Jacob Bernoulli
Ars conjectandi

Basel 1713 (postumo)
a cura di R. Haussner
Verlag Harri Deutsch, Frankfurt am Main 2002

Quarta parte

Applicazioni della precedente teoria ai rapporti civili, morali e scientifici

Cap. I

Osservazioni introduttive su certezza, probabilità, necessità e casualità delle cose.

La *certezza* di qualunque cosa si può trattare o *oggettivamente*, cioè in sé – e in questo caso non mostra altro che la reale, presente o futura, presenza di quella cosa – o *soggettivamente*, cioè in rapporto a noi – e in questo caso consiste nella misura della nostra conoscenza di quella realtà.

Tutto ciò che esiste o sorge sotto il sole – il passato, il presente e il futuro – ha in sé la massima certezza. Rispetto alle cose presenti o passate questa affermazione si chiarisce da sé, dato che quelle cose, essendo presenti o essendo state, escludono per ciò stesso la possibilità che non esistano o non siano esistite. Anche rispetto alle cose future non c'è da dubitare che saranno presenti, anche se non con l'ineluttabile necessità di una certa fatalità, come nella previsione e predeterminazione divine. Infatti, se ciò che è futuro non avvenisse sicuramente, non si capirebbe perché l'altissimo creatore dovrebbe godere della fama incondizionata di onnisciente e onnipotente. Ma su come questa certezza dell'essere futuro si concili con la casualità e l'indipendenza dalle cause efficienti, altri possono discutere. Noi non vogliamo addentrarci nella questione, la cui discussione non rientra nei nostri scopi.

La certezza soggettiva delle cose – quella relativa a noi – non è in tutti uguale, ma varia di molto in più e in meno. Le cose di cui siamo certi per rivelazione, riflessione, percezione sensoriale, esperienza, autopsia [*sic*] o per altri modi, cioè le cose per le quali non dovremmo avere dubbi sulla loro esistenza presente o futura, possiedono per noi la certezza massima e assoluta. Tutte le altre cose contengono, a seconda della nostra conoscenza, una misura incompleta di certezza, che può essere più grande o più piccola, a seconda che esistano maggiori o minori probabilità che una certa cosa sia stata, sia o sarà.

La *probabilità* è, infatti, un grado di certezza e si differenzia dalla certezza come la parte dal tutto. Quando, per esempio, indichiamo con a o con 1 la certezza piena e assoluta, composta per ipotesi da cinque probabilità o parti, di cui tre sono a favore del verificarsi di un certo evento presente o futuro, mentre le due rimanenti depongono contro, allora l'evento possiede la certezza $3/5a$ o 3/5.

Di due cose è *più probabile* quella che possiede una parte di certezza maggiore, anche se nell'uso linguistico comune viene detto realmente *probabile* ciò la cui probabilità è sensibilmente superiore a metà della certezza. Ho detto *sensibilmente*. Infatti, una cosa la cui probabilità è solo approssimativamente uguale a metà della certezza è detta *dubbia* o fluttuante. Ciò che ha 1/5 di certezza è più probabile di ciò che ne ha solo 1/10, anche se nessuno dei due è effettivamente probabile.

Possibile è ciò che ha una parte di certezza, anche se molto piccola. *Impossibile* è ciò che non possiede alcuna parte di certezza o ne possiede una infinitamente piccola. Possibile è per esempio ciò che 1/20 o 1/30 di certezza.

Moralmente certo è ciò che si avvicina alla certezza piena, in modo tale che la differenza sia impercettibile. Per contro *moralmente impossibile* è ciò che possiede tanta probabilità quanto manca alla certezza morale per essere certezza piena. Se si tratta ciò che ha certezza di 999/1000 come moralmente certo, ciò che ha solo 1/1000 di certezza è moralmente impossibile.

Necessario è ciò che deve essere, essere stato in passato o essere in futuro.[1] La *necessità* può essere

fisica – per esempio, è necessario che il fuoco brucia; il triangolo ha tre angoli che sommano a due retti; l'eclissi di luna deve verificarsi quando al plenilunio la luna si trova in un nodo;

ipotetica, come conseguenza di cose che sono o sono supposte essere state in passato o essere in futuro, che esistono o devono essere esistite – in questo senso è necessario che Peter, che so o suppongo che scriverà, scriva effettivamente;

stabilita per accordo – per esempio, il giocatore di dadi che faccia "sei" con un dado, deve necessariamente vincere, se i giocatori si sono preventivamente accordati che la vittoria va assegnata al lancio del "sei".

Casuale (sia in dipendenza dall'arbitrio di un essere dotato di ragione sia in dipendenza da un evento contingente o del destino) è ciò che poteva non essere, non essere stato nel passato o non essere nel futuro, beninteso come conseguenza di una possibilità lontana, non immediata.[2] Infatti, non sempre la casualità esclude del tutto la necessità limitatamente a cause di importanza secondaria, come chiarirò con esempi. Certamente, in funzione del luogo, della velocità, della distanza del tavolo da gioco e del momento in cui lascia la mano, il dado non può cadere in modo diverso da come effettivamente cade. Analogamente il tempo per una determinata condizione atmosferica – quantità, dislocazione, movimento, direzione del vento, della nebbia e delle nuvole, nonché certe leggi meccaniche, secondo cui questi fattori interagiscono tra di loro – domani mattina non può fare altro che il tempo che farà. Questi eventi seguono le loro cause prossime non meno necessariamente dell'eclissi il movimento degli astri. E tuttavia ci si attiene alla consuetudine di considerare solo l'eclisse come evento necessario, mentre i lanci del dado e le configurazioni future del tempo atmosferico sono classificati come casuali. Ciò dipende esclusivamente dal fatto che non ci è

---

[1] [All'interno del calcolo delle probabilità non si può definire l'operatore *necessario*, perché non ha senso l'equazione *necessario = impossibile che non*. N.d.T.]

[2] [Questa definizione di casualità include la nozione moderna di contingente, inteso come non impossibile *e* non necessario, in pratica come ciò che è possibile che sia ed è possibile che non sia. Ma nella nozione di casualità c'è qualcosa di più del puramente logico che la rende problematica. Bernoulli non sapeva che quella di casualità è una pseudozione, un misto di oggettivo e soggettivo, che non può essere definita in modo rigoroso. Inoltre Bernoulli non sapeva distinguere tra determinismo e meccanicismo. Il determinismo è l'azione della causa ed esclude che ci possano essere effetti in assenza di causa, "spontanei". Il determinismo censura l'esperienza del "falso", che si fa quotidianamente in laboratorio: da una parte i "falsi positivi" (effetti senza causa) e dall'altra i "falsi negativi" (causa senza effetti). Il meccanicismo è indipendente dall'indeterminismo, in quanto si limita a introdurre delle simmetrie nello spazio degli eventi, come la leva di Archimede che sta in equilibrio se ha bracci uguali e pesi uguali. Così due eventi complementari di probabilità pari a p e $1–p$ sono simmetrici rispetto a ½, quindi sono meccanici, anche se non sono deterministici. N.d.T.]

sufficientemente noto ciò che è assunto come dato per la determinazione degli eventi successivi e in realtà è effettivamente dato. Ma se ci fosse sufficientemente noto, lo studio della matematica e della fisica è sufficientemente avanzato da consentirci di calcolare, a partire da date cause, gli effetti successivi, come dalle leggi astronomiche possiamo prevedere le eclissi.[3] Prima che l'astronomia giungesse a tale perfezione si dovevano classificare le eclissi come gli altri eventi citati nel novero degli eventi casuali del futuro. Ne consegue che a un certo uomo e in un determinato momento può sembrare casuale quel che a un altro uomo (magari lo stesso) in un altro momento, può sembrare necessario, una volta note le cause. La causalità dipende soprattutto dalla nostra ignoranza [delle cause],[4] nella misura in cui non possiamo formulare nessun argomento, che deponga contro il fatto che qualcosa non sia o non sarà, benché sulla base delle cause immediate a noi ancora sconosciute è o sarà necessariamente.[5]

Si chiama *fortuna* o *sfortuna* quel che di buono o di cattivo ci capita, ma non un buono o un cattivo qualunque, bensì solo ciò che probabilmente o almeno altrettanto probabilmente avrebbe potuto non toccarci in sorte. Perciò la fortuna e la sfortuna sono tanto maggiori quanto minore è la probabilità dell'evento buono o cattivo. Così è baciato dalla fortuna chi scavando una fossa per terra trova un tesoro, un caso che non si verifica neppure una volta su migliaia di casi. Se di venti disertori uno viene estratto a sorte per essere impiccato, come esempio dimostrativo per gli altri, i diciannove che restano, con i quali il destino si è mostrato benevolo, non possono dirsi particolarmente fortunati, ma il ventesimo che salito sul patibolo, quello è stato il più sfortunato di tutti. Se il tuo amico torna dalla battaglia in cui sono caduti in pochi, senza neppure un graffio, non lo puoi dire fortunato, anche se forse lo vorresti, perché si è distinto solo per la fortuna di mantenersi in vita.

Cap. II

Sapere e congetturare. L'arte della congettura. Argomenti per congetture. Teoremi generali.

Di ciò che è certo e indubitabile diciamo che lo *sappiamo* o lo *conosciamo*, ma di tutto il resto diciamo solo che lo *congetturiamo* o lo *supponiamo*.

---

[3] [Potenza della mentalità eziologia aristotelica! L'aristotelismo non molla la presa sull'intelletto umano neppure nel momento in cui il soggetto della scienza apre nuovi territori di ricerca, come qui nel caso di Bernoulli. Bernoulli dimentica che le leggi astronomiche individuano il moto dei corpi celesti, a meno di una componente rettilinea uniforme. Grazie al principio di relatività di Galilei, secondo cui le leggi della fisica si scrivono allo stesso modo in tutti gli universi inerziali, la scienza esordisce in modo sostanzialmente indeterministico. Esordisce su un teatro dove va in scena il determinismo aristotelico del principio di ragion sufficiente. Le due drammaturgie convivono conflittualmente fino ad oggi. N.d.T.]

[4] [Sembra che Bernoulli proponga una definizione epistemica di casualità come assenza di "cognizione di causa". Casuale = acausale. N.d.T.]

[5] [In autori come Laplace e Bernoulli il determinismo eziologico del principio di ragion sufficiente trionfa proprio nel momento in cui essi fondano un discorso scientifico nuovo, sostanzialmente indeterministico. Esempio clamoroso di resistenza alla scienza degli stessi operatori scientifici? O divisione dell'Io alla Freud, dove una parte dell'Io ammette la castrazione e l'altra la nega? N.d.T.]

Congetturare una certa cosa significa misurarne la probabilità.[6] Perciò designiamo con *arte della congettura* (*ars coniectandi sive stochastice*), l'arte di misurare il più esattamente possibile la probabilità delle cose allo scopo di poter costantemente scegliere quel giudizio e seguire quella condotta, che ci sembrano migliori, più consoni, più sicuri o più consigliabili. Solo in ciò consiste la saggezza del filosofo e il senno dell'uomo di Stato.

Le probabilità vengono valutate in funzione del *numero* e del *peso degli argomenti*, (*Beweisgründe*) che in un certo qual modo dimostrano o segnalano che una certa cosa è stata, è o sarà. Con *peso* intendiamo la forza dimostrativa.

Gli *argomenti* sono o

interni, essenzialmente artificiali, supposti a partire dai punti dimostrativi delle cause, degli effetti, dei soggetti, dei collegamenti, degli indizi o di qualsivoglia altra circostanza, che sembri avere una qualche connessione con la cosa da dimostrare, o

esterni, o non artificiali, assunti dall'autorità o dalle testimonianze degli uomini.

Un esempio. Tizio viene trovato morto ammazzato per strada. Mevio viene accusato dell'omicidio.[7] Gli argomenti dell'accusa sono:

1. Notoriamente Mevio odiava Tizio. (Argomento ripreso dalla possibile causa, in quanto l'odio di Mevio avrebbe potuto portare alla morte di Tizio).
2. Mevio impallidì alla notizia e rispose impaurito. (Argomento ripreso dagli effetti, in quanto pallore e paura possono essere suscitati dalla coscienza dell'omicidio commesso.)
3. In casa di Mevio fu trovato un pugnale sporco di sangue. (Questo è un argomento indiziario.)
4. Nello stesso giorno in cui Tizio fu ucciso, anche Mevio passò dalla stessa strada. (Questa è una circostanza di tempo e luogo.)
5. Caio afferma che nel giorno del delitto ci fu una lite tra Tizio e Mevio. (Questa è una testimonianza.)

Prima di affrontare il nostro compito specifico, che è quello di come si devono utilizzare gli argomenti per misurare le probabilità delle congetture, mi sia concesso di premettere alcune regole generali, o assiomi, che la pura ragione detta all'uomo dotato di sano intelletto e che l'uomo riflessivo osserva sempre nella vita quotidiana.

1.*Sono inammissibili congetture su cose su cui si può ottenere la certezza totale*. Sarebbe un controsenso che un astronomo, il quale sa che annualmente si verificano due o tre eclissi di luna, congetturi prima del plenilunio se ci sarà eclissi oppure no, dato che

6 [Bernoulli non sembra sensibile ad argomenti di matematica qualitativa, nonostante fosse in rapporto con Leibniz, che gli fece conoscere il triangolo di Pascal e fu promotore di un calcolo logico-sintattico. La matematica del Nostro è ancora euclidea, cioè è una matematica della misura di grandezze. Nel calcolo di Bernoulli non c'è posto per congetture puramente qualitative come il *dubito* cartesiano. Tuttavia, le sue considerazioni sulle congetture furono riprese nel secolo scorso da de Finetti e messe alla base del suo approccio soggettivistico al calcolo *delle* probabilità, inteso come logica dell'incerto. Non mi risulta che de Finetti abbia mai dichiarato il suo debito nei confronti di Bernoulli. N.d.T.]

7 [Il fatto che Bernoulli dia a questo punto un esempio giuridico e usi sistematicamente il verbo *ermitteln*, con cui si indica il procedere istruttorio nelle indagini giudiziarie preliminari, dimostra la sua regressione dal discorso scientifico, fatto in assenza di cause, al discorso giuridico, fatto a partire da cause. Il discorso scientifico contempla eventi spontanei – i moti inerziali, i decadimenti radioattivi, le speciazioni biologiche – ai quali si applica benissimo il calcolo delle probabilità senza scomodare il discorso prescientifico delle cause.]

può sapere la verità con calcoli sicuri. Analogamente il giudice, che alla domanda sulla provenienza dei beni rubati dal ladro ottenesse la risposta di averli comprati da Sempronio, agirebbe da folle se volesse dedurre la probabilità di questa risposta da congetture sui gesti e i toni del ladro o dallo stato della merce o da altre circostanze, quando Sempronio è lì presente e può molto semplicemente venire a sapere tutto da lui.

2. *Non basta ponderare l'uno o l'altro argomento, ma bisogna esaminare tutti gli argomenti di cui possiamo venire a conoscenza, i quali sembrino essere in qualche modo adatti a dimostrare la cosa*.[8]

Un esempio. Tre navi salpano dal porto. Dopo qualche tempo viene data la notizia che una delle tre è affondata dopo aver fatto naufragio. Congetturiamo quale delle tre. Considerando solo il numero delle navi, dovremmo concludere che la sfortuna può essere toccata con uguale probabilità a ciascuna delle tre. Ma ricordando che una delle tre navi era vecchia, guasta e mal corredata di vele e sartiame, e inoltre aveva un timoniere giovane e inesperto, allora è per noi più probabile che sia affondata questa nave, piuttosto che una delle altre due.

3. *Non si deve tener conto solo di tutti gli argomenti a favore di una cosa, ma anche di quelli che possono essere portati contro di essa, affinché sia possibile stabilire in tutta evidenza dopo ponderazione accurata quali prevalgano*.

Dopo un lungo periodo di assenza dell'amico da casa, ci si chiede se sia possibile darlo per morto. A favore del sì depongono i seguenti argomenti. Nonostante tutti gli sforzi in vent'anni non si è riusciti a ottenere sue notizie. Viaggiando si è esposti a diversi pericoli, da cui si è al riparo rimanendo a casa. Forse è scomparso tra le onde, o ucciso per strada o in battaglia, o ha perso la vita per una malattia o per qualche disgrazia in un luogo ignoto. Se fosse ancora in vita, avrebbe raggiunto un'età che raramente è concesso agli uomini di toccare, anche rimanendo a casa. Avrebbe scritto, anche dalle più remote coste dell'India, sapendo che a casa era in vista un'eredità.E ulteriori considerazioni. Ma non ci si deve accontentare di questi argomenti, ma si devono loro contrapporre i seguenti che negano la questione sollevata. È noto che era un tipo sventato, distratto, che malvolentieri prendeva la penna in mano e non teneva in gran conto gli amici. Forse è stato fatto prigioniero dai selvaggi e portato lontano, dove non può scrivere. Forse ha scritto qualche volta dalle Indie, ma le lettere sono andate perse o per distrazione del messaggero o per naufragio della nave. Infine, molti sono rimasti ancora più a lungo lontano da casa e poi sono rimpatriati incolumi.

4. *Per giudicare di cose generali bastano argomenti generali e generici. Ma per formulare congetture su cose singole si devono impiegare argomenti particolari e individuali, ammesso che siano disponibili.*

Se si tratta di indicare in modo del tutto generale di quanto sia più probabile che un giovane uomo di vent'anni sopravviva a un vecchio di sessanta, come questi a quello qualsiasi, allora oltre alla differenza di età e agli anni, non occorre considerare altro. Ma se si tratta di due persone determinate, il giovane Peter e il vecchio Paul, allora bisogna considerare anche lo stato di salute di entrambi e la cura che ciascuno dedica alla propria salute. Infatti, se Peter è malato, lascia correre le proprie passioni a briglia sciolta e vive senza misura, allora Paul, benché più vecchio, può a pieno titolo sperare di sopravvivere a Paul.

5. *In cose incerte e dubbie bisogna rimandare l'azione finché si sia fatta più luce. Ma se l'occasione favorevole all'azione non tollera rinvii, allora tra due cose bisogna*

[8] [L'esaustività è richiesta anche dal *Discorso sul metodo* di Cartesio. N.d.T.]

sempre scegliere quella che sembra più adatta, più sicura, più vantaggiosa e più probabile dell'altra, anche se nessuna delle due possiede effettivamente queste qualità.

Così, se scoppia un incendio, al quale puoi sfuggire saltando dal tento o da uno dei piani inferiori, è per te meglio scegliere la seconda possibilità, perché più sicura, anche se nessuna delle due è senza pericolo di farsi del male.

6. *Ciò che in un caso può essere vantaggioso e in nessun caso svantaggioso, va preferito a ciò che in nessun caso è vantaggioso o svantaggioso*.

Ciò corrisponde al nostro proverbio "Se non giova, almeno non danneggia". Questa tesi segue immediatamente dalla precedente. Infatti, ciò che può essere vantaggioso, può risultare in circostanze altrimenti uguali migliore, più sicuro e più desiderabile di ciò che non è vantaggioso.

7. *Il valore delle azioni umane non può essere valutato dal loro successo*.

Infatti le azioni più folli hanno a volte il migliore successo, mentre le più appropriate il peggiore. Dice il poeta:

Careat successibus, opto, quisquis ab eventu
Facta notando putat.[9]

Se, per esempio, un giocatore vuole fare un triplo sei con tre dadi, di lui si dirà che ha operato da folle, anche se per caso ha vinto. Al contrario sono i giudizi del popolo, a cui un uomo sembra eccellente, ogni volta che è baciato dalla fortuna, del quale anche malaffare andato a buon fine spesso sembra anche un'azione buona.

8. *Nei nostri giudizi dobbiamo stare attenti ad attribuire alle cose più peso di quanto spetti loro e considerare ciò che è solo probabile come affatto sicuro, o imponendolo ad altri come tale*.

Dobbiamo, infatti, porre la fiducia che riponiamo in una cosa proporzionale al grado di certezza o tanto meno quanto minore è la probabilità della cosa.

9. *Poiché solo raramente si può pretendere la certezza assoluta, la necessità e la tradizione impongono di considerare come incondizionatamente certo ciò che è solo moralmente certo*.

Sarebbe utile fissare d'autorità limiti ben definiti alla certezza morale, per esempio esigere il 99/100 o 999/1000 della certezza, in modo che il giudice non sia di parte, ma abbia un punto di vista solido, da tener presente in ogni caso di giudizio.

Ognuno di noi, che abbia esperienza di vita quotidiana, può scrivere di proprio pugno altri assiomi del genere. Non possiamo tenerli a mente tutti, tanto più che ci manca l'occasione opportuna per applicarli.

Cap. III

Diversi tipi di argomenti dimostrativi. Valutazione del loro peso per il calcolo delle probabilità delle cose.

Chi voglia comprovare gli argomenti dimostrativi di un giudizio o di una congettura, ne distinguerà di tre tipi diversi. Alcuni sono *necessariamente presenti e mostrano la casualità della cosa*; altri sono *casualmente presenti e mostrano la necessità della cosa*;

[9] Possa fare a meno del successo, chi pensa di dover valutare le cose dal successo.

altri ancora sono *casualmente presenti e mostrano la casualità della cosa*. Chiarirò le differenze con esempi.

È tanto tempo che mio fratello non mi scrive. Sono in dubbio se si tratti della sua pigrizia o di affari, ma temo anche che sia morto. Ecco tre motivi per l'assenza di lettere: pigrizia, morte, affari.

Il primo motivo è necessariamente presente, come conseguenza di una necessità ipotetica, in quanto conosco e presumo la pigrizia di mio fratello, ma mostra solo la casualità del mancato arrivo di lettere, perché la pigrizia di mio fratello non avrebbe necessariamente ostacolato la scrittura.

Il secondo motivo è casualmente presente, infatti mio fratello può essere ancora in vita, ma segnala necessariamente il mancato arrivo delle lettere, perché un morto non può scrivere.

Il terzo motivo è casualmente presente e segnala la casualità del mancato arrivo delle lettere, perché mio fratello può essere occupato in affari oppure no e nel primo caso non sono affari così assorbenti da impedirgli di scrivere.

Un ulteriore esempio è il seguente.

Per accordi presi un giocatore vince se fa sette con due dadi. Io congetturo se ha speranza di vincere. In questo caso l'argomento dimostrativo per la vittoria del lancio di un sette, che segnala necessariamente la vittoria (in forza degli accordi presi), ma questo argomento è solo casualmente presente, perché oltre al sette possono uscire altri numeri.

Oltre a queste diversità tra argomenti dimostrativi, se ne possono registrare altre, nella misura in cui alcuni sono *puri* e altri *misti*. Chiamo *puri* quegli argomenti dimostrativi che in certi casi dimostrano una cosa, mentre in altri non la dimostrano positivamente [D*p* *vs* ¬D*p*]. Chiamo *misti* quegli argomenti che in certi casi dimostrano una cosa, mentre in altri dimostrano l'esatto contrario [D*p* *vs* D¬*p*].[10]

Ad esempio, in un gruppo di contendenti, uno perisce di spada. Secondo la testimonianza di un uomo degno di fede l'omicida portava un mantello nero. Siccome tra i contendenti Gracco e tre altri portavano un mantello nero, questo diventa un argomento contro Gracco, ma è solo un argomento misto. Infatti, in un caso il mantello nero dimostra la sua colpa, ma negli altri tre la sua innocenza, a seconda che l'omicidio sia stato perpetrato da lui o da uno degli altri tre e non può essere imputato a quest'ultimo senza che Gracco risulti innocente. Ma nel successivo interrogatorio Gracco impallidisce e tale pallore costituisce un argomento dimostrativo puro. Infatti, dimostra la colpa di Gracco, se proviene da cattiva coscienza, ma non dimostra positivamente la sua innocenza, se il pallore è determinato da un'altra causa cui Gracco è soggetto. Gracco può impallidire facilmente ed essere l'omicida.

Da quanto detto risulta evidente che la forza dimostrativa di un argomento dipende dal numero dei casi in cui esso è o può essere presente, mostrando una cosa o non mostrandola o potendo addirittura mostrare il suo contrario. In base a questi casi si può calcolare, con l'aiuto della teoria esposta nella prima parte, calcolare esattamente il grado di certezza o di probabilità, che un argomento dimostrativo porta con sé, nonché la speranza di vincita di un partecipante a un gioco d'azzardo.

Per dimostrarlo assumiamo che il numero di casi in cui un argomento possa essere casualmente presente sia b, il numero di casi in cui non possa essere presente sia c, e che il totale dei casi sia $b + c = a$. Inoltre sia β il numero di casi in cui un argomento dimostrativo mostra casualmente una certa cosa, e sia γ il numero di casi in cui non mostra quella cosa o mostra il suo contrario, e sia $\beta + \gamma = \alpha$. Assumiamo inoltre che tutti i casi siano egualmente possibili, cioè che ogni caso possa verificarsi con la stessa

[10] [Con queste distinzioni Bernoulli anticipa la moderna differenza tra probabilità assolute, $p(A)$ e probabilità condizionate, $p(A|B)$, probabilità di A dato B. N.d.T.]

facilità degli altri. In altri casi adottiamo la modifica per cui, al posto di un caso che si verifica più facilmente contiamo tanti casi che nell'insieme si verificano tanto facilmente quanto quello. Per esempio, invece di un caso tre volte più probabile contiamo tre casi che si verificano con la stessa facilità.

1. Se un argomento dimostrativo può essere *casualmente presente* e *mostrare necessariamente una cosa*, allora per le convenzioni precedenti esistono b casi in cui è presente e può mostrare anche la cosa e c casi in cui non è presente e quindi non può mostrare la cosa. Ciò dà il peso $(b.1 + c.0)/a = b/a$ (per il teorema II della prima parte), così che l'argomento dimostra b/a della cosa o possiede b/a della certezza della cosa.

2. Se un argomento dimostrativo è *necessariamente presente* e *può mostrare casualmente la cosa*, per le assunzioni precedenti sono presenti β casi in cui l'argomento può mostrare la cosa e γ casi in cui non la mostra o mostra addirittura il suo contrario. Ne consegue che la forza dimostrativa dell'argomento è $(\beta.1 + \gamma.0)/\alpha = \beta/\alpha$ e dimostra quindi β/α della certezza della cosa. Se l'argomento dimostrativo è di tipo misto, consegue per la certezza del suo contrario $(\gamma.1 + \beta.0)/\alpha = \gamma/\alpha$.

3. Se un argomento dimostrativo è *casualmente presente* e *può mostrare casualmente la cosa*, assumiamo innanzitutto che esso sia presente. In tal caso, come abbiamo visto, dimostra β/α della cosa e, se l'argomento è di tipo misto, dimostra γ/α del contrario della cosa. Poiché ora ci sono b casi, in cui l'argomento dimostrativo è presente, e c casi, in cui non è presente, l'argomento dimostrativo ha per la cosa il peso $(b.(\beta/\alpha) + c.0)/a = b\beta/a\alpha$. Se l'argomento è di tipo misto il valore per la dimostrazione del contrario è $(b.(\gamma/\alpha) + c.0)/a = b\gamma/a\alpha$.

4. Se per la stessa cosa sono presenti più argomenti e indichiamo con

Argomenti	1	2	3	4	5	…
Numero casi	a	d	g	p	s	...
Numero casi dimostrativi	b	e	h	q	t	…
Numero casi non dimostrativi o dimostrativi del contrario	c	f	i	r	u	…

allora la forza dimostrativa risultante da tutti gli argomenti si valuta nel modo seguente.

Primo caso. Tutti gli argomenti dimostrativi sono *puri*. Il peso del primo argomento da solo è $b/a = (a–c)/c$, come abbiamo già visto.[11] Ora si presenta un secondo argomento dimostrativo che in $e = d–f$ casi dimostra la cosa e in f casi non la dimostra. In questi ultimi f casi rimane, come residuo efficiente, il peso del primo argomento già trovato $(a–c)/a$. I due argomenti insieme hanno quindi il peso

$$((d–f).1 + f(a–c)/a)/d = (ad–cf)/ad = 1 – cf/ad.$$

Assumiamo ora un terzo argomento dimostrativo. Sono presenti $h = g–i$ casi in cui dimostra la cosa e i casi in cui non la dimostra. In questi ultimi casi rimangono effettivi i due primi argomenti dimostrativi con il loro peso $(ad–cf)/ad$. Di conseguenza i tre argomenti dimostrativi, presi insieme, hanno il peso:

[11] Invece di b/a dovremmo scrivere β/α, quando l'argomento mostra casualmente la cosa, e $b\beta/a\alpha$, quando l'argomento dimostrativo è casualmente presente.

$((g-i).1 + i(ad-cf)/ad)/g = (adg-cfi)/adg = 1 - cfi/adg.$

E, se sono presenti altri argomenti dimostrativi, si può procedere allo stesso modo. Chiaramente la probabilità che complessivamente porgono, dista dalla certezza, cioè dall'unità, solo di una quantità pari alla frazione data dal prodotto di tutti i casi non dimostrativi diviso per il prodotto di tutti i casi di tutti gli argomenti.

5. *Secondo caso.* Tutti gli argomenti sono *misti*. Poiché il numero di tutti i casi dimostrativi del primo argomento è pari a b, del secondo è e, del terzo h, … e dei casi dimostrativi del contrario rispettivamente $c, f, i, \ldots$, allora in forza del primo argomento da solo la probabilità della cosa da dimostrare rispetto al suo contrario è b/c, in forza del secondo da solo è e/f, in forza del terzo da solo h/i, ecc. Evidentemente alla forza dimostrativa complessiva, risultante dall'effetto congiunto di tutti gli argomenti dimostrativi, basta il contributo di tutte le forze dimostrative dei singoli argomenti dimostrativi, In altri termini, il rapporto della probabilità della cosa sta alla probabilità del contrario come $beh\ldots$ sta a cfi. Pertanto la probabilità della cosa sarà $beh\ldots/(beh\ldots+cfi\ldots)$ e la probabilità del suo contrario $cfi\ldots/(beh\ldots+cfi\ldots)$.

6. *Terzo caso*. Alcuni argomenti sono *puri* (per esempio, i primi tre di cinque) e alcuni sono *misti* (per esempio, gli ultimi due di cinque). Tratto dapprima i casi puri che, per il numero precedente, porgono $(adg-cfi)/adg$ come probabilità della cosa, tale che manca cfi/adg all'unità. Abbiamo pertanto $adg-cfi$ casi in cui i tre argomenti puri dimostrano insieme la cosa e cfi casi in cui non la dimostrano ma in cui ora prendono posto gli argomenti misti. Per il numero precedente, questi ultimi pesano per la cosa $ru/(qt+ru)$ e per il suo contrario $ru/(qt-ru)$. Perciò sulla base di tutti gli argomenti la probabilità per la cosa sarà

$$((adg-cfi).1 + cfi.(qt/(qt+ru)))/adg = (adg(qt+ru)-cfiru)/(adg(qt+ru)) =$$
$$= 1-(cfi/adg).(ru/(qt+ru)).$$

Questa probabilità è di $cfi/adg\ .\ ru/(qt+ru)$ inferiore alla certezza o all'unità. La prima frazione di questo prodotto è esattamente la frazione che esprime di quanto è inferiore alla certezza la probabilità risultante in forza di tutti gli argomenti puri (secondo il punto 4), mentre la seconda frazione esprime la probabilità della cosa contraria, risultante da tutti gli argomenti misti (secondo il punto 5).

7. Quando, oltre agli argomenti a favore della cosa da dimostrare, si presentano anche argomenti che depongono per la cosa contraria, i due tipi di argomenti vanno soppesati a uno a uno in base alle regole precedenti, per stabilire il loro rapporto, confrontando le probabilità della cosa e le probabilità della sua contraria. In proposito va notato che quando gli argomenti delle due parti sono abbastanza forti, le due probabilità possono superare sensibilmente la metà della certezza, cioè ciascuna delle possibilità contrapposte è probabile, anche se una è meno probabile dell'altra. Così è possibile che una cosa abbia i 2/3 della certezza e la sua contraria i ¾. Infatti, ognuna delle possibilità contrapposte è probabile, pur essendo la prima meno probabile della seconda, avendo un rapporto di 2/3 : 3/4 = 8 : 9.[12]

[12] [Non se ne deduca che Bernoulli ignora la legge di additività semplice: $p(A)+p(\neg A)=1$. Infatti, la distribuzione binomiale positiva, la cosiddetta bernoulliana, deriva dallo sviluppo della potenza del binomio $(p+q)^n$, dove $p+q=1$. In questo caso Bernoulli opera con probabilità diversamente condizionate, che non necessariamente sommano a 1, rispettivamente $p(A|B_1)$ e

Prevedo, in quanto non posso negarlo, che si presenteranno diverse circostanze in cui, applicando a casi speciali queste regole, si possa spesso incorrere in sottili errori per non aver accuratamente elaborato le differenze tra gli argomenti. Infatti, talvolta gli argomenti dimostrativi sembrano diversi, mentre rappresentano un unico argomento. Viceversa, argomenti dimostrativi effettivamente diversi sembrano formare un unico argomento. Si arriva ad applicare argomenti tali da rendere impossibile dimostrare il contrario e così via.

Per chiarire questi rapporti riporto qualche esempio.

Nell'esempio soprariferito di Gracco assumo che persone degne di fede, osservando i contendenti, abbiamo visto che l'assassino aveva i capelli rossi, come Gracco e due altri, ma che nessuno di loro portava un mantello nero. Se ora dagli indizi che, a parte Gracco, tre altri portavano un mantello nero e due altri avevano come lui capelli rossi, qualcuno concludesse che la probabilità che Gracco sia colpevole stia alla probabilità che sia innocente come 1: (2.3) = 1: 6 (secondo il punto 5), e che è più probabile che Gracco sia innocente che colpevole, ebbene costui sbaglierebbe. Infatti, in questo caso non sono presenti due argomenti dimostrativi, ma uno solo, sostenuto contemporaneamente da due circostanze differenti: il colore del mantello e il colore dei capelli. Poiché tali circostanze coincidono solo in Gracco, ciò dimostra che solo costui può essere stato l'assassino.

Un contratto scritto solleva dubbi sulla data del documento. Si sospetta che sia stata anticipata a fini di dolo. Un argomento che non sia così è che il documento fu sottoscritto di proprio pugno da un notaio, cioè da un personaggio pubblico giurato. È improbabile che questi abbia esercitato una frode senza mettere a repentaglio onore e posizione. Su cinquanta notai non se ne trova uno che osi spingersi tanto in là nella disonestà. A favore della cosa parlano argomenti circa la cattiva reputazione di questo notaio. Grazie all'inganno il notaio si aspettava un grosso guadagno per sé e avrebbe anche testimoniato qualcosa di estremamente improbabile (per esempio, che qualcuno avrebbe prestato a qualcun altro 10.000 pezzi d'oro quando poteva possederne al massimo 100). Se in questo caso consideriamo solo l'argomento della posizione ufficiale del sottoscrittore, valuteremmo la probabilità dell'autenticità della data a 49/50 della certezza. Ma se consideriamo le probabilità del contrario dobbiamo ammettere che è difficile che il documento non sia stato falsificato e che il dolo raggiunge la certezza morale, diciamo 999/1000 della certezza. Da qui non dobbiamo dedurre che la probabilità dell'autenticità del documento sta alla probabilità della falsificazione come 49/50 a 999/1000, cioè che le due probabilità siano quasi uguali. Infatti, assumendo che il notaio goda di cattiva fama, assumiamo al tempo stesso che non sia dei 49 notai onesti su 50, che aborrono l'inganno, ma che sia il cinquantesimo, il quale non si fa scrupoli a mancare al proprio mandato. A questo punto perde ogni forza dimostrativa e non dice più nulla l'argomento a favore dell'autenticità del documento.[13]

Cap. IV

$p(\neg A|B_2)$, dove B_1 e B_2 sono argomenti diversi. Sarebbe come valutare le probabilità di avere la malattia A o di non averla ($\neg A$) in presenza di agenti patogeni diversi, B_1 e B_2. N.d.T.]

[13] [È curioso constatare che Bernoulli, come in seguito Laplace, senta il bisogno di agganciare il nuovo sapere scientifico intorno alle probabilità al vecchio sapere prescientifico, in particolare giuridico. (Nel caso di Laplace si tratterà della probabilità che una testimonianza in tribunale dica il vero). Il *refrain* è sempre quello: c'è una resistenza alla scienza anche negli uomini di scienza. N.d.T.]

Sui due modi di determinare i casi. Validità della determinazione attraverso l'osservazione. Problema principale e altro.

Nel capitolo precedente si è visto come in base ai numeri dei casi, in cui possono essere o non essere presenti argomenti dimostrativi per una certa cosa, essi possano mostrare o non mostrare o mostrare il contrario,[14] essendo possibile determinare e valutare le loro forze dimostrative e le probabilità a loro proporzionali. Quindi siamo arrivati al punto in cui, per formare corrette congetture su qualcosa, non si richiede altro che determinare dapprima esattamente il numero di questi casi e poi stabilire quanto è più facile che un caso si verifichi rispetto a un altro. La difficoltà – ci sembra – sta proprio qui. Infatti, ciò è possibile per pochissimi fenomeni e quasi esclusivamente per i giochi d'azzardo. Affinché i partecipanti al gioco avessero uguali prospettive di vincita, gli inventori dei giochi d'azzardo li hanno predisposti in modo che il numero di casi in cui si dà guadagno o perdita siano noti e determinabili *a priori* e che tutti i casi si possano verificare con uguale probabilità.[15] Nella stragrande maggioranza dei casi, dipendenti dalle forze della natura o dall'arbitrio dell'uomo, ciò non si verifica.

Per esempio, nei dadi il numero dei casi è noto, perché per ogni dado esistono tanti casi quante facce. Tutti questi casi sono egualmente facilmente possibili. Data l'uguale forma delle facce e la simmetria del peso dei dadi, non ci sono ragioni per cui una faccia debba cadere più facilmente dell'altra, come sarebbe il caso se le superfici del dado avessero forma diversa e una parte del dado fosse più pesante dell'altra.[16]

Stesso discorso vale per il numero dei casi di possibili estrazioni di una pallina bianca o nera da un'urna nota, da cui tutte le palline possono essere estratte con eguale

---

[14] [Sembra che Bernoulli proponga una logica epistemica, o dell'argomentazione, di tipo intuizionista, dove non vale il principio del terzo escluso. Un sistema modale, adeguato al calcolo di Bernoulli, potrebbe essere il sistema T di Lewis, dove non vale il principio del terzo escluso in forma modale: $Lp \vee \neg Lp$, atteso che L sia l'operatore "necessario", regolato dall'assioma $Lp \rightarrow p$. In questo caso il sistema bernoulliano sarebbe un sistema epistemico cognitivo, dove il sapere (Lp) implica la verità (p). Esistono sistemi epistemici non cognitivi dove la verità implica il sapere – prima o poi – come è il caso del sistema epistemico inconscio. Cfr. Il mio *Una matematica per la psicanalisi* alla pagina http://www.sciacchitano.it/Tempo/Logica%20epistemica.html in www.sciacchitano.it . N.d.T.]

[15] [La richiesta dell'equiprobabilità dei casi elementari è un'eredità della geometria euclidea, che tratta grandezze misurate con opportune unità di misura, eventualmente suoi sottomultipli. Il numero di volte in cui l'unità di misura è contenuta nella misura di una grandezza, ripartisce quest'ultima in parti uguali, *quindi* equiprobabili nel senso che hanno lo stesso peso. Ci si sgancerà dalla richiesta di equiprobabilità quando il pensiero matematico si sarà svincolato da considerazioni metriche, o quantitative, e avrà approntato le nozione di numero reale e di limite, valide in topologie qualitative. Occorreranno ancora almeno altri 150 anni. N.d.T.]

[16] [Il calcolo delle probabilità trasferisce il meccanicismo dal piano ontologico al piano epistemico. Inizialmente il trasferimento sfrutta simmetrie meccaniche: la simmetria cubica del dado come la simmetria della leva di Archimede, che avendo bracci uguali è in equilibrio se ha pesi uguali agli estremi. Ma la simmetria del calcolo delle probabilità è più riposta e specifica. Essa poggia sul fatto che la probabilità di un evento A e del suo contrario $\neg A$ hanno probabilità simmetriche rispetto a ½ (legge di additività semplice). Il calcolo delle probabilità offre l'esempio di un calcolo meccanicistico, perché simmetrico, ma non deterministico. Come il calcolo meccanico, il calcolo delle probabilità usa la stessa nozione di "peso" che si usa in statica. N.d.T.]

facilità, in quanto è noto quante palline di ogni tipo sono presenti nell'urna e non si possono addurre ragioni perché sia più facile estrarre una pallina o l'altra.

Quale mortale potrebbe determinare il numero di malattie (cioè il numero di casi) che possono capitare al corpo dell'uomo in ogni sua parte e in ogni età e portarlo alla morte – la peste come l'idropisia, l'idropisia come la febbre – per trarne una congettura sul rapporto tra vita e morte nelle generazioni future? O chi potrebbe enumerare gli innumerevoli casi di cambiamento cui giornalmente va incontro l'aria, volendo già da oggi congetturare che tempo farà tra un mese o tra un anno? O ancora, chi potrebbe aver indagato tanto a fondo la natura della mente umana e la meravigliosa costruzione del nostro corpo da pretendere di determinare il numero dei casi in cui vincerà questo giocatore o perderà quell'altro in giochi dipendenti in tutto o in parte dall'acutezza dell'intelletto o dalla destrezza del corpo? Poiché queste e altre cose simili dipendono da cause totalmente nascoste,[17] le quali sfuggono alla nostra conoscenza attraverso la molteplicità infinita dei loro effetti combinati, sarebbe affatto insensato voler fare della ricerca in questo modo.

Ma qui ci si apre un'altra via per trovare ciò che si cerca e per determinare almeno *a posteriori*, cioè stabilendolo dall'esito osservato in numerosi casi simili, quel che non possiamo determinare *a priori*. A tal fine bisogna supporre che ogni singolo evento possa verificarsi o non verificarsi in altrettanti casi quanti sono stati osservati verificarsi o non verificarsi in precedenza nelle stesse circostanze. Per esempio, se si è osservato che di 300 uomini della stessa età e costituzione di Tizio 200 sono morti entro 10 anni, mentre gli altri sono sopravvissuti, allora si può ragionevolmente dedurre che c'è il doppio di casi in cui anche Tizio debba nel prossimo decennio pagare alla natura il tributo dovuto, rispetto ai casi che in questo frangente possono sopravvivere. Analogamente, se qualcuno ha da lungo tempo osservato il tempo atmosferico e notato quante volte è stato sereno e quante piovoso o se qualcuno ha osservato spesso due giocatori e notato quante volte questo o quello abbia vinto, può proprio a partire da ciò determinare il rapporto che probabilmente avranno in seguito i casi in cui gli stessi eventi si possono verificare o non verificare come in precedenza alle stesse condizioni.

Questo modo empirico di determinare il numero di casi attraverso l'osservazione non è né nuovo né fuori dal comune.[18] Infatti, già il famoso autore dell'opera *l'arte di pensare"*,[19] uomo acuto e talentuoso, ha descritto nel cap. 12 e sg. dell'ultima parte del

[17] [In Bernoulli è chiaro che la scienza moderna, a differenza dell'antica, si sviluppa "senza cognizione di causa". Il calcolo delle probabilità, essendo un calcolo epistemico, presuppone l'ignoranza delle cause. Si potrebbe dire "casuale = non causale". La scienza moderna tratta fenomeni "spontanei". N.d.T.]

[18] [Questo è più che un errore di Bernoulli. È un voluto misconoscimento. Bernoulli misconosce che il modo di trattare l'empiria della scienza moderna è diverso da quello della scienza antica. La scienza antica guarda all'esperienza per confermare il sapere scritto nel libro canonico. La scienza moderna tratta un "sapere nel reale", che non è ancora scritto nel libro, attraverso manipolazioni del reale stesso – osservazioni programmate o interventi strumentali. La macchina programmata dell'esperimento, una macchina epistemica, scava il sapere nel reale, attraverso conferme *e* confutazioni, e produce "cose epistemiche". Le conferme sono l'apporto di nuovo materiale grezzo. Le confutazioni sono il setacciamento del materiale grezzo. Quel che resta è l'oro della certezza scientifica. N.d.T.]

[19] [Si tratta della *Logica di Port Royale* (trattato comparso anonimo nel 1644), di Antoine Arnauld e Peter Nicole, i quali scrissero un trattato eclettico, aristotelico-cartesiano, fondamentalmente logocentrico (congruenza pensiero/linguaggio + universalità della struttura logica, uguale in tutti i pensanti), che dimostra solo la difficoltà della transizione dalla scienza antica alla moderna. N.d.T.]

proprio lavoro un processo del tutto analogo, come tutti gli uomini possono osservare nella vita quotidiana.

Inoltre è evidente a tutti che *non basta fare questa o quella osservazione per giudicare empiricamente un evento, ma è necessario un gran numero di osservazioni.* Talvolta un uomo veramente ingenuo, dotato di un proprio istinto naturale e senza precedente formazione, ha fatto l'esperienza – davvero meravigliosa – che, tanto più numerose sono le osservazioni pertinenti, tanto minore è il pericolo di deviare dalla verità. Sebbene tutti riconoscano che ciò stia nella natura della cosa, la dimostrazione fondata su principi scientifici non è a portata di mano e mi tocca produrla qui. Ma credo di far poca cosa soffermandomi a dimostrare questi pochi punti, noti a tutti.

[*Teorema dei grandi numeri*]

Bisogna prendere in considerazione molto di più di ciò a cui forse nessuno abbia finora pensato. Infatti, resta da esaminare se l'aumento delle osservazioni porti ad aumentare in modo continuo la probabilità che il numero di osservazioni favorevoli raggiunga il vero rapporto rispetto alle sfavorevoli. E precisamente nella misura in cui o questa probabilità supera definitivamente qualunque grado arbitrario di certezza oppure il problema ha, per così dire, il proprio asintoto,[20] cioè è presente un determinato grado di certezza – il vero rapporto dei casi da trovare – che non si può mai superare comunque si aumentino le osservazioni. Per esempio, non possiamo mai raggiungere una certezza superiore a ½, 2/3, ¾ per determinare il vero rapporto dei casi.

Un esempio chiarirà quel che intendo.

Supponiamo che in un'urna, senza che tu lo sappia, ci siano 3000 sassolini bianchi e 2000 neri. Procedendo per tentativi tu vuoi determinare il loro rapporto, estraendo un sassolino dopo l'altro (e rimettendolo nell'urna dopo averlo estratto in modo che il numero dei sassolini non diminuisca) e osservando quante volte esce un sassolino bianco e quante uno nero. Ci si chiede se tu puoi ripetere l'operazione tanto spesso, da far sì che diventi dieci, cento, mille volte più probabile (fino a raggiungere la certezza morale) che il numero delle volte in cui estrai un sassolino bianco rispetto al numero delle volte in cui estrai un sassolino nero, assuma lo stesso rapporto di 1½ che hanno tra loro i sassolini (o i casi), o questi numeri non formino un tutt'altro rapporto diverso da quello. Se non si verifica il primo caso, ammetto che il nostro tentativo di determinare il numero dei casi attraverso osservazioni è mal congegnato. Ma se si verifica e in questo modo [sperimentale] posso ottenere la certezza morale (che sia effettivamente così lo dimostrerò nel prossimo capitolo), allora posso trovare *a posteriori* il numero di casi esattamente come se lo conoscessi *a priori*. E questo per la vita civile, dove la certezza morale è considerata come assoluta (per l'assioma 9 del cap. 2), è sufficiente per orientare le nostre congetture in qualunque regime di casualità altrettanto scientificamente che nei giochi di azzardo. Infatti, se al posto dell'urna pensiamo all'aria o al corpo umano, che alberga in sé un'immensità di variazioni e di malattie diverse, come sassolini nell'urna, allora possiamo determinare allo stesso modo la facilità con cui anche in questo campo risulta un evento piuttosto che un altro.

[20] [Il concetto di asintoto (dal greco *a* (*alfa* privativo) e *sunpipto*, cado insieme) è una delle prime formulazioni della nozione topologica di punto limite o di accumulazione. Ogni intorno dell'asintoto contiene un punto diverso dall'asintoto, ma l'asintoto non appartiene all'insieme per cui funziona da asintoto. La nozione di asintoto – una sorta di *noli me tangere* – è entrata anche nella cultura umanistica. Freud parla di processo infinito di "guarigione asintotica" (*Vie della terapia psicanalitica*, GW, vol. XII, p. 192). N.d.T.]

A scanso di fraintendimenti, va precisato che *il rapporto tra i numeri dei casi [favorevoli e contrari], da determinare attraverso le osservazioni, non viene ottenuto con precisione assoluta,*[21] *ma solo con una certa approssimazione, cioè racchiuso tra due limiti, che tuttavia possono essere presi arbitrariamente vicini*.[22] Se nell'esempio precedente dell'urna piena di sassolini otteniamo due rapporti, per esempio 301/200 e 299/200 (o 3001/2000 e 2999/2000), di cui uno è maggiore e l'altro minore di 1½, allora si dimostra che con qualsivoglia probabilità è più probabile che, dopo osservazioni ripetute un gran numero di volte, il rapporto trovato giaccia all'interno di questi limiti, posti intorno a 1½, piuttosto che all'esterno.[23]

Questo è il problema che qui mi sono proposto di pubblicare, dopo essermene occupato per vent'anni. La sua novità, insieme alla sua straordinaria utilità, unitamente alla notevole difficoltà [concettuale[24]], supera in importanza tutti gli altri capitoli di questa teoria.

Prima di dimostrare il teorema, voglio brevemente confutare alcune obiezioni sollevate da alcuni dotti, [principalmente da Leibniz].

---

[21] Infatti, si verificherebbe il contrario, cioè sarebbe tanto più improbabile trovare il vero rapporto, quanto più numerose osservazioni si fanno. [L'osservazione di Bernoulli è ellittica. Forse vuol dire che, se dopo k osservazioni si ottenesse per caso il valore esatto, è molto probabile che le osservazioni successive portino lontano da quel valore. N.d.T.]

[22] [Qui bisogna correggere Koyré. Il passaggio dalla scienza antica alla moderna non avviene per transizione "dal mondo del pressappoco all'universo della precisione", ma abbandonando la cattiva approssimazione dottrinaria e adottando quella buona scientifica. Il termine "approssimativamente" (*näherungsweise*) è esplicitamente adottato da Bernoulli nella formulazione del suo teorema. Il soggetto della scienza sa approssimare bene. Il soggetto cognitivo no, perché è fortemente binario. Sa solo adeguarsi o non adeguarsi alla cosa. Il soggetto della scienza è libero di operare in un certo intorno della verità. Il soggetto cognitivo è servo della dottrina del maestro, che lo obbliga ad accettare o rifiutare valori ben precisi. N.d.T.]

[23] [Detto in termini moderni, comunque sia fissata una probabilità P e due limiti inferiore e superiore di p – qui 1½ – tali che $l_1 < p < l_2$, allora esiste un numero n di osservazioni tale che la frequenza relativa osservata f/n – qui f = numero osservato di sassolini bianchi – ha probabilità P di cadere dentro l'intervallo $l_1 < p < l_2$ e probabilità $1–P$ di cadere fuori. Questa è la cosiddetta convergenza in probabilità della frequenza alla probabilità. Il teorema di Bernoulli è un caso particolare del teorema centrale del limite. N.d.T.]

[24] [La difficoltà concettuale sta nel doppio riferimento alla probabilità. Nell'enunciato del teorema interagiscono due probabilità: p e P. p è la probabilità dell'evento: "estrarre un sassolino bianco dall'urna". P è la probabilità dell'evento: "la frequenza osservata di sassolini bianchi dopo n estrazioni dall'urna giace tra due limiti prestabiliti l_1 e l_2. p è una probabilità ontologica. P è una probabilità epistemica. La probabilità epistemica permette di formulare congetture sulla probabilità ontologica. C'è da credere che la difficoltà abbia impegnato Bernoulli per vent'anni. Non meno tempo occorse a Galilei per comprendere che nel moto uniformente accelerato la velocità è proporzionale al tempo, non allo spazio. Entrambe sono difficoltà a staccarsi dall'aristotelismo, dove non esistono né nozioni di probabilità – confusa con l'assenza di intenzionalità – né di velocità istantanea – confusa con la velocità media. N.d.T.]

1. In primo luogo si obietta che il rapporto tra sassolini ha un altro statuto rispetto a quello tra malattie e cambiamenti atmosferici, essendo il primo determinato, mentre i secondi sarebbero indeterminati e incerti.

Rispondo che, rispetto alla nostra conoscenza, sono entrambi incerti e indeterminati. D'altra parte, che una cosa sia in sé e per natura incerta e indeterminata è per noi altrettanto poco comprensibile come ammettere che dio abbia al tempo stesso creato e non creato qualcosa. Infatti, tutto quel che dio ha creato, proprio creandolo, l'ha anche determinato.[25]

2. In secondo luogo si obietta che il numero dei sassolini è finito, il numero delle malattie infinito.

Ribatto che il numero delle seconde, piuttosto che infinito, è grande in modo stupefacente. Ma, anche ammesso che sia infinitamente grande, [dal calcolo infinitesimale] si sa che anche tra numeri infinitamente grandi può esistere un rapporto, esprimibile con numeri finiti o in modo esatto o con l'approssimazione che si desidera.[26] Così la circonferenza di un cerchio ha sempre un determinato rapporto con il proprio diametro, che può essere espresso solo un numero infinito di decimali con i procedimenti di Ludolph o di Archimede, ma rimane compreso entro limiti determinati, che sono pienamente sufficienti per l'uso comune. Quindi, *nulla osta a determinare con un numero finito di osservazioni il rapporto tra due numeri infinitamente grandi, rappresentabile con buona approssimazione come rapporto tra numeri finiti*.[27]

---

[25] [Qui si vede molto bene la resistenza, di marca religiosa, del soggetto della scienza ad accettare l'indeterminismo, proprio nel momento – e questo è paradossale – in cui sta elaborando la teoria per trattarlo, cioè il calcolo delle probabilità. Sembra che per accettare l'indeterminismo il soggetto debba prima negarlo, come in psicanalisi accetta il rimosso inconscio, negandolo coscientemente. Di fatto, come gli uomini di scienza fino ad Einstein e Gödel, Bernoulli non sa distinguere tra meccanicismo e determinismo. Il primo basato sull'azione di simmetrie, il secondo sull'azione di cause. Il primo scientifico, il secondo prescientifico. Per la nozione moderna di simmetria, che nella scienza moderna incrina il predomino della quantità sulla qualità, rimando al programma di Erlangen di Felix Klein, (1872) riportato in questo stesso sito alla pagina http://www.sciacchitano.it/Scienziati/F%20Klein/Felix%20Klein.html . N.d.T.]

[26] [Qui si vede molto bene come il "giovane" soggetto della scienza tenti di sottrarsi al fascino della matematica euclidea della grandezza, contrapponendole la nozione ancora acerba di numero e di limite. Di fatto, la grandezza infinitamente grande non esiste, perché non sarebbe una grandezza, non essendo misurabile, mentre un numero infinitamente grande esiste e può stabilire rapporti con altri numeri infinitamente grandi. Il merito dell'avvenuto distacco dalla dottrina euclidea (in realtà aristotelica) sarà dovuto all'efficacia della nozione topologica di approssimazione "vicina quanto si vuole". A testimonianza dell'urgenza storica della transizione dalla grandezza al numero, ricordo che, non disponendo della matematica giusta per la propria cinematica, Galilei ingenuamente pretendeva riformare il V libro degli *Elementi*, quello che abbastanza paradossalmente tratta delle proporzioni tra grandezze incommensurabili, libro che per altro è in sé perfetto e immodificabile (merito di Eudosso di Cnido, contemporaneo di Platone). La tradizione storica rende parziale giustizia a Bernoulli, chiamando il suo teorema "legge dei grandi numeri". Non è propriamente una legge nel senso deterministico del termine, ma riguarda numeri e non grandezze. N.d.T.]

[27] [Il gioco tra finito e infinito ha un valore che va oltre la contingenza della determinazione di qualche costante fisica. Il gioco tra finito e infinito è il gioco esistenziale della modernità. Esso si gioca tra due attori: soggetto finito e oggetto infinito. È un gioco che gli antichi non

3. In terzo luogo si obietta che il numero delle malattie non è fisso ma cresce di giorno in giorno.

A ciò oppongo che nel corso del tempo le malattie possono ben essere aumentate di numero – non lo nego. Sicuramente chi volesse in base a osservazioni attuali trarre conclusioni sui tempi antidiluviani, sbaglierebbe della grossa. Ma da ciò non consegue che non si debbano eseguire nuove osservazioni. Sarebbero necessarie nuove osservazioni anche nel caso dell'urna di sassolini, qualora si debba supporre che il loro numero è cambiato.

Cap. V

Soluzione del problema precedente

Per abbreviare al massimo la dimostrazione estesa e renderla più chiara, tento di esporre tutto in termini puramente matematici. A tale scopo premetto i seguenti lemmi. Una volta dimostrati, il resto della dimostrazione consiste semplicemente nell'applicare i lemmi.

Lemma 1. Sia dato [il tratto iniziale del] la serie dei numeri naturali

$$0, 1, 2, \ldots, r-1, r, r+1, \ldots, r+s,$$

dove r è un numero qualsiasi intermedio e $r-1$ e $r+1$ i numeri contigui a r rispettivamente a sinistra e a destra. Si prolunghi la serie finché l'ultimo termine abbia un rapporto intero arbitrario con $r+s$, diciamo $nr+ns$. La nuova serie consisterà dei termini

$$0, 1, 2, \ldots, nr-n, nr, nr+n, \ldots, nr+ns.$$

Al crescere di n cresce il numero di termini compresi tra nr e $nr+n$, rispettivamente tra n e $nr-n$, come anche aumenta il numero di termini compresi tra $nr+n$, rispettivamente $nr-n$, e i termini più esterni $nr+ns$ e 0. Tuttavia, per quanto grande si scelga n, il numero di termini superiori a $nr+n$ non supera mai di più di $(s-1)$ volte il numero di termini compresi tra nr e $nr+n$ e il numero di termini minori di $nr-n$ non supera mai di $(r-1)$ volte il numero di termini compresi tra $nr-n$ e n.[28]

---

conoscevano, forse non volevano neppure conoscere. Un gioco che prende le mosse, come dimostra il trattato di Bernoulli, dalla teoria dei giochi d'azzardo. N.d.T.]

[28] [Esempio con r=4, s=2, n=3.

			r-1	r	r+1	r+s			nr-n			nr			nr+n			nr+ns
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
$n(r$-1) = 3(4-1) = 9 termini									3 termini				3 termini			$n(s$-1)=3 termini		

La dimostrazione insiemistica a destra è la seguente.
$\lvert\{x \mid ns+nr \geq x > nr+n\}\rvert = \lvert\{x \mid ns \geq x > n\}\rvert = ns - n = n(s-1)$.
La dimostrazione insiemistica a sinistra è
$\lvert\{x \mid nr-n > x \geq 0\}\rvert = \lvert\{x \mid nr \geq x > n\}\rvert = nr - n = n(r-1)$. N.d.T]

Dimostrazione. Il numero di termini maggiori di $nr+n$ è pari a $n(s-1)$ e il numero di termini inferiori a $nr-n$ è pari a $n(r-1)$. Il numero di termini compreso tra nr escluso e uno dei due limiti incluso è pari a n. Ma si verifica sempre

$$n(s-1) : n = s-1 : 1$$
$$n(r-1) : n = r-1 : 1.$$

Da cui segue ecc.

Lemma 2. Elevando il binomio $r+s$ a una potenza intera arbitraria, lo sviluppo ha sempre un termine in più della potenza.

Infatti, lo sviluppo del quadrato ha tre termini, del cubo quattro ecc.

Lemma 3. Nello sviluppo della potenza del binomio $r+s$, la cui potenza sia un multiplo intero di $r+s = t$, per esempio $n(r+s) = nt$, il termine M assume il valore massimo[29] solo se il numero dei termini che lo precedono sta al numero di quelli che lo seguono come s sta a r o, equivalentemente, se gli esponenti di r ed s stanno tra loro come r ed s e ciascuno dei termini contigui a M a destra o a sinistra è maggiore di ogni termine più distante dallo stesso lato. Inoltre M ha con un termine contiguo un rapporto inferiore di questo con un termine più distante – a parità di distanza dei termini.

Dimostrazione. 1. I matematici sanno che la nt-esima potenza del binomio $r+s$ si rappresenta mediante la serie [finita]

$$(r+s)^{nt} = r^{nt} + \binom{nt}{1} r^{nt-1}s + \binom{nt}{2} r^{nt-2}s^2 + \ldots + \binom{nt}{2} r^2 s^{nt-2} + \binom{nt}{1} r s^{nt-1} + s^{nt},$$

in cui gli esponenti di r progressivamente diminuiscono mentre quelli di s crescono, mentre i coefficienti [binomiali] del primo e dell'ultimo termine, del secondo e del penultimo, ecc. coincidono. Poiché il numero dei termini, tolto M, è uguale a $nt = nr+ns$ (lemma 2) e per ipotesi [necessaria] il numero di termini che precedono M sta al numero di quelli che lo seguono come s a r, allora il numero dei termini che precedono M deve essere ns e il numero di quelli che lo seguono nr. Per la regola di formazione della serie si può scrivere

$$M = \binom{nt}{ns} r^{nr}s^{ns} = \binom{nt}{nr} r^{nr}s^{ns}.$$

Indicando con $L_1, L_2, L_3, \ldots$ i termini a sinistra di M e con $R_1, R_2, R_3, \ldots$ i corrispondenti termini a destra, si può scrivere

$$L_1 = \binom{nt}{ns-1} r^{nr+1}s^{ns-1}, \qquad R_1 = \binom{nt}{nr-1} r^{nr-1}s^{ns+1};$$

[29] [A rigore Bernoulli "dimentica" di dimostrare che *in un insieme X, finito e totalmente ordinato, esiste il massimo M ed è unico*. L'esistenza si dimostra per assurdo. Ammettiamo che $M \notin X$. Sia $x_1 \in X$. Per ipotesi $x_1 < M$. Poiché X è totalmente ordinato, per ogni coppia di elementi x_i e x_j di X è possibile stabilire se $x_i < x_j$ o $x_i > x_j$. Passiamo in rassegna tutti gli elementi di X, cosa possibile essendo X finito. Non possono essere tutti inferiori a x_1, che risulterebbe il massimo e apparterrebbe a X. Allora esiste un elemento di X, diciamo x_2, tale che $x_1 < x_2 < M$. Procedendo in modo analogo si determina un x_3 tale che $x_2 < x_3 < M$. E così via. Si conclude che X è infinito, contro l'ipotesi. Quindi $M \in X$. Lascio per esercizio la dimostrazione dell'unicità di M. N.d.T.]

$$L_2, = \binom{nt}{ns-2} r^{nr+2} s^{ns-2}, \qquad R_2 = \binom{nt}{nr-2} r^{nr-2} s^{ns+2};$$
$$\ldots \qquad \ldots$$

Dividendo si ottiene

$$M/L_1 = (nr+1)s/nsr, \qquad M/R_1 = (ns+1)r/nrs;$$
$$M/L_1 = (nr+2)s/(ns-1)r, \qquad M/R_1 = (ns+2)r/(nr-1)s;$$
$$\ldots \qquad \ldots$$

Ma è

$$nr+1)s > nsr, \qquad (ns+1)r > nrs,$$
$$(nr+2)s > (ns-1)r, \qquad (ns+2)r > (nr-1)s,$$
$$\ldots \qquad \ldots$$

Di conseguenza

$$M > L_1, \qquad M > R_1;$$
$$L_1 > L_2, \qquad R_2 > R_1;$$
$$\ldots \qquad \ldots \qquad \textit{C.V.D.}$$

2. È evidente che

$$(nr+1)s/nsr < (nr+2)s/(ns-1)r, \qquad (ns+1)\mathrm{r}/nrs < (ns+2)\mathrm{r}/(nr-1)$$

o, equivalentemente,

$$M/L_1 < L_1/L_2, \qquad M/R_1 < R_1/R_2.$$

Analogamente si dimostra che

$$L_1/L_2 < L_2/L_3 < \ldots, \qquad R_1/R_2 < R_2/R_3 < \ldots$$

Conseguentemente il termine massimo M ha con il termine contiguo un rapporto inferiore rispetto a quello di un termine omolaterale alla stessa distanza. *C.V.D.*

Lemma 4. Nella potenza del binomio con esponente nt il numero n può essere preso così grande che il rapporto del massimo M con i due termini L_n e R_n, che stanno a destra e a sinistra di M nello sviluppo in serie, può assumere un valore maggiore di qualunque valore stabilito.[30]

Dimostrazione. Per il lemma precedente M ha il valore:

$$M = \frac{nt\,(nt-1)\,(nt-2)\,\ldots\,(nr+1)}{1\,.\,2\,.\,3\,.\,\ldots\,ns} r^{nr} s^{ns} = \frac{nt\,(nt-1)\,(nt-2)\,\ldots\,(ns+1)}{1\,.\,2\,.\,3\,.\,\ldots\,nr} r^{nr} s^{ns}.$$

Per la regola di formazione della serie i membri L_n *e* R_n hanno i valori

[30] [Chissà se Boltzmann conosceva questo lemma. L'avrebbe aiutato a dimostrare il suo teorema H dell'aumento probabilistico dell'entropia. N.d.T.]

$$L_n = \frac{nt\,(nt-1)\,(nt-2)\,\ldots\,(nr+n+1)}{1\,.\,2\,.\,3\,.\,\ldots\,(ns-n)}\, r^{nr+n} s^{ns-n},$$

$$R_n = \frac{nt\,(nt-1)\,(nt-2)\,\ldots\,(ns+n+1)}{1\,.\,2\,.\,3\,.\,\ldots\,(nr-n)}\, r^{nr-n} s^{ns+n}.$$

Eliminando per divisione i fattori comuni, si ottiene

$$\frac{M}{L_n} = \frac{(nr+n)\,(nr+n-1)\,(nr+n-2)\,\ldots\,nr\, s^n}{(ns-n+1)\,(ns-nn+2)\,(ns-n+3)\,\ldots\,ns\, r^n},$$

$$\frac{M}{R_n} = \frac{(ns+n)\,(ns+n-1)\,(ns+n-2)\,\ldots\,ns\, r^n}{(nr-n+1)\,(nr-n+2)\,(nr-n+3)\,\ldots\,nr\, s^n}$$

o, dopo aver ripartito r^n *e* s^n sui singoli fattori, cosa sempre possibile, poiché il numeratore e il denominatore dei coefficienti hanno già n fattori:

$$\frac{M}{L_n} = \frac{(nrs+ns)\,(nrs+ns-s)\,(nrs+ns-2s)\,\ldots\,(nrs+s)}{(nrs-nr+r)\,(nrs-nr+2r)\,(nrs-nr+3r)\,\ldots\,nrs},$$

$$\frac{M}{R_n} = \frac{(nrs+nr)\,(nrs+nr-r)\,(nrs+nr-2r)\,\ldots\,(nrs+r)}{(nrs-ns+s)\,(nrs-ns+2s)\,(nrs-ns+3s)\,\ldots\,nrs}.$$

Ma tali rapporti assumono un valore infinitamente grande quando n diventa infinitamente grande. Allora, infatti, spariscono i numeri 1, 2, 3, … rispetto a n e i fattori $nr(+/-)n(-/+)1,2,3,\ldots$ hanno gli stessi valori come $nr(+/-)n$ e i fattori $ns(-/+)n(+/-)1,2,3,\ldots$ hanno gli stessi valori come $ns(-/+)n$, così che, dividendo numeratore e denominatore per n si ottiene

$$\frac{M}{L_n} = \frac{(rs+s)\,(rs+s)\,(rs+s)\,\ldots\,rs}{(rs-r)\,(rs-r)\,(rs-r)\,\ldots\,nrs},$$

$$\frac{M}{R_n} = \frac{(rs+r)\,(rs+r)\,(rs+r)\,\ldots\,rs}{(rs-s)\,(rs-s)\,(rs-s)\,\ldots\,rs}.$$

Chiaramente le due grandezze sono formate dallo stesso numero di frazioni $(rs+s)/(rs-r)$, rispettivamente $(rs+r)/(rs-s)$, presenti come fattori nel numeratore (o denominatore), il cui numero è uguale a n, cioè è infinitamente grande. Perciò entrambi i rapporti sono una potenza infinitamente alta delle frazioni $(rs+s)/(rs-r)$, rispettivamente $(rs+r)/(rs-s)$, quindi è essa stessa infinitamente grande.

Chi volesse mettere in dubbio questo risultato prenda due successioni geometriche infinitamente decrescenti a ragione $(rs-r)/(rs+s)$ e $(rs-s)/(rs+r)$. Il rapporto del primo termine al terzo, quarto, quinto, … ultimo termine è pari al rapporto $(rs-r)/(rs+s)$, rispettivamente $(rs-s)/(rs+r)$, moltiplicato per se stesso due, tre, quattro… infinite volte.

Chiaramente il rapporto del primo all'ultimo termine, che in una serie geometrica infinitamente decrescente deve essere pari a zero, è infinitamente grande. Ne consegue che anche la potenza infinitamente elevata di $(rs+s)/(rs-r)$ e $(rs+r)/(rs-s)$ ha un valore infinitamente grande. Con ciò si è dimostrato che nello sviluppo della potenza infinitamente elevata di un binomio il rapporto del termine massimo M rispetto ai termini L_n e R_n, ha un valore maggiore di qualunque valore dato. C.V.D.

Lemma 5. Nella potenza di un binomio con esponente nt il numero n può essere scelto tanto grande che il rapporto tra la somma bilaterale di tutti i termini compresi tra il massimo M e i termini L_n e R_n (compresi) e la somma dei termini rimanenti, che giacciono al di fuori dei limiti L_n e R_n, assuma un valore superiore a qualunque valore dato.

Dimostrazione. Per la seconda affermazione del lemma 3 valgono le seguenti diseguaglianze

$$\frac{M}{L_1} < \frac{L_n}{L_{n+1}}, \quad \frac{L_1}{L_2} < \frac{L_{n+1}}{L_{n+2}}, \quad \frac{L_2}{L_3} < \frac{L_{n+2}}{L_{n+3}}, \ldots$$

quindi valgono anche le diseguaglianze

$$\frac{M}{L_n} < \frac{L_1}{L_{n+1}} < \frac{L_2}{L_{n+2}} < \frac{L_3}{L_{n+3}} < \ldots$$

Per il lemma 4 per un valore di n infinitamente grande il valore di M/L_n diventa infinitamente grande e di conseguenza i rapporti L_1/L_{n+1}, L_2/L_{n+2}, $L_3/L_{n+3}\ldots$ hanno valori tanto infinitamente più grandi. Ne consegue che

$$\frac{L_1 + L_2 + L_3 + \ldots + L_n}{L_{n+1} + L_{n+2} + L_{n+3} + \ldots + L_{2n}} = \infty,$$

Cioè la somma di tutti i termini compresi tra il termine maggiore M e il termine n-esimo L_n incluso è di infinite volte superiore alla somma di altrettanti termini che stanno a sinistra di L_n. Ma per il lemma 1 il numero di tutti i termini a sinistra di L_n supera di sole $(s-1)$ volte (cioè di un numero finito di volte) il numero di termini compresi tra L_n e M, e per il lemma 3 i termini diventano tanto minori quanto più sono lontano a sinistra da L_n, allora, complessivamente considerati, tutti i termini compresi tra L_n (incluso) e M (eventualmente escluso) superano di infinite volte tutti i termini a sinistra di L_n.

Analogamente si dimostra che, complessivamente considerati) tutti i termini compresi tra R_n (incluso) e M (eventualmente escluso) superano di infinite volte i termini a destra di R_n, il cui numero supera di sole $(r-1)$ i suddetti (lemma 1).

Concludendo, la somma di tutti i termini compresi tra i limiti L_n e R_n (inclusi), magari senza considerare il termine massimo, supera di infinite volte la somma dei termini che stanno all'esterno di tali limiti. A maggior ragione vale il teorema, includendo nella prima somma il termine massimo. *C.V.D.*[31]

[31] [Come fa giustamente notare il curatore dell'edizione tedesca dell'opera di Jakob Bernoulli, che la traduce dal latino nel 1899, la dimostrazione di Bernoulli è perfettamente rigorosa e

Osservazione. Chi non si fosse familiarizzato con il trattamento dell'infinito potrebbe formulare la seguente obiezione ai lemmi 4 e 5.[32]

Anche nel caso di un valore infinito del numero n i fattori delle espressioni, che rappresentano i rapporti M/L_n e M/R_n, precisamente nr (+/–) n (–/+) 1,2,3,… e ns (–/+) n (+/–) 1,2,3,… hanno lo stesso valore di nr (+/–) n e ns (–/+) n perché i numeri 1,2,3,… spariscono nei singoli fattori rispetto alla parte rimanente. Pertanto tutti questi numeri moltiplicati tra loro producono come risultato (anche a causa del numero infinito di fattori) un numero infinitamente grande. Quindi dalla potenza infinitamente elevata delle frazioni $(rs+s)/(rs-r)$ e $(rs+r)/(rs-s)$ viene sottratto infinitamente molto, con il risultato che si potrebbe ottenere un numero finito.

A queste elucubrazioni non posso contrapporre niente di meglio che eseguire ora effettivamente il calcolo per un valore finito di n. Dimostrerò che anche in una potenza finita del binomio la somma dei termini compresi tra L_n e R_n (inclusi) conserva rispetto alla somma di tutti i termini rimanenti un rapporto, che supera ogni valore c comunque grande sia prefissato. Dimostrato questo, l'obiezione dovrà necessariamente decadere.

A tale scopo assumo per i termini a sinistra di M un valore arbitrario inferiore a $(rs+s)/(rs-r)$, per esempio $(rs+s)/rs = (r+1)/r$ e lo moltiplico per se stesso tante volte (diciamo m volte) finché il prodotto non eguagli o superi $c(s-1)$, cioè sia

$$(r+1)^m/r^m \geq c(s-1).$$

Per determinare m si prendano i logaritmi

$$m \log (r+1) - m \log r \geq \log [c(s-1)].$$

Allora si scelga m in modo tale che

$$m \geq \log [c(s-1)] /[\log (r+1) - \log r].$$

Ora nel lemma 4 si determini il rapporto M/L_n come prodotto delle frazioni

$$\frac{nrs+ns}{nrs-nr+r},\quad \frac{nrs+ns-s}{nrs-nr+2r},\quad \frac{nrs+ns-2s}{nrs-nr+3r},\ \ldots,\ \frac{nrs+s}{nrs}.$$

Ognuno di questi singoli fattori è inferiore a $(rs+s)/(rs-r)$ e si avvicina tanto più a tale frazione, quanto maggiore è n. Pertanto per un n opportuno una di queste frazioni eguaglia $(r+1)/r$. Indicando con m il posto di tale frazione nella serie si ha:

$$\frac{r+1}{r} = \frac{nrs+ns-(m-1)s}{nrs - nr + mr}.$$

---

rispetto alle dimostrazioni successive, più compatte perché utilizzano l'approssimazione asintotica di Stirling al fattoriale, ha il vantaggio di essere elementare. L'argomento di questa osservazione è la vera dimostrazione del teorema di Bernoulli. N.d.T.]

[32] [Questa osservazione è storicamente importante. Bernoulli sa bene che l'oggetto infinito è una novità scientifica, che bisogna rendersi amica (*sich befreunden*). Per trattare l'oggetto infinito non basta la vecchia geometria euclidea. Occorre familiarizzarsi con la topologia dei limiti. N.d.T.]

Di conseguenza

$$n = m + (ms–s)/(r+1)$$
$$nt = mt + (mst–st)/(r+1).$$

Ora affermo che il valore trovato per nt dà l'esponente della potenza alla quale va elevato il binomio $(r+s)$ affinché il termine massimo M dello sviluppo superi il termine limite L_n di almeno $c(s–1)$ volte. Infatti, a questa condizione la frazione m-esima nel prodotto di cui sopra è uguale a $(r+1)/r$ e per ipotesi vale $(r+1)^m/r^m \geq c(s–1)$. Ma tutte le frazioni che precedono l'm-esima nel prodotto sono maggiori di $(r+1)/r$ e tutte le successive sono almeno superiori a 1. Di conseguenza il prodotto di tutti i termini supera certamente $(r+1)^m/r^m$ e a maggior ragione $c(s–1)$. Poiché tale prodotto è uguale a M/L_n segue

$$M > c(s–1)L_n.$$

Inoltre, come è stato dimostrato

$$\frac{M}{L_n} < \frac{L_1}{L_{n+1}} < \frac{L_2}{L_{n+2}} < \frac{L_3}{L_{n+3}} < \dots < \frac{L_n}{L_{2n}},$$

e di conseguenza vale anche

$$L_1 > c(s–1)L_{n+1},$$
$$L_2 > c(s–1)L_{n+2},$$
$$L_3 > c(s–1)L_{n+3},$$
$$\dots$$
$$L_n > c(s–1)L_{2n},$$

e sommando

$$L_1+L_2+L_3+ \dots +L_n > c(s–1)[L_{n+1}+L_{n+2}+L_{n+3}\dots].$$

Ma poiché i termini da M in su decrescono costantemente e il numero dei termini a sinistra di L_n non supera di $(s–1)$ volte il numero dei termini $L_1, L_2, L_3, \dots, L_n$, segue che

$$L_1+L_2+L_3+ \dots +L_n > c[L_{n+1}+L_{n+2}+L_{n+3}\dots],$$

dove nelle parentesi di destra si trovano ora tutti i termini a sinistra di L_n.

Allo stesso modo procedo con i termini a destra di M. Assumo ora il rapporto

$$(s+1)/s < (rs+r)/(rs–r)$$

e, determinando il valore di m, trovo che

$$(s+1)^m/s^m \geq c(r–1)$$

ossia

$$m \geq \log [c(r-1)]/[\log(s+1) - \log s].$$

Quindi, nella successione delle frazioni

$$\frac{nrs+nr}{nrs-ns+s}, \quad \frac{nrs+nr-r}{nrs-ns+2s}, \quad \frac{nrs+nr-2r}{nrs-ns+3s}, \ldots, \frac{nrs+r}{nrs},$$

che determinano il rapporto M/R_n, metto l'm-esima

$$\frac{s+1}{s} = \frac{nrs+nr-(m-1)r}{nrs - ns + ms},$$

da cui si ottiene

$$n = m + (mr-r)/(s+1)$$

e

$$nt = mt + (mrt-rt)/(s+1).$$

Da qui si dimostra esattamente come prima che elevando il binomio $r+s$ a questa potenza nt il termine massimo M supera di più di $c(r-1)$ volte il termine limite di destra R_n e che la somma di tutti i termini compresi tra M (escluso) e R_n (incluso) e di più di c volte superiore al numero di tutti i termini rimanenti, quando il binomio r+s sia elevato alla potenza il cui esponente sia uguale al maggiore dei due termini

$$mt + (mst-st)/(r+1) \text{ e } mt + (mrt-rt)/(s+1).$$

Si è così trovata la potenza finita che possiede la caratteristica desiderata. *C.V.D.*

A questo punto finalmente consegue il teorema, la cui dimostrazione richiede solo l'applicazione dei lemmi stabiliti alle condizioni stabilite. Per evitare penose trascrizioni dico *favorevoli* (o *fruttuosi*) i casi in cui si può verificare un evento e *sfavorevoli* (o *infruttuosi*) i casi in cui non si può verificare. Così dico *favorevoli* i tentativi e le osservazioni in cui si verifica uno dei casi favorevoli e *sfavorevoli* quelli in cui si osserva un caso sfavorevole.

Teorema [***debole dei grandi numeri o della buona approssimazione***]. *Il numero dei casi favorevoli rispetto al numero di casi sfavorevoli sia esattamente o approssimativamente come r/s, quindi rispetto al numero di tutti i casi come* $r/(r+s) = r/t$, *dove* $t = r+s$, *essendo tale rapporto compreso tra i limiti* [superiore] $(r+1)/t$ *e* [inferiore] $(r-1)/t$. *Dimostriamo che si possono fare osservazioni in numero tale che diventa più probabile quanto si vuole, diciamo c volte, che il rapporto tra le osservazioni favorevoli e tutte le osservazioni disponibili stia all'interno di tali limiti piuttosto che all'esterno, cioè non risulti né maggiore di* $(r+1)/t$ *né minore di* $(r-1)/t$.[33]

[33] [In notazione moderna il teorema si scrive: $\lim_{t\to\infty} P\{|(r/t) - p|\}=1$ (Cfr. B.V. Gnedenko, *Teoria della probabilità*, trad. G. Grosso e C. Maffei, Editori Riuniti Roma 1979, p. 94). A parole De Finetti formula il teorema così: *la media aritmetica dei guadagni aleatori non correlati ha previsione quadratica → 0 e probabilità → 0 di avere valore assoluto*

Dimostrazione. Si ponga uguale a nt il numero delle osservazioni programmate e ci si chieda quanto grande sia la speranza che siano favorevoli rispettivamente tutte le osservazioni, tutte le osservazioni meno una, meno due, meno tre, meno quattro ecc. Poiché per ipotesi in ogni osservazione sono possibili t casi, di cui r favorevoli e s sfavorevoli, e ogni caso della prima osservazione si può combinare con ogni caso della seconda e di nuovo i casi combinati si possono associare con ogni caso della terza, quarta… osservazione, è evidente che devono essere applicate la regola che segue all'osservazione all'esercizio XII della prima parte e il suo secondo comma.[34] In questo modo si trova che la speranza di nessuna osservazione sfavorevole è pari a r^{nt}/t^{nt}, di un'osservazione sfavorevole $\binom{nt}{1}r^{nt-1}s/t^{nt}$, di due osservazione sfavorevoli $\binom{nt}{2}r^{nt-2}s^2/t^{nt}$, di tre osservazioni sfavorevoli $\binom{nt}{3}r^{nt-3}s^3/t^{nt}$, ecc. Tralasciando il denominatore comune t^{nt}, il grado di probabilità o il numero di casi in cui si può verificare che tutte le osservazioni, tutte meno una, tutte meno due, tutte meno tre… siano favorevoli è rispettivamente

$$r^{nt},\ \binom{nt}{1}r^{nt-1}s,\ \binom{nt}{2}r^{nt-2}s^2,\ \binom{nt}{3}r^{nt-3}s^3\ \ldots$$

Queste espressioni sono i termini nella nt-esima potenza del binomio $r+s$, già trattato nei lemmi precedenti, e quindi quel che ne consegue è chiaro.

Per il modo di formazione di questa serie è evidente che il numero di casi in cui nr osservazioni sono favorevoli e le restanti ns sfavorevoli è esattamente uguale al massimo M (lemma 3), dove ns termini lo precedono e nr lo seguono. È altrettanto evidente che il numero di casi in cui di tutte le nt osservazioni $nr+n$, rispettivamente $nr-n$, sono favorevoli e le restanti sfavorevoli, sono dati dai termini L_n e R_n rispettivamente, in quanto distano dal termine massimo M di n termini da entrambi i lati. Di conseguenza il numero di tutti i casi in cui tra tutte le nt osservazioni sono presenti non più di $nr+n$ e non meno di $nr-n$ osservazioni favorevoli è pari alla somma di tutti i termini dello sviluppo di $(r+s)^{nt}$, che stanno tra L_n e R_n (limiti inclusi). Il numero di tutti i casi rimanenti, in cui sono presenti più di $nr+n$ o meno di $nr-n$ osservazioni favorevoli, è pari alla somma dei termini rimanenti dello sviluppo della potenza, che stanno fuori dall'intervallo compreso tra L_n e R_n. Poiché [l'esponente del] la potenza può essere scelto così grande che la somma dei termini compresi tra L_n e R_n (limiti inclusi) risulti superiore di c volte alla somma dei termini che stanno fuori dai suddetti limiti, segue il principio:

Si possono effettuare tante osservazioni che il numero di casi in cui il rapporto tra casi favorevoli e casi totali non superi i valori limite di $(nr+n)/nt$ e $(nr-n)/nt$ (o $(r+1)/t$

$> \varepsilon$ *(positivo prefissato arbitrario)*. (Cfr. B. de Finetti, *Teoria delle probabilità*, vol. II. Einaudi, Torino 1970, p. 386). Si noti la contrapposizione singolare/plurale, corrispondente all'approccio rispettivamente oggettivistico/soggettivistico al calcolo probabilistico. Essendo un teorema di buona approssimazione, quello di Bernoulli è un enunciato soggettivistico, fortemente connotato in senso epistemico. N.d.T.]

[34] "Si formi il prodotto dei numeri di casi in cui lo scopo da raggiungere è raggiunto nei prescritti lanci e dei casi in cui non viene raggiunto, perché non può essere raggiunto, e lo si divida per il numero di tutti i casi nei lanci complessivi: il quoziente è il valore della speranza cercata". Qui Bernoulli anticipa la nozione di spazio prodotto. Per la nozione di speranza riporto il primo teoroma del *De ludo aleae* di Christian Huyghens, commentato da Bernoulli. "Se mi aspetto la somma a o la somma b, di cui la prima è altrettanto facile da ottenere della seconda, allora il valore della mia speranza è $(a+b)/2$.

*e (r–*1*)/t) sia di c volte maggiore della somma dei casi rimanenti, cioè sia c volte più probabile che il rapporto tra casi favorevoli e casi totali non superi i limiti (r+*1*)/t e (r–*1*)/t, piuttosto che li superi. C.V.D.*

Applicando il teorema a casi numerici si riconosce facilmente che tanto più grandi si prendono i numeri r, s e t (con r/s costante), tanto più si avvicinano i limiti $(r+1)/t$ e $(r-1)/t$. Se per esempio il rapporto $r/s = 3/2$, non pongo $r = 3$ e $s = 2$, ma $r = 30$ e $s = 20$, con $t = r+s = 50$, o $r = 300$ e $s = 200$, con $t = 500$. Nel primo caso i limiti sono:

$$(r+1)/t = 31/50 \text{ e } (r-1)/t = 29/50.$$

Prendo ora $c = 1000$. m e nt si determinano come nell'osservazione precedente per i termini a sinistra di M:

$$m \geq \log [c(s-1)]/[\log (r+1) - \log r] = 4{,}2787536/0{,}0142405 < 301,$$

$$nt = mt + (mst-st)/(r+1) < 24728$$

e per i termini a sinistra di M:

$$m \geq \log [c(r-1)]/[\log (s+1) - \log s] = 4{,}4623980/0{,}0211893 < 211,$$

$$nt = mt + (mrt-rt)/(s+1) < 25550.$$

Per il teorema appena dimostrato su 25550 osservazioni è 1000 volte più probabile che il rapporto tra casi favorevoli e totali cada tra i limiti 31/50 e 29/50 piuttosto che fuori. Allo stesso modo, ponendo $c = 10.000$, si trova che sono necessarie 31258 osservazioni perché sia 10.000 volte più probabile che il rapporto cada tra i suddetti limiti anziché fuori. Con $c = 100.000$ sono necessarie 36966 osservazioni perché le probabilità siano di 100.000 a 1 e così via all'infinito, aggiungendo 5708 osservazioni per ogni unità logaritmica di probabilità.

Se si osservassero per tutta l'eternità tutti gli eventi – nel qual caso la probabilità dovrebbe trasformarsi in certezza completa – si troverebbe che tutto nel mondo avviene su determinate basi e secondo una determinata conformità a leggi, tanto da essere costretti ad assumere una certa necessità o per così dire un fato anche per cose che avvengono apparentemente in modo casuale. Non so se a questo proposito Platone parlerebbe nella propria dottrina di ciclo generale delle cose, in cui dopo un percorso di innumerevoli anni tutto ritorna nello stato originario.[35]

Traduzione dal tedesco di Antonello Sciacchitano

---

[35] [Così, alla fine della propria notevole innovazione, l'uomo di scienza dimostra tutta l'innata resistenza alla scienza, da cui era probabilmente già abitato, tornando a essere l'uomo prescientifico che crede al determinismo assoluto e alle frottole sul "grande anno", secondo Empedocle o Eraclito. (Il termine usato da Berrnoulli per il "grande ritorno" è *apocatastasi*, che compare negli scritti pseudoplatonici di Axioco.) Torna alla mente il teorema di Althusser, secondo il quale la filosofia spontanea dell'uomo di scienza è un rozzo idealismo. Abbandonata l'epistemologia, l'uomo di scienza si offre a servizio del potere come ontologo tecnologicamente aggiornato. N.d.T.]