CIME Summer Schools

18

Bruno de Finetti (Ed.)

Induzione e statistica

Varenna, Italy 1959

Springer

FONDAZIONE CIME ROBERTO CONTI

Bruno de Finetti (Ed.)

Induzione e statistica

Lectures given at the
Centro Internazionale Matematico Estivo (C.I.M.E.),
held in Varenna (Como), Italy,
June 1-10, 1959

Springer

FONDAZIONE CIME ROBERTO CONTI

C.I.M.E. Foundation
c/o Dipartimento di Matematica "U. Dini"
Viale Morgagni n. 67/a
50134 Firenze
Italy
cime@math.unifi.it

ISBN 978-3-642-10931-7 e-ISBN: 978-3-642-10934-8
DOI:10.1007/978-3-642-10934-8
Springer Heidelberg Dordrecht London New York

Reprint of the 1st ed. C.I.M.E., Florence 1959
With kind permission of C.I.M.E.

Printed on acid-free paper

Springer.com

CENTRO INTERNATIONALE MATEMATICO ESTIVO
(C.I.M.E)

Reprint of the 1st ed.- Varenna, Italy, June 1-10, 1959

INDUZIONE E STATISTICA

CENTRO INTERNAZIONALE MATEMATICO ESTIVO
(C.I.M.E.)

BRUNO DE FINETTI

LA PROBABILITA' E LA STATISTICA NEI RAPPORTI CON L'INDUZIONE, SECONDO I DIVERSI PUNTI DI VISTA

ROMA - Istituto Matematico dell'Università - 1959

SOMMARIO

L'induzione secondo Bayes, la speranza morale di Daniele Bernoulli; motivi delle critiche cui diedero luogo. Teorie non-bayesiane: costruzioni di Fisher, di Neyman-Pearson, di Wald. Concezione neo-bayesiana e neo bernoulliana su basi soggettive: fondamenti della probabilità, dell'induzione, della statistica, della teoria delle decisioni. Schemi particolari di ragionamento induttivo: eventi scambiabili e parzialmente scambiabili, in particolare casi di tipo markoviano.

INDICE

LA PROBABILITA' E LA STATISTICA NEI RAPPORTI CON L'INDUZIONE, SECONDO I DIVERSI PUNTI DI VISTA

di

BRUNO DE FINETTI

§1. INTRODUZIONE

Ragionare per induzione altro non vuol dire che ***imparare dall'esperienza***: un fatto ovvio comune a tutti noi, ai bambini, agli animali (1), che è ben noto a tutti nei suoi aspetti sommari dall'osservazione quotidiana, e che in modo più approfondito viene studiato dagli psicologi sperimentali [8].

Parlare di "ragionamento" induttivo significa però, evidentemente, attribuire una certa validità a tale modo di apprendere, considerandolo non come il frutto di una capricciosa reazione psicologica, ma come un processo mentale suscettibile di venire analizzato interpretato giustificato.

Disgraziatamente una tale discussione porta in genere nel

---

(1) Ciò che fa dire con ragione a Good [29] che la teoria delle probabilità è molto più antica del genere umano.

ginepraio delle concezioni filosofiche, dove raramente l'essenza concreta dei problemi riesce a salvarsi dalla nebbia delle invenzioni verbali.

Nel nostro caso riesce particolarmente pregiudizievole la tendenza a sopravvalutare - spesso addirittura in modo esclusivo - la *ragione*, che, a mio avviso, è invece utilissima solo a patto di venir considerata come un complemento atto a perfezionare tutte le altre facoltà istintive intuitive psicologiche (ma non - guai! - a surrogarle).

Conseguenza di tale stortura è infatti l'erezione del ragionamento deduttivo a modello (benchè tutte le verità non vuotamente tautologiche siano basate su altro!). Così il ragionamento induttivo viene generalmente considerato come qualcosa di appartenente a un livello più basso, da accogliere con riserva e diffidenza,o, peggio, quando si tenta di dargli dignità, si cerca di snaturarlo facendolo apparire come qualcosa che possa quasi farsi rientrare nel ragionamento deduttivo [2]

Come - nelle grandi linee - è stato chiaramente riconosciuto fin dall'analisi dell'idea di causa per opera di Hume [34], il ragionamento induttivo deriva anzitutto dall'associazione delle idee legata all'impressione di "analogia" fra certi fatti,e quindi a quella derivante dalla impressione di un'associazione tra fatti diversi ("causa ed effetto", nella terminologia più primitiva).

---

(2) Postulando qualcosa come la "necessaria esistenza di leggi naturali" con particolarità più o meno prestabilite, si può naturalmente camuffare da "deduttivo" un ragionamento induttivo. Tutte le concezioni che considerano il determinismo come postulato necessario per la possibilità stessa della scienza si basano sostanzialmente su tale presupposto e negano quindi il ragionamento induttivo nella sua natura genuina.

E' a questo punto che, al fine di render preciso ciò che può esser reso preciso in siffatte considerazioni apparentemente del tutto vaghe, si deve far intervenire la matematica.

Qual'è l'apporto della matematica? Può essa portare chiarimenti alle questioni concettuali? Strumenti per la loro formulazione logica? Metodi generali o particolari per le applicazioni o per certi tipi di applicazioni?

Cominciamo col rilevare che anche nella matematica, che pure è il regno delle verità tautologiche, non è affatto escluso ed è anzi necessario il ragionamento induttivo. E' necessario nel momento creativo, perchè nessuno si accingerebbe a cercar di dimostrare un teorema se non vi attribuisse una certa verosimiglianza. Come dice P.Levy ([46], prefaz.),chi vuole arrivare a un certo punto deve pur vederlo con gli occhi (l'intuizione) prima di raggiungerlo usando altrettanto necessariamente i piedi (la logica). Esemplificazioni profonde al riguardo sono date in particolare da H.Poincaré [65],[66] e G.Polya [67].

Ma queste non sono che considerazioni incidentali, utili solo per illustrare l'ampiezza del campo d'applicazione del ragionamento induttivo; quello che ci interessa effettivamente è viceversa il ruolo della matematica nel ragionamento induttivo, nella teorizzazione di esso, ossia nella sua esatta formulazione e impostazione.

E' chiaro (a parte dubbiosità cui fra poco accenneremo) che l'argomento dell'induzione diverrà suscettibile di trattazione matematica non appena la matematica troverà (al di là dell'algebra di Boole per tradurre la logica del certo) uno strumento adeguato per padroneggiare la logica dell'incertezza; e tale strumento è la teoria delle probabilità. Le dubbiosità di cui occorre far cenno sono quelle provenienti da certe tendenze a limitare il campo della teoria delle probabilità a zone ristrette, ove ragioni di simmetria (come per i dadi) o regolarità statistiche (come per il sesso di un nascituro) facilitano delle valutazioni di probabilità e la concordanza in esse fra diversi individui. Non è questa la sede per discutere espressamente tali questioni preliminari; mi limito a dire che, secondo il mio punto di vista (ripetutamente illustrato: v.p.es. [14], [16], [20]), nessuna discriminazione è giustificata: fra il campo totale dei fatti incerti e i campi cui si pretenderebbe attribuire un ruolo privilegiato non esiste alcuna differenza sostanziale che modifichi il senso della nozione di probabilità (3). Incidentalmente, anche discutendo di induzione, ci capiterà comunque spesso di discutere questioni strettamente connesse alla concezione della probabilità.

Il ruolo della teoria della probabilità nell'impostazione della logica induttiva consiste nell'indicare come debba modificarsi la valutazione di probabilità relativa ad eventi futuri in seguito al risultato di eventi osservati (4): è questo il senso

---

(3) Fra le esposizioni di idee analoghe da parte di altri AA. va segnalata quella breve ed efficace di Good in [29]. Opere in tale indirizzo sono quelle di Ramsey [68], B.O.Koopman [43], L.J.Savage [70] e dello stesso Good [27].

(4) Il termine "modificarsi", qui usato per comodità dato che rende breve e apparentemente chiara l'espressione, è tuttavia inesatto se lo s'interpreta come "correggere". La "probabilità di un evento subordinata a un certo risultato" è *un'altra* probabilità e *non una migliore valutazione* dello stesso ente "probabilità di quell'evento". [Cfr. analoga citazione da Keynes nel §5, con riferimento nella nota in calce (1)]. Chi gioca una cinquina al lotto ha probabilità $1/\binom{90}{5}$ di vincere; se assiste all'estrazione, la sua probabilità cresce a $1/\binom{89}{4}$, $1/\binom{88}{3}$, $1/\binom{87}{2}$ 1/86, 1, man mano che vede uscire numeri giocati, o cade a zero non appena esce un numero diverso. Ma queste successive probabilità non sono, evidentemente, correzioni della prima. Ciascuna è la probabilità relativa allo stato d'informazione specificato; ogni valutazione è "provvisoria" nel senso che in definitiva la probabilità sarà 1 o 0 quando sapremo che l'evento si è verificato o non si è verificato.

che ha nell'impostazione matematica la frase "imparare dall'esperienza".

Da tale punto di vista, la logica induttiva si riduce sostanzialmente al teorema delle probabilità composte o alla sua variante appena un pò più elaborata chiamata ***teorema di Bayes*** :

$$P\left(\frac{H}{E}\right) = P(H).\frac{P\left(\frac{E}{H}\right)}{P(E)}. \tag{1}$$

Esso mostra in che modo, apprendendo che si è verificato E, si passa dalla valutazione ***iniziale*** (cioè: anteriore a tale apprendimento) della probabilità di H, che è P(H), a quella ***finale*** (cioè: posteriore a tale apprendimento) che è P(H/E) (5). In sostanza, ciò non fa che precisare l'atteggiamento spontaneo di chiunque accresce o diminuisce il credito a un'ipotesi a seconda che apprende fatti che essa renderebbe più o meno verosimili o spiegabili.

Non aggiungiamo precisazioni, che sarebbero premature dato il fine meramente illustrativo delle considerazioni che stiamo svolgendo in questo momento, e che incontreremo del resto nel seguito. Ciò che si è detto basterà senz'altro per concludere la presente introduzione pervenendo a indicare come il nesso ora stabilito fra induzione e probabilità conduca in particolare al nesso tra induzione e statistica, che è più specificamente l'argomento del presente ciclo.

A tal fine non rimane che da specificare alquanto la natura

---

(5) Qualche osservazione critica andrebbe fatta anche contro tale interpretazione di P(H/E), che più propriamente andrebbe descritta come probabilità attribuita "ora" ad H sotto la condizione ipotetica che E dovesse verificarsi. Cfr. [17], §31, e cenni nel §5 (nel "punto *4)*").

dell'evento E che si suppone di aver osservato o di poter osservare. In generale esso consiste di un insieme di fatti distinti più o meno numerosi e almeno press'a poco "indipendenti" nel senso della teoria delle probabilità, subordinatamente all'ipotesi H da saggiare o a ciascuna delle impotesi H_j da confrontare.

Ciò significa che, scrivendo $E = E_1 E_2 \ldots E_q$, come prodotto logico dei fatti "singoli" che vogliamo distinguere, la probabilità P(E/H) (o le probabilità $P(E/H_j)$) si esprime (esattamente se si suppone l'indipendenza, approssimativamente se la si suppone valida press'a poco) come prodotto delle probabilità $P(E_i/H)$ (rispettivamente delle $P(E_i/H_j)$). E dalla (1) scende facilmente che i rapporti tra le probabilità delle varie ipotesi vengono allora successivamente alterati nel modo rispondente all'osservazione di ciascuno singolarmente dei fatti $E_1, \ldots, E_q$.

E' specialmente facile ammettere l'indipendenza quando tali fatti E_i sono i più disparati. Ciò non fa che esprimere in sostanza la massima intuitiva secondo cui riteniamo praticamente accertata una tesi quando è suffragata da numerosi indizi a favore, anche se ciascuno di essi da solo non ci sembrerebbe probativo (principio del Cardinale Newman [58]).

Tre esempi del genere chiariranno il concetto.

Se un individuo è sospettato di un delitto, e si raccolgono dati e testimonianze su fatti comunque in relazione alla possibilità che egli sia il colpevole, l'insieme E di tutti i fatti E_i accertati può rendere praticamente certi della colpevolezza o dell'innocenza se quelli che ci spingono in un senso sono sufficientemente prevalenti, tenuto conto dell'opinione iniziale.

Così in guerra (o circostanze analoghe), se informazioni disparate tendono ad avvalorare una certa ipotesi riguardo ai piani del nemico, potremo giungere - sempre partendo da un giudizio iniziale di verosimiglianza - a considerare quella ipotesi come praticamente sicura.

Lo stesso si può ripetere con riferimento a teorie scientifiche, e l'esempio più adatto mi sembra quello della teoria della "deriva dei continenti" di Wegener. Nella sua opera [80] intesa a provare che i continenti si sono formati staccandosi e allontanandosi, mentre inizialmente costituivano un unico blocco (e in particolare l'America del Sud costituisce la parte che combaciava con la rientranza dell'Africa), vengono elencati e coordinati fatti di natura diversissima, tratti dalla geografia, dalla geodesia, dalla geofisica, dalla geologia, dalla paleontologia, dalla biologia, dalla paleoclimatologia, ed insieme considerazioni sulla possibilità di spiegazioni fisiche circa le forze che produrrebbero le traslazioni continentali.

Dalla prefazione a tale libro del Wegener, riproduciamo le sue considerazioni che mi sembrano condensare i canoni del ragionamento induttivo, e in modo tanto più degno di apprezzamento perchè dovuto a persona impegnata ad applicarlo in un campo concreto e difficile di ricerche, e pertanto aliena da possibili deformazioni professionali o pregiudizi filosofici da cui può non essere immune chi si pone il problema in generale e in astratto, come me e come altri che si trovano in situazioni analoghe, non importa se in accordo o in disaccordo con le mie vedute.

Scrive dunque il Wegener (op.cit., p.16): " in un dato momento la terra può aver avuto un solo aspetto. Su ciò non si hanno notizie dirette. Noi ci troviamo di fronte ad essa come il giudice dinanzi all'accusato che si rifiuta di dare spiegazioni e abbiamo il dovere di stabilire la verità in base ad indizi. Tutt le prove che possiamo addurre hanno il carattere fallace degli indizi. *Come giudicheremmo il giudice che emette il suo giudizio fondandosi solamente su una parte degli indizi a sua disposizione?*"

"E' solo abbracciando tutti i rami della scienza geologica che possiamo sperare di giungere alla verità, di tracciare cioè un quadro che rappresenti con ordine la totalità dei fatti noti e perciò possa pretendere di avere il maggior fondamento; e anche in questo caso dobbiamo tener presente che ogni nuova scoperta, da qualunque scienza essa provenga, può modificare i risultati ottenuti".

E veniamo finalmente al caso della statistica.

Esso non differisce da quello illustrato nei precedenti esempi se non per il fatto che gli eventi da osservare, $E_1, \ldots, E_q$, anzichè disparati, sono *analoghi*, o addirittura (secondo un linguaggio che giudico inammissibile) *uguali* (6); secondo l'uso, potremo dire che si tratta di "prove di uno stesso fenomeno".

Spesso tale "analogia" viene presentata non solo come la circostanza esteriore che caratterizza il caso trattato dalla statistica, ma anche come la ragione fondamentale della validità delle conclusioni. Ma, se ragionamenti e conclusioni sono basati sulla teoria delle probabilità (e non su tentativi di traduzione

---

(6) Logicamente, due eventi sono uguali solo se si tratta del medesimo evento (ossia, volendo parlare di "prove", di quel dato risultato in quella data prova). Altrettanto priva di senso è, a rigore, la nozione di "prove di uno stesso fenomeno" (lo sono due colpi qualunque a testa e croce? o solo se si usa una stessa moneta? o moneta di un dato conio? o se il lancio è effettuato da una stessa persona?). Tuttavia la locuzione può essere accettata e usata pur di avvertire che non ha alcun significato ma che la si usa per indicare collettivamente degli eventi di cui faccia comodo sottolineare una qualche analogia.

in regolette macchinali rese autonome da ogni criterio di sen - ta applicabilità), la natura dei fatti osservati è irrilevante, importando solo le relazioni tra le probabilità. L'analogia può favorire il giudizio di ugual probabilità, che porta a semplificazioni; viceversa però l'indipendenza va accolta con molto maggiori precauzioni che nel caso di fatti disparati. Ma queste sono cose da vedersi caso per caso [19].

Una circostanza utile, che spesso si presenta nel caso statistico, risiede nella possibilità di moltiplicare a volontà "prove" su quel dato tipo di eventi "analoghi" (mediante ripetute osservazioni, o a volte mediante esperimenti espressamente predisponibili).

Comunque, secondo il punto di vista che seguiremo, il caso della statistica non è che un caso particolare del ragionamento induttivo impostato secondo la teoria delle probabilità: caso caratterizzato da particolarità interessanti dal punto di vista pratico e applicativo, ma che non comportano alcunchè di nuovo e diverso dal punto di vista concettuale. Non si nega, tutt'altro, l'esistenza di sviluppi interessanti anche teoricamente, ma riguardano l'aspetto tecnico-matematico della teoria, non i fondamenti concettuali.

Questa presa di posizione, che potrebbe apparire rispondente a mere questioni di punti di vista filosofici, risulterà invece essenziale al fine di chiarire le divergenze reali sui metodi e criteri da seguire nella trattazione di problemi concreti (7)

---

(7) Ci sono molte disparità nell'uso della parola statistica (talvolta estesa a includere il calcolo delle probabilità, o ristretta alla parte descrittiva o poco più che descrittiva dei dati collettivi). La distinzione qui proposta mi sembra la più rispondente a un'interpretazione moderna basata sul concetto tradizionale,

./.

Fra i molti aspetti che affioreranno in tali discussioni, il motivo principale (e forse l'unico cui tutti gli altri si possono ricondurre) è proprio quello riportato in corsivo nella citazione di Wegener : *la necessità di tener conto di tutto ciò che si sa, non importa con quale metodo e da quale fonte*. Tale affermazione non si può certo dire nuova nell'ambito della Statistica: basti rammentare che R.A.Fisher ha affermato giustamente con molta enfasi che, *mentre nella logica deduttiva utilizzando una parte delle premesse si potrà avere un minore insieme di conclusioni ma pur sempre ESATTE, nella logica induttiva utilizzando solo una parte dell'informazione si può giungere invece a conclusioni FALSIFICATE* (8) (come avviene se si sopprimono le testimonianze a favore, o quelle contrarie).

Tuttavia, i metodi statistici, per amore di scheletricità e meccanicità, o di apparente eliminazione di aspetti concettuali e soggettivi, ricorrono spesso sistematicamente a una perdita d'informazione per usarne una parte meglio addomesticata per particolari elaborazioni.

L'argomento sarà approfondito non soltanto nel seguito del presente corso, ma anche in quello parallelo del Prof.Savage [71]

---

nonchè la più utile per sottolineare una particolarità significativa. Si tratta comunque di una mera questione di terminologia. - La definizione della statistica proposta da Savage nel corso parallelo ([71],§2) è diversa, senza che ciò significhi alcun disaccordo sostanziale.

(8) Citiamo da [23], pag.55 : "Although in the deduction of statements of certainty it is legitimate to draw inferences from some of the axioms available while ignoring others, or, in other words to base a valid argument on a chosen subset only of the available axioms, no such liberty can be taken with statements of uncertainty, where it is essential to take the whole of the data into account, though some part of it may be shown on examination to be irrelevant, and not to affect the result."

dove numerose esemplificazioni renderanno particolarmente evidente l'importanza e il significato (fondamentalmente unico ma ricco di svariate apparenze) di siffatte manchevolezze e della loro eliminazione, in vari tipi di problemi e per le più diverse applicazioni.

B.de Finetti

SGUARDO STORICO E COMPARATIVO

§2. DAGLI INIZI ALLA CRISI DELL'IMPOSTAZIONE CLASSICA

La teoria delle probabilità, sorta come è noto nel '600 dallo studio di problemi sui giochi (dadi, carte, sorteggi, e simili), era subito pervenuta a stabilire in quel campo, come criterio di decisione, quello consistente nel cercar di massimizzare la speranza matematica del guadagno.

Per giungere a quello che si può considerare attualmente come il criterio di decisione più accreditato, mancavano soltanto due aggiunte, entrambe sopravvenute ben tosto, nel '700 (1):

- un allargamento dell'obbiettivo delle decisioni, raggiunto da Daniele Bernoulli [2] con l'introduzione, sotto il nome di *speranza morale*, di quella che ora diremmo l'*utilità;*
- un allargamento del campo delle probabilità, mettendole in relazione coll'osservazione statistica e quindi stabilendo il nesso che lega ad esse il ragionamento induttivo, e ciò fu realizzato da Bayes [1].

Se non che, quel traguardo che in tal modo si potrebbe considerare raggiunto d'impeto fin dagli inizi, si è dovuto invece faticosamente riconquistare due secoli dopo (ed anzi è tuttora controverso). Perchè? Racconteremo la storia che, secondo la nostra interpretazione, consiste delle seguenti tre fasi.

Prima fase: rapida affermazione della teoria bayesiana (e di quella bernoulliana), con qualche leggerezza nell'interpretazio-

(1) Come ha recentemente fatto rilevare Guilbaud [39], anche il concetto del valore Minimax della "teoria dei giochi" (nel senso di Borel [4] e von Neumann-Morgenstern [57])era già stato scoperto fin dal 1712 [56] (sia pure con riferimento ad un semplice esempio)!

ne e applicazione, e conseguente discredito e abbandono indiscriminato di tutta la concezione.

Seconda fase: ricerca di altre vie per affrontare e impostare i problemi dell'induzione statistica facendo a meno degli elementi e concetti su cui si era allargato il discredito.

Terza fase: graduale affioramento delle insufficienze di tali tentativi; revisione del giudizio sulle teorie bayesiana e bernoulliana con separazione degli elementi validi da quelli equivoci; ritorno alla posizione di partenza opportunamente rettificata (posizione spesso designata come neo-bayesiana neo-bernoulliana).

In un rapido sguardo storico cercheremo di illustrare questo sviluppo, non certo sotto tutti gli aspetti e con molti dettagli, ma solo con riferimento alle circostanze preminenti agli effetti delle questioni d'impostazione concettuale che c'interessano. Ci soffermeremo sulla prima fase in questo paragrafo, sulla seconda e la terza nei due successivi.

Dei due elementi costitutivi della moderna teoria della decisione, quello riguardante la speranza morale di D.Bernoulli ha una parte piuttosto secondaria nelle vicissitudini di cui ci dobbiamo occupare. Egli ebbe il torto di precorrere di un secolo i concetti marginalisti dell'economia [36] e di due la loro reinterpretazione probabilistica [68], [57], [53], [69], [70] cosicchè fu poco compreso e avvenne che alla generalità della sua impostazione si preferisse la particolare esemplificazione in cui suggerì di misurare l'utilità di un patrimonio mediante il logaritmo del valore.

Importanza decisiva ebbero invece le discussioni sull'impostazione data da Thomas Bayes al problema da lui formulato come segue: "cosa si può dire della probabilità di un evento *di cui non si sa nulla* quando si conosce il risultato di n eventi ad esso analoghi (favorevole per m, sfavorevole per n-m)?". Non soffermiamoci sul fatto che tale formulazione è poco soddisfacente (su tali questioni dovremo discutere abbondantemente in seguito); vediamo per ora di distinguere tre elementi costitutivi dell'impostazione, che andranno esaminati separatamente.

Bayes suppone :

(1) che la "probabilità incognita" p abbia probabilità dx di esser compresa in un qualunque intervallo (x, x+dx) in (0,1),

(2) che gli eventi considerati siano indipendenti, per ogni ipotesi p=x sul valore di p ,

(3) che quindi, dopo le osservazioni indicate, la probabilità che p cada tra x e x+dx diviene

(2) $$Kx^{m}(1-x)^{n-m}dx$$ (2).

L'ultimo punto è fuori discussione: esso non è che il *teorema di Bayes* già menzionato nel §1 (formula (1)), ossia, come osservato, il teorema delle probabilità composte in una forma un pò più elaborata. C'è qualcosa di controverso, come avremo ampia occasione di vedere, ma riguarda solo la vastità del campo di applicazione che si restringe se si vuole limitare a un significato restrittivo la nozione di probabilità.

Il punto *(2)* è un'ipotesi occorrente per descrivere il caso che si considera, e non avremo che a chiarirne il vero significa-

---

(2) Indicheremo sempre con K la costante occorrente, in una qualunque formula, per normalizzarla (nel senso che di norma apparirà ovvio e sarà sottinteso: qui, deve valere 1 l'integrale da 0 a 1, cosicchè $K=(n+1)\binom{n}{m}$). Si badi che il valore di K non solo è in genere diverso da una formula all'altra ma potrebbe cambiar valore entro uno stesso calcolo.

to in seguito (§7).

Il punto (*i*) invece è quello che costituì il punto nero in tutte le discussioni. Esso viene spesso designato come "postulato di Bayes": il postulato secondo cui, quando "non si sa nulla", si deve adottare la distribuzione uniforme per p su (0,1). Esso non ha nessuna relazione necessaria con l'impostazione del problema precedente secondo il teorema di Bayes: si potrebbe benissimo assumere una distribuzione iniziale per p di forma qualunque, p.es. con probabilità $\phi(x)dx$ che p cada tra x e x+dx. Esso poi non ha neppure in realtà alcun significato, dato che è tutt'altro che chiaro cosa si debba intendere, in un caso concreto, per "non saper nulla". Se vuol dire che attribuisco a p distribuzione uniforme, il "postulato" non è che una tautologia; se intendo "nulla" alla lettera il postulato è assurdo (se non so nulla degli eventi E_i non so nulla neppure dei prodotti E_iE_j e quindi anche p^2 dovrebbe avere distribuzione uniforme); a parte che "non saper nulla" a rigore dovrebbe significare allora che non so neppure di che evento si sta parlando.

Tutto sommato non si tratta quindi neppure di qualcosa che ha un significato matematicamente parlando, e possa quindi ragionevolmente entrare nell'impostazione per determinarla o per inficiarla: è una frase vuota dall'apparenza metafisica, che può avere solo il senso deteriore di invitare a far uso in ogni applicazione, senza precauzioni e senza criterio, della distribuzione uniforme.

E' questo che è stato fatto largamente, e che ha condotto a mettere in dubbio e a condannare in blocco l'intera impostazione.

La colpa non si può far risalire a Bayes, che su questo punto aveva i suoi bravi dubbi, tanto che la sua memoria non fu presentata che postuma (due anni dopo la sua morte avvenuta nel 1761) dal suo amico Richard Prize. Questi non condivideva, anzi mostra di trovare strane, le esitazioni di Bayes, ed analogo atteggiamento sembre essere quello di Laplace [45] (3).

Laplace (e la maggior parte degli autori contemporanei o di poco successivi), pur conoscendo l'impostazione più generale con distribuzione iniziale $\phi(x)dx$ e distribuzione finale

(3) $$K\ \phi(x).x^{m}(1-x)^{n-m}dx,$$

sembra infatti accolgano per acritici motivi aprioristici come privilegiata la distribuzione uniforme (ossia, il cosiddetto postulato di Bayes). Da notare che nulla si potrebbe obiettare se, in luogo di partire da pregiudizi aprioristici, essi dicessero che, sotto condizioni molto poco restrittive per la ϕ, e per n abbastanza grande, è praticamente lecito sostituire approssimativamente il caso effettivo con quello $\phi\equiv 1$.

Soffermiamoci anzitutto un momento a indicare come il passaggio dalla distribuzione iniziale a quella finale non sia che un'applicazione del concetto premesso nel n.1: ogni risultato favorevole fa moltiplicare per x la probabilità dell'ipotesi $p=x$ (e ogni risultato sfavorevole per $1-x$), e l'indipendenza delle varie prove subordinatamente ad ogni ipotesi (punto (2) dell'impostazione di Bayes) fa sì che si abbia semplicemente a moltiplicare fra loro tutti questi fattori.

Conviene inoltre soffermarci ancora a richiamare i semplici risultati (stabiliti da Bayes) per il caso speciale della distribuzione uniforme da lui considerato.

(3) Notizie storiche più diffuse e con ampie citazioni si trovano in R.A.Fisher [23], Cap.II.

Inizialmente (cioè: per $\phi(x)\equiv 1$), la probabilità di una qu - lunque frequenza su N colpi è la stessa, cioè $1/(N+1)$ (le frequenze possibili essendo M/N con $M=0,1,2,\ldots,N$); in particolare è $1/(N+1)$ la probabilità che i risultati di N prove siano tutti favorevoli (oppure: che siano tutti sfavorevoli); più in particolare ancora, la probabilità in una prova è 1/2.

Dopo osservate n prove, di cui m con risultato favorevole ed n-m sfavorevole, ossia quando la distribuzione è divenuta $Kx^m(1-x)^{n-m}$ la probabilità in una prova successiva è $(m+1)/(n+2)$: è cioè la frequenza osservata corretta nel senso di pensare aggiunte due prove in più con risultati uno favorevole e uno sfavorevole. Questo risultato è conosciuto sotto la denominazione di "regola della successione"; in particolare, se m=0, risulta $1/(n+2)$ la probabilità di un evento dopo n prove tutte sfavorevoli (o viceversa: del non verificarsi dopo n tutte riuscite). Quanto alle probabilità delle varie frequenze M/N su N prove, esse non sono più naturalmente uguali, ma proporzionali a

$$\binom{m+M}{m}\cdot\binom{n-m+N-M}{n-m} \tag{4}$$

Le formule relative a questo caso vennero applicate meccanicamente ad ogni tipo di esemplificazioni, tra cui celebre quello della probabilità che il sòle sorga domani dato il numero di giorni da cui ci è tramandato che il sole è sempre sorto.

Questo uso acritico e generale del caso particolare corrisponde al "postulato di Bayes", e la debolezza logica di tale "postulàto" in sé, condussero col tempo, nonostante il grande prestigio di Laplace, al sorgere di voci discordi. La giustificazione del ragionamento induttivo secondo la traccia di Bayes e Laplace era ed appariva difettosa, ma, anzichè pensare che per eliminare il difetto occorreva basarsi su qualcosa di perfezionato, i critici sembra trovassero sufficiente sopprimere senz'altro il ragionamento difettoso per far diventare accettabile almeno empiricamente senza giustificazione alcuna il metodo che la giustificazione imperfetta non bastava a giustificare. Come chi dicesse che, essendo pericoloso costruire sulla sabbia, basta levar via la sabbia e costruire sul vuoto per eliminare ogni pericolo.

Sono sorte in tal modo le tendenze che ritengono lecito ricondurre in qualche modo *per definizione* la probabilità alla frequenza, eludendo così la necessità di spiegare il fatto che erigono a verità, e cioè di chiedersi perchè mai si sia indotti a valutare delle probabilità in base alle frequenze ossia a prevedere che probabilmente certe frequenze non varieranno molto.

Ma, più che in tale campo della definizione "statistica" della probabilità, l'eliminazione dell'impostazione bayesiana comporta rivolgimenti vasti e complessi nella formulazione dei concetti base e dei metodi di lavoro della statistica matematica. Di ciò ci occuperemo espressamente illustrando in che modo i concetti bayesiani furono sostituiti con criteri di tipo "oggettivistico".

§3. IL SOPRAVVENTO DELLE CONCEZIONI OGGETTIVISTICHE

Spieghiamo anzitutto perchè designiamo come "oggettivistiche" le concezioni di tipo non-bayesiano: perchè, come vedremo, per ricostruire l'impostazione bayesiana occorrerà interpretarla in senso soggettivo.

Ed ecco in cosa consiste la caratteristica essenziale delle concezioni oggettivistiche: nell'introdurre qualcosa che non è nè logica del certo nè logica del probabile.

Mentre nell'impostazione bayesiana "all expressions of uncertain knowledge must have the same logical form, namely that of a statement of probability", come ben dice Fisher ([23], p.44), egli propende invece per l'uso di metodi che "do not generally lead to any probability statements about the real world, but to a rational and well-defined measure of reluctance to the acceptance of the hypothesis they test" (ibid.). E se, nonostante l'esplicito diniego, tale "reluctance" può forse somigliare a una probabilità soggettiva, più radicale ancora è Neyman ([61],p.235) nel sottolineare che i metodi statistici conducono a "stating that..." ma che occorre guardarsi dall'attribuire a tale parola un qualunque significato logico ("conclude" that...) o probabilistico ("believe" that...).

Si abbandona in tal modo ogni idea di una interpretazione sistematica e significativa del problema dell'inferenza statistica, per ridursi ad escogitare caso per caso dei "test" per "confermare" delle ipotesi, o dei metodi per "stimare" dei parametri, configurando così come una questione autonoma e in larga misura arbitraria il problema di trarre dall'esperienza un qualchecosa che si suggerisce di utilizzare come se fosse una conclusione o