: Selected Papers by Lotfi A. Zadeh* (Advances in Fuzzy Systems - Applications and Theory Vol. 6). World Scientific Pub Co.

Approvato su parere favorevole di Antonio Maturò e Luca Nicotra