Paradossi sulle medie (*)

Sull'argomento delle medie si possono dire molte cose interessanti e istruttive. Interessanti sono, come vedremo sia pure su pochi e semplici esempi, molte questioni ove trovano applicazione. Ma più ancora sarà istruttivo, e su ciò insisteremo maggiormente, osservare come e perchè sia facile cadere in errori, che sembrano cose naturali e che sembra paradossale dover correggere nel senso esatto.

Un semplice problema (dato in una gara matematica, a Trieste) chiedeva la velocità media su un certo percorso date le velocità mantenute nella prima e nella seconda metà di esso. Per semplicità, supponiamo che siano di 10 e di 50 km/h, e che il percorso totale (che del resto è irrilevante) sia di 100 km. Purtroppo, la risposta più « naturale » risulta 30 km/h (perchè 30 è la semisomma di 10 e 50, $30=(10+50)/2$, ossia la media aritmetica, che è in genere considerata la *media* senz'altro, media per antonomasia, di due o più numeri). Basta invece riflettere che il tempo occorrente è di 5 ore per percorrere i primi 50 km e di un'ulteriore ora per gli altri 50; in tutto sono 6 ore per 100 km e la velocità è di $(100/6)\mathrm{km/h}=16{,}66\ \mathrm{km/h}$.

Si può osservare che il calcolo così eseguito consiste nel fare la media (aritmetica) non sulle velocità ma sui reciproci:

$$\frac{1}{2}\left(\frac{1}{10}+\frac{1}{50}\right)=\frac{1}{2}\cdot\frac{6}{50}=\frac{6}{100},$$

di cui il reciproco è $100/6=16{,}66$.

(*) Questo articolo riproduce il contenuto di una conversazione tenuta dal prof. B. de Finetti a studenti delle scuole secondarie del « Club Matematico Mathesis » di Roma il 1° aprile 1966.

Questa media (reciproco della media aritmetica dei reciproci) si dice media armonica. Al nostro problema, così come è stato posto, risponde la media armonica; risponderebbe invece effettivamente la media aritmetica (velocità media = 30 km/h) se le due velocità di 10 e di 50 km/h fossero mantenute nella prima e seconda metà di *durata* del viaggio (anzichè del *percorso*): ad es. in una prima e una seconda *ora*.

Analogamente, date delle resistenze (elettriche), la loro media è la media aritmetica se si pensa di usarle in serie e quella armonica se si pensa di usarle in parallelo; nei due diversi casi il passaggio di corrente è infatti lo stesso che se esse avessero tutte la medesima resistenza, ma data dall'una o dall'altra media a seconda del modo di impiego. Occorre quindi sempre fare attenzione, caso per caso, al significato che deve avere la media e alla forma che deve di conseguenza applicarsi. Quanto sia facile, altrimenti, cadere in errore, lo provano molti episodi, come il seguente. Uno dei metodi impiegati (specie all'epoca in cui mancavano rilevazioni statistiche più dirette) per stimare la « ricchezza » di un paese consisteva nel basarsi sui dati della ricchezza trasmessa per eredità in un singolo anno, moltiplicandoli per l'intervallo devolutivo medio (durata tra due successivi trapassi per eredità); ma a tal fine la media dev'essere quella armonica, e solo dopo qualche tempo è stato scoperto che l'aver usato, in precedenza, la media aritmetica costituiva un errore.

Si può anche dire subito che l'errore è sempre nello stesso verso: è infatti vero sempre che (come risultava nel nostro esempio) la media armonica è minore della media aritmetica. Più in generale, si può dimostrare che (con riferimento ad altre tre forme di media che accenneremo), le medie

armonica - geometrica - aritmetica - quadratica - cubica

vanno (in tale ordine) dalla più piccola alla più grande. Non vogliamo qui svolgere la pur semplice trattazione richiesta per simili dimostrazioni, che potrebbero del resto facilmente estendersi a casi più generali (che accenneremo). Per attirare l'attenzione su poche cose che potranno forse rimanere impresse dopo rapidi cenni in una conversazione come questa, sembra preferibile illustrare la cosa in modo euristico sull'esempio più semplice possibile e scritto nella forma più semplice possibile, dicendo poi se e come le conclusioni valgano, esattamente o approssimativamente, più in generale.

Consideriamo cioè sempre la media di due grandezze positive (due sole e non più; medie semplici, cioè con «pesi» uguali e non «ponderate» come ad es. se alcuni valori si dovessero pensare ripetuti più volte), e le indicheremo $1-x$ ed $1+x$ (quindi x sarà positivo e minore di 1); ciò vuol dire semplicemente che prendiamo come unità di misura la media aritmetica $([(1-x)+$ $+(1+x)]/2=1)$.

La media armonica è allora

$$1 : \left[\frac{1}{2}\left(\frac{1}{1-x}+\frac{1}{1+x}\right)\right]=\frac{2(1-x)(1+x)}{(1+x)+(1-x)}=(1-x)(1+x)=1-x^2.$$

La media quadratica è quella che si ottiene facendo la media aritmetica dei quadrati ed estraendone la radice quadrata; analogamente è definita la media cubica (cubi, radice cubica); la media geometrica si ottiene facendo il prodotto ed estraendo la radice (quadrata se, come nel nostro esempio, i termini sono due; in genere, n-esima se sono n). Per il momento bastino queste definizioni formali, tanto per ottenere subito formule analoghe a quella del primo caso; naturalmente riprenderemo subito le considerazioni sul significato concettuale, le sole che contino realmente.

Le formule che seguono valgono però solo approssimativamente, se x è piccolo; impiegheremo infatti le formule approssimate

$$\sqrt{1+a}=1+\frac{1}{2}a, \qquad \sqrt[3]{1+a}=1+\frac{1}{3}a$$

(si verifica subito che $(1+a/2)^2$ e $(1+a/3)^3$ danno $1+a$, a meno di termini in a^2 e a^3, trascurabili se a è piccolo (1)).

(1) Nel nostro caso l'approssimazione è anche migliorata dal fatto che la x vi entra al quadrato. Per averne un'idea basti confrontare i valori esatti e approssimati per $x=1/2$, che è già un valore alto per lo scostamento tra i due valori $1-x$ ed $1+x$, e tener conto che l'approssima-

Per la media quadratica si ha:

$$\frac{1}{2}[(1-x)^2+(1+x)^2]=1+x^2, \qquad \sqrt{1+x^2}\cong 1+\frac{1}{2}x^2;$$

per la media cubica:

$$\frac{1}{2}[(1-x)^3+(1+x)^3]=1+3x^2, \qquad \sqrt[3]{1+3x^2}\cong 1+x^2;$$

per la media geometrica:

$$(1-x)(1+x)=1-x^2, \qquad \sqrt{1-x^2}\cong 1-\frac{1}{2}x^2.$$

Le cinque medie considerate risultano quindi, in base alle formule approssimate, non solo crescenti ma equidistanti (differenza tra ciascuna e la successiva, $x^2/2$). Come ciò valga in generale (anzichè solo nel caso dell'esempio, con $1-x$ ed $1+x$) accenneremo poi.

Quanto al significato delle nuove medie introdotte, se ne intuisce subito l'adeguatezza a rispondere a problemi geometrici. Se abbiamo dei quadrati, o dei cubi, o dei dischi, o delle sfere, di diversa grandezza (data dal lato, risp. dal raggio, o da una qualunque misura lineare per una qualunque collezione di figure o solidi simili), e ci chiediamo quale sia il valore comune che dovrebbero avere tutti i lati (o raggi, ecc.) affinchè rimanesse invariata la somma delle aree, è chiaro che a ciò risponde la media quadratica. Se invece ci interessiamo dei volumi (nei casi ove ha senso anche il volume) interverrà la media cubica. Per la media geometrica (di due o di tre termini) si ha l'ovvio significato di lato del quadrato di area uguale a quella del rettangolo di lati dati, risp. di lato del cubo di volume uguale a quella del prisma (rett.) di lati dati.

Per la media quadratica va anche notata un'altra interpretazione, importante in meccanica ma anche atta a illustrare l'importanza che le viene attribuita nella statistica. Se consideriamo diverse masse (che supponiamo allineate, su una retta, p. es. infilzate su una sbarretta), il momento d'inerzia di ciascuna di esse,

---

zione migliora rapidamente quanto più x diminuisce:

media quadratica	valore esatto $=1{,}11803$,	approssimato	$=1{,}125$.
» cubica	» » $=1{,}20507$	»	$=1{,}25$.
» geometrica	» » $=0{,}86603$,	»	$=0{,}875$.

rispetto a un punto qualunque, è il prodotto della massa per il quadrato della distanza, e, per il sistema, è la somma di tali prodotti; se dividiamo per la massa totale ed estraiamo la radice otteniamo la media quadratica delle distanze, ossia quella distanza (detta «raggio giratore») alla quale potrebbe venir concentrata tutta la massa conservando inalterato il momento d'inerzia. La cosa riesce più espressiva ricordando che, se si fa rotare il sistema attorno al punto scelto, la forza viva (a parità di velocità di rotazione, o velocità angolare) è proporzionale al momento di inerzia. Si può quindi dire che, agli effetti della forza viva (cioè pensando al sistema in funzione di «volano»), sarebbe indifferente concentrare tutta la massa a detta distanza (raggio giratore). Particolare interesse ha il momento rispetto al baricentro (e si ricorderà che esso è il minimo); se indichiamo con σ il giratore rispetto al baricentro (di cui indichiamo l'ascissa con m), possiamo dire in particolare che, collocando la massa metà in $m-\sigma$ e metà in $m+\sigma$, si ha un sistema che è identico al precedente come caratteristiche di baricentro e momenti d'inerzia e proprietà meccaniche che comunque ne dipendono.

Nella statistica, questo stesso numero σ, che è la media quadratica delle distanze dalla media aritmetica, si dice *scarto quadratico medio*, o brevemente *scarto standard* (e σ^2 si dice *varianza*); è chiaro che dà una misura (e sotto molti aspetti la misura più opportuna) della variabilità, ossia del minore o maggore addensamento dei diversi valori di una distribuzione attorno al valor medio (media aritmetica, baricentro). Per molti aspetti la conoscenza di m e σ può sostituire la conoscenza completa della distribuzione (cioè di tutti i singoli valori e rispettivi pesi); in particolare, se σ non è troppo grande (rispetto ad m) e sotto altre condizioni generiche su cui non è il caso di soffermarsi qui, risultano valide — sempre come approssimazioni — le formule stabilite precedentemente per il caso di due valori $1-x$ ed $1+x$. Poichè in generale la massa sarà invece concentrata nei due valori $m-\sigma$ ed $m+\sigma$, dovremo sostituire 1 con m ed x con σ/m (per avere $m \pm \sigma = m(1 \pm x)$); le espressioni saranno allora

$$m-\sigma^2/m, \qquad m-\frac{1}{2}\sigma^2/m, \qquad m, \qquad m+\frac{1}{2}\sigma^2/m, \qquad m+\sigma^2/m,$$

per le medie

armonica - geometrica - aritmetica - quadratica - cubica [2].

La media geometrica risponde al caso in cui i termini abbiano il significato, ad es., di fattori di accrescimento (o diminuzione). Se ad es. una popolazione si è accresciuta, in successivi anni, nei rapporti da 1 a 1,0231, 1,0164, 1,0193, il fattore medio (quello che, conservandosi invariato nei tre anni, avrebbe dato il medesimo accrescimento nel triennio) sarebbe dato ovviamente dalla radice cubica del prodotto dei tre fattori.

---

(2) Si potrebbe osservare che tutte queste medie rientrano nel caso più generale delle «medie di potenze», consistenti, per ogni dato esponente α, nel fare la media aritmetica di valori elevati ad α, e quindi elevando il risultato alla potenza $1/\alpha$. Per $\alpha = 1, 2, 3$ abbiamo le medie aritmetica, quadratica, cubica; per $\alpha = -1$ si ha la media armonica; la media geometrica si ha come caso limite per $\alpha = 0$ (ponendo $\alpha = 0$ la cosa non ha senso; per $\alpha \simeq 0$ si ha però $x^\alpha \simeq 1 + \alpha \log x$ e a ciò corrisponde il fatto che la media geometrica si può pensare ottenuta facendo la media aritmetica dei logaritmi e prendendone poi l'antilogaritmo). Più in generale, si possono ottenere medie trasformando analogamente la media aritmetica con funzioni quasi ad arbitrio (p. es., prendendo una funzione esponenziale, si può fare la media aritmetica degli esponenziali e prenderne poi il logaritmo); si ottengono così tutte le medie «associative», ma ne esistono anche di «non associative» (come la media antiarmonica, che si ottiene dividendo per la media aritmetica il quadrato della media quadratica, e che dà la lunghezza ridotta del pendolo composto; si dice «non associativa» perchè non si può calcolarla sostituendo alcuni valori con la loro media).

Nel caso delle medie associative (come le 5 specificamente considerate) il confronto si riduce a una semplice ed elegante questione sul grafico della funzione che le definisce (p. es., in detti casi, $1/x$, $\log x$, x, x^2, x^3): la media è maggiore per la funzione col diagramma «più concavo in senso relativo». Se, per semplicità di confronto, modifichiamo inessenzialmente le funzioni affinchè i diagrammi passino per il punto (1,1) e siano ivi tangenti alla bisettrice ($y = x$), esse saranno $2 - 1/x$, $1 + \log x$, x, $(1 + x^2)/2$, $(2 + x^3)/3$; in generale (per esponente α qualunque) $1 + (x^\alpha - 1)/\alpha$ (da considerare per $x \geq 0$). Disegnando i diagrammi si vedrebbe subito come essi siano concavi (verso l'alto) per $\alpha > 1$, e tanto più quanto più α è grande, convessi per $\alpha < 1$ (e tanto più quanto minore è α: quando α diviene negativo, ciò significa tanto più quanto più esso è grande in valore assoluto). Il fatto segnalato ora, circa il diagramma, non esprime però tutto ciò che andrebbe detto per completare la spiegazione della condizione; serve solo a «dare un'idea» della sua natura.

Ma è stato anche osservato (in certo senso da GALILEO, e spesso dai pratici, cui comunque in genere è noto) che la media geometrica è la più indicata per ricavare un valore attendibile da valori stimati di grandezze (positive). È noto ad es. agli artiglieri (3) che, se la distanza del bersaglio viene stimata ad occhio da alcuni tra loro, e si prende per buona la media aritmetica delle stime, la valutazione risulta in genere errata per eccesso; perciò usano correggerla alquanto in meno. Come si spiega tale esigenza? È naturale vedere una giustificazione nel fatto che la probabilità di un errore dipende non tanto dalla sua grandezza assoluta quanto piuttosto dalla grandezza relativa (come appunto sosteneva GALILEO). Se una distanza è, in realtà, di 10 km, potrà essere circa ugualmente probabile un errore di un km in più o in meno (9 o 11), ma non di 5 km in più o in meno (5 o 15; grosso modo si potrà ritenere ugualmente probabile un errore di 5 in meno o 10 in più — stime di 5 o 20 — trattandosi sempre di errore nel rapporto di 1 a 2 in un verso o nell'altro); se poi pensiamo a un errore in meno di 10 km (il bersaglio è qui!) o più (il bersaglio è dalla parte opposta, dietro di noi rispetto alla direzione in cui lo guardiamo!), è ovvio che ciò è da escludere (mentre errori in più, anche di 10 o 20 km ecc., non sono impossibili). Ammettendo una distribuzione di errori simmetrica in scala relativa (meglio: in scala *logaritmica*, per chi conosce tale termine), risulta appunto che per la valutazione appropriata si richiede la media geometrica. Volendo ottenerla applicando alla media aritmetica la «correzione in meno» secondo la pratica degli artiglieri, si potrebbe far ricorso alla formula m. geom.$\cong m-\sigma^2/2m$, sia pure stimando grossolanamente σ come «ordine di grandezza degli scarti fra le singole stime e la loro media» (un po' più della media aritmetica degli scarti assoluti). P. es., date quattro stime in km 6, 8, 11, 15, la m. aritm. è $m=10$ km, quella geom. 9,4 km, lo scarto standard $\sigma=3{,}4$ km ($\sigma^2=11{,}5$ km^2), e la m. geom. data dalla formula approssimata $10-(11{,}5/20)=10-0{,}6=9{,}4$ (ed anche stimando grossolanamente $\sigma=3$ o $\sigma=4$, cioè $\sigma^2=9$ o $\sigma^2=16$, la correzione sarebbe stata di $-0{,}45$ o $-0{,}8$, sempre sufficientemente indicativa dell'ordine di grandezza).

---

(3) Questo ed altri esempi mi sono stati comunicati, come cose di cui ebbe direttamente conoscenza o esperienza, dal prof. G. TAGLIACARNE.

Abbiamo visto così, sia pure sfiorandole di sfuggita, diverse osservazioni matematicamente e praticamente interessanti, e molte altre si potrebbero aggiungere se s'intendesse andare più a fondo. Ma altre osservazioni, ancor più fondamentali, si possono fare dal punto di vista concettuale, sia specificamente con riferimento alla nozione di media, sia più generalmente come avvertimenti per fare attenzione a ben ragionare.

Pensando a una « media », non si deve pensare a una convenzione matematica, a una formuletta, al risultato di certe operazioni, scelte chissà perchè. Si deve invece pensare a un problema ben determinato, in cui interessa considerare un aspetto ben determinato, e conviene sapere quale valore comune si potrebbe dare a certe grandezze (più o meno differenti tra loro) volendo che, *per riguardo all'aspetto che interessa, il risultato rimanesse invariato*. È questa, concettualmente, la definizione di media data da CHISINI; è ad essa che si sono ispirate tutte le precedenti esemplificazioni, come si sarà notato, e comunque conviene ora notare esplicitamente come fatto fondamentale. Esprimendo tale concetto in forma matematica, tutta la trattazione sulle medie diventa significativa ed elegante (4).

Ma da cosa deriva l'impressione di qualcosa di « paradossale » che in genere fanno le osservazioni riferite? Gli è che, in genere, l'abitudine ad usare e sentir usare la parola « media » come se fosse dotata di per sè di significato suo proprio, intrinseco, assoluto, induce a pensare come se un tale significato esistesse davvero. Rinunciare a siffatta illusione « nominalistica », ad una fiducia così cieca nell'esistenza di un significato cui sembrava inutile pensare perchè l'esistenza della parola si ritiene spesso sufficiente garanzia al riguardo, pensare che invece sia

(4) L'articolo originale di O. CHISINI, « Sul concetto di media », apparve in *Periodico di Matematiche* (1929), n. 2; l'argomento fu sviluppato in un articolo dallo stesso titolo da B. DE FINETTI in *Giorn. Ist. Ital. Attuari* (1931), pp. 369-396. Una presentazione elementare (destinata intenzionalmente a studenti di Ist. Tecnici Comm.) si trova in B. DE FINETTI e F. MINISOLA, *La matematica per le applicazioni economiche*, ed. Cremonese, Roma, 1961, (Cap. V, Nozioni sulle distribuzioni statistiche). Non sempre la nozione di media viene intesa in senso così specifico; a volte è un valore « da cui le grandezze date si scostano poco »; anche in tal caso esso dipende però dallo scopo se la misura di tale « discostarsi » è scelta appropriatamente in nesso all'aspetto che si vuole studiare.

necessario pensare, ed anzi che risulti necessario sempre pensarci e ripensarci caso per caso, in relazione al problema e allo scopo, è uno choc contro le abitudini di pigrizia mentale istillate dall'abitudine al linguaggio comune e al suo uso incontrollato. Occorre invece sempre diffidare, chiedersi se veramente una parola ha un senso ben definito (e non tanto isolatamente, chè isolata non serve, ma come ingrediente in affermazioni di tipo da precisare). Di tale esigenza così dice PAPINI in un profilo di CALDERONI: « a lui premeva insegnare con quali cautele e quali accorgimenti si possa giungere a costruire proposizioni che abbiano un senso ». E nella stessa direzione ci viene l'avvertimento di GOETHE: « l'uomo ritiene per solito, quando sente nient'altro che parole, che vi sia dentro qualcosa su cui ci sia da pensare » (5).

Per aggravare queste deficienze della mente, giova la lettura di scritti « filosofici » (almeno in genere); per tentare di affrancarsene giovano letture e riflessioni di critica scientifica o filosofico-scientifica (almeno in genere). Giova ad es. riflettere sul significato delle definizioni, in matematica e in fisica; in particolare, nel campo delle scienze fisiche, sull'esigenza del carattere operativo delle definizioni. Pensare al paradosso della « simultaneità » nell'ambito della teoria della relatività (se ne è stata acquisita notizia).

Nel campo della probabilità e della statistica è particolarmente facile trovare apparenti paradossi atti a stimolare utili riflessioni al riguardo. La nozione di media, oltre a quelli del tipo già accennato con qualche esemplificazione, ce ne fornisce altri, riconducibili a un altro concetto che va inteso con maggior cautela di quanto non si pensi: quello di « scelta a caso ».

Qual'è il numero medio di componenti di una famiglia? Si tratti di tutte le famiglie di un paese (p. es., Italia, ultimo censimento) oppure di quelle abitanti attualmente in un dato caseggiato, la risposta è facile: basta dividere il numero degli individui che vi appartengono per il numero delle famiglie. In altro

---

(5) Il saggio sul filosofo pragmatista MARIO CALDERONI fa parte del volume di GIOVANNI PAPINI, *Stroncature*; del CALDERONI cfr. due scritti sul Pragmatismo (in collaborazione con G. VAILATI) in GIOVANNI VAILATI, *Scritti*, Seeber, Firenze, 1911 (nn. CCX e CCXI, pp. 920-932 e 933-941). La citazione di GOETHE (credo dal *Faust*) nel testo originale suona:

« Gewöhnlich glaubt der Mensch, wenn er nur Worte hört,
es müsse sich dabei auch etwas denken lassen ».

modo, se è noto il numero di famiglie composte di un solo individuo, o di due, di tre, ecc., fino al massimo (sia p. es. 9), lo stesso risultato si otterrà col calcolo che risulta chiaro sull'esempio illustrato nella tabella che segue.

N. di comp.	N. famiglie	% fam.	N. dei comp.	% comp.	Id. × N. di comp.
1	9	9 %	9	3 %	3
2	24	24 %	48	16 %	32
3	41	41 %	123	41 %	123
4	21	21 %	84	28 %	112
6	3	3 %	18	6 %	36
9	2	2 %	18	6 %	54
Tot.	100	100 %	300	100 %	370

Il numero medio di componenti per famiglia è 300/100=3 (naturalmente, il fatto che i numeri 100 e 300 siano tondi è accidentale; in effetti l'esempio è stato costruito così per semplicità).

Possiamo dire, di conseguenza, che ogni individuo di tale gruppo ha in media altri due componenti nella sua famiglia? A prima vista la deduzione sembra lapalissiana, ma si tratta ancora di un abbaglio sul senso assoluto del termine « media ». Se pensiamo non più alla famiglia ma ai componenti, è ovvio (e si vedano i dati nella colonna « % comp. ») che i componenti di famiglie numerose sono in proporzione assai più che le famiglie, e viceversa (p. es. al 9 % di famiglie con un componente appartiene solo il 3 % degli individui, mentre al 2 % di famiglie con 9 componenti appartiene il 6 % degli individui). Riferirsi agli individui significa, insomma, contare ogni famiglia tante volte quanti sono i suoi membri, e in tal modo esse sono pure 300 e gli individui (contati ripetutatmente allo stesso modo) sono 1110; il numero medio è 1110/300=3,7 (o indifferentemente il calcolo si può eseguire sulle percentuali, come mostra la tabella). Ogni individuo ne ha in media altri 2,7 (non 2) in famiglia.

Dicevamo che l'errore sta nel non far attenzione al significato relativo di « scelta a caso »: se s'intende (come si deve intendere) scelta che *dia uguale probabilità ad ogni « unità statistica »*, il senso varia a seconda di quale unità statistica s'intenda considerare. Se è la famiglia, va bene il primo calcolo (media=3); ma se si sceglie « a caso » un individuo, ossia la famiglia cui appartiene un « individuo scelto a caso », la scelta della famiglia non è più fatta « a caso » bensì in modo *distorto* a favore delle famiglie numerose. È come estrarre a sorte imbussolando per ogni famiglia o un solo gettone (primo caso) oppure tanti quanti sono i suoi componenti (secondo caso); e altri casi (tutti diversi e diversamente distorti l'uno rispetto all'altro) si avrebbero ad es. scegliendo una famiglia per ogni casa, oppure una per ogni iniziale del cognome, ecc. (il che favorirebbe le famiglie di case con uno o pochi appartamenti, oppure con nomi dall'iniziale rara, ecc.).

Lo stesso paradosso s'incontra in esempi apparentemente lontani.

Pensiamo a un segmento suddiviso « a caso » in un certo numero n di parti, p. es. in $n=8$; ammaestrati dalle osservazioni precedenti spieghiamo che la « suddivisione a caso » s'intende ottenuta mediante $n-1$ punti scelti indipendentemente a caso sul segmento, cioè ciascuno scelto in modo che ogni intervallo abbia probabilità di contenerlo proporzionale alla sua lunghezza, quale che sia la posizione degli altri. Nel nostro esempio, ne avremo scelti $8-1=7$. Ciascuna delle 8 parti, scegliendola « a caso », avrà come lunghezza media $1/8$ (della lunghezza del segmento; per semplicità poniamola $=1$); ciò varrà ad es. per la 3^a da sinistra, per « quella che segue immediatamente a destra la più breve » (6), ... ma attenzione. Non si dovrà mai astenersi dal chiedersi se veramente una certa modalità di « scelta a caso » è o non è tale, se è o non è *informativa* nel senso di fornire qualche indizio sulla lunghezza della parte scelta. È chiaro ad es. che se dicessi di scegliere *la più lunga*, oppure *la più breve*, delle 8 parti, la lunghezza della prima (lunghezza media) sarebbe certamente superiore a $1/8$ (e l'altra inferiore). Ma cosa avverrà se scelgo una delle 8 parti scegliendo a caso un punto del segmento totale, e prendendo la parte che lo contiene? Probabilmente questo sembra, a prima vista, un ottimo metodo di scelta a caso, ma è analogo a quello della scelta di una famiglia scegliendo a caso uno degli individui, e ciò dovrebbe mettere in sospetto. In tal modo

(6) Beninteso, se la più breve fosse l'ultima dovremmo considerare come « seguente a destra » la prima a sinistra.

un pezzo lungo ha maggior probabilità di essere scelto, perchè ha maggior probabilità che il punto « scelto a caso » vi appartenga. Ed infatti basta un ragionamento sorprendentemente semplice per dare la risposta precisa.

Il nuovo punto scelto a caso per individuare la parte da scegliere altro non è che un 8° punto « a caso » dopo e come gli altri 7; con esso otteniamo una suddivisione in 9 parti, ciascuna di lunghezza media 1/9; la parte che viene scelta (fra la 8 precedenti) consta di due delle nuove parti, ed ha quindi lunghezza media 2/9 (cioè 0,222 anzichè $1/8 = 0{,}125$ come se fosse scelta « a caso »). Quasi il doppio; per n maggiore è sempre più vicina al doppio, essendo $2/(n+1)$ anzichè $1/n$ (ed $1/n$ e $1/(n+1)$ divengono praticamente uguali per n grande).

Al limite abbiamo il caso in cui tale lunghezza media diviene esattamente doppia. È il caso di avvenimenti che si succedono « a caso nel tempo », con probabilità uguali in intervalli di tempo uguali, costituendo quello che si chiama un « processo poissoniano » (da Poisson, matematico francese dell'800). Si può pensare agli arrivi di persone a uno sportello, di chiamate a un apparecchio telefonico, di gocce di pioggia su una data area, qualora si escludano interdipendenze quali arrivi a grappolo (p. es. persone giunte col medesimo treno) o regolarità (come passaggi di tram a intervalli uguali salvo per ritardi) o variazioni d'intensità (come se si comprende nello studio periodi « di punta » e no, di pioggia più o meno forte ecc.). Il tempo d'attesa tra due ripetizioni successive dell'avvenimento considerato è sempre uguale (diciamo $T = 1/\mu$, se μ è l'intensità, cioè il numero medio di avvenimenti per unità di tempo). Se uno comincia ad osservare il processo in un certo momento, quanto dovrà attendere in media per vedere il primo arrivo?

Appaiono spontanei due modi di ragionare, che sono però in contraddizione tra loro. Il primo consiste nel dire che l'inizio dell'osservazione divide a metà un intervallo tra due arrivi successivi, di lunghezza media T, e quindi sarà in media $T/2$ il tempo di attesa (e altrettanto quello trascorso dall'ultimo arrivo precedente all'inizio dell'osservazione). L'altro consiste nel dire che il tempo d'attesa è sempre, in media, T: il processo non ha memoria, e nulla cambia che il tempo si cominci a contare dall'ultimo arrivo o da un momento qualunque, p. es. quello da cui si è iniziata l'osservazione. E allora la lunghezza dell'intervallo fra due arrivi, individuato dall'aver iniziato l'osservazione (o dal contenere un dato istante) non è più T bensì $2T$. La risposta giusta è quest'ultima, e non val la pena di ripetere le considerazioni che la provano, identiche a quelle del caso precedente (suddivisione di un intervallo).

B. DE FINETTI