I PERICOLI DELLA STATISTICA (1939)

Corrado Gini

Introduzione
di Sandro Graffi

Corrado Gini (Motta di Livenza, Treviso, 1884 - Roma 1965) fu statistico, economista e sociologo di prominenza mondiale. L'indice di Gini, ad esempio, è l'indicatore quantitativo universalmente impiegato per misurare le disuguaglianze nella distribuzione del reddito interno nei vari paesi.

Fu Presidente dell'ISTAT, Professore all'Università di Roma dove fondò l'attuale Facoltà di Scienze statistiche, fondatore di varie riviste scientifiche e della Società Italiana di Statistica, e molto altro ancora.

L'itinerario culturale di Corrado Gini ne fa senz'altro una delle figure ispiratrici principali di questa rivista. Si laureò in Giurisprudenza a Bologna nel 1905, discutendo una tesi in Statistica, disciplina che a quei tempi era obbligatoria per quella laurea, e per quella sola[1]. Oltre ai prescritti esami sulle materie giuridiche sostenne quelli fondamentali del biennio di scienze matematiche nonché quelli del biennio di Scienze naturali. Da un lato, questa solida preparazione scientifica lo mise in grado di affrontare i problemi delle scienze sociali ed economiche anche sul piano strettamente quantitativo; dall'altro, la sua formazione giuridica, economica e sociologica lo rese

[1] La Statistica rimase disciplina obbligatoria per la laurea in Giurisprudenza fino alla riforma Gentile, che introdusse al suo posto la Filosofia del diritto.

A mio parere questa sostituzione condensa come meglio non si potrebbe lo spirito profondo della riforma Gentile. Le conseguenze sul cambiamento di prospettiva culturale della classe dirigente italiana, in larga maggioranza di formazione giuridica, sono del tutto evidenti.

ben conscio dei rischi che si corrono quando si utilizzano acriticamente nelle scienze umane i concetti matematici elaborati nelle scienze esatte. In ultima analisi, il rischio principale è attribuire ad opinioni valide quanto si voglia, ma pur sempre opinioni, valore di affermazione scientifica incontrovertibile.

Questo è il tema dominante del saggio che qui in parte ripubblichiamo, *I pericoli della statistica*, che costituì il discorso introduttivo alla prima riunione della Società Italiana di Statistica tenutasi a Roma il 9 ottobre 1939[2].

In tempi in cui i sondaggi d'opinione, basati sui campionamenti statistici, a loro volta fondati sulla sfuggente nozione di probabilità, rappresentano il fondamento di ogni scelta che in qualsiasi modo riguardi la collettività, comprendere i loro limiti è necessità ineliminabile. In tempi in cui tutti i cittadini sono chiamati a decidere sulla sorte dei loro simili, affermarono congiuntamente J.L. Lagrange e P.S. Laplace[3], occorre soprattutto insegnare loro a diffidare delle intuizioni, fossero anche le più verosimili; ed a questo scopo, continuarono, nulla è più adatto delle considerazioni probabilistiche e statistiche, in cui spesso risultati rigorosi ed intuizione si trovano in contraddizione. Nelle parole di Gini medesimo:

> *La morale della favola* [...] *è che il calcolo delle probabilità, e la statistica a mezzo del calcolo delle probabilità, anche se applicate a masse di casi, non possono mai portare a conclusioni sicure, ma solo a conclusioni probabili. Possono legittimare dei dubbi più o meno forti – e questa è certamente una funzione utile – ma non possono mai scioglierli in modo definitivo. Possono fornire non «testi di significatività» ma «elementi di sospetto». Per quanto alto sia il grado di probabilità delle loro conclusioni, di fronte ad un dato sicuro della coscienza e dell'esperienza essi devono ritirare le corna.*

[2] Il saggio *I pericoli della statistica* è stato ristampato nel volume di scritti di Corrado Gini dal titolo *Statistica e induzione*, Bologna, CLUEB, 2001, al quale si rinviano i lettori interessati. Ringrazio il curatore del libro Prof. Italo Scardovi, primo Preside della Facoltà di Scienze statistiche dell'Università di Bologna, che fu in gioventù collaboratore diretto di Gini, per averci accordato il permesso di riprodurlo qui (sono state tolte, per motivi di spazio, solo alcune pagine non essenziali).

[3] Nel discorso introduttivo al primo corso di matematica mai tenutosi presso l'École Normale Supérieure di Parigi. Il discorso fu pronunciato il 1° Piovoso dell'anno III (20 gennaio 1795).

I PERICOLI DELLA STATISTICA (1939)*

Gli inconvenienti e i pericoli, a cui dà luogo la rilevazione dei fenomeni naturali e sociali eseguita senza il sussidio del metodo statistico, sono stati ripetutamente segnalati e possono dirsi ormai soddisfacentemente sistemati: impossibilità di cogliere certe regolarità di massa; conclusioni erronee sopra altre, sia per il fatto che taluni casi cadono sotto la nostra osservazione più di frequente che taluni altri, sia per il fatto che, pur cadendo sotto la nostra osservazione con la stessa frequenza, colpiscono maggiormente la nostra attenzione e lasciano una traccia più profonda nella nostra memoria; nella migliore delle ipotesi,conclusioni approssimative, sia sulle regolarità di massa, che, con l'osservazione comune, è possibile cogliere solo qualitativamente, sia sulla intensità di fenomeni singoli, che è bensì possibile rilevare quantitativamente, ma senza la desiderabile precisione. È appunto da tali inconvenienti e pericoli che sorge l'utilità, e spesso la necessità, del metodo statistico[1].

Ma non ha, d'altra parte, anche il metodo statistico i suoi inconvenienti e i suoi pericoli?

«Con la statistica» – si afferma anzitutto – «si dimostrano le tesi più contraddittorie».

«Con la statistica» – si afferma pure – «si raggiungono talvolta conclusioni destinate poi ad essere smentite da ricerche successive, condotte con lo stesso metodo statistico o con altri metodi».

Gli statistici non si sono mai curati – che io sappia – di prendere in considerazione tali accuse. E a torto. Non serve nascondere la testa sotto l'ala, a mo' dello struzzo, per cancellare una realtà sgradevole. La realtà è che tali accuse sono diffusissime nel grosso pubblico, né sono limitate al grosso pubblico, trovando anzi talvolta eco presso scienziati di valore indiscusso.

Né persuade il dire che divergenze si verificano in tutte le scienze e che è appunto dalla loro discussione che emerge la verità, e che simil-

* Discorso inaugurale alla Prima Riunione Scientifica della Società Italiana di Statistica (Pisa, 9 ottobre 1939); apparso negli «Atti della I Riunione Scientifica della S.I.S.», Roma 1939.

[1] Cfr., in proposito, gli *Appunti di statistica metodologica* raccolti alle nostre lezioni tenute nelle Università di Padova e di Roma a partire dal 1914, e tradotti, dopo revisione, in ispagnolo sotto il titolo *Curso di Estadística*, Barcelona, Editorial Labor, S.A., 1935, pp. 10-14.

mente in tutte le scienze i risultati raggiunti sono continuamente soggetti a revisione e correzione, donde scaturisce, appunto, il progresso scientifico. Non persuade, perché l'accusa di contraddittorietà mossa ai risultati della statistica si basa proprio sul fatto che molte volte le divergenze, a cui essa conduce, non si appianano o, quanto meno, non si appianano col metodo statistico, e l'accusa di insufficienza e di fallacia deriva dal fatto che i risultati ottenuti dalla statistica spesso non vengono integrati e precisati, ma addirittura cancellati o capovolti, da indagini successive.

Né, infine, basta il dire che non si può far colpa al metodo degli errori di chi lo applica. Il metodo sarebbe corretto; e non corrette, invece, le sue applicazioni. Se non che i metodi di ricerca scientifica non sono che strumenti, i quali hanno un valore in quanto si possono praticamente applicare, e applicare, ben inteso, dai comuni mortali. D'altra parte, i risultati contraddittori o falsi, che alla statistica si rinfacciano, non sempre sono stati raggiunti da persone inesperte, ma spesso da specialisti, talvolta da personalità scientifiche di primissimo ordine. D'accordo che tali risultati divergenti od erronei non devono togliere la visione dei molti risultati veridici e concordanti, a cui col metodo statistico si è pervenuti. Ma, se i primi, secondo una opinione molto diffusa tra i non statistici, assumono di fronte ai secondi una certa importanza, bisogna riconoscere che il metodo statistico è, pur nelle mani degli statistici, un metodo pericoloso. Conviene dunque cercare di diminuirne la pericolosità, tenuto anche conto della normale capacità delle persone che se ne servono.

D'altra parte, la stessa affermazione che i metodi statistici in uso, sono, sempre e tutti, corretti, e gli inconvenienti, che si lamentano nelle conclusioni raggiunte, dipendono, tutti e sempre, da chi meno correttamente li applica è, al postutto, gratuita e potrebbe ben essere troppo ottimista. Essa merita, in ogni modo, di venire esaminata criticamente.

Comunque, dipendano la contraddittorietà e la fallacia dei risultati, a cui la statistica talvolta perviene, da deficienza dei metodi o da insufficienza di chi li applica, la natura e le cause di tale deficienza o insufficienza meritano di venire ricercate, individuate e segnalate per evitare di incorrervi e per avvisarne, se possibile, i rimedi.

Procederemo nella nostra indagine per via di esempi particolari, generalizzando poi le conclusioni raggiunte.

Vediamo un esempio di conclusioni contraddittorie.

A e *B* sono due paesi. Si domanda in quale dei due la mortalità è più elevata. Il rapporto dei morti alla popolazione è più elevato in *A* che in *B* (supponiamo che la differenza relativa sia del 10%) e quindi molti statistici concludono che la mortalità è più elevata in *A*.

Ma altri dicono: No, tale conclusione non è autorizzata; anzi è erronea. In *A* vi sono molti più bambini che in *B*, e, tra i bambini, come è noto, la mortalità è particolarmente elevata; conviene eliminare l'influenza dell'età, col metodo, per esempio, della popolazione tipo. Si trova allora che la mortalità è invece più elevata in *B* per una differenza relativa – poniamo – dell'8%. Un momento! – obbietta un terzo gruppo di statistici – *A* è un paese di forte emigrazione e il numero dei maschi vi resta molto al di sotto di quello delle femmine, ciò che invece non succede in *B*; dobbiamo eliminare anche l'influenza del sesso, poiché è certo che la mortalità, a parità di altre circostanze, è più alta per il sesso maschile. Eliminata anche tale influenza, i coefficienti di mortalità risultano eguali per i due paesi. Ma non basta – osserva un quarto gruppo –: la sproporzione dei sessi rende più rari i matrimoni e quindi meno alta la percentuale dei coniugati in *A*: si deve eliminare anche l'influenza dello stato civile. Fatta questa terza eliminazione, la mortalità risulta di nuovo più elevata in *B*, ma leggermente (differenza relativa del 4%). E perché non eliminare anche l'influenza della professione, che tanto influisce sulla mortalità? Eliminato anche questo fattore, la differenza relativa di mortalità a svantaggio del paese *B* sale al 15%. Ma perché non eliminare anche l'influenza della ricchezza? Dopo questa ulteriore eliminazione, la mortalità risulta di nuovo più elevata in *A*, ma in modo appena sensibile (differenza relativa dell'1%).

Se non che il clima è molto diverso nei due paesi. Ammettiamo che si possa eliminare anche l'influenza di questo fattore, e alla fine la mortalità risulti pressoché uguale nei due paesi con una differenza relativa del 2% a svantaggio di *B*. Così, in una questione tanto semplice quanto quella di sapere se la mortalità è più elevata in *A* o in *B*, siamo passati, coi metodi statistici, attraverso sette conclusioni più o meno diverse. La differenza relativa a vantaggio di *B* è risultata infatti successivamente: +10%, –8%, ±0%, –4%, –15%, +1%, –2%. Né può dirsi che si sia trattato di successive approssimazioni verso il risultato finale; il valore più approssimato (±0%) al risultato finale (–2%), si è ottenuto, infatti, dopo la seconda eliminazione, e il risultato più divergente (–15%) dopo la quarta eliminazione. In realtà, i successivi valori ottenuti si dispongono irregolarmente intorno al risultato finale.

Quale è dunque la conclusione esatta?

In realtà, tutte sono in un certo senso esatte e tutte possono essere in un certo senso errate. La divergenza dipende dal fatto che l'espressione «mortalità» viene intesa in senso diverso. Noi possiamo, infatti, voler misurare, mediante la mortalità, la resistenza vitale della popolazione considerata nella sua composizione effettiva per età, sesso, stato civile, professione, ricchezza e nell'ambiente in cui essa effettivamente vive; e, per taluni scopi, è proprio quello che ci occorre. Oppure possiamo voler misurare la sua resistenza vitale prescindendo dalla circostanza, transitoria per ogni individuo, dell'età; e anche ciò può essere rispondente a certi scopi. O ancora possiamo voler misurare la resistenza vitale a prescindere anche dal sesso e dallo stato civile, ciò che può essere pienamente giustificato quando, ad esempio, si voglia fare una previsione sulla resistenza vitale delle generazioni future. Per altri scopi ancora, può interessare di confrontare la resistenza vitale delle due popolazioni a prescindere da ogni influenza sociale, eliminando pertanto anche l'influenza della professione e della ricchezza; e può infine rispondere al fine della ricerca di eliminare anche l'influenza del diverso ambiente fisico, per esempio quando interessi di giudicare quale delle due popolazioni sia meglio adatta per colonizzare un nuovo territorio, il cui clima sia intermedio tra quelli dei due paesi considerati.

Definire con precisione l'oggetto della ricerca è una necessità che non vale solo per le indagini statistiche, ma per ogni genere di indagini scientifiche, e quindi si può dire che inconvenienti simili si verifichino in tutte le discipline. – Sì, ma si verificano nella stessa misura? Dobbiamo riconoscere che ben difficilmente noi troveremmo, nella definizione di un fenomeno non statistico, la molteplicità di significati che abbiamo visto potersi attribuire alla parola «mortalità», e che analogamente si potrebbero attribuire alle parole «natalità», «nuzialità», «morbilità» e via dicendo. Gli è che i fenomeni collettivi, di cui la statistica si occupa, dipendono da una molteplicità di fattori, tra i quali, secondo i casi, può interessare di eliminarne qualcuno o di eliminarne parecchi o di eliminarli tutti salvo uno o infine di non eliminarne nessuno. La possibilità di conclusioni contraddittorie, derivante da concetti divergenti del fenomeno in esame, si presenta dunque in tutte le ricerche scientifiche, ma nelle ricerche statistiche assume particolare importanza.

D'altra parte, noi abbiamo finora supposto di disporre di tutti i dati necessari per eliminare l'influenza dei fattori che ci interessava di escludere. In pratica, di solito, le cose stanno ben diversamente.

Talvolta dipende dalla natura stessa del fenomeno che i dati siano insufficienti a tale scopo, come quando si tratti di eliminare l'influenza del clima. Noi dovremmo per ciò disporre di dati, sulla mortalità del paese *A*, in anni in cui il clima vi è stato, per eccezione, analogo a quello che normalmente è nel paese *B*; e, viceversa, di dati, sulla mortalità del paese *B* in anni in cui il clima vi è stato, per eccezione, analogo a quello che è normalmente nel paese *A*. Ma il solo fatto che ciò non può verificarsi se non per eccezione rende difficile di avere una massa di dati sufficiente, poiché è da ricordarsi che, trattandosi di fenomeni statistici, non solo è necessario avere dei dati, ma anche dei dati sufficientemente numerosi per essere autorizzati a riguardare come neutralizzate le circostanze perturbatrici di carattere accidentale.

Al fine di eliminare altri fattori, la natura può offrire bensì una esperienza sufficientemente ampia, ma le rilevazioni statistiche possono essere inadeguate. Quale è il paese che permette di classificare i morti e i viventi secondo le combinazioni di età, sesso, stato civile, professione e ricchezza, così da poter eliminare ad un tempo l'influenza di tutti questi fattori dai coefficienti di mortalità? Che io sappia nessuno. Lo statistico deve pertanto in tal caso accontentarsi di eliminare l'influenza di alcuni soltanto dei fattori che vorrebbe escludere, giungendo così a conclusioni approssimate. Questa è anzi la regola.

Altre volte si dà il caso che il materiale statistico permetta bensì l'eliminazione di determinati fattori, ma unicamente in base a certe ipotesi, valide solo in via approssimativa, per esempio in base all'ipotesi della sovrapposizione degli effetti. È l'ipotesi che sta alla base del metodo delle correlazioni parziali, a cui il materiale statistico permette di far ricorso in certi casi, nei quali i procedimenti della popolazione tipo o dei coefficienti tipo non si possono invece applicare.

Altre volte, infine, il materiale statistico non permette di eliminare tutti i fattori γ, δ, ε ma solo i due fattori γ e δ oppure δ ed ε oppure γ ed ε. È arbitrario eliminare questa o quella coppia di fattori. In tale circostanza, vi è chi preferisce eliminare una e chi l'altra. La conclusione sarà sempre approssimata, ma l'approssimazione potrà compiersi in senso diverso a seconda che si sia eliminata l'una o l'altra coppia. Nell'esempio sopra considerato, se si elimina età, sesso, stato civile e professione, la mortalità risulta più elevata nel paese *B* che nel paese *A*; se si elimina età, sesso, stato civile e ricchezza, essa risulta al contrario, più elevata nel paese *A* che nel paese *B*. Ora, supponiamo che non si abbia la classificazione dei dati necessari per eliminare contemporaneamente, oltre che età, sesso e stato civile, anche ricchezza e professione. Chi elimini allora la ricchezza giungerà a conclusioni diverse da chi elimini la professione. Nell'impossibilità di eliminare tutti i fattori di perturbazione, è affidato all'intuito statistico del ricercatore di eliminare l'uno o l'altro di essi. Questo caso si presenta in moltissime, direi anzi in quasi tutte le ricerche statistiche, e sta qui la ragione fondamentale per cui l'intuito ha così grande importanza nella statistica applicata.

Vero è che esiste la possibilità di eliminare, l'uno dopo l'altro, tutti i gruppi di fattori eliminabili: per esempio, prima i due fattori γ e δ e poi i due fattori δ ed ε e infine i due fattori γ ed ε. Ciò mostrerebbe che i dati disponibili non permettono una conclusione sicura, ma eviterebbe che una conclusione, ritenuta sicura, venga successivamente contraddetta. Se non che, bisogna pur fare i conti con la laboriosità dei calcoli che tale ricerca esigerebbe e che entra spesso in conflitto con le disponibilità di tempo e di mezzi.

Concludendo, la insufficienza dei dati che la natura stessa fornisce e, non meno spesso, la inadeguatezza delle rilevazioni statistiche, impediscono di solito alla statistica di giungere a conclusioni sicure e, aggiungendosi alle limitazioni di tempo e di mezzi, lasciano ampio campo all'intuito personale che porta inevitabilmente a conclusioni divergenti da parte dei vari ricercatori. È quanto si verifica in tutti i campi di ricerca, in cui il risultato, a cui si mira, dipende da una complessa molteplicità di fattori impossibili a disciplinare oggettivamente: tale è il caso nella medicina, nelle arti belle, nelle arti militari, in tutte – si può dire – le arti. La larga parte lasciata all'intuito soggettivo è appunto ciò che distingue le arti dalle scienze. La statistica applicata è in buona parte ancora, e sempre in un certo grado sarà, un'arte, nel cui successo l'intuito, o, se più vi piace, il genio del ricercatore, ha e avrà sempre una parte importante. Vi è chi potrà riguardare tale suo carattere come un vantaggio, altri come uno svantaggio, di fronte alle altre discipline. Esso ci spiega, in ogni modo, molte delle divergenze e contraddizioni a cui i suoi cultori pervengono.

Chi, ponendosi da un punto di vista esclusivamente scientifico, giudichi tale carattere uno svantaggio, deve, d'altra parte, tener presente che i risultati, sia pur spesso approssimati ed incerti, raggiunti con la statistica, non si potrebbero generalmente ottenere con altri metodi. La statistica è una scienza marginale, che giunge là dove, coi dati e coi mezzi disponibili, le altre non giungono. Talvolta le altre scienze vi arrivano dopo, quando i dati dell'esperienza si sono accresciuti e i mezzi tecnici sviluppati. In tal caso, la statistica ha anzitutto una funzione di avanscoperta, funzione che è sempre pericolosa, ma preziosa. Tale è stata, ad esempio, la funzione della statistica nello studio della variabilità e dell'eredità dei caratteri. In altri campi, le altre scienze non arrivano mai. La statistica sfrutta allora terreni che altrimenti non sarebbero sfruttati. Talvolta si tratta – è vero – di terreni di minor rendimento, e s'intende, pertanto, come, in certi argomenti, quali s'incontrano, ad esempio, nell'economia e nella biologia, i risultati della statistica appaiano di secondaria importanza di fronte a quelli che si possono raggiungere con altri metodi. Ma altre volte è alla statistica che le costruzioni scientifiche devono le rifiniture che le rendono apprezzate, come è avvenuto in molteplici branche delle scienze fisiche, biologiche e sociali. Senza la statistica, esse sarebbero spesso costrette a mantenere le loro conclusioni nel campo qualitativo o accontentarsi di approssimazioni insufficienti.

Vediamo un altro esempio in cui la statistica sembra aver condotto a risultati contraddittori.

Classificando in uno stesso Stato le varie circoscrizioni territoriali secondo gli indici della prosperità economica (reddito, ricchezza, risparmio), si constata che la natalità risulta più bassa nelle zone più prospere. Anche il confronto tra Stati ricchi e Stati poveri mostra, generalmente, una più bassa natalità negli Stati più ricchi. E la stessa popolazione presenta, di solito, una natalità più alta all'inizio della sua storia che negli stadi successivi, quando la sua organizzazione economica si è sviluppata ed è cresciuta la ricchezza accumulata. Infine, in molti paesi, le classi sociali elevate presentano una natalità più bassa del complesso della popolazione. Se ne trae la conclusione che il benessere economico costituisce un fattore sfavorevole alla riproduzione.

Il confronto tra il numero dei nati in mesi e in anni successivi ha, d'altra parte, mostrato che, nei periodi di crisi economica, la natalità si contrae e in quelli di prosperità si espande. In una medesima classe sociale, le famiglie con mezzi economici più larghi risultano, in generale, più numerose. Ciò ha fatto pervenire altri ricercatori alla opposta conclusione che il benessere economico favorisce la riproduzione.

Divergenze consimili porgono buon gioco ai detrattori della statistica.

Un esame accurato mostra però che, in realtà, esse sono in parte apparenti.

Non è certo nuova agli studiosi di fatti biologici la constatazione che le reazioni dell'organismo umano all'azione momentanea di un fattore possono risultare opposte a quelle che provoca una sua azione prolungata. Questa è anzi la norma per le reazioni provocate dagli eccitanti. Ciò può spiegare, almeno fino ad un certo punto, come, ad una variazione delle condizioni economiche, la natalità reagisca diversamente quando si tratta di oscillazioni mensili od annuali o, invece, di modificazioni evolutive a lunga scadenza o di differenze permanenti che si verificano secondo il territorio o la classe sociale.

È pure di osservazione comune che, per moltissimi fenomeni fisici, biologici e sociali, gli effetti non risultano proporzionali alla loro intensità e talvolta, secondo il livello iniziale o secondo l'intensità delle variazioni, differiscono anche per verso. Per ciò che riguarda la temperatura, l'alimentazione, l'esercizio fisico, noi ne facciamo quotidiana esperienza. Il principio della sovrapposizione degli effetti, largamente ammesso nelle scienze fisiche e che, in statistica, sta alla base – come ricordavo – del metodo delle correlazioni parziali e, più in generale, del metodo dei residui, è generalmente ammissibile, come prima approssimazione, ma solo entro brevi intervalli della variabile. Non vi è, pertanto, da meravigliarsi se le variazioni della riproduzione umana si presentano talvolta diverse per intensità e per verso al variare del benessere economico. Ciò può, ad esempio, spiegare come, mentre all'elevarsi della classe sociale, la natalità diminuisce, tra le classi basse, il numero dei figli risulta meno elevato negli strati economici infimi.

Anche in questi casi, le contraddizioni – chi ben guardi – sorgono dal fatto che, in realtà, le espressioni «benessere economico» e «miglioramento economico» sono espressioni vaghe, che hanno bisogno di venire precisate per ciò che riguarda sia la durata del benessere, sia il suo livello iniziale e l'intensità delle sue variazioni. Inconvenienti consimili possono ovviamente verificarsi, e si verificano, in ogni campo della ricerca scientifica e non sono una peculiarità della statistica.

D'altra parte, però, le considerazioni esposte non spiegano completamente le accennate contraddizioni. Vi è in queste, evidentemente, un altro fattore.

Quando noi confrontiamo dati relativi a mesi, anni, circoscrizioni territoriali, stadi evolutivi, categorie di reddito, classi sociali, e, nella stessa classe sociale, gradi gerarchici, che differiscono per ciò che concerne il benessere economico, ci riferiamo a modalità che in realtà non differiscono solo per questo carattere. Mesi ed anni di crisi differiscono, spesso, da mesi ed anni di prosperità, anche per ciò che riguarda le condizioni climatiche e, generalmente, per ciò che riguarda la mortalità. I contribuenti delle varie categorie di reddito, del pari che gli appartenenti ai vari gradi gerarchici di una stessa classe sociale, differiscono per ciò che riguarda l'età. Circoscrizioni territoriali, ricche e povere, differiscono generalmente per ciò che riguarda la densità della popolazione, il grado di urbanesimo, le occupazioni professionali. Le diverse classi sociali e le popolazioni che si trovano ad un diverso stadio evolutivo differiscono sotto molteplici aspetti psicologici, per esempio per ciò che riguarda la conoscenza dei mezzi contraccettivi. Fino a che punto le osservate relazioni tra variazioni della natalità e variazioni del benessere economico secondo mesi, anni, circoscrizioni territoriali, classi sociali, gradi gerarchici, categorie di reddito, stadi evolutivi, dipendono dalle variazioni del benessere economico, e fino a che punto, invece, da altri fattori differenziali che talvolta accentuano, talvolta attenuano o neutralizzano, talvolta dominano l'influenza delle variazioni del benessere economico?

Difficoltà consimili derivano dal fatto che i fenomeni che si esaminano dipendono da un complesso di variabili numerose e difficilmente isolabili in modo completo; se non sono assenti in nessun campo delle ricerche scientifiche, esse sono più frequenti, più gravi e meno facilmente eliminabili nel campo della ricerca statistica, particolarmente quando questa si riferisce ai fenomeni collettivi sociali o biologici relativi alla specie umana, che non si possono sottoporre a sperimentazione.

Bisogna quindi riconoscere che la statistica, per la stessa natura dei fenomeni che tratta, è più esposta ad incorrere in contraddizioni di quanto non lo siano le altre discipline. Ciò impone una particolare cautela nelle applicazioni, ma non infirma l'esattezza dei metodi.

Io ho già accennato però, che, di fronte alle contraddizioni che alla statistica si rimproverano e alle smentite che le sue conclusioni hanno talvolta subito, noi non possiamo esimerci dall'esaminare se certi metodi generalmente usati dagli statistici siano veramente corretti. E qui entro in un campo più delicato, poiché si tratta di scuotere la fiducia nella portata di procedimenti a cui sono legati nomi tra i più autorevoli della nostra disciplina e che vengono, dalla generalità degli statistici, riguardati come l'ultima e più sicura espressione del progresso scientifico.

Sono dubbi lungamente maturati. Al qual proposito richiamo la vostra attenzione sulla seguente circostanza. Mentre la generalità degli statistici non calcola una costante, caratteristica di una serie statistica (media, rapporto, indice di variabilità, indice di concordanza o di connessione, e via dicendo), senza determinarne l'errore probabile o l'errore quadratico medio o la probabilità che il suo errore raggiunga una certa intensità, e crederebbe anzi di poter essere accusata di ignorare i più sicuri procedimenti della metodologia statistica se non lo facesse, vi è una minoranza che non calcola tale errore se non in casi particolari o accompagnandolo con restrizioni e riserve. Il più curioso si è che, in questa minoranza, vi sono per lo meno alcuni specialisti che allo studio della dispersione hanno dedicato studi approfonditi. Tra questi miscredenti io mi pongo incondizionatamente. Tra i colleghi che serbarono un'attitudine analoga, desidero ricordare il compianto Lucien March, del quale so che analogo era il punto di vista. È un peccato che Lucien March sia passato senza averne, per quanto mi consta, dettagliatamente esposto le motivazioni. Ma io lo comprendo: trattasi di argomenti così sottili in cui in realtà è difficile dare al proprio pensiero una forma definitiva e prendersi la responsabilità, in tali condizioni, di scuotere ciò che è generalmente riguardato come il fondamento dell'attendibilità dei risultati statistici. Io stesso sono stato a lungo incerto prima di manifestare questi dubbi, sorti sin dall'inizio dei miei studi e progressivamente consolidatisi; ma, al postutto, ritengo che, non facendolo, mancherei al mio dovere di maestro verso i discepoli e colleghi che mi onorano della loro fiducia e talvolta a me ricorrono per avere direttive nelle loro ricerche.

Un primo elemento di incertezza nel calcolo dell'errore che si può attribuire con una certa probabilità ad una costante statistica o della probabilità che questa sia affetta da un certo errore, è generalmente riconosciuto: esso deriva dal fatto che, in tale calcolo, si sostituisce al valore esatto – ignoto – della costante, il valore approssimativo – più o meno inesatto – che si conosce.

Dal punto di vista strettamente logico, il circolo vizioso è evidente, ché si ammette che non sia affetta da errore una quantità per cui si vuole determinare proprio la probabilità di un certo errore; ma, dal punto di vista pratico, si ritiene che la sostituzione non abbia importanza quando il numero dei casi è grande, e di essa, d'altra parte, si è cercato, per le principali costanti, di tener conto quando il numero dei casi è piccolo. Credo che qualche osservazione ci sarebbe ancora da fare a questo proposito.

Osservazione che credo importante, già fatta altre volte a proposito dell'inversione del teorema di Bernoulli[2], ma di portata generale, è

[2] Cfr. le nostre osservazioni alle pp. 382-386 dell'articolo *Sulle vaccinazioni antitifiche nell'esercito italiano durante la guerra*, preparato in collaborazione col dott. L. De Berardinis e pubblicato in «Metron», Vol. III, n. 3-4, I-II-1924.

che, nel sostituire al valore esatto il valore osservato della costante, si suppone che questo possa divergere da quello indifferentemente in un senso piuttosto che in un altro, e quindi che, indipendentemente dalla conoscenza del valore osservato della costante, non si abbia alcuna nozione sul suo valore effettivo. Quando ciò non avvenga – e generalmente ciò non avviene – si potrà tutt'al più determinare solo un limite, secondo i casi inferiore o superiore, della probabilità che il valore osservato della costante sia affetto da un certo errore accidentale.

In molte applicazioni, oltre a sostituire al valore vero il valore osservato della costante, si sostituisce anche, al valore teorico, il valore empirico del suo indice di variabilità.

Questa sostituzione non è necessaria per le medie di grandezze intensive, che esprimono la frequenza dei fenomeni, poiché, se p è la frequenza media in questione, il valore dello scostamento quadratico medio si calcola, com'è noto, mediante la formula $\sqrt{p\,(1-p)/n}$; ma è generalmente necessaria per le medie di grandezze estensive (quali la statura, il reddito, ecc.), che esprimono l'ammontare dei fenomeni. Di queste grandezze, si conosce, invero, solo lo scostamento quadratico medio empirico, nel quale entra spesso una componente sistematica, che generalmente non si elimina e spesso, anzi, sarebbe impossibile eliminare. Ciò tende a dare una misura errata per eccesso della probabilità che la media sia affetta da un errore superiore ad un dato limite.

Si sia determinata, ad esempio, la statura media degli n coscritti di un paese, e si voglia determinare la probabilità che essa sia affetta da un errore accidentale di intensità superiore ad x. Ci si basa, per farlo, sulla proprietà che, nella curva degli errori accidentali, vi è una relazione costante tra la probabilità di un errore di intensità superiore ad x e l'errore quadratico medio, e, come errore quadratico medio, si assume generalmente lo scostamento quadratico medio σ' delle stature rilevate negli n coscritti. Senonché questo scostamento quadratico medio non rappresenta l'errore quadratico medio accidentale, poiché le stature osservate, oltre che per cause accidentali, possono differire, ed effettivamente differiscono, per cause sistematiche, quali la classe sociale, la professione, il reddito, i fattori ereditari, ecc. In realtà, noi dovremmo basarci sullo scostamento quadratico medio σ delle stature di individui che non fossero soggetti a simili differenze sistematiche, scostamento quadratico medio che costituirebbe la componente accidentale di σ'. Solo basandoci su σ, noi potremmo determinare la probabilità che cerchiamo, con cui la statura media degli n individui differisce, per una quantità superiore ad x, dalla statura media che si otterrebbe per un numero infinito di individui, i quali si discostassero dagli individui osservati per pure differenze accidentali. Basandoci invece su σ', noi otteniamo una probabilità che ha ancora un senso, ma un senso diverso.

Otteniamo la probabilità con cui la statura media degli n individui osservati differisce, per una quantità superiore ad x, dalla statura media che si otterrebbe per un numero infinito di individui, i quali differissero dagli individui osservati, non solo per differenze accidentali, ma anche per differenze sistematiche della stessa importanza di quelle per cui differiscono tra loro gli n individui osservati.

Per rendersi conto della diversità dei risultati, a cui si perviene basandosi su σ' o su σ, conviene ricorrere ad un esempio relativo a grandezze intensive, in cui σ' si è determinato empiricamente e σ si può determinare teoricamente.

Prendiamo, per esempio, i rapporti dei sessi, relativi alle 10690 famiglie che hanno denunciato un dodicesimo figlio in Sassonia nel decennio 1876-85. Il rapporto dei maschi alle nascite verificatesi in tali famiglie[3] risulta $p_{12} = 0{,}51683$. Si domanda quale è la probabilità P_a che p_{12} sia affetto da un errore accidentale uguale o superiore a $\pm 0{,}001$.

Lo scostamento quadratico medio effettivo dei 10690 rapporti dei maschi alle nascite nelle famiglie considerate risulta $\sigma' = 0{,}1523$; la sua componente accidentale, in base alla formula $\sqrt{p\,(1-p)/n}$, risulta $\sigma = 0{,}1443$. In base a σ, l'errore quadratico medio di p_{12} risulta $= \dfrac{0{,}1443}{\sqrt{10690}} = 0{,}001395$, da cui si ricava $P_a = 0{,}473$, che può ritenersi il

[3] *Il sesso dal punto di vista statistico*, «Biblioteca del Metron», Istituto di Statistica della R. Università di Roma, 1908, p. 384, e *Considerazioni sulle probabilità a posteriori e applicazioni al rapporto dei sessi nelle nascite umane*, in «Studi economico-giuridici pubblicati a cura della Facoltà di Giurisprudenza della R. Università di Cagliari», Anno III, Volume III, p. 111.

valore corretto. In base a σ', l'errore quadratico medio di p_{12} sarebbe risultato $= \frac{0,1523}{\sqrt{10690}} = 0,001473$, da cui si sarebbe ricavato $P'_a =$ 0,496, che in realtà dà la probabilità che il valore osservato di p_{12} differisca per almeno ±0,001 dal valore che si sarebbe verificato qualora, in luogo delle 10690 famiglie osservate, si fosse osservato un numero infinitamente grande di famiglie che, rispetto alla tendenza a produrre i due sessi, differissero, dalle osservate, non solo accidentalmente, ma anche sistematicamente nella stessa misura in cui le osservate differiscono tra di loro.

In questo caso, la differenza tra P_a e P'_a è poco rilevante, perché la componente sistematica della variabilità dei rapporti dei sessi nelle singole famiglie è piccola; ma ben altrimenti notevole risulta la differenza per altri fenomeni, per cui la componente sistematica è forte. Consideriamo, ad esempio, la mortalità infantile al di sotto di un anno, studiata dal Lexis[4]. Nel Belgio, per i 19 anni 1847-65, il coefficiente medio di mortalità maschile sotto l'anno risultava $p = 0,1711$; per lo scostamento probabile effettivo dei 19 coefficienti annuali di mortalità, il Lexis trovava un valore uguale a 0,008315, al quale contrapponeva il valore dello scostamento probabile teorico uguale a 0,00095. A questi corrisponde uno scostamento quadratico medio effettivo $\sigma' =$ 0,01236 e rispettivamente uno scostamento quadratico medio teorico $\sigma = 0,00141$. In base a σ, l'errore quadratico medio di p risulterebbe uguale a $\frac{0,00141}{\sqrt{19}} = 0,000323$, da cui si ricava la probabilità $P_a =$ 0,002 che il valorevero di p non sia compreso nei limiti 0,1721 e 0,1701. In base a σ', l'errore quadratico medio di p sarebbe stato invece uguale a $\frac{0,01236}{\sqrt{19}} = 0,00284$, da cui si sarebbe ricavata una probabilità P'_a uguale a 0,724 (362 volte maggiore di P_a) che p non fosse compreso fra codesti limiti. In realtà, P'_a dà la probabilità che il valore di p non resti compreso entro i detti limiti in un numero infinito di nati, la cui probabilità di morte differisca solo accidentalmente dalla

[4] W. Lexis, *Zur Theorie der Massenerscheinungen in der menschlichen Gesellschaft*, Freiburg B.,Wagner, 1877, pp. 79-81.

probabilità di morte dei nati nei 19 anni considerati, mentre P_a dà la probabilità che il valore di p non resti compreso dentro i detti limiti in un numero infinito di nati la cui probabilità di morte differisca, oltre che accidentalmente, anche sistematicamente, da quella dei nati nel periodo 1847-65, come differiscono anche sistematicamente fra loro i coefficienti annuali di mortalità che si sono verificati nel periodo suddetto.

In conclusione, in base allo scostamento quadratico medio empirico σ', altro non si può calcolare se non un limite superiore della probabilità che l'errore accidentale di uno scostamento raggiunga od ecceda un certo limite. Ciò è certo meglio di niente; senonché il limite superiore può essere così lontano dal valore vero che l'utilità pratica della sua determinazione riesce molto limitata.

Prescindiamo tuttavia dalle anzidette obbiezioni e ammettiamo di conoscere la probabilità P_a che l'errore accidentale di una data costante ecceda un certo limite e quindi la probabilità $1 - P_a$ che esso resti entro tale limite. Poter dire che il valore vero della costante cada, con una certa probabilità, in un certo intorno sarebbe già qualche cosa. Ma potremo noi anche dire che, se l'errore della costante in questione eccede tale limite, esso ha la probabilità P_a di essere puramente accidentale e quindi la probabilità $1 - P_a$ di essere significativo?

È quanto generalmente si fa. Vi è una probabilità contro 22 che l'intensità dell'errore accidentale ecceda il doppio, e una contro 370 che ecceda il triplo dell'errore quadratico medio; quindi si conclude: un errore che ecceda il doppio dell'errore quadratico medio ha 1 su 22 probabilità di essere accidentale e 21 contro 1 di essere significativo, e un errore che ecceda il triplo dell'errore quadratico medio ha 1 su 370 probabilità di essere accidentale e quindi 369 contro 1 di essere significativo[5].

[5] Citerò, per tutti, due trattati di statistica meritatamente diffusissimi: quelli di G.U. Yule e di R.A. Fisher. Si legge nel primo: «...the great majority of (the) members (of the binomial distribution) lie within a range $\pm 3\sigma$ on each side of the mean, i.e. of $\pm 3\sqrt{npq}$ on each side of the value np. If the distribution is exactly normal, 0,9973 of the curve lies within this range. We can therefore say that, if a particular sample gives a value of p outside this range, the deviation from the expected value is most unlikely to have arisen from fluctuations of simple sampling. If n is large the

Sta qui il fondamento di tutta la teoria dei così detti «testi di significatività» diretti a misurare il grado di attendibilità che si deve attribuire a un risultato statistico o a una differenza fra due risultati statistici. È la teoria che ha segnato, nel pensiero della maggior parte degli statistici moderni, il coronamento dell'edificio della metodologia statistica.

Ora, o io m'inganno – ma proprio mi pare di non ingannarmi – oppure si commette in tale procedimento un errore logico grossolano, scambiando quello che in statistica si chiama un rapporto di derivazione con quello che si chiama un rapporto di composizione.

Un ragionamento analogo in altro campo varrà a mettere in luce la fallacia del procedimento. Supponiamo di aver stabilito che, dei bari che circolano in Roma, l'80% capita, ad una data ora, in un certo pubblico locale, per esempio da Aragno. Voi andate a quell'ora da Aragno e alla prima persona in cui vi imbattete dite, in nome del calcolo delle probabilità: vi è l'80% di probabilità che voi siate un baro. L'assurdo di tale affermazione è evidente. Per calcolare la probabilità che una persona incontrata a caso da Aragno, a quell'ora, sia un baro, voi dovreste conoscere non solo la probabilità (che abbiamo supposto dell'80%) che in quell'ora un baro si trovi da Aragno, ma anche la probabilità (che possiamo supporre, ad esempio, del 2%) che a quell'ora ci si trovi una persona che non sia baro, e, infine, il rapporto di frequenza (che possiamo supporre di 1 contro 10000) dei bari sulla popolazione di Roma. La probabilità che a quell'ora da Aragno una persona in cui ci si imbatte sia un baro, si potrebbe allora determinare, e sarebbe di 80 contro 20080 e non di 80 contro 20. Una differenza apprezzabile, come vedete! Se non si conosce la probabilità, da una parte, che una persona che non sia un baro si trovi a quell'ora da

chances are about three in a thousand that it arose in that way» (*An Introduction to the Theory of Statistics* by G.U. Yule and M.G. Kendall, Eleventh edition, London, Griffin, 1937, p. 352). E, più in generale il Fisher scrive: «Twice the standard deviation is exceeded only about one in 22 trials, thrice the standard deviation only once in 370 trials, while... to exceed the standard deviation sixfold would need a thousand milion trials... It is convenient to take (the) point (for which $P = 0{,}05$ or 1 in 20) as a limit in judging whether a deviation is to be considered significant or not. Deviations exceeding twice the standard deviation are thus firmally regarded as significant. Using this criterion, we would be led to follow up a false indication only once in 22 trials, even if the statistics are the only guide available» (*Statistical Methods for Research Workers*, Fourth edition, Oliver and Boyd, Edinburgh, 1932, p. 45).

Aragno e, dall'altra, la frequenza dei bari in Roma, la probabilità che un baro si trovi da Aragno non ci dice nulla sulla probabilità che una persona che si trova da Aragno sia un baro.

Identico è il ragionamento per la probabilità π_a che un errore di una certa intensità (o superiore ad una certa intensità) sia accidentale. Per determinarla, è necessario che noi conosciamo, non solo la probabilità P_a che un errore accidentale abbia quella data intensità (o sia superiore all'intensità considerata), ma è anche necessario che conosciamo la probabilità p_s che un errore non sia accidentale, e, ad un tempo, la probabilità P_s che un errore, che non sia puramente accidentale, presenti l'intensità considerata (o sia superiore all'intensità considerata). In tal caso, la probabilità che l'errore avente l'intensità in questione (o un'intenità superiore al limite considerato) sia accidentale, si ricava dalla formula $\pi = \dfrac{P_a}{P_a + fP_s}$ dove è $f = \dfrac{p_s}{1 - p_s}$.

Ma se noi non abbiamo tali conoscenze – e, in pratica, salvo casi particolari, non le abbiamo – è assurdo pretendere di calcolare π_a ed è un grave errore scambiare P_a, rapporto di derivazione, per π_a, rapporto di composizione. Il fatto che generalmente lo si fa sulla scorta dell'insegnamento, concorde e ripetuto, dei più accreditati cultori di calcolo della probabilità e di statistica, sta ad attestare, da una parte, la possibilità che anche le menti più elette cadano in equivoci grossolani e, dall'altra, la terrificante pedissequità della mente umana nell'applicare senza critica gli insegnamenti ricevuti.

E bisogna proprio dire che la mente umana sia talvolta divisa in compartimenti stagni. Poiché non è naturalmente da credere che tutti coloro che hanno introdotto o consigliato o applicato tali metodi ignorassero la differenza che passa tra i rapporti di derivazione e i rapporti di composizione. Chi sa quanti, anzi, avranno richiamato l'attenzione, nei loro scritti e nel loro insegnamento, su di essi, mettendo in guardia, ad esempio, contro l'errore che si commetterebbe deducendo la probabilità di morte di un bambino dalla frequenza con cui i bambini figurano tra i morti! Eppure poche pagine o poche lezioni dopo, se non pure nella stessa pagina o nella stessa lezione, essi commettevano, nel senso inverso, proprio il medesimo errore!

Sarà bene illustrare la diversità fra P_a e π_a mediante un esempio concreto, che cercherò di costruire in modo che restino, per quanto possibile, eliminate le difficoltà prospettate ai precedenti paragrafi.

Supponiamo di scegliere a caso, in una vasta tenuta, due appezzamenti di terreno e di determinare il rendimento medio di ciascuno durante n anni, diversi per i due appezzamenti. Tra il rendimento medio dei due appezzamenti si riscontri una certa differenza. Si pone il problema di determinare la probabilità che tale differenza sia significativa di una diversa fertilità del terreno e non dipenda, invece, da fattori variabili da anno ad anno, che, rispetto alla fertilità del terreno, si possono considerare come accidentali.

Supponiamo che, sulla base della conoscenza dei rendimenti di ciascun appezzamento nei singoli anni, sia possibile determinare la variabilità dovuta ai fattori accidentali rispetto alla fertilità del terreno e quindi calcolare la probabilità P_a che la differenza accidentale tra i rendimenti medi in n anni dei due appezzamenti raggiunga la intensità osservata. Sia $P_a = 0{,}9$. Questo valore di P_a ci dice che sarebbe da attendersi 9 volte su 10 che la differenza tra i rendimenti medi di n anni, relativi allo stesso appezzamento di terreno (che si suppone mantenuto nelle stesse condizioni di fertilità), uguagli o superi la intensità in questione.

La quantità $P_a = 0{,}9$ esprime dunque la probabile relazione di frequenza tra due categorie di differenze (le une inferiori e le altre superiori ad un dato limite) relative allo stesso appezzamento e accidentali rispetto alla fertilità del terreno. Essa non è per nulla atta ad esprimere la probabile relazione di frequenza tra due categorie di differenze (le une dipendenti dalla fertilità del terreno e le altre accidentali rispetto alla fertilità stessa), relative a due appezzamenti diversi, scelti a caso ed entrambe superiori al limite considerato. S'intende infatti che, a parità di P_a, questa seconda relazione deve risultare differente a seconda che è più o meno probabile che i due appezzamenti scelti a caso abbiano fertilità diversa, a seconda, cioè, che è più o meno alto il valore di p_s e quindi quello di $f = \frac{p_s}{1 - p_s}$. Si intende pure che, a parità di P_a e di f, la seconda relazione deve risultare diversa a seconda che è più o meno probabile che la differenza di fertilità, eventualmente esistente fra i due appezzamenti considerati, uguagli o superi un dato limite, a seconda, cioè, che sia più o meno alto il valore di P_s. La quantità π_a, che esprime la seconda relazione, non dipende dunque solo da P_a, ma anche da f e da P_s.

Ma, in casi siffatti, non potremo proprio dire nulla sulla maggiore o minore probabilità che un certo errore o una certa differenza sia accidentale o invece significativa?

Se noi non conosciamo la frequenza dei bari in Roma e la probabilità che una persona che non sia un baro si trovi a quell'ora da Aragno, tutto quello che potremo dire (fondandoci sulla cognizione che i bari frequentano i locali pubblici più dei non bari) è che vi è maggiore probabilità di trovare un baro in una persona che si incontra a caso da Aragno che in una persona che si incontra a caso fuori Aragno.

Se noi non conosciamo, come generalmente non conosciamo, la probabilità che un errore sia puramente accidentale o, invece, sia in parte sistematico, e, ad un tempo, la curva di probabilità degli errori in parte sistematici, la sola cosa che possiamo dire (fondandoci sulla nozione che, in mancanza di compensazione tra componente accidentale e componente sistematica dell'errore, gli errori puramente accidentali sono in generale meno intensi di quelli che hanno anche una componente sistematica) è che gli errori più piccoli saranno puramente accidentali più di frequente degli errori più grandi. Nell'esempio, fatto sopra, dei due appezzamenti, potremo dire che, più la differenza tra i loro rendimenti medi è elevata, e più è probabile che la fertilità dei loro terreni sia diversa. È ben poco davvero; e molti troveranno che non vi era proprio bisogno, per arrivare a tale ovvia conclusione, di mettere in moto il calcolo delle probabilità.

E, in ogni modo, è chiaro che non vi è bisogno, per arrivarvi, di conoscere il valore di P_a. Ma io, in verità, non vedo che cosa si possa dire di più. Questa stessa conclusione, anzi, non è accettabile se non con certe restrizioni, di cui diremo più innanzi.

Ma la teoria della dispersione – si domanderà – non ci dice nulla sulla questione? No – si risponde: essa non ci dice proprio nulla. Data, non una sola costante o indice o dato statistico, ma una intera serie di costanti o indici o dati, la teoria della dispersione ci insegna a calcolare se queste quantità differiscono tra di loro più o meno di quanto dovrebbero differire per effetto di puri errori accidentali e quindi ci impegna a scindere, nella dispersione complessiva, la parte accidentale dalla parte sistematica; ma essa non ci dice quale è la probabilità che una differenza sia anche sistematica, né quale è la distribuzione delle intensità delle differenze che sono anche sistematiche, cosicché non ci permette di dedurre quale è la probabilità che una tra le differenze riscontrate abbia carattere puramente accidentale.

Nell'esempio fatto sopra, sarà certo interessante conoscere, per un numero adeguato di anni, il rendimento di tutti gli appezzamenti della tenuta e confrontarne l'indice di dispersione effettivo con quello che sarebbe stato da attendersi se la fertilità del terreno fosse la stessa per tutti gli appezzamenti. Tale confronto ci darà la misura della eterogeneità della tenuta, per ciò che ha riguardo alla fertilità del terreno. Ma una certa divergenza fra l'indice di dispersione effettivo e il teorico potrà corrispondere a situazioni diversissime. Potrà darsi che essa dipenda dalla circostanza che una metà degli appezzamenti è alquanto più fertile e una metà alquanto meno fertile e potrà darsi, invece, che dipenda dal fatto che gli appezzamenti hanno, nella grande maggioranza, la stessa fertilità, salvo alcuni che sono quasi sterili. È chiaro che la probabilità p_s che due appezzamenti scelti a caso non abbiano la stessa fertilità sarebbe molto differente nei due casi; ed è pure chiaro che del pari differente sarebbe la probabilità P_s che, se i due appezzamenti scelti a caso non hanno la stessa fertilità, il loro rendimento medio in alcuni anni differisca per più di una data quantità.

La teoria della dispersione è dunque fondata e fruttifera. Ma non bisogna domandarle più di quanto essa possa dare.

Torniamo alla consueta determinazione dell'errore quadratico medio (o dell'errore probabile) delle costanti statistiche per mettere in guardia contro un altro pericolo a cui essa espone. Non solo, invero, essa non è in grado di dirci quale probabilità ha un certo errore di essere accidentale e quale di essere significativo, ma essa rischia di farci prendere per significativi errori che sono puramente accidentali.

Supponiamo di aver determinato le medie di una cinquantina di caratteri di due popolazioni, oppure le medie di un carattere nelle 50 circoscrizioni territoriali di uno Stato. Troviamo che la differenza tra le medie ottenute per due caratteri A e B, nelle due popolazioni, oppure la differenza tra la media generale dello Stato e le medie dei caratteri nelle due circoscrizioni α e β eccede il doppio dello scostamento quadratico medio, e quindi concludiamo, secondo quella che viene riguardata la buona norma del metodo statistico, che si tratta di differenze che si possono riguardare come significative. L'antropologo o il demografo o il sociologo, secondo che si tratti di caratteri antropologici, demografici o sociali, sarà così autorizzato a ricercare le cause sistematiche di tali differenze in circostanze razziali o economiche o professionali o di altro genere. Senonché 2 volte su 50 (anzi un po' più di 2 volte, ammesso che la distribuzione delle medie sia normale) è da attendersi che le differenze superino l'errore quadratico medio per puro effetto del caso. Se noi consideriamo, non le medie dei singoli caratteri o delle singole circoscrizioni isolatamente presi, come si fa nella consueta determinazione dell'errore probabile delle costanti, ma le intere serie di tali medie, come si fa nella teoria della dispersione, dovremo correttamente concludere che le differenze riscontrate si possono attribuire all'effetto del caso e sarebbe inseguire un fantasma il correre dietro alle cause sistematiche di tali differenze.

«In fondo» – si dirà – «è questa una verità lapalissiana. Quando ci si comporta in base ad una probabilità, si corre inevitabilmente un rischio. Se noi riguardiamo come accidentale una differenza che sappiamo verificarsi accidentalmente solo 48 volte su 50, sappiamo già che ci esponiamo a sbagliare 2 volte su 50». Ciò è esatto; ma sono precisamente le verità lapalissiane che talvolta si dimenticano più facilmente. D'altra parte, se nella vita, che è azione, è giustificato di assumersi un rischio, ché il non assumerlo condurrebbe all'inerzia, nella scienza, che è ricerca della verità, può apparire discutibile se sia da preferirsi una conclusione affetta da rischio a un prudente riserbo.

Quante differenze tra i risultati statistici, ritenute significative, furono contraddette da ricerche successive! La causa può farsi risalire molte volte al rischio inevitabilmente insito nell'applicazione degli errori probabili alle singole costanti statistiche.

Il pericolo che, applicando a singoli casi il calcolo della probabilità di un errore si sia, senza fondamento, indotti a negare il carattere accidentale della differenza, quando questa supera un certo multiplo dell'errore quadratico medio, o, in generale, un valore che sarebbe da attendersi solo eccezionalmente, è molto più grave quando l'esempio che si considera non venga scelto a caso, ma in dipendenza del suo stesso carattere eccezionale. È ciò che si fa generalmente quando si studia l'eredità sulla base della casistica. Per esempio, si cercano le famiglie in cui l'eredità di una certa malattia si manifesta in modo più spiccato, la malattia presentandosi, oltre che nei padri, in tutti i figli, e poi si dice: «Voi non potete negare che questa malattia sia ereditaria. Considerate, invero, che, nel complesso della popolazione, essa si presenta, poniamo, nell'1% dei casi; la probabilità che essa si presenti in tutti gli n figli di una famiglia, come avviene nelle nostre osservazioni, è pertanto di $(1/100)^n$; in una famiglia, ad esempio con quattro figli, vi è da scommettere 100 milioni contro uno che non si tratta di un caso».

Applichiamo questo criterio al fine di decidere se un certo giuoco si svolge sinceramente. Passando l'altro giorno davanti a un botteghino del lotto, notavo che, in una delle ruote, erano sortiti, l'uno dopo l'altro, al secondo e al terzo estratto, due numeri successivi: mi pare fossero il 63 e il 64. Vediamo come, in simili casi, ragionerebbe uno dei tanti raccoglitori di casistiche che si dilettano di sottometterle alla prova del calcolo delle probabilità. «Vi era» – egli direbbe – «una probabilità di 1/89 che, nel secondo estratto, uscisse il 63 e di 1/88 che, nel terzo, uscisse il 64: e quindi una probabilità di $1/89 \times 1/88$ (vale a dire di circa 1/7800) che i due numeri uscissero di fila. Ci sarebbe quindi da scommettere 7800 contro 1 che non si tratta di un effetto del caso. Prima però di sospettare qualcuno, andiamo a vedere i risultati di altre estrazioni avvenute in passato». Cerca, cerca, e alfine trova. Che cosa trova? Trova che in un anno, in una ruota, in una estrazione, i numeri usciti sono stati 1, 11, 21, 31, 41. «Qui l'imbroglio è evidente» – egli pensa «vi è da scommettere 1 contro 5 miliardi che non si tratta dell'effetto del caso, perché vi è solo una probabilità su oltre 5 miliardi che tale risultato si verifichi». Il nostro probabilista più non esita, ed eccolo a reclamare presso il Direttore Generale del lotto perché metta in istato di accusa chi presiede alle estrazioni, disponga per una inchiesta severa, prenda le misure per assicurare la sincerità del giuoco. Il Direttore Generale del lotto, perplesso, chiama un altro probabilista, un super-probabilista se così si può dire. «Ma no» – questo risponde – «perché? La sequenza 63, 64 ha – è vero – una probabilità di verificarsi su 7800 e la sequenza 1, 11, 21, 31, 41, ha sì una probabilità di verificarci su 5 miliardi, ammesso che tutti i 90 numeri abbiano la stessa probabilità di sortire. Ma che per ciò? Ciò significa appunto che talvolta si possono verificare ed è dunque naturale che qualche volta si verifichino. Voi mi direte che in tutte le ruote del Regno, dacché c'è il giuoco del lotto, non si sono effettuate ancora 5 miliardi di estrazioni. Ma tenete presente che qualunque sequenza di due numeri diversi, non solo quella dei numeri 63, 64 ha una probabilità su 7800 e che analogamente qualunque sequenza di cinque numeri diversi, non solo la sequenza 1, 11, 21, 31, 41, ha una probabilità su 5 miliardi di verificarsi. Una certa sequenza deve pure verificarsi. Non dovete fissare la vostra attenzione su un'estrazione speciale e su una sequenza determinata; dovete considerare tutte le estrazioni avvenute in tutte le possibili sequenze. Se trovate che la distribuzione di tutte le possibili sequenze diverge sistematicamente da quella che il calcolo delle probabilità farebbe attendere, allora ritornate e ne riparleremo». Il probabilista ha chinato la testa e se n'è andato riconoscendo il suo torto.

Così dovrebbero riconoscere il loro torto tutti i probabilisti che, scelti dalle cartelle cliniche o dalle genealogie certi casi caratteristici dell'eredità dei caratteri, hanno preteso poi di calcolare la probabilità con cui potevano escludere che si trattasse di una coincidenza casuale.

Indizio che non si tratta di coincidenza casuale, ma che il carattere veramente si eredita, sarebbe che la combinazione caratteristica considerata si verificasse più di frequente di quanto dal calcolo delle probabilità sarebbe da attendersi per puro effetto del caso. Quante volte non abbiamo sentito fare la seguente osservazione: Guardate quella coppia: 5 figli, e 5 maschi; e l'altra: 6 figli, tutte femmine. Che cosa devono esse avere in corpo per non produrre che figli di un sesso? In realtà, molte volte, esse non hanno in corpo nulla di speciale: l'uniformità del sesso dei loro figli è l'effetto del caso. Altre volte, però, non è così; vi è, in alcune coppie, una tendenza, più o meno spiccata, a produrre un sesso. In quanti casi, noi non possiamo dire, e neppure possiamo dire con quanta probabilità l'uniformità per sesso di una figliolanza sia l'effetto del caso piuttosto che di una tendenza sistematica; ma possiamo plausibilmente dire che dal 10 al 15% dello scostamento quadratico medio, che si verifica nella composizione per sesso delle famiglie, dipende da tendenze sistematiche delle coppie a produrre maschi o femmine; il resto dipende invece dal caso[6].

Quando si abbia a che fare, non con caratteri discreti (come i numeri del lotto o le combinazioni dei sessi nelle singole figliuolanze), ma con caratteri continui o che devono praticamente trattarsi come tali (quale la statura o il rapporto dei sessi nelle nascite di una intera popolazione), la determinazione della probabilità che in un caso singolo si presenti un certo errore accidentale va incontro a un'altra difficoltà ed espone ad un altro pericolo a cui difficilmente ci si sottrae. La probabilità ${}_xP_a$ che la determinazione del carattere sia affetta da un errore accidentale di intensità x, è, in tal caso, infinitesima. Si può, però, calcolare la probabilità ${}_{\geq x}P_a$ che si presenti un errore di intensità uguale o superiore a x. Volendo e non potendo calcolare ${}_xP_a$, si calcola, in sua vece, ${}_{\geq x}P_a$, che è manifestamente tutt'altra cosa. Nell'esempio dei due appezzamenti, io pure ho operato tale sostituzione – probabilmente senza che nessuno di voi l'avesse avvertito, tanto la sostituzione è consuetudinaria. Spesso poi, come si è detto, la probabilità ${}_{\geq x}P_a$ si può determinare solo approssimativamente, sostituendo, al valore reale della costante, il suo valore osservato e, al valore teorico del suo errore quadratico medio, il rispettivo valore empirico che include una componente sistematica. Generalmente, infine, si ritiene di avere in tal modo calcolato la probabilità che l'errore riscontrato x sia accidentale. Si commettono così quattro errori ad un tempo: *a)* sostituzione a π_a di ${}_xP_a$; *b)* sostituzione a ${}_xP_a$ di ${}_{\geq x}P_a$; *c)* determinazione inesatta di ${}_{\geq x}P_a$; *d)* applicazione di ${}_{\geq x}P_a$ a un caso individuale, spesso scelto in vista del suo carattere eccezionale.

Ho spiegato più sopra l'equivoco da cui nasce il primo e le deficienze dei dati disponibili che spesso rendono inevitabile il terzo errore e ho testé richiamato l'attenzione su le circostanze da cui sorge il secondo; ma è soprattutto sul quarto che intendo insistere a conclusione dei paragrafi immediatamente precedenti.

[6] Cfr. *Il sesso dal punto di vista statistico*, *op. cit.*, Cap. X: *La variabilità individuale nella tendenza a produrre i due sessi*, pp. 371 e ss.

La conclusione è che il calcolo delle probabilità non si può correttamente applicare in casi singoli, ma solo in masse di casi. All'indirizzo del calcolo delle probabilità è stato lanciato una volta un epigramma che voleva forse essere una critica feroce. «Il calcolo delle probabilità» – ha detto Adolfo Thiers – «è l'arte di precisare ciò che si ignora». La critica, che io mi sappia, frequentemente riprodotta, non ha trovato risposta. Rileviamola e rispondiamo che il calcolo delle probabilità insegna a precisare nella massa ciò che si ignora nei casi singoli. Ciò si può fare e si fa col sussidio del calcolo delle probabilità. In altri casi, ciò si fa, forse con un grado di precisione minore, anche senza il suo sussidio. Col sussidio del calcolo delle probabilità, il fisico, pur ignorando i movimenti delle molecole, precisa le proprietà risultanti dei corpi, e il demografo, pur ignorando il sesso dei singoli nascituri, precisa la probabilità che, nella massa delle nascite, i maschi eccedano in una certa misura sulle femmine, e l'attuario, pur ignorando la durata di vita delle singole persone assicurate, ne prevede la durata media. Senza il sussidio del calcolo delle probabilità, il Ministero delle Finanze, pur ignorando l'attività economica dei singoli contribuenti, prevede il gettito delle imposte sugli affari, e il Capo del Governo, pur ignorando i sentimenti individuali di ogni cittadino, prevede gli orientamenti politici della nazione. Ma, se il calcolo delle probabilità pretende insegnare alla statistica, o la statistica si illude di imparare dal calcolo delle probabilità, l'arte di precisare per i casi singoli ciò che per i casi singoli si ignora, statistica e calcolo delle probabilità si ingannano a partito e non possono divenire che fomite di errori e fonte di delusioni.

«Dopo tutto, però, qualche cosa di solido rimane» – voi direte; «rimangono le applicazioni del calcolo delle probabilità a masse di casi mediante la teoria della dispersione».

Se non che le riserve non sono finite. V'è un altro punto a cui conviene fare attenzione. Gli è che ogni massa di casi può a sua volta considerarsi come un caso singolo di ordine superiore. Un rapporto statistico o una media rappresenta un caso singolo entro la serie di rapporti o di medie a cui si applica la teoria della dispersione; ma la serie a sua volta diviene un caso singolo per l'applicazione dei metodi «χ^2» o «z».

Confrontando la distribuzione effettiva di una serie di rapporti statistici o di medie con la rispettiva distribuzione teorica, si potrà dire se i rapporti e le medie della serie presentano una dispersione supernormale, normale o subnormale, ma non si potrà determinare, mediante il metodo χ^2 o z, quale è la probabilità che la divergenza tra la distribuzione effettiva considerata e la teorica sia accidentale. Né, qualora la divergenza tra la distribuzione effettiva e la teorica sorpassi in una serie il doppio dell'errore quadratico medio, o altro limite convenzionale, si sarà autorizzati a dichiarare che i rapporti o le medie della serie in questione differiscono tra di loro sistematicamente, ciò che, d'altronde, diverrebbe manifesto se, avendo fatto applicazioni di χ^2 o di z a molte serie, si fosse riscontrato che i valori ottenuti per χ^2 o per z non eccedono il doppio dell'errore quadratico medio, a altro limite convenuto, più di frequente di quanto sarebbe da attendersi per effetto del caso.

D'altra parte, ottenuta così tutta una serie di valori di χ^2 o di z, questa serie diviene a sua volta un caso singolo ai fini di determinare se la sua dispersione è, o meno, superiore alla teorica: si potrà stabilire se la sua dispersione risulti, nel caso considerato, supernormale, normale o subnormale, ma non si potrà determinare la probabilità con cui la divergenza tra la distribuzione effettiva riscontrata e la teorica è da attribuirsi al caso.

E qui mi incorre l'obbligo di chiarire un'apparente contraddizione, in cui si potrebbe credere che io fossi caduto. Poiché, da una parte, io ho dianzi affermato che la teoria della dispersione è corretta e può condurci a determinare la componente sistematica della dispersione e, d'altra parte, ho testé mostrato come non si sia in grado di stabilire con quale probabilità la dispersione trovata possa considerarsi significativa o accidentale.

Tale questione non è tuttavia che un caso particolare di quella che abbiamo considerato sotto forma generale per tutte le costanti.

A parte, infatti, la difficoltà di determinare la probabilità che, per effetto del caso, si presenti una certa dispersione quando questa venga misurata mediante un indice che varia con continuità, e i pericoli di applicare tale probabilità a un caso singolo, dobbiamo tener presente che non basta conoscere tale probabilità per conoscere la probabilità che la dispersione riscontrata possa considerarsi come accidentale. Bisognerebbe conoscere altresì – e noi generalmente non conosciamo – la probabilità che la serie di dati considerata sia perturbata da cause anche sistematiche, e la probabilità che, essendolo, la dispersione osservata si verifichi per effetto di dette cause anche sistematiche.

In generale, possiamo dire che la statistica insegna a determinare molti indici: medie, rapporti, indici di variabilità, di concentrazione, di dispersione, di transvariazione, di dissomiglianza, di connessione, di concordanza, e via dicendo, tutti relativi a masse di casi. Ora, in quanto le masse di casi su cui si opera interessano non per se stesse, ma come campioni di masse più comprensive, ogni indice deve riguardarsi come approssimato. Tutti gli indici sono, invero, affetti da errori accidentali, che a certi fini interesserebbe eliminare; vi è poi, in molti casi, la eventualità che essi siano affetti altresì da errori sistematici. Se non che (a parte il caso che tale eventualità risulti *a priori* o *a posteriori* esclusa, nel qual caso la questione non si pone), noi siamo nell'impossibilità, salvo particolari cognizioni derivanti da altre fonti, di determinare la probabilità che la differenza riscontrata tra due indici, o tra uno di essi e un suo valore teorico, sia puramente accidentale. Di solito – come ho già detto – tutto quello che possiamo affermare è che, più la differenza è piccola, e più è probabile che essa sia accidentale, e più essa è grossa, e più è probabile che in essa entri una componente sistematica.

Anche per poter affermare ciò, è però necessario che il confronto sia fatto tra differenze relative a dati della stessa serie o di serie concernenti lo stesso fenomeno e ad ogni modo comparabili dal punto di vista che ci occupa. Così, in una serie di coefficienti annuali di mortalità, saremo autorizzati a dire che gli scarti lievi della media, che si verificano in alcuni anni, hanno carattere puramente accidentale più probabilmente degli scarti forti che si verificano in altri; ma, se, in una serie di coefficienti annuali di mortalità, lo scarto resta, un dato anno, entro i limiti dell'errore quadratico medio teorico della serie, e, in una serie di rapporti annuali dei sessi delle nascite, lo scarto, un dato anno, eccede, invece, il limite dell'errore quadratico medio teorico della serie, sarebbe arbitrario affermare che il primo scarto ha maggiore probabilità del secondo di essere puramente accidentale.

Ritorniamo, per illustrare questo punto, alla nostra formula $\pi = \dfrac{P_a}{P_a + fP_s}$. Quando si confrontano dati relativi alla stessa serie o a serie relative allo stesso fenomeno, è da ritenere che il valore di f rimanga costante e che, col crescere dell'errore, decresca P_a e, meno fortemente, decresca P_s. Perciò potremo dire che π_a, decresce col decrescere di P_a e col crescere dell'errore, senza d'altra parte, poter dire nulla, né sulla proporzione tra la diminuzione di π_a e quella di P_a, né sulla proporzione tra la diminuzione di π_a e l'aumento dell'errore, e ancor meno sull'altezza assoluta di π_a. Quando, invece, si tratta di dati reltivi a fenomeni diversi, potremo dire soltanto che P_a risulterà più basso per il dato relativo – poniamo – al fenomeno α, per cui l'errore risulta più lieve, che per quello relativo al fenomeno β. Ma, d'altra parte, f può essere, e normalmente è, diverso per i due fenomeni: né è escluso che, se P_a risulta più basso per il dato relativo ad α, anche P_s risulti per questo più basso nella stessa proporzione o in proporzione maggiore. Per modo che non si può escludere affatto che π_a risulti uguale o anche più basso per il dato relativo a β che per quello relativo ad α.

E potremo dire qualche cosa sul variare di π_a, al variare del numero delle osservazioni?

P_a, come è noto, decresce col crescere del numero delle osservazioni; precisamente, l'intensità dell'errore che ha una certa probabilità P_a di essere accidentale, varia in ragione inversa della radice quadrata del numero delle osservazioni. Ma come variano, col crescere delle osservazioni, f e P_s? Non si può fare in proposito, una affermazione di carattere generale. In alcuni casi, la componente sistematica dell'errore è costante, e quindi indipendente dal numero delle osservazioni. Ma, in altri, essa decresce col crescere del numero delle osservazioni, come quando dipende dalla inesperienza iniziale dell'esperimentatore, che si viene via via correggendo. In altri ancora, essa cresce, come quando dipende dalla progressiva stanchezza dell'esperimentatore o del soggetto in osservazione, oppure dal fatto che il fenomeno osservato (per esempio: statura di un individuo, posizione di un corpo) varia col tempo. Similmente, f può essere costante o, invece, crescere o diminuire.

Di conseguenza, nessuna affermazione di carattere generale è autorizzata sul variare di π_a col variare del numero delle osservazioni, mentre affermazioni fondate, più o meno sicure, possono farsi in casi particolari basandosi sulla conoscenza della struttura del fenomeno e delle condizioni in cui si svolge la rilevazione o l'esperimento. Per esempio, se tale conoscenza porta ad ammettere che f e P_s rimangono costanti col variare del numero delle osservazioni si potrà dire che π_a, col crescere di tale numero, diminuisce e diminuisce in ragione meno che proporzionale al crescere della radice quadrata di tale numero. Non sarà possibile però precisare – a meno che non si conoscano i valori di f e di P_s – la legge che lega la diminuzione di π_a all'aumento del numero delle osservazioni.

Queste osservazioni sapranno di forte agrume ai moltissimi che hanno fondato i risultati delle loro ricerche sui così detti «testi di significatività» e più a coloro che intorno ad essi hanno lavorato di cesello; ma in realtà, esse non possono tornare gradite a nessun cultore di statistica, poiché portano a concludere che noi non siamo in grado di misurare l'attendibilità di un risultato statistico o di una differenza fra due risultati statistici, in quanto quello o questi si considerino come rappresentativi del risultato o dei risultati che si dovrebbero ottenere in una massa di casi più vasta, salvo il caso di risultati che si sanno affetti unicamente da errori accidentali.

Risalire dai risultati ottenuti in una massa, necessariamente limitata, di casi, ai risultati che si sarebbero ottenuti in una massa di casi più vasta, o, al limite, infinita, non è certo indispensabile in tutte le indagini statistiche e a tutti gli scopi – come taluni erroneamente pretendono; ma in talune indagini o a taluni scopi è veramente indispensabile e in altri, se non indispensabile, è certo utile, al fine di ottenere la precisione delle conclusioni desiderabile. Io mi rendo conto pertanto di tutta la gravità della suddetta conclusione; ma scopo della scienza è, non di raggiungere conclusioni gradite, ma di raggiungere conclusioni fondate.

Ma ecco il nostro probabilista che ritorna in grande eccitazione. Che cosa vorrà? Voi vi ricorderete come egli avesse dovuto abbassare la testa davanti alle osservazioni del super-probabilista e, sia pur masticando amaro, riconoscere il suo torto. Ma, tenace come ogni buon scienziato, egli ci ha poi ripensato, e prima di darsi per vinto, ha voluto esaminare la dispersione delle estrazioni del lotto in cui sospettava che ci fosse un imbroglio. Calcoli laboriosi, fatti e rifatti, lo hanno portato alla conclusione che la dispersione è super-normale. Le applicazioni di χ^2 e z attestano che, per effetto del caso, una tale dispersione si verificherebbe soltanto una volta su mille. Egli ritorna ora alla carica presso il Direttore Generale del lotto, e questa volta è spalleggiato dal super-probabilista. «Vedete» – egli esclama trionfalmente, sciorinando sotto il naso del Direttore Generale il quaderno dei suoi calcoli – «io avevo ragione: neppure una volta su mille potrebbe verificarsi, per effetto del caso, la dispersione che nelle vostre estrazioni si verifica. Mille contro uno potrei scommettere che le estrazioni non sono affidate alla sorte; l'imbroglio c'è, ve lo dicevo». E il super-probabilista interviene di rincalzo: «Signor Direttore Generale, vi assicuro: io ho verificato ad uno ad uno codesti calcoli e non vi è proprio nulla da ridire. Alla luce del calcolo delle probabilità egli ha ragione. Che ci sia sotto un imbroglio, come egli dice, o che le estrazioni siano fatte in modo da favorire inconsciamente certi numeri o certe combinazioni, io non so; ma bisogna riconoscere che le estrazioni non dipendono puramente dalla sorte». Il Direttore Generale li ascoltò, rifletté un momento, poi alzò le spalle e secco secco rispose: «Io conosco come le estrazioni si svolgono e posso garantire che avvengono in regime di pura sorte», e mandò a carte quarantotto probabilista e super-probabilista, i loro calcoli e tutto il calcolo delle probabilità che conduceva a simili conclusioni.

Chi aveva ragione? Dal punto di vista logico, almeno, aveva ragione il Direttore Generale. Perché? Perché è vero bensì che il calcolo delle probabilità ci dice che vi è una probabilità su mille che per cause accidentali si verifichi la dispersione osservata, ma esso non ci dice la probabilità P_s che, verificandosi le perturbazioni sistematiche sospettate, esse darebbero luogo alla dispersione in questione (tutto ciò che noi sappiamo in proposito è che P_s è compreso tra zero e uno) e neppure ci dice la frequenza f di tali perturbazioni sistematiche di fronte alle oscillazioni puramente accidentali. In sostanza, nessun giudizio si può formulare sulla probabilità che la dispersione osservata abbia carattere accidentale senza conoscere il sistema di estrazione. Se si può escludere, come il Direttore Generale, in base alla conoscenza del sistema di estrazione, credeva di poter sicuramente escludere, che perturbazioni sistematiche si verifichino, in modo che sia $f = 0$, o la probabilità π_a che la dispersione osservata sia accidentale, non è di 1/1000, ma è di $\dfrac{1/1000}{1/1000 + 0P_s}$ vale a dire è uguale a 1. È un caso particolare della divergenza tra π_a e P_a, che già avevamo messo in luce in forma generale. Ma è un caso particolarmente significativo in quanto illustra un altro aspetto delle applicazioni del calcolo delle probabilità che non converrebbe mai dimenticare.

Quale è dunque la morale della favola? La morale della favola – o per meglio dire di quest'ultima parte della favola – è che il calcolo delle probabilità, e la statistica a mezzo del calcolo delle probabilità, anche se applicati a masse di casi, non possono mai portare a conclusioni sicure, ma solo a conclusioni probabili. Possono legittimare dei dubbi più o meno forti – e questa è certamente una funzione utile – ma non possono mai scioglierli in modo definitivo. Possono fornire non «testi di significatività», ma «elementi di sospetto». Per quanto alto sia il grado di probabilità delle loro conclusioni, di fronte a un dato sicuro della coscienza o dell'esperienza, essi devono ritirare le corna. Riconosciamolo esplicitamente e dichiariamolo francamente agli sperimentatori prima che questi ci colgano in fallo.

Non sarebbe la prima volta. Siamo stati colti in fallo flagrante nella applicazione degli schemi teorici.

* * *

Concludendo, dobbiamo riconoscere che, non solo il terreno della statistica è seminato di pericolo, ma anche che chi in esso si spinge, spesso ci cade. Effettivamente, nel fissare il titolo di questo mio discorso, io sono stato incerto tra *I pericoli della Statistica* e *I peccati della Statistica*.

Una bella donna, che spesso si avventura in paesi inesplorati e non ha obbiezioni a far comunella con ogni sorta di gente e talvolta si accompagna a giovanotti intraprendenti, è inevitabilmente esposta a pericoli e difficilmente va immune da peccato. Ora la statistica è indubbiamente una scienza molto seducente, come è provato dal fatto che ad essa si accostano e aspirano, non solo i molti che la conoscono a fondo, ma anche quelli – e sono molti di più – che la conoscono appena di vista. È una scienza che, per alcuni campi, ha una funzione di avanscoperta e, in altri, è addiritttura la sola ad avventurarsi; è una scienza che si associa senza difficoltà ad ogni altra branca dello scibile, e spesso poi gradisce o cerca la compagnia di un attraentissimo, ma pericolosissimo, soggetto, denominato calcolo delle probabilità. Pare-va un «*enfant prodige*» questo calcolo delle probabilità e quante speranze non aveva fatto sorgere nella puritana famiglia delle matematiche a cui appartiene! E non è che le speranze sieno tramontate; ma quanti grattacapi non ha esso frattanto procurato ai suoi parenti sempre inclini, del resto, come tutti i parenti, a dar la colpa delle sue malefatte alle cattive compagnie! Le cronache narrano di uno scandalo da lui provocato quando ha voluto mettere il naso nelle decisioni delle assemblee e nelle sentenze dei giudici, pretendendo di dettare norme generali per la composizione dei corpi deliberanti, e si mormora persino che abbia seriamente pensato – audacia quasi incredibile – a calcolare la probabilità che il padre Sole una mattina più non si alzasse dal talamo nuziale. Frattanto fratelli geometri e cugini statistici sono di quando in quando costretti ad arrovellarsi il cervello per spiegare certe paradossali contraddizioni del suo comportamento.

In tale compagnia, con siffatte attrattive e con tante tentazioni, non è da meravigliarsi se la condotta della statistica non sia sempre stata e non sempre sia di una rigidità a tutta prova. Certamente, col passare degli anni, la sua individualità di forma e il suo potere di autocritica si sviluppa, cosicché viene ad essere più cauta nelle iniziative, più guardinga nelle amicizie, più pratica nei fini, più coerente nelle direttive, più contegnosa nel comportamento. La natura, la varietà e la complessità delle sue manifestazioni sono però tali, e tante sono, d'altra parte, le invidie da cui è circondata in questo basso mondo, che difficilmente potremo aspettarci che la sua condotta vada mai esente da critiche. Ma, d'altra parte, anche se queste potessero essere in parte giustificate, non dovrà mai dimenticarsi, e noi certo non dimenticheremo, che molte volte nessuno potrà arrivare dove essa arriva, vedere o intravedere quello che essa vede o intravede, raccogliere i frutti, sia pure talvolta immaturi, che essa raccoglie, che nessuno potrà, in una parola, sostituirla, e, d'altra parte, che, se vi è una disciplina a cui il motto nitzschiano *gefährlich leben*[7] si applica come una necessità di vita, questa è appunto la statistica.

7 «Den glaubt er mir! – das Geheimniss, um die grösste Fruchtbarkeit und den grössten Genuss vom Dasein einzuernten, heisst: *gefährlich leben!*». Fr. Nietzsche, *Die fröliche Wissenschaft* (La gaya scienza), Leipzig, C.G. Naumann, 1900, p. 215.