SUL CONCETTO DI MEDIA

In: *«Giornale dell'Istituto Italiano degli Attuari»*, Roma, 1931, Anno II, n. 3,
pp. 369-396

BRUNO DE FINETTI

SUL CONCETTO DI MEDIA

Estratto dal *Giornale dell'Istituto Italiano degli Attuari*
Anno II, n. 3, luglio 1931-IX

ROMA
ISTITUTO ITALIANO DEGLI ATTUARI
17, VIA MARCO MINGHETTI
1931-IX

Roma, 1931-IX. — Tipografia del Senato del dott. G. Bardi.

SUL CONCETTO DI MEDIA

BRUNO DE FINETTI.

SUNTO. — L'A. si occupa del concetto di media dal punto di vista del Chisini, della sua importanza e del suo significato essenzialmente relativo. Studia le proprietà di tipi particolari di medie: medie interne, medie monotone, medie associative, medie omogenee, medie traslative. Considera l'espressione generale delle medie associative data da un teorema di Nagumo-Kolmogoroff, e dà dei complementi a tale teorema. Determina quindi l'espressione generale delle medie associative omogenee (sono le medie di potenze e la media geometrica) e delle medie associative traslative (sono le medie esponenziali e la media aritmetica).

1. È strano come il concetto di media, di uso comune e generale, non solo in molti campi della scienza, ma anche nella vita pratica, non sia stato per lungo tempo analizzato nella sua vera essenza. Potrebbe anche stupire che il primo a metterne in luce il profondo significato sia stato un insigne cultore di geometria, che ebbe ad occuparsene in via puramente incidentale, mentre che tale significato pur tanto semplice e importante, era sfuggito a tutti coloro che del concetto di media ebbero a trattare di proposito e diffusamente, o a farne uso continuo nella pratica.

Ciò è forse dovuto al fatto che l'interesse per talune questioni concettuali non è ancora tanto diffuso tra i pratici quanto meriterebbe e quanto sarebbe necessario.

La visione del concetto di media, cui alludo, è dovuta al Chisini che la espose in una sua Nota del 1929 [1]. Scopo del presente lavoro

---

[1] O. CHISINI, *Sul concetto di media*, « Periodico di Matematiche », n. 2, 1929.
Durante la pubblicazione della presente memoria, ne è uscita una di P. MARTINOTTI (*Le medie relative*, « Giornale degli Economisti e Rivista di Statistica », Anno XLVI, n. 4, aprile 1931) in cui dalla definizione del Chisini, modificata secondo me inopportunamente, si prendono le mosse per varie considerazioni, illustrate da esempi, di cui parecchi (che rientrano anche nella mia concezione) interessanti.

è di richiamare l'attenzione sulla definizione del Chisini, e di rilevare come ne sia ovvia e significativa l'estensione al caso dei numeri aleatorî (variabili casuali). Si mostrerà poi come, in quest'ordine di idee, un teorema recente di Nagumo–Kolmogoroff, cui sarà apportato qualche complemento, venga ad assumere un posto fondamentale, e come ad esso si possano ricollegare le ricerche di G. Bemporad sulla media aritmetica.

Chiedersi, con spirito critico, quale sia il significato del concetto di media, come di un qualunque altro concetto, vuol dire analizzare i motivi profondi e essenziali che hanno costituito, sia pure inconsciamente, lo scopo per cui quel concetto è stato introdotto e che spiegano la ragione intima della sua utilità. Non si eseguisce certamente tale analisi quando si pretende di definire «media fra più quantità date una nuova quantità compresa tra la più piccola e la più grande delle quantità considerate», e la si evita o trascura quando si preferisce definire direttamente, volta per volta, le singole specie di medie che s'incontrano abitualmente, facendo così – come osserva il Chisini – «opera bensì esatta, ma puramente formale e antifilosofica, che può servire, e male, solo per un uso empirico».

Bisogna cominciare invece, come fa il Chisini, mettendo in rilievo che *la ricerca di una media ha come scopo quello di semplificare una data questione* «sostituendo, in essa, a due, o più, quantità date *una quantità sola che valga a sintetizzarle, senza alterare la visione d'insieme del fenomeno considerato*», e si noterà allora anzitutto, di conseguenza, che «non ha senso parlare di *media di due* (o più) *quantità*, ma ha senso parlare di media di esse *all'effetto della valutazione sintetica di un'altra grandezza che ne dipende*». Cercheremo in seguito di esprimere i medesimi concetti sotto forma più generale; vediamo intanto il modo preciso con cui tale definizione è tradotta matematicamente dal Chisini, insieme a qualche esempio che chiarisca il concetto informatore, e persuada della sua naturalezza.

Abbiansi delle grandezze omogenee $x_1, x_2, \cdots, x_n$, di cui interessa considerare la funzione

$$f(x_1, x_2, \cdots, x_n);$$

se per un dato valore x è

$$f(x, x, \cdots, x) = f(x_1, x_2, \cdots, x_n),$$

vuol dire che, agli effetti del calcolo della funzione f che interessa, tutto va come se le n variabili avessero il medesimo valore x. Si esprimerà tale fatto dicendo, col Chisini, che x *è la media di* $x_1, x_2, \cdots, x_n$ *agli effetti del calcolo di* f.

Consideriamo un esempio – quello stesso da cui parte il Chisini – che renderà più chiara la questione. Un'automobile percorre 225 chilometri, i primi 120 alla velocità di 60 km. all'ora, e gli altri 105 alla velocità di 105 km. all'ora. Qual'è la velocità media? Qui è spontaneo rispondere che, avendo l'automobile impiegato complessivamente tre ore a percorrere i 225 km., la velocità media è di 75 km. all'ora ($225 : 3 = 75$). Ma supponiamo che l'automobile sia noleggiata e il noleggiatore si interessi del consumo della benzina. Egli troverà che il consumo fatto per i 225 km. non è quello che si sarebbe verificato se l'automobile avesse mantenuto per tutto il percorso una velocità costante di 75 km., ma quello invece corrispondente a una velocità costante diversa, ad es., di 80 km. all'ora (come si ricaverebbe da una formula empirica che esprime il consumo della benzina alle diverse velocità, formula che è qui inutile ricordare). La velocità media per quanto riguarda il consumo di tempo è dunque di 75 all'ora, perchè per riguardo al consumo di tempo è come se l'automobile avesse mantenuto una velocità costante di 75 all'ora; per riguardo al consumo di benzina la velocità media è invece di 80 all'ora, è cioè come se l'automobile avesse costantemente marciato a 80 all'ora. Potevamo anche preoccuparci di una circostanza diversa: in ogni caso, per riguardo a una certa circostanza, la velocità media sarà di N km. all'ora se, per riguardo ad essa, è come se l'automobile avesse marciato sempre alla velocità di N km. all'ora.

In generale, secondo il concetto chiarito da quest'esempio: se, per riguardo a una data circostanza, è *come se* alcune grandezze $x_1, x_2, \cdots, x_n$ avessero tutte il valore comune x, tale valore x si chiamerà, per definizione, la media di $x_1, x_2, \cdots, x_n$ per riguardo a quella circostanza.

Il concetto viene anche chiarito, da tutt'altro punto di vista, illustrandone un'interpretazione grafica coi metodi della nomografia (cfr. la nota cit. del Chisini). Basti un cenno relativo al caso di due sole variabili x e y. Consideriamole come coordinate cartesiane nel piano, e sia $f(x, y) = f(x_1, y_1)$ la linea di livello della funzione f passante per il punto di coordinate x_1, y_1. Tale linea incontri la bisettrice del primo quadrante (la retta $x = y$) in un punto $x = y = m$; allora $f(m, m) = f(x_1, y_1)$, ed m è la media di x_1 e y_1 per riguardo alla circostanza f.

2. È facilmente spiegabile come nella maggior parte dei problemi pratici si presentino sempre e soltanto pochi tipi di medie più semplici: ciò corrisponde al fatto che la funzione f di cui interessa occuparsi è quasi sempre di un tipo semplice, come la somma $f(x_1, x_2, \cdots, x_n) = = x_1 + x_2 + \cdots + x_n$ delle x_i, che conduce a considerare la media aritmetica, il prodotto $f(x_1, x_2, \cdots, x_n) = x_1 x_2 \cdots x_n$ delle x_i, che conduce a considerare la media geometrica, o la somma degli inversi o dei quadrati delle x_i, che conducono rispettivamente alla media armonica o quadratica. Espressioni un po' più generali della f, ricavabili con ovvie modificazioni da queste, danno poi luogo alle corrispondenti medie ponderate.

Ecco qualche esempio che, oltre a mostrare come naturalmente si presentino, in problemi pratici, i tipi ricordati di medie, chiarirà ancora il concetto generale di media, e il fatto che esso abbia un valore puramente relativo: relativo a una data circostanza.

L'accrescimento naturale di una popolazione sia, in tre anni successivi, del 12 °/oo, 2 °/oo, 8 °/oo. Qual'è l'accrescimento annuo medio nel triennio? Se la circostanza che interessa è l'accrescimento della popolazione entro il triennio, la media sarà quel tasso d'accrescimento che avremmo dovuto avere in ciascun anno per arrivare, alla fine del triennio, al medesimo accrescimento complessivo. Sarà cioè il numero r tale che

$$(1 + r)^3 = 1{,}012 \times 1{,}002 \times 1{,}008 = 1{,}02214,$$

e avremo

$$1 + r = \sqrt[3]{1{,}02214} = 1{,}00729,$$

ossia un accrescimento medio del 7,29 °/oo. In questo caso quella che serve è la media geometrica, essendo

$$f(r_1, r_2, r_3) = (1 + r_1)(1 + r_2)(1 + r_3);$$

se invece gli accrescimenti del 12 °/oo, 2 °/oo, 8 °/oo, fossero quelli verificatisi, nello stesso anno, in tre paesi diversi, le cui popolazioni ammontino ad esempio a 30, 50, 20 milioni, la loro media, per riguardo all'aumento della popolazione complessiva, sarà la media aritmetica ponderata, e cioè di 6,20 °/oo. Infatti la prima popolazione aumenterebbe di 360.000 individui (30 milioni $\times$ 0,012), la seconda e la terza di 100 e 160 mila; l'accrescimento totale sarebbe quindi di 620 mila individui su 100 milioni, e cioè del 6,20 °/oo. In generale, l'aumento avrebbe l'espressione

$$f(r_1, r_2, r_3) = a_1 r_1 + a_2 r_2 + a_3 r_3$$

che è lineare, e perciò appunto si è condotti a una media aritmetica.

Abbiansi dei conduttori elettrici, ad esempio tre, per fissare le idee, e sia r_1, r_2, r_3 la loro resistenza. Qual'è la resistenza media? La domanda non ha senso se non si precisa per riguardo a quale circostanza; consideriamo il caso più naturale, e tale circostanza sarà l'intensità di corrente. Ma i conduttori si possono disporre in serie o in parallelo: nel primo caso la legge di Ohm dà, detta E la differenza di potenziale, l'intensità di corrente

$$I = f(r_1, r_2, r_3) = \frac{E}{r_1 + r_2 + r_3} = \frac{E}{r + r + r} = f(r, r, r)$$

se

$$r = \frac{1}{3}(r_1 + r_2 + r_3),$$

e nel secondo caso dà invece

$$I = f(r_1, r_2, r_3) = \frac{E}{r_1} + \frac{E}{r_2} + \frac{E}{r_3} = \frac{E}{r} + \frac{E}{r} + \frac{E}{r} = f(r, r, r)$$

se

$$r = \frac{3}{\frac{1}{r_1} + \frac{1}{r_2} + \frac{1}{r_3}}.$$

La resistenza comune r che dovrebbe avere ciascuno dei tre conduttori perchè l'intensità di corrente rimanesse la stessa è dunque nel primo caso la media aritmetica, e nel secondo invece la media armonica, di r_1, r_2, r_3.

Si immagini un pendolo composto, costituito di un segmento rigido, fisso a uno degli estremi, e portante n punti materiali, di massa uguale a, alle distanze $x_1, x_2, \cdots, x_n$ dal centro di sospensione. Per riguardo al momento statico

$$f(x_1, x_2, \cdots, x_n) = a(x_1 + x_2 + \cdots + x_n)$$

la media x è data dalla media aritmetica; relativamente al momento d'inerzia

$$f(x_1, x_2, \cdots, x_n) = a(x_1^2 + x_2^2 + \cdots + x_n^2)$$

la media x è data invece dalla media quadratica; se interessa, infine, la durata delle (piccole) oscillazioni

$$f(x_1, x_2, \cdots, x_n) = \frac{\pi}{\sqrt{g}} \sqrt{\frac{x_1^2 + x_2^2 + \cdots + x_n^2}{x_1 + x_2 + \cdots + x_n}}$$

(g = accelerazione di gravità), la media x è data da un'espressione meno semplice, detta media antiarmonica, e cioè da

$$x = \frac{x_1^2 + x_2^2 + \cdots + x_n^2}{x_1 + x_2 + \cdots + x_n}.$$

Le due ultime medie considerate sono note in meccanica sotto i nomi di *giratore* e di *lunghezza ridotta.*

Un altro tipo notevole di medie è quello delle medie esponenziali, e cioè delle medie di $x_1, x_2, \cdots, x_n$ per riguardo a una funzione del tipo

$$f(x_1, x_2, \cdots, x_n) = e^{ax_1} + e^{ax_2} + \cdots + e^{ax_n}.$$

Ne vedremo (n. 9) un'importante applicazione alla matematica attuariale, che ha luogo quando valga la legge di mortalità di Makeham; trattando lo stesso problema in generale, e cioè per una legge di mortalità qualunque, vedremo poi come spontaneamente anche problemi pratici possano condurre a medie che non rientrano affatto nei tipi usuali. Ciò dimostra che, se le medie dei tipi usuali sono in pratica le più utili, non sono però le sole utili, e che una concezione generale delle medie, come quella che sostengo, risponde quindi non solo a esigenze critiche e filosofiche, ma anche ad effettive necessità, che s'incontrano in problemi pratici.

3. Prima d'entrare nel campo analitico, liberiamoci subito da una restrizione evidentemente estranea alla natura del problema e ai concetti cui s'ispira la definizione generale di media. Riprendendo, per riferirci a un caso pratico, l'ultimo esempio trattato, è naturale di considerare, più in generale, un pendolo composto in cui la distribuzione di masse sia qualunque; la definizione precedente di «giratore» e di «lunghezza ridotta» mantiene intatta la sua validità anche sotto queste ipotesi più larghe. Allo stesso modo, per il concetto di velocità media, nel problema dell'automobile, sia in relazione al consumo di tempo che al consumo di benzina, nulla c'è da modificare passando dal caso che sui diversi tratti che costituiscono il percorso siano state mantenute delle velocità costanti, al caso generale di un moto comunque vario.

Siamo condotti così spontaneamente alla seguente più generale formulazione del concetto di media: *si definisce media di una grandezza in una data distribuzione* (di qualunque natura essa sia) *per rapporto a un'assegnata circostanza quell'unico valore della grandezza che si può sostituire alla distribuzione senza alterarvi la circostanza in parola.* Prendendo come immagine una distribuzione di masse su di una retta, un punto ne rappresenterà la media, relativamente a una data circostanza, se, agli effetti di essa, tutto va come se la distribuzione fosse concentrata in quel punto.

Sotto questa forma la definizione si applica senz'altro anche ai numeri aleatorî, dove si ha appunto a considerare una distribuzione di probabilità; la media, rispetto a una data circostanza, di un numero aleatorio, sarà quel numero fisso cui è, rispetto a quella circostanza, equivalente.

4. Una distribuzione è caratterizzata, nel modo più semplice, dalla sua funzione di ripartizione $\Phi(\xi)$ esprimente (quasi dappertutto) la misura (frequenza, massa, durata di tempo, probabilità, lunghezza o altro, a seconda dei problemi) della parte della distribuzione in cui il valore della data grandezza è minore di ξ; secondo la convenzione usuale, si potrà poi supporre, nei punti di discontinuità, Φ uguale alla semisomma dei limiti destro e sinistro. E $\Phi(\xi)$ risulterà, come è noto, una funzione non decrescente, che tende al limite inferiore 0 e al limite superiore 1 quando ξ tende rispettivamente a $-\infty$ e $+\infty$.

La più generale espressione di una media, nel senso considerato, si avrà traducendo in linguaggio matematico, per via analoga a quella seguita nel n. 1, la circostanza di cui interessa tener conto. Anzichè una generica funzione f di $x_1, x_2, \cdots, x_n$, interverrà qui un generico funzionale $\mathfrak{F}[\Phi(\xi)]$, e cioè un numero reale dipendente, secondo una legge qualunque, dalla forma della funzione Φ, ossia dalla forma della distribuzione. Se si indica col simbolo $\mathbf{1}(\xi)$ la funzione uguale rispettivamente a 0, 1/2, 1 a seconda che ξ è minore, uguale, maggiore di zero (ossia la funzione di ripartizione di una distribuzione concentrata nel punto zero), la funzione di ripartizione di una distribuzione concentrata nel punto x (di un numero aleatorio certamente uguale ad x) si scrive $\Phi(\xi) = \mathbf{1}(\xi - x)$. La x è la media della distribuzione $\Phi(\xi)$ agli effetti della valutazione di $\mathfrak{F}$ se è

$$\mathfrak{F}[\Phi(\xi)] = \mathfrak{F}[\mathbf{1}(\xi - x)]$$

De Finetti. - 2.

e, supponendo – come faremo sempre, escludendo i casi in cui tale proprietà non varrebbe, e che per ciò stesso sono poco significativi – che la funzione

$$F(x) = \mathfrak{F}[1(\xi - x)]$$

sia biunivoca (almeno nell'insieme dei numeri che interessa considerare), si ha subito l'espressione esplicita della più generale media sotto la forma

$$[1] \qquad F^{-1}\{\mathfrak{F}[\Phi(\xi)]\}$$

(con F^{-1} funzione inversa di F). Benchè il significato di tale espressione sia evidente, sarà forse bene tradurlo in parole per far notare con quanta immediatezza esso coincida colla definizione data del concetto di media. $F(x) = \mathfrak{F}[1(\xi - x)]$ rappresenta il valore assunto dalla circostanza che interessa quando si supponga la distribuzione concentrata nel punto x; data una distribuzione qualunque $\Phi(\xi)$, dire che $\mathfrak{F}[\Phi(\xi)] = F(x)$ significa dunque asserire che, per riguardo a quella circostanza, è indifferente che la distribuzione sia quella effettiva $\Phi(\xi)$ o che sia invece concentrata in $\xi = x$; ciò significa dunque, secondo la data definizione, che x è la media della distribuzione $\Phi(\xi)$ per riguardo a quella circostanza. La media è dunque quel valore x che rende soddisfatta la precedente uguaglianza, e di cui si dà l'espressione esplicita introducendo la funzione inversa F^{-1} [$F^{-1}(x) = y$ se $F(y) = x$; $F^{-1}(x)$ indica cioè quel valore y per cui $F(y)$ assume il valore x].

Il caso più elementare di n variabili rientra nel precedente osservando che la funzione di ripartizione, quando si hanno n soli valori $x_1, x_2, \cdots, x_n$ (che supporremo già disposti in ordine crescente), è la funzione $\Phi(\xi)$ che nel tratto precedente x_1 e in quello successivo a x_n ha il valore 0 e 1, e nei tratti intermedi, fra x_h e x_{h+1}, ha il valore h/n (o, più in generale, $p_1 + p_2 + \cdots + p_h$, se si hanno dei pesi diversi p_i per i diversi valori x_i). Nel caso di n variabili il risultato precedente si scrive quindi in modo più semplice (come si trova nell'articolo del Chisini)

$$F(x) = f(x, x, \cdots, x)$$

e

$$[1'] \qquad F^{-1}\{f(x_1, x_2, \cdots, x_n)\}$$

è l'espressione esplicita di un qualunque tipo di media. E risulta necessariamente che f è funzione simmetrica, nel caso che le x_i intervengano con pesi uguali, e che, in ogni caso, la media di n grandezze uguali ad x è uguale ad x. Sono queste anzi le sole restrizioni da imporre ad una funzione perchè la si possa, formalmente, considerare come una media; non verrà però in mente, è sperabile, di voler così *definire* le medie, trascurando il significato concettuale che ne giustifica l'interesse.

È curioso che, passando al caso di una distribuzione qualsiasi, anzichè dare come espressione generale di una media quella soprascritta, che non costituisce se non la pura e semplice estensione delle formule precedenti al caso di una distribuzione qualunque, il Chisini dia come espressione generale una formula che costituisce solo un caso particolare, sia pure notevolissimo [2]. Tradotta nel nostro simbolismo, la forma più generale di media considerata dal Chisini nel caso delle distribuzioni qualunque, corrisponde in sostanza al tipo particolare seguente di funzionali

$$\mathfrak{F}[\Phi(\xi)] = \int \psi(\xi)\, d\Phi(\xi)$$

esprimibili mediante un integrale (di Stieltjes, quando si dia, come qui, la formulazione più generale), esteso, come sempre sottintenderemo, a tutto il campo in cui Φ varia da 0 a 1, e con ψ funzione monotona. Essendo ovviamente

$$\int \psi(\xi)\, d\mathbf{1}(\xi - x) = \psi(x)$$

si ha

$$F(x) = \mathfrak{F}[\mathbf{1}(\xi - x)] = \psi(x)\,,$$

e quindi la più generale media di questo tipo ha la forma

$$\psi^{-1}\left\{\int \psi(\xi)\, d\Phi(\xi)\right\}.$$

Faremo presto vedere (n. 7) che questa è effettivamente la più generale espressione della media che soddisfi un'interessante e significativa condizione.

5. Il poco che precede esaurirebbe, in sostanza, lo studio delle medie, se non fosse necessario o utile approfondire l'esame di

---

[2]) **L'errore proviene dall'aver tacitamente ammesso che il contributo portato al funzionale da un elemento $d\Phi$ della distribuzione sia indipendente dalla distribuzione stessa.**

alcuni tipi speciali di medie, caratterizzati da proprietà notevoli e importanti.

Cominciamo dalla proprietà che – già l'abbiamo accennato in principio – è stata da molti assunta come proprietà caratteristica delle medie: la media di più quantità è una quantità compresa fra la più grande e la più piccola di queste quantità. Secondo la definizione che seguiamo, la proprietà ora accennata non è affatto necessaria; essa è naturalmente verificata in quasi tutti i problemi, ma non è difficile tuttavia immaginare degli esempi, e nemmeno poi tanto artificiosi, ove cade in difetto. Fissiamo ad esempio un angolo α di vertice in O, e sui due lati consideriamo i due punti P e Q a distanza rispettivamente $\overline{OP} = x$ e $\overline{OQ} = y$; il terzo lato $\overline{PQ} = z$ del triangolo OPQ è dato, come è noto, dalla formula

$$z^2 = x^2 + y^2 - 2xy \cos \alpha \,,$$

e la «media» di x e y agli effetti del calcolo di z è

$$m(x\,, y) = \frac{1}{\sqrt{2(1 - \cos \alpha)}} \sqrt{x^2 + y^2 - 2xy \cos \alpha} \,.$$

Si vede subito che quando $\alpha < \pi/3$ (ossia di 60^0) non sempre m è compreso fra x e y (in particolare è infatti in tal caso $m(x\,, 0) = x : \sqrt{2(1 - \cos \alpha)} > x$); tale fatto riesce del resto evidente sia dal significato del problema (è intuitivo facendo la figura), sia dal nomogramma delle curve di livello di z nel piano $x\,, y$ (ellissi coassiali).

Usualmente invece la media di due o più grandezze è compresa fra la minima e la massima; un'espressione della media che gode incondizionatamente di questa proprietà si dirà una *media interna*. Geometricamente, la condizione caratteristica delle medie interne è (nel caso di due variabili) che la curva di livello passante per un punto generico $(m\,, m)$ della bisettrice, non penetri mai nel quadrante $(x > m\,, y > m)$ nè in quello opposto $(x < m\,, y < m)$; è tale condizione che non è soddisfatta nel precedente esempio, e che permette di riconoscere dal nomogramma che si tratta di una media non interna. La condizione che definisce le medie interne si esprime poi, nel caso generale di una distribuzione (di cui Φ è la funzione di ripartizione), dicendo che per $\xi > m$ è $\Phi(\xi) > 0$, e per $\xi < m$ è $\Phi(\xi) < 1$.

Una condizione sufficiente perchè una media sia interna è che sia una media *monotona*: che, cioè, nel caso di n variabili, la media di $x_1, x_2, \cdots, x_n$ cresca sempre (o almeno non diminuisca) quando anche una sola delle x_i aumenta; la condizione non è però necessaria, ed è facile dare esempi di medie interne che non sono medie monotone. Tale la $m(x, y)$ del precedente esempio quando α è compreso tra $\pi/3$ e $\pi/2$ (fra 60° e 90°); tale è la media antiarmonica, che dà la lunghezza ridotta del pendolo composto. Basta osservare che, posto

$$m(x, y) = (x^2 + y^2) : (x + y),$$

risulta

$$m(2, 0) = 2 > m(2, 1) = 5/3;$$

anche questi casi risultano evidenti considerando il nomogramma. Geometricamente, infatti, la condizione di monotonia si esprime (nel caso di due variabili) imponendo alle curve di livello del nomogramma di non avere in nessun punto un andamento ascendente.

Nel caso generale di una distribuzione, la condizione che definisce le medie monotone è la seguente: date due funzioni di ripartizione Φ_1 e Φ_2, se è, per ogni ξ, $\Phi_1(\xi) \leq \Phi_2(\xi)$, allora si ha, per le medie corrispondenti, $m_1 \geq m_2$. L'espressione generale della media antiarmonica è ad esempio

$$F^{-1}\{\mathfrak{F}[\Phi(\xi)]\} = \frac{\int \xi^2 \, d\Phi(\xi)}{\int \xi \, d\Phi(\xi)};$$

per quanto abbiamo osservato, possiamo concludere che essa non soddisfa la detta condizione.

6. Una terza condizione, anzi la più interessante di tutte, è quella che caratterizza le medie che propongo di chiamare ***medie associative***. Osserviamo che una distribuzione può sempre pensarsi scomposta, e in infiniti modi, in due (o anche più) parti; se Φ_1 e Φ_2 sono le funzioni di ripartizione della prima e seconda parte, e esse contengono rispettivamente le parti λ_1 e λ_2 ($\lambda_1 + \lambda_2 = 1$) della distribuzione totale, questa ha come funzione di ripartizione

$$\Phi = \lambda_1 \Phi_1 + \lambda_2 \Phi_2 .$$

La condizione che caratterizza le medie associative è questa: che *rimangono invariate alterando una parte della distribuzione senza alterarne la media parziale*. La media della distribuzione $\Phi = \lambda_1 \Phi_1 + \lambda_2 \Phi_2$

è cioè uguale a quella della distribuzione $\Phi' = \lambda_1 \Phi'_1 + \lambda_2 \Phi_2$ purchè Φ_1 e Φ'_1 abbiano la stessa media.

Più chiaro riuscirà il concetto riferendoci al caso particolare di n variabili $x_1, x_2, \cdots, x_n$; era tuttavia necessario partire dal caso generale, dovendosi ora usare insieme le medie di n grandezze con valori diversi di n: non basta quindi supporre (come nei casi precedenti) di avere una funzione f di n variabili, ma occorre una legge generale che determini una successione di funzioni $f_2, f_3, \cdots, f_n, \cdots$ esprimenti la media di $2, 3, \cdots, n, \cdots$ variabili, e ciò si può avere appunto e solo partendo dal concetto generale. La condizione si esprime allora nel modo seguente: una media è associativa se nella sua espressione relativa a n variabili

$$f_n(x_1, x_2, \cdots, x_n)$$

si può sostituire un numero qualunque di variabili con la loro rispettiva media: se si ha cioè, per ogni $m < n$,

$$f_n(x_1, x_2, \cdots, x_m, x_{m+1}, \cdots, x_n) = f_n(x, x, \cdots, x, x_{m+1}, \cdots, x_n)$$

quando

$$f_m(x, x, \cdots, x) = f_m(x_1, x_2, \cdots, x_m).$$

Il significato di questa proprietà si può facilmente rilevare riferendoci agli esempi già considerati. La velocità media di un'automobile (nell'uno e nell'altro dei due sensi nominati) si può evidentemente calcolare conoscendo la velocità media (nel medesimo senso) relativa a due o più parti (di lunghezza nota) costituenti il percorso; il giratore di un pendolo composto si può calcolare conoscendo il giratore di due o più pezzi (di massa nota) costituenti il pendolo composto; in tutti questi problemi e quelli che si comportano in modo analogo – e sono la grande maggioranza – abbiamo a che fare con medie associative. Se invece vogliamo, di un pendolo composto, avere la lunghezza ridotta, non è affatto sufficiente conoscere le lunghezze ridotte di due o più pezzi, di massa nota, costituenti il pendolo: la lunghezza ridotta che si cerca dipende infatti dalla distribuzione della massa internamente a ciascuno dei singoli pezzi, e non soltanto dalle rispettive lunghezze ridotte. La media antiarmonica non è dunque una media associativa; vedremo anzi che (salvo casi praticamente inconcepibili) ogni media associativa è monotona, e quindi, sapendo che la media antiarmonica non è monotona, si ritroverà allora tale conclusione come caso particolare.

Ma esistono anche medie monotone non associative: tale ad esempio una assai nota: la mediana. Essa rientra – come è facile persuadersi – nel tipo generale considerato, ed ha precisamente l'espressione

[2] $$F^{-1}\{\mathfrak{F}[\Phi(\xi)]\} = \Phi^{-1}(1/2)$$

ove Φ^{-1} è la funzione inversa di Φ, che si estende in modo univoco ai valori in cui non è definita direttamente purchè le si imponga d'essere monotona, e che si può, volendo, rendere univoca anche quando in tutto un intervallo $a < \xi < b$ è $\Phi(\xi) = \eta =$ costante, ponendo ad esempio $\Phi^{-1}(\eta) = (a + b)/2$. Se per ogni ξ è $\Phi_1(\xi) \leq \Phi_2(\xi)$, essendo Φ_1 e Φ_2 non decrescenti, è anche $\Phi_1^{-1}(\eta) \geq \Phi_2^{-1}(\eta)$ per ogni η $(0 < \eta < 1)$, e in particolare quindi $\Phi_1^{-1}(1/2) \geq \Phi_2^{-1}(1/2)$; la mediana è pertanto una media monotona, come è del resto noto anche intuitivamente. Non è però associativa, come ben si sa; sarebbe assurdo ad esempio chiedere quale sia la statura mediana del complesso della popolazione quando sia nota separatamente la statura mediana dei maschi e delle femmine (pur supponendo nota, naturalmente, la percentuale dei maschi e delle femmine nella popolazione totale).

7. Fondamentale è a questo proposito il seguente *teorema di Nagumo-Kolmogoroff* [1]) che dà *la forma generale di ogni media associativa monotona*. Noi dimostreremo poi che *ogni media associativa è monotona*, cosicchè il teorema di Nagumo-Kolmogoroff ci darà *l'espressione generale di ogni media associativa.*

Tale espressione non è se non quella data erroneamente dal Chisini come espressione generale di una media nel caso di una distribuzione qualunque, e che abbiamo rammentata al n. 4; cioè

[3] $$\psi^{-1}\left\{\int \psi(\xi)\, d\Phi(\xi)\right\}$$

con ψ funzione crescente.

Sotto forma più semplice si ha, per il caso di n variabili (quello cui si riferisce il Kolmogoroff):

---

1) La dimostrazione cui mi riferisco trovasi in A. KOLMOGOROFF, *Sur la notion de moyenne*, « Rend. Lincei », 2° sem., fasc. 9, 1930. Il Kolmogoroff stesso mi avverte, e desidera, per correttezza, sia reso noto che indipendentemente e quasi contemporaneamente tale teorema è stato dimostrato sotto simili ipotesi da MITIO NAGUMO in una Nota del « Japanese Journal of Mathematics », vol. VI, n. 1, pag. 71, pubblicata mentre la sua Nota ai Lincei era in istampa; non ho però avuto occasione di vedere questo lavoro, e non posso quindi che riportarne la citazione, così come mi è stata data.

$$[3'] \quad f_n(x_1, x_2, \cdots, x_n) = \psi^{-1}\left(\frac{\psi(x_1) + \psi(x_2) + \cdots + \psi(x_n)}{n}\right);$$

l'operazione è quindi, in tal caso, una *trasformata* della media aritmetica, e anche nel caso più generale è una trasformata della media aritmetica

$$[4] \qquad \int \xi d\Phi(\xi)$$

come appare senz'altro ricordando che ψ è funzione crescente [4].

Riassumiamo succintamente la dimostrazione del Kolmogoroff, colle varianti opportune per adattarla alla nostra trattazione. Si consideri una certa media associativa, e si fissi un certo intervallo $a \leq \xi \leq b$; indichiamo con $g(p)$ la media di b col peso p e di a col peso $1 - p$. La media di $g(p)$ col peso r e $g(q)$ col peso $1 - r$ è, per la proprietà associativa, la media di b con peso $rp + (1 - r)q$ e di a con peso $r(1 - p) + (1 - r)(1 - q)$, ed è quindi $g\{rp + (1 - r)q\}$. In particolare, osserviamolo per fissare le idee, la media «semplice» di $g(p)$ e $g(q)$, e cioè coi pesi $r = 1 - r = 1/2$, è data da $g\left(\frac{p+q}{2}\right)$.

Supponiamo la media sia monotona, e anzi crescente (non solo non decrescente); lo stesso varrà allora per $g(p)$, e non potrà essere discontinua [5]. Posto quindi $\psi = g^{-1}$, la media di x con peso r e y con peso $1 - r$ (x e y fra a e b) sarà data da

$$\psi^{-1}\{r\psi(x) + (1 - r)\psi(y)\}$$

---

[4]) Il tipo di medie [3′] era già stato considerato e studiato da tempo, ad es. da C. E. Bonferroni, *La media esponenziale in matematica finanziaria*, «Annuario del R. Ist. Sup. di Scienze Ec. e Comm. di Bari», anno Acc. 1923-24, e la stessa formula [3′] è anche stata talvolta assunta come *definizione* della media. Benchè, dandola *a priori*, indipendentemente dal teorema di Nagumo-Kolmogoroff, sfugga quanto c'è di più interessante nel suo significato, mi sembra che, tra le definizioni *formali*, questa, pur limitandosi al caso particolare delle medie associative, sia forse la meno peggiore.

Aggiungerò anche che in una lettera al prof. Chisini, segnalandogli l'importanza della ricerca della forma generale di una media associativa, avevo affermato che, verosimilmente, tale espressione doveva essere la [3′], poco dopo effettivamente dimostrata dal Kolmogoroff, e che mi proponevo io stesso di dimostrare.

[5]) Il Kolmogoroff lo dimostra come segue. Esistono certo, essendo g crescente, i limiti $g(p - 0)$ e $g(p + 0)$; suppongasi $g(p - 0) < g(p + 0)$. Se x tende a p crescendo e y tende a p decrescendo, la media $g\left(\frac{x+y}{2}\right)$ di x e y tende alla media di $g(p - 0)$ e $g(p + 0)$, media che, per l'ipotesi che si tratti di una media cre-

e l'espressione generale della media da

$$\psi^{-1}\left\{\int\psi(\xi)\,d\Phi(\xi)\right\},$$

perchè la definizione stessa dell'integrale di Stieltjes mostra che tale espressione più generale è conseguenza della precedente nell'ipotesi che la media sia monotona.

Togliamo ora la restrizione che la media sia monotona, dimostrando che essa è conseguenza (almeno sotto larghe condizioni generiche) della proprietà associativa. Abbiamo visto che la media di $g(p)$ e $g(q)$ coi pesi r e $1-r$ è $g\{rp+(1-r)q\}$; se $g(p)=g(q)$ la media considerata è essa stessa uguale a $g(p)=g(q)$, e ne scende che se in due punti p, q, la g assume valori uguali, essa è costante (e ha quello stesso valore) in tutto l'intervallo p, q. Basta quindi che g sia funzione continua di p, o anche solo che assuma tra p e q tutti i valori compresi tra $g(p)$ e $g(q)$, per concludere che è monotona [6]. Un'altra condizione che è sufficiente supporre per giungere alla stessa conclusione è che la media sia interna; allora la stessa formula mostra senz'altro che, per ogni s compreso tra p e q, $g(s)$ è compreso tra $g(p)$ e $g(q)$.

---

scente, sarà maggiore di $g(p-0)$. Possiamo però far variare x e y in modo che sia sempre

$$\frac{x+y}{2}<p\quad,\quad g\left(\frac{x+y}{2}\right)<g(p-0),$$

e ciò contraddice il risultato precedente.

Ma, se non mi sfugge qualche circostanza, tale dimostrazione è insufficiente: non vedo infatti come possa affermarsi *a priori* che la media di $g(x)$ e $g(y)$ tenda alla media di $g(p-0)$ e $g(p+0)$. Dò quindi una dimostrazione diversa. Sia z compreso tra $g(p-0)$ e $g(p+0)$; la media di z e $g(q)$ (q generico) è compresa tra $g\left(\frac{p+q}{2}-0\right)$ e $g\left(\frac{p+q}{2}+0\right)$. Scegliamo q in modo che g sia continua in $\frac{p+q}{2}$ (ciò che è certo possibile, essendo g, funzione crescente, discontinua al più in un insieme numerabile); ne segue che la media di $g(q)$ e z è $g\left(\frac{p+q}{2}\right)$, qualunque sia z compreso tra $g(p-0)$ e $g(p+0)$. La media non sarebbe quindi crescente, contro l'ipotesi.

[6] La conclusione è anzi proprio che è crescente, come occorre per la validità del teorema di Nagumo–Kolmogoroff, e non solo che è non decrescente. Supposto infatti $g(p)=g(q)$, scende, qualunque sia s, che la media di $g(s)$ e $g(p)$ è $g\left(\frac{s+p}{2}\right)=g\left(\frac{s+q}{2}\right)$, e ciò implicherebbe che la g si riducesse a una costante.

Per quanto riguarda le pratiche applicazioni, non è certo il caso di preoccuparsi delle restrizioni generiche ora introdotte; si può però facilmente vedere che esse sono essenziali, non eliminabili. Sia infatti ψ una qualunque funzione biunivoca discontinua fra numeri reali e numeri reali: allora

$$f_n(x_1, x_2, \cdots, x^n) = \psi^{-1}\left(\frac{\psi(x_1) + \psi(x_2) + \cdots + \psi(x_n)}{n}\right)$$

definisce una media associativa che, per quanto precede, è certamente discontinua e non interna.

8. Il procedimento seguìto per dimostrare l'esistenza di una funzione ψ che trasforma ogni assegnata media associativa nella media aritmetica non si presta però praticamente per l'effettiva e comoda costruzione della funzione ψ. Indichiamone pertanto un altro che si applica immediatamente e richiede soltanto la conoscenza dell'espressione $f(x, y) = f_2(x, y)$ – della media, cioè, di due grandezze x e y – purchè derivabile (che f debba essere simmetrica va naturalmente sottinteso, trattandosi di media non ponderata).

Poniamo

$$\beta(x, y) = \frac{\partial}{\partial x} f(x, y);$$

la condizione necessaria e sufficiente perchè $f(x, y)$ sia l'espressione di una media associativa è che

$$\frac{\beta(x, y)}{\beta(y, x)}$$

sia il prodotto di una funzione della sola x per una funzione della sola y [7]; posto allora

$$\gamma(x) = \frac{\beta(x, y_0)}{\beta(y_0, x)}$$

con y_0 costante arbitraria, risulta che

$$\frac{\beta(x, y)}{\beta(y, x)} = \frac{\gamma(x)}{\gamma(y)}$$

---

[7] Il criterio differenziale per stabilire che una funzione $z(x, y)$ gode di tale proprietà è

$$\frac{\partial^2}{\partial x \partial y} \log z = 0 \qquad \text{ossia} \qquad \frac{\partial^2 z}{\partial x \partial y} = \frac{\partial z}{\partial x} \frac{\partial z}{\partial y}.$$

ed è

[5]
$$\psi(x) = a\int\gamma(x)\,dx + b\,,$$

con a a b costanti arbitrarie ($a \neq 0$).

Dimostriamo quanto sopra. Sia per ipotesi

$$f(x\,,y) = \psi^{-1}\left(\frac{\psi(x)+\psi(y)}{2}\right) = \psi^{-1}(z)$$

ove per brevità si indichi

$$2z = \psi(x) + \psi(y)\,.$$

Derivando

$$\beta(x\,,y) = \frac{\partial}{\partial x}f(x\,,y) = \frac{d\psi^{-1}(z)}{dz}\cdot\frac{\partial z}{\partial x} = \frac{1}{2}\,\frac{d\psi^{-1}(z)}{dz}\,\psi'(x)$$

e

$$\beta(y\,,x) = \frac{1}{2}\,\frac{d\psi^{-1}(z)}{dz}\,\psi'(y)$$

da cui

$$\frac{\beta(x\,,y)}{\beta(y\,,x)} = \frac{\psi'(x)}{\psi'(y)}\,.$$

Esempi:

1°. Sia $f(x\,,y) = \sqrt{xy}$ (media geometrica); si ha

$$\beta(x\,,y) = \frac{\partial}{\partial x}\sqrt{xy} = \frac{1}{2}\sqrt{\frac{y}{x}}\quad,\quad \frac{\beta(x\,,y)}{\beta(y\,,x)} = \sqrt{\frac{y}{x}}:\sqrt{\frac{x}{y}} = \frac{y}{x} = \frac{1/x}{1/y}\quad,\quad \gamma(x) = \frac{1}{x}\,,$$

$$\psi(x) = \log x\quad,\quad \psi^{-1}(x) = e^x\quad,\quad f(x\,,y) = e^{\frac{1}{2}(\log x + \log y)}\,.$$

La media geometrica è dunque la media associativa con $\psi(x) = \log x$.

2°. Sia $f(x\,,y) = \dfrac{x^2+y^2}{x+y}$ (media antiarmonica); si ha

$$\beta(x\,,y) = \frac{x^2 - y^2 + 2xy}{(x+y)^2}\,.$$

La media antiarmonica non è dunque (come già sapevamo) una media associativa.

9. Si osservi che la [5] determina $\psi(x)$ a meno di due costanti arbitrarie: un'arbitraria costante moltiplicativa (non nulla), e un'arbitraria costante additiva. Ma si verifica immediatamente che tale grado di arbitrarietà è insito nel problema: se una funzione ψ trasforma una data media associativa nella media aritmetica, ogni funzione $a\psi + b$ gode della stessa proprietà, e inversamente ogni tale funzione è del tipo $a\psi + b$. Dicendo *equivalenti* due funzioni ψ_1 e ψ_2 che danno luogo alla stessa espressione della media

$$\psi_1^{-1}\left(\frac{\psi_1(x)+\psi_1(y)}{2}\right) = \psi_2^{-1}\left(\frac{\psi_2(x)+\psi_2(y)}{2}\right),$$

avremo cioè che ψ_1 è equivalente a ψ_2, $\psi_1 \equiv \psi_2$, se e soltanto se $\psi_1 = a\psi_2 + b$ con a e b costanti, $a \neq 0$.

Va anche notato che, se ψ è funzione crescente, $a\psi + b$ è crescente o decrescente a seconda che $a > 0$ o $a < 0$. Restringersi alle funzioni crescenti, come abbiamo fatto enunciando il teorema di Nagumo–Kolmogoroff, non è quindi essenziale; necessario è solo che la funzione sia monotona, ma in questo caso essa può sempre supporsi crescente, essendo equivalente a una funzione crescente.

Più generalmente, possiamo proporci di determinare le condizioni perchè la media associativa determinata dalla funzione ψ_1 sia sempre non minore della media associativa determinata dalla funzione ψ_2; in particolare si dovranno così ritrovare le disuguaglianze note fra le medie aritmetica, geometrica, armonica e quadratica. Suppongasi (per ogni x, y di un campo di variabilità u)

$$\psi_1^{-1}\left(\frac{\psi_1(x)+\psi_1(y)}{2}\right) \geq \psi_2^{-1}\left(\frac{\psi_2(x)+\psi_2(y)}{2}\right);$$

operando a sinistra con ψ_2 si ha

$$[6] \quad \psi_2\psi_1^{-1}\left(\frac{\psi_1(x)+\psi_1(y)}{2}\right) \geq \frac{\psi_2(x)+\psi_2(y)}{2} = \frac{\psi_2\psi_1^{-1}[\psi_1(x)] + \psi_2\psi_1^{-1}[\psi_1(y)]}{2}$$

supponendo ψ_2 crescente (mentre la disuguaglianza, e quindi le conclusioni che seguono, vanno invertite, se ψ_2 è decrescente).

Tale condizione esprime che la funzione $\psi_2\psi_1^{-1}$ deve essere convessa verso l'alto (in tutto il campo in cui varia $\psi_1(x)$ quando x percorre il campo u): è questa dunque la condizione necessaria e sufficiente perchè la media determinata da ψ_1, sia sempre non minore di quella determinata da ψ_2. Il caso opposto è caratterizzato dal comportamento opposto, e cioè dalla convessità verso il basso; come caso limite, se entrambe le condizioni sono soddisfatte, e $\psi_2 \psi_1^{-1}$ è cioè lineare, le due medie coincidono, e $\psi_1 \equiv \psi_2$. Si ritrova così la condizione di equivalenza soprascritta: da

$$\psi_2 \psi_1^{-1}(x) = ax + b$$

scende infatti

$$\psi_2 [\psi_1^{-1}(x)] = a\psi_1 [\psi_1^{-1}(x)] + b \quad , \quad \psi_2 = a\psi_1 + b \, .$$

Una funzione può però essere convessa in certi intervalli verso l'alto e in altri verso il basso, o anche non presentare in nessun intervallo una convessità di verso determinato; verificandosi per $\psi_2 \psi_1^{-1}$ uno di tali casi, nessuna conclusione *a priori* è possibile circa la disuguaglianza delle medie dedotte da ψ_1 e ψ_2, che non potrà avere un verso determinato, indipendentemente dalla particolare distribuzione che si considera.

Ponendo, in particolare, $\psi_1(x) = x$, si ha che la media associativa determinata da una funzione crescente ψ_2 è minore o uguale della media aritmetica se ψ_2 è convessa verso l'alto, maggiore o uguale nel caso opposto [8]. Nel caso delle medie di potenze si vede poi che la media cresce col crescere dell'esponente.

10. A un esempio molto generale di media associativa conduce il problema di matematica attuariale di cui demmo un cenno in fine del n. 2. Dato un gruppo d'individui di età $x_1, x_2, \cdots x_n$, si tratti di determinare l'età comune ξ che dovrebbero avere perchè, suppostili coetanei, rimanga invariata la probabilità che tutti sopravvivano dopo un tempo z. Indicando con $l(x)$ la funzione di sopravvivenza (e supponendo al solito che le probabilità di morte degli n individui siano indipendenti), la definizione si traduce in forma analitica nell'equazione

$$[7] \qquad \left(\frac{l(\xi + z)}{l(\xi)}\right)^n = \frac{l(x_1 + z)}{l(x_1)} \, \frac{l(x_2 + z)}{l(x_2)} \cdots \frac{l(x_n + z)}{l(x_n)}$$

[8]) Questo caso particolare era già stato considerato da C. E. BONFERRONI, *Sulle medie dedotte da funzioni concave*, «Giornale di Matematica Finanziaria», n. 1, 1927.

e ξ è quindi la media associativa definita dalla funzione

$$\psi_z(x) = \log l(x+z) - \log l(x) \cdot$$

In generale, l'età media di un dato gruppo d'individui dipenderà da z: l'età media per riguardo alla sopravvivenza dopo un anno non è necessariamente quella che si avrebbe considerando la sopravvivenza dopo due o dieci anni. Si può però cercare se e per quali forme speciali della funzione $l(x)$ avvenga che l'età media sia indipendente da z, ed esista cioè un'età ξ tale che un gruppo di n individui di età $x_1, x_2, \cdots, x_n$ abbia *per qualsiasi istante futuro* la stessa probabilità di sopravvivere al completo come se tutti gli n individui avessero l'età comune ξ. Tale condizione è quella che nel teorema di Quiquet definisce le *leggi di sopravvivenza di 1° ordine*; illustrarne – come abbiamo fatto, e come ulteriormente faremo nel n. 12 – il significato, alla luce del concetto generale di media, mi sembra possa essere istruttivo.

Per determinare le leggi $l(x)$ che godono della voluta proprietà, si osservi che la condizione necessaria e sufficiente perchè ciò avvenga è che sia $\psi_z(x)$ equivalente a una funzione $\psi(x)$ indipendente da z, ossia

$$\psi_z(x) \equiv \psi(x) = a(z)\,\psi(x) + b(z)$$

con a, b, ψ funzioni arbitrarie; per la ψ possiamo poi sempre supporre $\psi(0) = 0$, disponendo ancora di una costante moltiplicativa arbitraria. Posto $f(x) = \log l(x)$, si ha

$$f(x+z) - f(x) = a(z)\,\psi(x) + b(z)$$

e, per $x = 0$, $b(z) = f(z) - f(0) = f(z)$ essendo $f(0) = \log l(0) = = \log 1 = 0$.

Quindi

$$f(x+z) = f(x) + f(z) + a(z)\,\psi(x) = f(x) + f(z) + k\,\psi(x)\,\psi(z)$$

dovendosi avere, per la simmetria,

$$a(z)\,\psi(x) = a(x)\,\psi(z)$$

ossia

$$a(z) = k\psi(z).$$

Sviluppando

$$f(x+y+z) = f(x) + f(y) + f(z) + k\,[\psi(x)\,\psi(y) + \psi(x+y)\,\psi(z)],$$

si vede che $\psi(x)\,\psi(y) + \psi(x+y)\,\psi(z)$ dev'essere funzione simmetrica di $x, y, z, 0$, il che è lo stesso, di $\xi = \psi(x), \eta = \psi(y), \zeta = \psi(z)$. Ma perchè $\xi\eta + F(\xi, \eta)\,\zeta$ risulti funzione simmetrica di ξ, η, ζ, è necessario e sufficiente risulti tale (e quindi costante, essendo indipendente da ζ) anche

$$\frac{F(\xi, \eta)}{\xi\eta} - \frac{1}{\xi} - \frac{1}{\eta}$$

che se ne ricava dividendo per la funzione simmetrica $\xi\eta\zeta$ e sottraendo la funzione simmetrica

$$\frac{1}{\xi} + \frac{1}{\eta} + \frac{1}{\zeta}.$$

Ne segue

$$F(\xi, \eta) = \xi + \eta + a\xi\eta\,;$$

è quindi

$$\psi(x+y) = \psi(x) + \psi(y) + a\psi(x)\,\psi(y)$$

ossia

$$\{a\psi(x+y) + 1\} = \{a\psi(x) + 1\}\{a\psi(y) + 1\}$$

$$a\psi(x) = e^{cx} - 1\,.$$

Possiamo disporre di a, e porremo $a = 1$.

È ora

$$f(x+z) = f(x) + f(z) + k(e^{cx} - 1)(e^{cz} - 1)$$

ossia

$$\{f(x+z) - k(e^{c(x+z)} - 1)\} = \{f(x) - k(e^{cx} - 1)\} + \{f(y) - k(e^{cy} - 1)\}$$

da cui

$$f(x) - k(e^{cx} - 1) = mx\,,$$

[8] $$f(x) = \log l(x) = mx + k(e^{cx} - 1)\,.$$

Si ritrova così (anche indipendentemente dalle ordinarie ipotesi di derivabilità) che la proprietà in questione è caratteristica della legge di Makeham, e che in tal caso (essendo $\psi(x) \equiv e^{cx}$) l'età media è data da una media esponenziale.

Si osservi ancora che oltre l'età media *per riguardo alla sopravvivenza completa* del gruppo si potrebbe considerare l'età media per rispetto *alla sua completa estinzione*: sarebbe questa ad esempio che interesserebbe a chi dovesse pagare un vitalizio che venisse a cessare solo col decesso dell'ultimo superstite di una famiglia o di altro gruppo qualsiasi. Allora l'età media ξ sarebbe data da

$$\left(1-\frac{l(\xi+z)}{l(\xi)}\right)^n=\left(1-\frac{l(x_1+z)}{l(x_1)}\right)\left(1-\frac{l(x_2+z)}{l(x_2)}\right)\cdots\left(1-\frac{l(x_n+z)}{l(x_n)}\right),$$

e sarebbe quindi la media associativa definita dalla funzione

$$\psi_z(x)=\log\left[l(x)-l(x+z)\right]-\log l(x)\,.$$

Si vede subito che la media in questo senso è sempre maggiore della precedente, solo che si supponga la probabilità di morte funzione crescente dell'età (precisamente: $l(x+z)/l(x)$ funzione decrescente di x per z assegnato); posto infatti $l(x+z)/l(x)=g(x)$, e indicate con ψ_1 e ψ_2 le ψ_z relative al primo e al secondo caso, si ha

$$\psi_1(x)=\log g(x)\quad,\quad \psi_2(x)=\log\left[1-g(x)\right]\quad,\quad \psi_1^{-1}(x)=g^{-1}(e^x)\,,$$

$$\psi_2\,\psi_1^{-1}(x)=\log\left[1-gg^{-1}e^x\right]=\log\left(1-e^x\right);$$

tale funzione è convessa verso l'alto, essendone la derivata seconda $-e^x(1-e^x)^{-2}<0$. Anche in relazione a questo caso si potrebbe poi cercare se e per quali funzioni $l(x)$ l'età media risulti indipendente da z.

Più in generale, essendo $0<m<n$, si potrebbe considerare l'età media per riguardo alla probabilità che, dopo un tempo z, degli n individui ne sopravvivano m, oppure m almeno, o m al più. Non si avrebbero però allora medie associative.

11. Due condizioni d'indole affatto diversa da quelle finora considerate sono le seguenti, che ci definiranno le *medie omogenee* e le *medie traslative*.

Diremo che una media è *omogenea* se il suo valore risulta moltiplicato per a quando tutte le variabili vengono moltiplicate per uno stesso numero a, o, nel caso generale, quando alla funzione di ripartizione $\Phi(\xi)$ si sostituisca la $\Phi_a(\xi)=\Phi\left(\frac{\xi}{a}\right)$. Deve essere cioè

[9]
$$f_n(ax_1, ax_2, \cdots, ax_n)=a f_n(x_1, x_2, \cdots, x_n)\,,$$

$$F^{-1}\{\mathfrak{F}[\Phi_a(\xi)]\}=a\,F^{-1}\{\mathfrak{F}[\Phi(\xi)]\}\quad\text{con}\quad \Phi_a(\xi)=\Phi\left(\frac{\xi}{a}\right).$$

Una media omogenea può essere associativa (come le medie aritmetica, geometrica, armonica, quadratica); può anche essere però monotona ma non associativa (come la mediana), interna ma non monotona (come la media antiarmonica), o addirittura non interna (come è il caso nell'esempio che si è dato nel n. 5 di media non interna). Geometricamente, la condizione caratteristica delle medie omogenee (nel caso di due variabili) è che le curve di livello del nomogramma siano omotetiche rispetto all'origine; l'espressione generale di una media associativa di due variabili x e y è quindi

$$z = \frac{x+y}{2} f(\log x - \log y)$$

con f funzione pari uguale a 1 nell'origine.

Interessante è cercare l'espressione più generale di una media omogenea associativa: si tratta, come vedremo, delle medie di potenze e della media geometrica.

Suppongasi infatti

$$\psi^{-1}\left(\frac{\psi(ax)+\psi(ay)}{2}\right) = a\psi^{-1}\left(\frac{\psi(x)+\psi(y)}{2}\right);$$

ne risulta che $\psi(ax)+\psi(ay)$ è funzione di $\psi(x)+\psi(y)$ e di a:

$$\psi(ax)+\psi(ay) = \varphi[\psi(x)+\psi(y)\,;a]\,.$$

Sfruttando l'arbitrarietà rilevata nel n. 9, potremo supporre $\psi(0) = 0$ e $\psi(1) = 1$; allora si ha, ponendo uguale a 0 prima y e poi x,

$$\psi(ax) = \varphi[\psi(x)\,;a] \quad , \quad \psi(ay) = \varphi[\psi(y)\,;a]$$

$$\psi(ax)+\psi(ay) = \varphi[\psi(x)+\psi(y)\,;a] = \varphi[\psi(x)\,;a] + \varphi[\psi(y)\,;a]\,.$$

Ma l'equazione funzionale

$$\varphi(\lambda+\mu) = \varphi(\lambda)+\varphi(\mu)\,,$$

sotto condizioni molto ampie, abbondantemente soddisfatte nel nostro caso ove φ è manifestamente una funzione crescente, ammette le sole soluzioni della forma $\varphi(\lambda) = k\lambda$; si ha pertanto

$$\psi(ax) = \varphi[\psi(x)\,;a] = k(a)\,\psi(x) = k(x)\,\psi(a)$$

e, per $a = 1$, $\psi(x) = k(x)\,\psi(1) = k(x)$ avendo supposto $\psi(1) = 1$.

Allora

$$\psi(ax) = \psi(a)\,\psi(x),$$

ma tale equazione funzionale (facilmente riducibile alla precedente) implica che si abbia

$$\psi(x) = x^m \quad \text{(con } m \text{ reale qualsiasi).}$$

Le medie associative omogenee sono quindi le medie di potenze

[10] $$\left\{\int \xi^m \, d\Phi(\xi)\right\}^{\frac{1}{m}},$$

di cui sono casi particolari:

la media aritmetica ($m = 1$);
la media armonica ($m = -1$);
la media quadratica ($m = 2$).

Al limite per $m \to 0$ si ha la media geometrica

[11] $$e^{\int \log \xi \, d\Phi(\xi)}, \quad \left(\psi(\xi) = \log \xi = \lim_{m \to 0} \frac{\xi^m - 1}{m}\right),$$

che non si ottiene col procedimento seguìto causa la singolarità della funzione logaritmo nell'origine, che impedisce di supporre $\psi(0) = 0$.

12. Diremo che una media è *traslativa* se il suo valore viene aumentato di a quando tutte le variabili vengono aumentate di una stessa quantità a, o, nel caso generale, quando alla funzione di ripartizione $\Phi(\xi)$ si sostituisca la $\Phi_a(\xi) = \Phi(\xi - a)$.

Dev'essere cioè

[12] $$\begin{gathered} f_n(x_1 + a, x_2 + a, \cdots, x_n + a) = f_n(x_1, x_2, \cdots, x_n) + a, \\ F^{-1}\{\mathfrak{F}[\Phi_a(\xi)]\} = F^{-1}\{\mathfrak{F}[\Phi(\xi)]\} + a \quad \text{con } \Phi_a(\xi) = \Phi(\xi - a). \end{gathered}$$

Una media traslativa può essere associativa, sia omogenea (come la media aritmetica) che non omogenea (come la media esponenziale); può essere però monotona ma non associativa (come la mediana), e anche interna ma non monotona o addirittura non interna. Ciò risulta evidente ricordando il significato geometrico di dette condizioni sul nomogramma (relativo al caso di due variabili) e tenendo presente che, geometricamente, la condizione che caratterizza le medie traslative (sempre nel caso di due variabili) è che le curve di livello del nomogramma si ottengano una dall'altra mediante una traslazione lungo la bisettrice $x = y$; l'espressione generale di una media traslativa z di due variabili x e y è quindi

$$z = \frac{x+y}{2} + f(x-y)$$

con f funzione pari nulla nell'origine.

Cerchiamo ora, analogamente a quanto abbiamo fatto per le medie omogenee, l'espressione più generale di una media traslativa associativa: si tratta come vedremo, delle medie esponenziali e della media aritmetica.

Sia infatti

$$\psi^{-1}\left(\frac{\psi(x+a)+\psi(y+a)}{2}\right) = \psi^{-1}\left(\frac{\psi(x)+\psi(y)}{2}\right) + a \ , \ \text{con } \psi(0) = 0 ;$$

ne risulta che $\psi(x+a) + \psi(y+a)$ è funzione di $\psi(x) + \psi(y)$ e di a, sia

$$\psi(x+a) + \psi(y+a) = \varphi[\psi(x) + \psi(y) ; a] .$$

Per $y = 0$, rispettivamente $x = 0$, si ha

$$\psi(x+a) + \psi(a) = \varphi[\psi(x) ; a] \ , \ \psi(y+a) + \psi(a) = \varphi[\psi(y) ; a] ,$$

e sommando

$$\varphi[\psi(x) + \psi(y) ; a] = \varphi[\psi(x) ; a] + \varphi[\psi(y) ; a] - 2\psi(a)$$

ossia

$$\varphi[\psi(x)+\psi(y) ; a] - 2\psi(a) = \{\varphi[\psi(x) ; a] - 2\psi(a)\} + \{\varphi[\psi(y) ; a] - 2\psi(a)\}$$

da cui

$$\varphi[\psi(x) ; a] - 2\psi(a) = k(a)\psi(x) = \psi(x+a) - \psi(a)$$

$$\psi(x+a) = \psi(a) + k(a)\psi(x) = \psi(a+x) = \psi(x) + k(x)\psi(a) .$$

Si ricava

$$\frac{1-k(a)}{\psi(a)} = \frac{1-k(x)}{\psi(x)} = \text{cost} = C \ , \ k(x) = C\psi(x) + 1 ,$$

e ci si riduce quindi all'equazione funzionale

$$\psi(x+a) = \psi(x) + \psi(a) + C\psi(a)\psi(x)$$

ossia (se $C \neq 0$)

$$C\psi(x+a) + 1 = \{C\psi(x) + 1\}\{C\psi(a) + 1\}$$

che dà

$$C\psi(x) + 1 = e^{cx} \quad , \quad \psi(x) \equiv e^{cx} ,$$

mentre nel caso, sopra escluso, $C = 0$, si ha senz'altro $\psi(x) \equiv x$.

Tale risultato si poteva anche immediatamente dedurre da quello analogo per le medie omogenee: è facile infatti trasformare i due problemi l'uno nell'altro, perchè se z è una media traslativa di x e y (così che $z + a$ sia la media di $x + a$ e $y + a$) si vede che e^z è una media omogenea di e^x e e^y (essendo $e^a e^z$ la media di $e^a e^x$ e $e^a e^y$), e viceversa.

Le medie associative traslative sono quindi le medie esponenziali

$$[13] \qquad \frac{1}{c} \log \int e^{c\xi} d\Phi(\xi) ,$$

e la media aritmetica [4]; soltanto quest'ultima è poi una media associativa traslativa omogenea. Che le proprietà associativa omogenea e traslativa bastino a caratterizzare la media aritmetica è del resto un risultato noto, dovuto a Giulio Bemporad [9], che vi giunse direttamente per tutt'altra via. E anche anteriormente il Bonferroni [10] era giunto al risultato che la media aritmetica è l'unica media del tipo [3'] traslativa e omogenea, il che è equivalente quando si conosca il teorema di Nagumo–Kolmogoroff. Invece il Bemporad considera esplicitamente – credo per il primo – la proprietà associativa.

Abbiamo visto che la media esponenziale è quella che si incontra cercando l'età media di un gruppo di individui, per rispetto alla loro completa sopravvivenza, nel caso che valga la legge di Makeham; in quel caso la media è dunque traslativa. Lo si può del resto comprendere *a priori* partendo dal significato stesso del problema pro-

---

[9] GIULIO BEMPORAD, *Sul principio della media aritmetica*, « Rend. Lincei », 1926, pag. 87.

[10] Cfr. loc. cit. [4]; ivi sono anche dati, sotto ipotesi di derivabilità di cui ci siamo liberati, i risultati di questo e del numero precedente sulle espressioni generali delle medie del tipo [3'] omogenee e traslative, dette ivi varianti *proporzionalmente* e varianti *uniformemente*.

postoci al n. 10, e che ci condusse appunto a caratterizzare la legge di Makeham. Abbiamo dalla [7]

$$\left\{\frac{l(\xi+z+t)}{l(\xi+z)}\right\}^n \left\{\frac{l(\xi+z)}{l(\xi)}\right\}^n =$$

$$= \left\{\frac{l(x_1+z+t)}{l(x_1+z)} \cdots \frac{l(x_n+z+t)}{l(x_n+z)}\right\} \left\{\frac{l(x_1+z)}{l(x_1)} \cdots \frac{l(x_n+z)}{l(x_n)}\right\}$$

ove ξ indichi l'età media di $x_1, x_2, \cdots, x_n$ relativa alla sopravvivenza completa dopo un tempo $z+t$. Ma se tale età media è la stessa di quella relativa alla sopravvivenza completa dopo un tempo z, il secondo fattore è uguale nei due membri, e l'eguaglianza dei primi, che ne consegue, dice che $\xi+z$ è l'età media fra $x_1+z, x_2+z, \cdots, x_n+z$ relativamente alla sopravvivenza completa dopo un tempo t. Se l'età media non dipende dal periodo che interessa considerare, vale quindi la proprietà traslativa.

Ciò implica che ξ sia una media esponenziale, e si arriva così direttamente alla

$$\psi_z(x) \equiv e^{cx} = a(z)\, e^{cx} + b(z)$$

da cui scende tosto la [8], e si evita il piuttosto lungo ragionamento diretto del n. 10.

13. Non è forse superfluo indicare esplicitamente come i risultati precedenti si esprimano per il caso di un numero aleatorio X, usando il simbolo $\mathfrak{M}$ per il valor medio (media aritmetica, o anche *speranza matematica*):

[4″] $$\mathfrak{M}(X) = \int \xi d\Phi(\xi)\,.$$

La più generale media associativa è

[3″] $$\mu(X) = \psi^{-1}\mathfrak{M}[\psi(X)] \quad [\text{ossia } \mu = \psi^{-1}\mathfrak{M}(\psi)]$$

con ψ funzione crescente; la più generale media associativa omogenea è

[10″] $$\mu(X) = [\mathfrak{M}(X^m)]^{\frac{1}{m}} \quad \text{con } m \text{ reale,}$$

oltre la media geometrica

[11″] $$\mu(X) = e^{\mathfrak{M}[\log X]}\,;$$

la più generale media associativa traslativa è

$$[13''] \qquad \mu(X) = \frac{1}{c} \log \mathfrak{M}(e^{cX}) \qquad \text{con } c \text{ reale,}$$

oltre la media aritmetica $\mu(X) = \mathfrak{M}(X)$, che è l'unica media associativa traslativa omogenea.

È facile poi, e dal punto di vista concettuale non aggiunge nulla di notevole, l'estensione al caso in cui l'ente di cui si ha a considerare la media non sia una grandezza, ma un vettore (es.: velocità media), un punto (es.: baricentro), o quel che altro si vuole, purchè sia applicabile, per rispetto a una data circostanza che interessa, il concetto informatore della definizione generale di media.