

10082
BIBLIOTECA
MINISTERO FINANZE

UMBERTO RICCI

SULLA

BIBLIOTECA
MINISTERO FINANZE

MEDIA ARITMETICA PONDERATA



TORINO

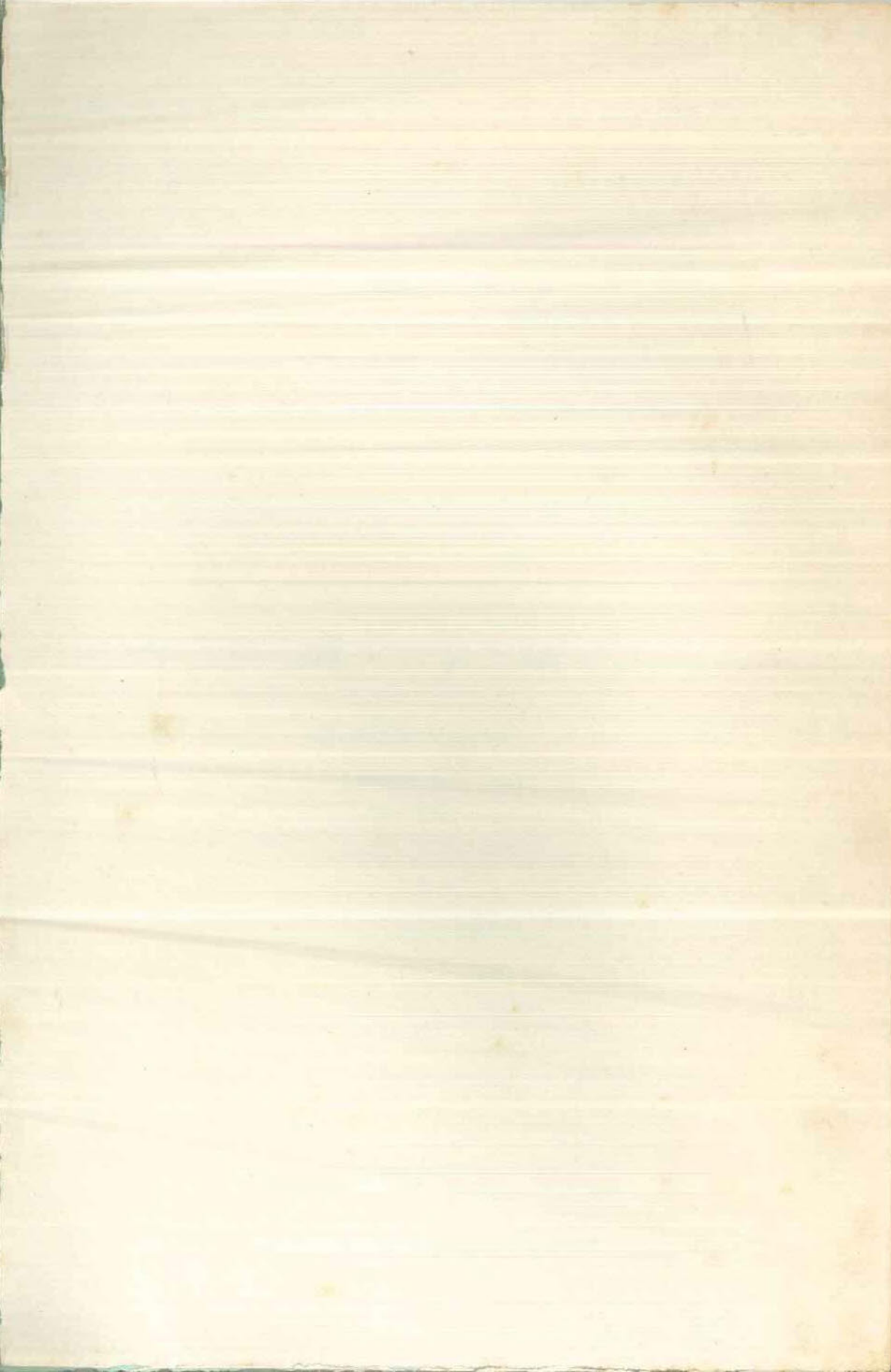
SOCIETÀ TIPOGRAFICO EDITRICE NAZIONALE

(già Roux e Viarengo e già Marcello Capra)

1912

BIBLIOTECA
MINISTERO DELLE FINANZE
MISCELL.
173
F
121

THE NEW YORK PUBLIC LIBRARY
ASTOR LENOX AND TILDEN FOUNDATIONS
500 N. 5TH ST. N. Y. C.



UMBERTO RICCI

SULLA

MEDIA ARITMETICA PONDERATA



TORINO

SOCIETÀ TIPOGRAFICO-EDITRICE NAZIONALE

(già Roux e Viarengo e già Marcello Capra)

1912

LIBRARY OF THE
UNIVERSITY OF TORONTO

PROPRIETÀ LETTERARIA

ATAZONER VILWELIV FONDERIV

ALFON

UMBERTO RICCI

SEZIONE I. — Concetto della media aritmetica ponderata.

§ I. — Definizione analitica della media aritmetica ponderata.

— Analiticamente è facile definire la media ponderata.

Date n grandezze, di cui:

p_1	tutte	eguali	fra	loro	e	uguali	ad	a_1
p_2	"	"	"	"	"	"	"	a_2
.
p_k	"	"	"	"	"	"	"	a_k

si chiama *media aritmetica ponderata* (μ_a) *delle k grandezze disuguali*:

$$a_1, a_2, \dots, a_k$$

la *media aritmetica semplice* delle n grandezze date.

Per fare questa media aritmetica dobbiamo addizionare indistintamente *tutti* i termini dati, uguali e disuguali, e dobbiamo dividere la somma per n .

Ora la somma di tutti i termini dati è uguale a:

$$a_1 p_1 + a_2 p_2 + \dots + a_k p_k = \sum_{i=1}^{i=k} a_i p_i$$

che scriveremo per brevità:

$$\sum a_i p_i.$$

Il numero n delle grandezze non è altro che la somma delle *frequenze*, intendendosi per frequenza il numero delle volte in cui ciascuna grandezza è ripetuta:

$$n = p_1 + p_2 + \dots + p_k = \sum_{i=1}^{i=k} p_i$$

che scriveremo per brevità:

$$\sum p_i.$$

Dunque la media aritmetica ponderata μ_a non è altro che:

$$(1) \quad \mu_a = \frac{\sum a_i p_i}{\sum p_i}.$$

Da questa formula si ricava l'altra:

$$(2) \quad \sum a_i p_i = \mu_a \sum p_i$$

che ci sarà utile in seguito.

Bisogna fare bene attenzione a non confondere la media aritmetica delle n grandezze date (uguali e disuguali) con la media aritmetica dei k termini disuguali, ossia non bisogna confondere μ_a con:

$$(3) \quad m_a = \frac{a_1 + a_2 + \dots + a_k}{k} = \frac{\sum_{i=1}^k a_i}{k}.$$

Quando si dice che una media aritmetica semplice di più termini è diversa dalla media aritmetica ponderata dei *medesimi termini*, si vuol appunto significare che m_a è diverso da μ_a .

Chiamiamo serie dei *termini a* la successione:

$$a_1, a_2 \dots a_k.$$

Chiamiamo *pesi* le frequenze:

$$p_1, p_2 \dots p_k.$$

Allora possiamo dire che *la media aritmetica ponderata di più termini si ottiene moltiplicando i termini per i rispettivi pesi e dividendo la somma dei prodotti per la somma dei pesi.*

Questa definizione è generale e sussiste anche per il caso che i vari pesi p_1, \dots, p_k siano numeri frazionari.

Nel corso del presente studio chiameremo per brevità *media ponderata* la media aritmetica ponderata.

§ 2. — **Costruzione geometrica della media ponderata.** — Supponiamo di avere una serie composta di due soli termini: a_1, a_2 e di volerne cercare la media ponderata. Sia a_1 il termine maggiore e a_2 il termine minore.

Costruiamo due rettangoli AC e BF, tali che la loro area sia rispettivamente uguale ai prodotti $a_1 p_1, a_2 p_2$. Le *basi* dei rettangoli avranno rispettivamente per misura i pesi p_1, p_2 , le *altezze* avranno rispettivamente per misura le grandezze a_1, a_2 . Facciamo in modo che le basi siano sulla medesima retta e l'una di seguito all'altra.

Prolungando DC ed EF in modo che s'incontrino in H veniamo a costruire un rettangolo, il quale ha per altezza l'altezza a_1 del rettangolo maggiore e per base la somma $p_1 + p_2$ delle due basi.

Prolungando FG in I veniamo a costruire un rettangolo, che ha per altezza l'altezza a_2 del rettangolo minore e per base la somma $p_1 + p_2$ delle due basi.

Congiungiamo infine i due nuovi vertici H e I. La congiungente taglierà CB in L. La misura del segmento LB è la media cercata.

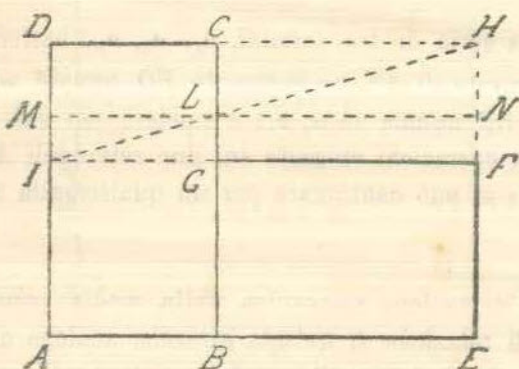


Fig. 1

Infatti, se per il punto L tiriamo MN parallela ad AE, veniamo ad avere due triangoli simili LIG , HLN , dai quali si ricava subito:

$$\overline{IG} : \overline{GL} = \overline{LN} : \overline{NH}$$

o anche:

$$\overline{ML} : \overline{GL} = \overline{LN} : \overline{CL}$$

da cui:

$$\overline{ML} \cdot \overline{CL} = \overline{GL} \cdot \overline{LN}$$

ossia i due rettangoli MC e GN sono uguali fra loro.

Ne consegue che:

$$\text{Rettangolo AN} = \text{Rett. AC} + \text{Rett. BF}$$

ossia:

$$\overline{AM} (\overline{AB} + \overline{BE}) = \overline{AB} \cdot \overline{BC} + \overline{BE} \cdot \overline{EF}$$

e infine:

$$\begin{aligned}\overline{AM} &= \frac{\overline{AB} \cdot \overline{BC} + \overline{BE} \cdot \overline{EF}}{\overline{AB} + \overline{BE}} \\ &= \frac{a_1 p_1 + a_2 p_2}{p_1 + p_2} = \mu_a.\end{aligned}$$

In conclusione

$$\overline{BL} = \mu_a$$

come dovevasi dimostrare.

Avendo una serie di tre termini, a_1, a_2, a_3 , basterebbe prendere sul prolungamento di \overline{AE} un segmento \overline{EO} uguale ad a_3 , costruire il rettangolo EP uguale ad $a_3 p_3$; e ripetere poi sui due rettangoli AN ed EP le operazioni eseguite sui due rettangoli AC e BF.

Uguualmente si può continuare per un qualsivoglia numero di termini.

§ 3. — **Illustrazione meccanica della media ponderata.** — Sia O il centro di rotazione di un'asta girevole, analoga al giogo di una bilancia. Quando l'asta è nella posizione di equilibrio la linea dei punti di sospensione AOM sia *orizzontale* (Cfr. Fig. 2).

A sinistra del punto O, e alle rispettive distanze

$$\overline{A'O} = a_1$$

$$\overline{A''O} = a_2$$

applichiamo i pesi p_1 e p_2 .

Perchè l'equilibrio non venga turbato dobbiamo attaccare il peso:

$$P = p_1 + p_2$$

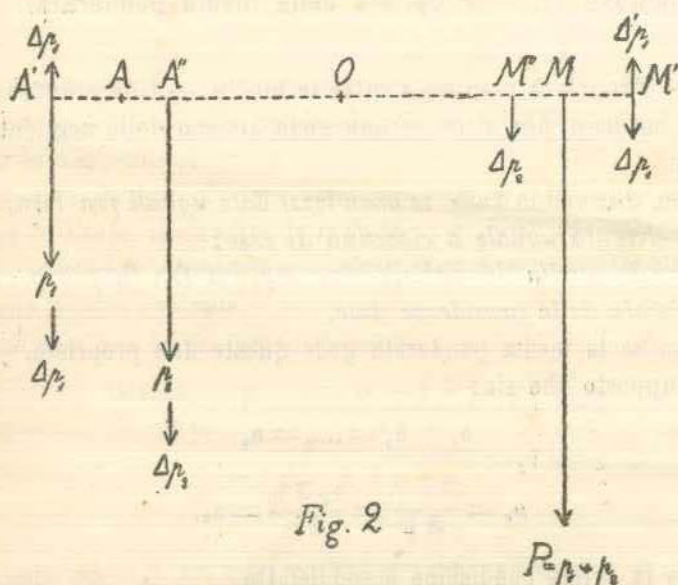
a destra di O e a una distanza \overline{OM} , tale che il *momento* di P rispetto al punto O sia uguale (e di segno contrario) alla somma dei momenti di p_1 e p_2 rispetto al medesimo punto O. In altre parole dev'essere:

$$\overline{OM} (p_1 + p_2) = \overline{A'O} \cdot p_1 + \overline{A''O} \cdot p_2 = a_1 p_1 + a_2 p_2.$$

Ne consegue che:

$$\overline{OM} = \frac{a_1 p_1 + a_2 p_2}{p_1 + p_2} = \mu_a$$

ossia la media ponderata di due termini non è altro che il braccio di leva che deve attribuirsi al peso totale P , perchè questo faccia equilibrio ai due pesi parziali p_1 e p_2 .



Il peso P , applicato nel punto M , si può decomporre in due nuovi pesi π_1 e π_2 , ciascuno dei quali fa rispettivamente equilibrio ai pesi p_1 e p_2 . Questi nuovi pesi si calcolano subito, riflettendo che deve essere

$$\mu_a \pi_1 = a_1 p_1$$

e

$$\mu_a \pi_2 = a_2 p_2$$

quindi:

$$(4) \quad \pi_1 = \frac{a_1 p_1}{\mu_a} = \frac{a_1 p_1 \sum p_i}{\sum a_i p_i}$$

$$\pi_2 = \frac{a_2 p_2}{\mu_a} = \frac{a_2 p_2 \sum p_i}{\sum a_i p_i}$$

Se invece di considerare due pesi e due bracci ne consideriamo un numero qualsivoglia, il ragionamento non cambia.

SEZIONE II. — Proprietà della media ponderata.

§ 4. — Proprietà comuni a tutte le medie. — Una grandezza non può dirsi media di più altre, se non gode almeno delle seguenti due proprietà:

1) che, divenendo tutte le grandezze date uguali fra loro, anche la media diventa uguale a ciascuna di esse;

2) che la media deve risultare compresa fra la più grande e la più piccola delle grandezze date.

Vediamo se la media ponderata gode queste due proprietà.

1) Supposto che sia:

$$a_1 = a_2 = \dots = a_k$$

abbiamo:

$$\mu_a = \frac{\sum a_k p_i}{\sum p_i} = \frac{a_k \sum p_i}{\sum p_i} = a_k.$$

Dunque la prima condizione è soddisfatta.

2) Di tutte le grandezze date sia a_1 la più grande e sia a_k la più piccola.

Calcoliamo le differenze tra a_1 e tutti gli altri termini. Queste differenze risulteranno tutte positive:

$$a_1 - a_2 = \sigma_2$$

$$a_1 - a_3 = \sigma_3$$

$$\dots \dots \dots$$

$$a_1 - a_k = \sigma_k.$$

Avremo:

$$\begin{aligned} \mu_a &= \frac{a_1 p_1 + (a_1 - \sigma_2) p_2 + \dots + (a_1 - \sigma_k) p_k}{\sum p_i} \\ &= \frac{a_1 \sum_{i=1}^k p_i - \sum_{i=2}^k \sigma_i p_i}{\sum_{i=1}^k p_i} \\ &= a_1 - \frac{\sum_{i=2}^k \sigma_i p_i}{\sum_{i=1}^k p_i}. \end{aligned}$$

Dunque μ_a è uguale ad a_1 meno una quantità positiva, ossia:

$$\mu_a < a_1.$$

Analogamente si dimostrerebbe che:

$$\mu_a > a_n.$$

§ 5. — **Proprietà comune alla media ponderata e alla media aritmetica.** — La media ponderata ha in comune colla media aritmetica la seguente proprietà:

Se tutti i termini di una serie vengono accresciuti o diminuiti di una costante, moltiplicati o divisi per una costante, anche la media ponderata viene accresciuta o diminuita, moltiplicata o divisa per quella costante.

Infatti, se k è la costante, abbiamo:

$$\frac{\sum (a_i + k) p_i}{\sum p_i} = \frac{\sum a_i p_i + k \sum p_i}{\sum p_i} = \mu_a + k$$

$$\frac{\sum (ka_i) p_i}{\sum p_i} = \frac{k \sum a_i p_i}{\sum p_i} = k \mu_a.$$

La sottrazione è un caso particolare dell'addizione e la divisione è un caso particolare della moltiplicazione (dividere per un numero significa moltiplicare per il suo reciproco).

Il principio può dunque ritenersi dimostrato.

§ 6. — **Proprietà specifiche della media ponderata.** — Le proprietà seguenti sono godute dalla media ponderata e solo da essa.

1) *La somma algebrica degli scostamenti moltiplicati per i rispettivi pesi è uguale a zero.*

Data la serie:

$$a_1 \ a_2 \ \dots \ a_k$$

calcoliamo gli scostamenti s_i rispetto alla media ponderata. Abbiamo:

$$a_1 - \mu_a = s_1$$

$$a_2 - \mu_a = s_2$$

$$\dots \dots \dots$$

$$a_k - \mu_a = s_k$$

e in generale:

$$(5) \quad a_i - \mu_a = s_i.$$

Allora la media ponderata si può mettere sotto questa forma:

$$\begin{aligned} \mu_n &= \frac{\sum a_i p_i}{\sum p_i} = \frac{\sum (\mu_n + s_i) p_i}{\sum p_i} \\ &= \frac{\mu_n \sum p_i + \sum s_i p_i}{\sum p_i} = \mu_n + \frac{\sum s_i p_i}{\sum p_i}. \end{aligned}$$

Si ricava subito:

$$(6) \quad \sum s_i p_i = 0.$$

Nella figura 1^a, gli scostamenti *positivi* sono rappresentati da quelle porzioni di altezza, analoghe a DM, le quali *sopravanzano la linea della media ponderata* ossia la linea MLN, e gli scostamenti *negativi* sono rappresentati dai *segmenti complementari*, analoghi a LG, i quali misurano la distanza tra la linea MLN della media ponderata e la base superiore dei rettangoli aventi un'altezza inferiore alla media. I prodotti degli scostamenti positivi per i rispettivi pesi, ossia i *prodotti s_i p_i positivi* sono rappresentati dai *rettangoli*, analoghi al rettangolo MC, *eccedenti* la linea MLN della media ponderata; e i prodotti degli scostamenti negativi per i rispettivi pesi, ossia i *prodotti s_i p_i negativi*, sono rappresentati dai *rettangoli complementari*, analoghi al rettangolo GN. Cosicchè il problema aritmetico di calcolare una media ponderata si traduce nel problema geometrico di tracciare una parallela MLN alla linea delle basi ABE, tale che *la somma dei rettangoli eccedenti sia uguale alla somma dei rettangoli complementari*.

Nella figura 2^a gli scostamenti sarebbero rappresentati dai bracci delle forze p_1 e p_2 , *rispetto non più al centro di rotazione O, ma al punto A di applicazione della risultante dei due pesi medesimi (OA = OM)*; cosicchè nel linguaggio della meccanica il teorema enunciato dalla (6) si tradurrebbe nell'altro che *la somma algebrica dei momenti dei pesi rispetto al punto di applicazione della risultante è uguale a zero*.

2) Prendendo gli scostamenti z_i da una grandezza qualsivoglia α , anzichè dalla media μ_n , la media ponderata degli scostamenti z_i , aggiunta al termine α , riproduce μ_n .

Infatti, ponendo:

$$\begin{aligned} a_1 - \alpha &= z_1 \\ \dots & \\ a_k - \alpha &= z_k \end{aligned}$$

abbiamo:

$$\mu_a = \frac{\sum a_i p_i}{\sum p_i} = \frac{\sum (a_i - \alpha + \alpha) p_i}{\sum p_i} = \frac{\sum (z_i + \alpha) p_i}{\sum p_i}$$

ossia:

$$(7) \quad \mu_a = \alpha + \frac{\sum z_i p_i}{\sum p_i}.$$

Questa formula è molto adoperata quando i vari termini della serie sono in progressione aritmetica (come è spesso il caso nelle seriazioni. Cfr. § 9). Allora si sceglie un termine che sia presso a poco a metà della serie e lo si pone uguale ad α . Gli scostamenti si calcolano facilmente e così anche il quoziente $\frac{\sum z_i p_i}{\sum p_i}$.

SEZIONE III. — Campo di applicazione delle medie ponderate.

§ 7. — Serie di fenomeni a due dimensioni. — Vediamo ora in quali casi si renda necessario o conveniente di adoperare in istatistica la media ponderata. Questi casi possono ridursi a tre. Le serie, che possono richiedere il calcolo di medie ponderate, sono:

- a) le serie di *fenomeni a due dimensioni*;
- b) le serie di *rapporti o di indici*;
- c) le *seriazioni*.

Chiamiamo *fenomeni a due dimensioni* tutti quei fenomeni che si possono decomporre in due altri, collegati fra loro in tal guisa, che l'uno esprima un certo carattere quantitativo posseduto da ciascuna unità dell'altro.

La misura di un fenomeno a due dimensioni si può mettere quindi sotto la forma di un prodotto di due grandezze. Se questo prodotto misura l'*estensione* del fenomeno composto, ossia del fenomeno a due dimensioni, in un determinato luogo o momento, la prima grandezza, che è la misura del carattere quantitativo posseduto da ciascuna unità dell'altro, indica l'*intensità* del fenomeno, mentre l'altra grandezza è il *peso*.

Noi chiameremo generalmente *termini* (per antonomasia) le prime grandezze, ossia quelle che misurano l'intensità del fenomeno composto (cfr. § 1 in fine).

Esempi di fenomeni a due dimensioni sono: il raccolto che si ricava da una certa coltura agraria, il quale è in ragione composta dall'area coltivata (*peso*) e del rendimento per ettaro (*intensità*); il valore di una certa partita di merce, il quale varia in ragione composta della quantità venduta (*peso*) e del prezzo dell'unità di merce (*intensità*).

Orbene, data una serie di termini:

$$a_1, a_2 \dots a_k$$

e dati i pesi corrispondenti:

$$p_1, p_2 \dots p_k$$

per modo che i dati composti sarebbero:

$$b_1 = a_1 p_1, \quad b_2 = a_2 p_2, \dots, b_k = a_k p_k$$

non si può trovare l'intensità media del fenomeno, se non calcolando la media ponderata dei vari termini a .

Essa si può mettere anche sotto la forma:

$$(8) \quad \mu_a = \frac{\sum b_i}{\sum p_i}$$

La grandezza che figura al numeratore misura l'estensione complessiva del fenomeno composto.

Le serie di fenomeni a due dimensioni possono essere *nel tempo* e *nello spazio*, vale a dire possono constare di dati relativi ad uno stesso fenomeno il quale si manifesta in successivi periodi di tempo in una medesima regione dello spazio, oppure di dati relativi ad uno stesso fenomeno il quale si manifesta contemporaneamente in diverse regioni dello spazio.

§ 8. — **Serie di rapporti e serie di indici.** — Bene spesso si calcolano, in statistica, i cosiddetti *valori proporzionali* o *rapporti per 100, per 1000, ecc.*

In generale, dato un fenomeno C, il quale abbia le seguenti misure:

$$c_1, c_2 \dots c_k$$

e dato un altro fenomeno D il quale abbia le misure corrispondenti:

$$d_1, d_2 \dots d_k$$

chiameremo valori proporzionali, o rapporti per mille, i rapporti:

$$p_1 = \frac{1000 d_1}{c_1}, p_2 = \frac{1000 d_2}{c_2}, \dots, p_k = \frac{1000 d_k}{c_k}$$

Le grandezze che figurano al denominatore le chiameremo *grandezze di riferimento*. Per dare un esempio di rapporti: C potrebbe indicare il fenomeno « ammontare della popolazione del tal paese, in tempi successivi », oppure il fenomeno « popolazione di paesi diversi in un tempo dato » e D potrebbe indicare il fenomeno « numero dei suicidi nelle popolazioni suindicate ».

Se ora vogliamo conoscere il valore proporzionale medio o rapporto medio, dobbiamo fare la *media ponderata dei vari rapporti*.

Infatti questo rapporto medio ρ_m non può essere se non:

$$\rho_m = \frac{1000 \sum d_i}{\sum c_i}.$$

Ma esso si può mettere sotto la forma:

$$\rho_m = \frac{1000 \sum \frac{\rho_i c_i}{1000}}{\sum c_i}.$$

ossia:

$$(9) \quad \rho_m = \frac{\sum \rho_i c_i}{\sum c_i}$$

che è evidentemente una media ponderata. Infatti basta confrontare la (9) con la (1) del § 1 per vedere che si tratta della stessa formula: i rapporti ρ_i rappresentano i *termini* (le *intensità*) e le grandezze di riferimento c rappresentano i *pesi*.

Un caso particolare dei rapporti sono gli *indici*, dei quali si fa molto uso nella statistica. Perchè si abbia un indice (numero-indice), occorre che entrambi i dati, fra i quali si istituisce il rapporto, si riferiscano allo stesso fenomeno: si confronta generalmente una certa misura del fenomeno in un dato luogo e tempo con la misura del fenomeno nello stesso luogo e in un altro tempo.

Siano:

$$e_1 e_2 \dots e_k$$

le misure di un certo fenomeno, per esempio il fenomeno « prezzo di alcune mercanzie determinate in un dato momento » o l'altro « superfici coltivate in vari paesi in un primo momento ».

Siano:

$$e'_1 e'_2 \dots e'_k$$

le misure corrispondenti del periodo successivo.

Gli indici sono i rapporti:

$$\frac{100 e'_1}{e_1} = i_1, \quad \frac{100 e'_2}{e_2} = i_2, \dots, \quad \frac{100 e'_k}{e_k} = i_k.$$

La *media degli indici*, ossia il nuovo indice:

$$\frac{100 \sum e'_i}{\sum e_i} = i_\mu$$

si può mettere sotto la forma:

$$i'_\mu = \frac{100 \sum i' e_i}{\sum e_i}$$

o anche:

$$(10) \quad i'_\mu = \frac{\sum i' e_i}{\sum e_i}.$$

Essa è quindi una vera media ponderata: gli indici i' rappresentano le intensità (*i termini*); le grandezze di riferimento e_i rappresentano i *pesi*.

Le medie ponderate di rapporti possono essere tanto di rapporti *coesistenti* quanto di rapporti *successivi*; le medie ponderate di indici sono generalmente medie di indici *coesistenti*.

§ 9. — **Seriazioni.** — Supponiamo di avere una collettività di oggetti od elementi graduabili secondo un medesimo carattere quantitativo: per esempio una moltitudine di coscritti che noi possiamo graduare secondo la statura. Questi elementi noi li possiamo distribuire in *classi*.

Ogni classe rimane identificata dalla *frequenza*, ossia dal numero degli elementi che vi appartengono, e dall'*intensità* di quel carattere quantitativo che è comune a tutta la collettività. Se la gradazione è *discontinua*, dobbiamo formare tante classi quante sono le intensità nettamente diverse del fenomeno e ogni classe comprenderà elementi che avranno esattamente la medesima grandezza. Se la gradazione è *continua*, la formazione delle classi è arbitraria e ogni classe comprenderà elementi di grandezza alquanto differenti l'una dall'altra.

La serie delle classi, ordinata secondo le frequenze, costituisce una *seriazione*.

Or bene, se noi vogliamo conoscere l'*intensità media* di una seriazione a elementi *continui*, dobbiamo fare una media *ponderata* delle intensità delle varie classi. Le intensità sono i *termini* e le frequenze i *pesi*.

Se la seriazione è invece a *elementi continui*, occorre prima ridurla alla forma di una seriazione a elementi discontinui. Ciò si può fare prendendo la media aritmetica dei termini massimo e minimo di ciascuna classe. Talora, però, ciò non basta e si rendono necessarie certe correzioni.

SEZIONE IV. — Differenza fra media ponderata e media aritmetica.

§ 10. — Calcolo della differenza δ . — Sia la solita serie:

$$a_1, a_2, \dots, a_k.$$

Chiamiamo δ la differenza fra media ponderata e media aritmetica:

$$\delta = \mu_a - m_a$$

x_i gli scostamenti dei termini a_i e y_i gli scostamenti dei pesi p_i dalle rispettive medie aritmetiche m_a e m_p , ossia poniamo genericamente:

$$x_i = a_i - m_a$$

$$y_i = p_i - m_p.$$

Ne consegue che la somma:

$$\sum a_i p_i$$

si può mettere sotto la forma:

$$\begin{aligned} & \sum (m_a + x_i)(m_p + y_i) \\ &= \sum m_a m_p + m_p \sum x_i + m_a \sum y_i + \sum x_i y_i. \end{aligned}$$

Ora tenuto conto che:

$$\sum m_a m_p = k m_a m_p = m_a \sum p_i$$

$$\sum x_i = 0 \quad \sum y_i = 0$$

si ottiene:

$$\sum a_i p_i = m_a \sum p_i + \sum x_i y_i.$$

Così la media ponderata diventa:

$$\mu_a = \frac{\sum a_i p_i}{\sum p_i} = m_a + \frac{\sum x_i y_i}{\sum p_i} = m_a + \frac{\sum x_i y_i}{k m_p}.$$

Infine la differenza fra le due medie è:

$$(11) \quad \delta = \mu_a - m_a = \frac{\sum x_i y_i}{k m_p}.$$

Se $\sum x_i y_i$ è positivo, allora:

$$\mu_a > m_a.$$

Se $\sum x_i y_i$ è negativo, allora:

$$\mu_a < m_a.$$

La differenza tra media ponderata e media aritmetica (di una medesima serie di termini) è uguale alla somma dei prodotti degli scostamenti dei termini per gli scostamenti dei pesi (dalle rispettive medie aritmetiche) divisa per il prodotto del numero dei termini per il peso medio.

§ 11. — In quali casi la differenza δ tende ad annullarsi. — Vediamo ora in quali casi la differenza tra media ponderata e media aritmetica sia trascurabile, ossia in quali casi il rapporto $\frac{\sum x_i y_i}{k m_p}$ tenda ad annullarsi.

Evidentemente questo rapporto è tanto più piccolo, quanto più piccolo è il numeratore o più grande il denominatore.

Il numeratore sarà tanto minore quanto più forti saranno le compensazioni fra i vari prodotti $x_i y_i$.

1) Tali prodotti sono di segno positivo, quando derivano da scostamenti dello stesso segno, negativi nel caso opposto. In realtà abbiamo quattro gruppi di prodotti:

il gruppo dei prodotti	(+ x)	(+ y)
"	"	" (+ x) (- y)
"	"	" (- x) (+ y)
"	"	" (- x) (- y).

Ora, quanto più lunga è la serie considerata, tanto più è probabile che i vari prodotti degli scostamenti si distribuiscano in egual numero nei quattro gruppi accennati, e quindi tanto più è probabile che le compensazioni si possano effettuare.

Resta naturalmente da escludere il caso eccezionalissimo di una forte correlazione fra i termini e i pesi: in tal caso, comunque lunga fosse la serie, la somma dei prodotti degli scostamenti non tenderebbe mai ad annullarsi, ma sarebbe sempre positiva o sempre negativa, secondo il senso della correlazione.

2) Ma non basta che i *segni* positivi o negativi si distribuiscano in egual numero. Perchè le compensazioni avvengano in misura considerevole, occorre che i *valori assoluti* dei prodotti $x_i y_i$ non siano molto diversi gli uni dagli altri. E perchè ciò si verifichi, si rende necessario che nè la serie dei termini, nè quella dei pesi, offrano forti oscillazioni attorno alla media; occorre cioè che la *dispersione* delle due serie sia piccola.

Il *denominatore* può essere grande:

1) O perchè il *numero* k dei *termini* sia *elevato*. Ecco già la seconda volta in cui vediamo essere desiderabile che la serie sia lunga affinchè la differenza δ risulti esigua.

2) O perchè la media aritmetica dei pesi raggiunga un alto valore, ciò che accade, naturalmente, quando il *valore numerico dei pesi* è *assai elevato*.

Riassumendo possiamo dire che, *sostituendo a una media ponderata la media aritmetica semplice, si commette un errore δ il quale probabilmente è tanto più piccolo:*

- 1) *quanto più lunga è la serie considerata,*
- 2) *quanto più piccola è la dispersione della serie dei termini e di quella dei pesi,*
- 3) *quanto più grandi sono i pesi (1).*

§ 12. — **Applicazione alla statistica agraria.** — Facciamo una applicazione di questo principio ad uno dei rami della statistica in cui più di frequente accade di doversi servire della media ponderata;

(1). Questo enunciato deve interpretarsi con qualche limitazione, perchè il numeratore e il denominatore della frazione $\frac{\sum x_i y_i}{k m_p}$ non sono indipendenti fra loro. Quanto più grandi sono i pesi, tanto più è probabile che siano grandi gli scostamenti dei pesi dalla loro media, e quindi, se cresce m_p , può darsi che crescano contemporaneamente il numeratore e il denominatore della frazione suddetta. Perciò l'enunciato del testo deve intendersi nel senso che, a *parità di m_p* , la quantità δ tanto più piccola 1°) quanto più piccola è la dispersione della serie dei termini e della serie dei pesi 2°) quanto più lunga è la serie considerata.

D'altra parte la lunghezza della serie influisce contemporaneamente sul numeratore e sul denominatore della frazione. Se k aumenta, è sicuro che il denominatore della frazione aumenti, ma non è ugualmente sicuro che il numeratore diminuisca, potendo darsi che le compensazioni dei prodotti $x_i y_i$ avvengano in un minor numero: perciò nel testo abbiamo fatto la riserva di escludere una forte correlazione tra i termini e i pesi.

vogliamo dire la statistica agraria. A rigore, volendo calcolare il rendimento per ettaro ottenuto *in media* da una certa coltura agraria — durante una successione di anni in una stessa regione o in uno stesso paese, oppure in un determinato anno ma in una moltitudine di regioni o di paesi — occorrerebbe sempre fare una media *ponderata*. Calcolando la media aritmetica invece della media ponderata, si commette un errore. Però la grandezza dell'errore è assai diversa, secondo che la media è tratta da una *serie nel tempo* o da una *serie nello spazio*.

Nel primo caso la dispersione delle serie è relativamente piccola, perchè non variano molto di anno in anno nè le superfici coltivate in un determinato paese, nè i rendimenti per ettaro.

Nel secondo caso si hanno serie a dispersione relativamente grande, perchè tanto i rendimenti per ettaro quanto le superfici possono essere assai diversi da un paese all'altro.

La media ponderata di una *serie nel tempo* è quindi assai più vicina alla media aritmetica, che non la media ponderata di una *serie nello spazio*.

Or bene questo ragionamento *a priori* viene confermato dall'esperienza, la quale dimostra che assai spesso è *praticamente indifferente* di calcolare la media aritmetica o la media ponderata *in una serie abbastanza lunga di rendimenti per ettaro succedentisi nel tempo*, mentre, sostituendo la media aritmetica alla ponderata in *una serie, anche abbastanza lunga, di rendimenti per ettaro distribuiti nello spazio*, si possono commettere *gravi errori*.

A titolo di esempio calcoliamo il rendimento per ettaro medio ottenuto dalla coltura del frumento in Argentina durante l'ultimo dodicennio. La differenza tra media aritmetica e media ponderata è praticamente trascurabile: le due medie cominciano a differire nella cifra dei centesimi.

Anno agrario	Rendimenti per ettaro	Superficie coltivata	Scostamenti dei rendimenti per ettaro	Scostamenti della superficie	Prodotti degli scostamenti
	a_i	p_i	$x_i = a_i - m_a$	$y_i = p_i - m_p$	$x_i y_i$
	Quintali	Ettari			
1899-1900	8,51	3.250.000	+ 1,26	- 1.485.373	- 1.871.570
1900-1901	6,02	3.379.749	- 1,23	- 1.355.624	+ 1.667.418
1901-1902	4,66	3.296.066	- 2,59	- 1.439.307	+ 3.727.805
1902-1903	7,64	3.695.343	+ 0,39	- 1.040.030	- 405.612
1903-1904	8,17	4.320.000	+ 0,92	- 415.373	- 382.143
1904-1905	8,37	4.903.124	+ 1,12	+ 167.751	+ 187.881
1905-1906	6,47	5.675.293	- 0,78	+ 939.920	- 733.138
1906-1907	7,46	5.692.268	+ 0,21	+ 956.895	+ 200.948
1907-1908	9,09	5.759.987	+ 1,84	+ 1.024.614	+ 1.885.290
1908-1909	7,01	6.063.100	- 0,24	+ 1.327.727	- 318.654
1909-1910	6,11	5.836.550	- 1,14	+ 1.101.177	- 1.255.342
1910-1911	7,49	4.953.000	+ 0,24	+ 217.627	+ 52.230
	$m_a =$ = 7,25	$m_p =$ = 4.735.373			$\sum x_i y_i =$ = + 2.755.113

$$\delta = \frac{\sum x_i y_i}{k m_p} = \frac{2.755.113}{12 \times 4.735.373} = + 0,05$$

$$\mu_a = m_a + \delta = 7,25 + 0,05 = 7,30.$$

Invece proviamoci a calcolare il rendimento medio ottenuto dalla coltura dell'orzo nell'anno agricolo 1910-11 nell'emisfero settentrionale (1).

(1) I dati sono presi dal *Bollettino di Statistica agraria* dell'Istituto Internazionale di Agricoltura per il mese di settembre 1911.

Paesi	Rendimenti per ettaro	Superfici coltivate	Scostamenti dei rendimenti per ettaro	Scostamenti delle superfici	Prodotti degli scostamenti
	s_i	p_i	$x_i = a_i - m_a$	$y_i = p_i - m_p$	$x_i y_i$
EUROPA					
	Quintali	Ettari			
Prussia	18,8	829.979	+ 3,75	— 440.159	— 1.650.596
Belgio	29,0	34.500	+ 13,95	— 1.235.638	— 17.237.150
Danimarca	23,3	233.000	+ 8,25	— 1.037.138	— 8.556.388
Spagna	13,5	1.446.138	— 1,55	+ 176.000	— 272.800
Gran Bretagna e Irlanda	18,4	762.471	+ 3,35	— 507.667	— 1.700.684
Ungheria (compresa la Croazia-Slavonia)	14,1	1.169.745	— 0,95	— 100.393	+ 95.373
Italia	9,5	247.000	— 5,55	— 1.023.138	+ 5.678.416
Lussemburgo	16,0	943	+ 0,95	— 1.269.195	— 1.205.735
Paesi Bassi	28,5	27.969	+ 13,45	— 1.242.169	— 16.707.173
Rumania	10,8	507.201	— 4,25	— 762.937	+ 3.242.482
Russia Europea	8,2	11.367.566	— 6,85	+ 10.097.428	— 69.167.382
Svizzera	19,3	5.200	+ 4,25	— 1.264.938	— 5.375.986
AMERICA					
Canadá	15,5	724.770	+ 0,45	— 545.368	— 245.416
Stati Uniti	10,9	2.848.189	— 4,15	+ 1.578.051	— 6.548.912
ASIA					
Giappone	16,0	1.285.000	+ 0,95	+ 14.862	+ 14.119
Russia Asiatica	6,5	352.813	— 8,55	— 917.325	+ 7.843.129
AFRICA					
Algeria	9,1	567.000	— 5,95	— 703.138	+ 4.183.671
Tunisia	3,5	453.000	— 11,55	— 817.138	+ 9.437.944
	$m_a =$ = 15,05	$m_p =$ = 1.270.138			$\sum x_i y_i =$ = — 98.173.088

$$\delta = \frac{\sum x_i y_i}{k m_p} = - \frac{98.173.088}{18 \times 1.270.138} = - 4,29.$$

$$\mu_a = m_a + \delta = 15,05 - 4,29 = 10,76.$$

L'errore che si commette calcolando la media aritmetica anzichè la media ponderata è assai forte.

SEZIONE V. — **Variazione dei pesi.**

§ 13. — **Variazioni in uno stesso senso dei soli pesi relativi ai termini superiori alla media, o dei soli pesi dei termini inferiori alla media.** — La media ponderata di una serie in cui rimanga costante il numero dei termini può variare, o perchè vengano a variare i termini, o perchè vengano a variare i pesi.

Nella presente sezione noi supponiamo che rimangano costanti i termini e che *crescano* o *diminuiscano* di valore i pesi. Nella sezione successiva supporremo che rimangano inalterati i pesi e aumentino o diminuiscano di valore i termini.

Tanto nel caso dell'aumento quanto in quello della diminuzione di valore dei pesi, dobbiamo suddividere due casi, secondo che i pesi accresciuti o scemati hanno il braccio *maggiore* o *minore* della media.

1) Prendiamo, per semplicità, una serie di due soli termini, a_1, a_2 e sia a_1 il termine maggiore, a_2 il termine minore. Possiamo rappresentare la serie mediante un sistema di forze e di pesi. Come abbiamo detto nel § 3, la media ponderata dei due termini a_1 e a_2 è rappresentata dal braccio di leva \overline{OM} che occorre al peso $P = p_1 + p_2$ per fare equilibrio ai due pesi p_1 e p_2 insieme considerati.

Si accresca ora, di una quantità Δp_1 , il peso p_1 , che ha il braccio *più lungo*. L'equilibrio rimane turbato e per ristabilirlo dobbiamo applicare un peso uguale a Δp_1 a destra di O e a una distanza $\overline{OM'}$ tale che sia:

$$\overline{OM'} = \overline{OA'}$$

La risultante dei due pesi P e Δp_1 applicati rispettivamente in M e M' avrà il braccio compreso fra $\overline{OM} = \mu_a$ e $\overline{OM'} = a_1$, e siccome $a_1 > \mu_a$ ne consegue che il braccio della detta risultante sarà superiore a μ_a .

In altri termini, aumentando il peso di braccio più lungo la media cresce.

È facile estendere questo principio al caso di più di due termini e di più di due pesi. Qualunque sia il numero dei termini della serie, se *aumenta il peso di uno o più termini superiori alla media, la media cresce.*

2) Aumenti ora, della quantità Δp_2 , il *peso p_2 di braccio più corto.* Per ristabilire l'equilibrio occorre applicare un peso uguale a Δp_2 in un punto M'' , che deve certamente trovarsi a sinistra di M . Quindi anche il punto di applicazione del peso risultante $P + \Delta p_2$ sarà a sinistra di M .

In generale, *aumentando il peso di uno o più termini inferiori alla media, la media diminuisce.*

Passiamo ora a studiare gli effetti di una diminuzione nel valore dei pesi.

1) Diminuisca il peso p_1 avente il *braccio più lungo.* Diminuire p_1 della quantità $\Delta' p_1$ è come applicare nel punto A' una forza uguale a $\Delta' p_1$, che abbia la stessa direzione di p_1 (ossia la direzione verticale) ma vada in senso opposto a p_1 (ossia vada dal basso in alto). L'equilibrio viene turbato e non lo si può ristabilire se non applicando nel punto M' una forza eguale e parallela a $\Delta' p_1$ e diretta nel medesimo senso. Sappiamo dalla meccanica che ai due pesi P e $\Delta' p_1$ possiamo sostituire un peso uguale alla loro differenza (peso risultante) il quale abbia il suo punto di applicazione sul *prolungamento* del segmento $\overline{MM'}$, dalla parte del peso maggiore P . In altre parole il peso risultante $P - \Delta' p_1$ avrà il suo punto di applicazione in un punto che sarà a sinistra di M , o ancora avrà un braccio minore di $\overline{OM} = \mu_n$. E siccome il braccio di questo peso risultante è appunto la nuova media, possiamo dire che, *diminuendo il peso di un termine superiore alla media, la media diminuisce.*

2) Analogamente si dimostra che, *diminuendo il peso di un termine inferiore alla media, la media aumenta.*

§ 14. — **Dimostrazione analitica.** — Di queste varie proposizioni possiamo offrire anche una dimostrazione analitica.

In generale, data la serie :

$$a_1, a_2, \dots, a_k$$

supponiamo che vari un peso qualunque: per esempio il peso p_h e vari della quantità Δp_h , la quale potrà essere positiva o negativa.

La risultante di questi due pesi addizionali avrà il suo punto di applicazione in M solo a patto che sia:

$$\overline{MM'} \cdot \Delta p_1 = \overline{MM''} \cdot \Delta p_2.$$

Solo in tal caso la media μ_a non cambia.

Se invece:

$$\overline{MM'} \cdot \Delta p_1 > \overline{MM''} \cdot \Delta p_2$$

il punto di applicazione della risultante dei due pesi addizionali cadrà fra M e M' e quindi anche il punto di applicazione della risultante di P e dei due pesi addizionali cadrà tra M e M', ossia la nuova media (il braccio della risultante di P e dei due pesi addizionali) risulterà superiore alla media precedente (il braccio di P).

Se invece:

$$\overline{MM'} \cdot \Delta p_1 < \overline{MM''} \cdot \Delta p_2$$

il punto di applicazione della risultante dei due pesi addizionali cadrà fra M e M'' e quindi la media diminuirà.

In conclusione, quando crescono entrambi i pesi p_1 e p_2 possono darsi tutti e tre i casi: o che la media cresca, o che rimanga inalterata, o che diminuisca.

Alla stessa conclusione si arriva, sempre servendosi di ragionamenti tolti dalla meccanica, qualora i pesi p_1 e p_2 , invece di crescere, diminuiscano di valore. Nè è difficile di estendere l'enunciato al caso di una serie di termini lunga quanto si voglia.

§ 17. — **Dimostrazione analitica.** — Possiamo dare anche una dimostrazione analitica del teorema relativo all'aumento (o alla diminuzione) generale dei pesi.

Abbiamo in primo luogo:

$$\sum a_i (p_i + \Delta p_i) = \sum a_i p_i + \sum a_i \Delta p_i.$$

Ricordando la (2) del § 1 e la (5) del § 6 abbiamo:

$$\begin{aligned} \sum a_i (p_i + \Delta p_i) &= \mu_a \sum p_i + \sum (\mu_a + s_i) \Delta p_i \\ &= \mu_a \sum (p_i + \Delta p_i) + \sum s_i \Delta p_i. \end{aligned}$$

E quindi:

$$\mu_a'' = \frac{\sum a_i (p_i + \Delta p_i)}{\sum (p_i + \Delta p_i)} = \mu_a + \frac{\sum s_i \Delta p_i}{\sum (p_i + \Delta p_i)}$$

ossia:

$$(13) \quad \mu_a'' - \mu_a = \frac{\sum s_i \Delta p_i}{\sum p_i}$$

essendosi posto:

$$p_i' = p_i + \Delta p_i.$$

La differenza $\mu_a'' - \mu_a$ ha lo stesso segno della quantità:

$$\sum s_i \Delta p_i.$$

1) Nel caso di un *aumento* generale dei pesi, i Δp_i sono tutti positivi. Degli scostamenti s_i , invece, alcuni debbono sempre essere positivi, altri negativi. Indicando:

con s' i valori numerici degli scostamenti positivi,

con $\Delta p'$ " " dei pesi corrispondenti,

con s'' " " degli scostamenti negativi,

con $\Delta p''$ " " dei pesi corrispondenti,

potremo dire che i prodotti $s' \Delta p'$ saranno positivi, i prodotti $s'' \Delta p''$ negativi e avremo:

$$(14) \quad \sum s_i \Delta p_i = \sum s' \Delta p' - \sum s'' \Delta p''.$$

Per conseguenza sarà:

$$\mu_a'' \begin{matrix} > \\ = \\ < \end{matrix} \mu_a$$

senonchè

$$\sum s' \Delta p' \begin{matrix} > \\ = \\ < \end{matrix} \sum s'' \Delta p''.$$

Ora tutti e tre i casi possono avverarsi, a seconda dei valori numerici che si assegnano ai vari incrementi dei pesi.

2) Nel caso di una *diminuzione* generale dei pesi, i Δp_i sono tutti negativi. I prodotti $s' \Delta p'$ saranno quindi negativi, i prodotti $s'' \Delta p''$ saranno positivi e avremo:

$$(15) \quad \sum s_i \Delta p_i = \sum s'' \Delta p'' - \sum s' \Delta p'.$$

Sicchè potrà dirsi che:

$$\mu_a'' \begin{matrix} > \\ = \\ < \end{matrix} \mu_a$$

secondo che:

$$\Sigma s'' \Delta p'' \begin{matrix} > \\ = \\ < \end{matrix} \Sigma s' \Delta p'.$$

Concludendo diremo che:

Tanto un aumento generale, quanto una generale diminuzione dei pesi, possono produrre ugualmente uno dei tre effetti seguenti:

- a) *di accrescere la media,*
- b) *di lasciarla inalterata,*
- c) *di scemarla.*

§ 18. — **Applicazioni.** — Merita di essere posto in rilievo il caso di un *aumento* generale dei pesi che abbia per conseguenza *una diminuzione della media (rimanendo inalterati i termini)*, o l'altro caso di una *diminuzione* generale dei pesi che abbia per conseguenza *un aumento della media (rimanendo inalterati i termini)*.

Per esempio, è teoricamente possibile che, rimanendo immutati o quasi da un anno all'altro i rendimenti per ettaro di una determinata coltura in più paesi, e aumentando dal primo al secondo anno la superficie coltivata in tutti i paesi, il rendimento per ettaro *medio*, nel complesso dei paesi considerati, diminuisca: ciò avverrà quando le superfici coltivate crescano in modo da soddisfare la disuguaglianza:

$$\Sigma s' \Delta p' < \Sigma s'' \Delta p''$$

dove i $\Delta p'$ rappresentano gl'incrementi di superficie nei paesi a rendimento superiore al medio, s' le eccedenze del rendimento di questi paesi sul rendimento medio, i $\Delta p''$ rappresentano gl'incrementi di superficie nei paesi a rendimento inferiore al medio, s'' le differenze in meno del rendimento di questi paesi per rapporto al rendimento medio.

Parimenti, dato che da un tempo all'altro i prezzi di varie qualità di una stessa merce rimangano inalterati o presso a poco, può darsi che il prezzo medio diminuisca, sebbene le quantità smerciate di ciascuna qualità siano tutte cresciute; ciò avverrà quando le quantità, vendute ad un prezzo inferiore al prezzo medio precedente, crescano relativamente più delle quantità vendute a un prezzo superiore al medio. Altri esempi potrebbero facilmente addursi.

Il fatto che, *rimanendo inalterati i termini* e variando i pesi, *la media può diminuire* è tale da suscitare sorpresa, ma vedremo fra breve che possono presentarsi casi in apparenza molto più strani.

SEZIONE VI. — **Variazione dei termini.**

§ 19. — **Accrescimento dei termini.** — Supponiamo che *aumenti* un termine della serie a_1, \dots, a_n . Poco monta che sia esso un termine superiore o inferiore alla media. La conseguenza è che la media *aumenta sempre*.

Dimostreremo questo assunto prima con considerazioni di meccanica e poi con considerazioni analitiche.

Siano i soliti due pesi p_1 e p_2 e i soliti due bracci a_1 e a_2 . Sappiamo che un peso uguale a P , avente per braccio la media ponderata μ_a , fa equilibrio ai due pesi p_1 e p_2 e sappiamo pure che la risultante P si può decomporre anche in due pesi π_1 e π_2 aventi entrambi il braccio $\overline{OM} = \mu_a$ e di cui il primo fa equilibrio al peso p_1 avente per braccio a_1 e il secondo fa equilibrio al peso p_2 avente per braccio a_2 .

Se aumentiamo della quantità Δa_1 il braccio a_1 dobbiamo aumentare anche il braccio del peso π_1 . Avremo così applicati a destra del punto O due pesi, uno π_2 col braccio $\overline{OM} = \mu_a$ e l'altro π_1 con un braccio più lungo di \overline{OM} . A questi due pesi ne possiamo sostituire un altro uguale alla loro somma, ossia uguale a P , e avente il suo braccio compreso fra μ_a e $\mu_a + \Delta a_1$, maggiore dunque di μ_a . E poichè questo nuovo braccio è la nuova media cercata, si vede dunque che la media aumenta.

Ripetendo il ragionamento per il braccio di p_2 si arriva alla medesima conclusione.

§ 20. — **Diminuzione dei termini.** — **Dimostrazione analitica del teorema sulla variazione dei termini.** — È pure facile dimostrare che, se uno dei bracci a_1, a_2 diminuisce, occorre spostare in dentro una parte del peso P e più propriamente il peso π_1 o π_2 , secondo che il braccio diminuito sia a_1 oppure a_2 . In ogni caso la media diminuisce.

È facile dare una dimostrazione analitica del principio generale che: *Aumentando uno o più termini della serie, la media ponderata*

aumenta, diminuendo uno o più termini della serie la media ponderata diminuisce.

Chiamando Δa_i gl'incrementi dei vari a_i , e μ_a''' la media che si ottiene quando i vari termini subiscono gl'incrementi Δa_i , avremo :

$$\begin{aligned}\mu_a''' &= \frac{\sum (a_i + \Delta a_i) p_i}{\sum p_i} \\ &= \mu_a + \frac{\sum p_i \Delta a_i}{\sum p_i}.\end{aligned}$$

Quindi:

$$(16) \quad \mu_a''' - \mu_a = \frac{\sum p_i \Delta a_i}{\sum p_i}.$$

Distinguiamo due casi:

1) I Δa_i sono positivi. Allora :

$$\mu_a''' > \mu_a.$$

2) I Δa_i sono negativi. Allora :

$$\mu_a''' < \mu_a.$$

In altre parole il segno della differenza :

$$\mu_a''' - \mu_a$$

è indipendente dai valori numerici dei pesi e dipende invece soltanto dal segno degl'incrementi Δa_i .

Appunto come si doveva dimostrare.

§ 21. — **Diversa sensibilità della media ponderata di fronte alle variazioni dei termini e alle variazioni dei pesi.** — Finora ci siamo occupati soltanto di vedere *in che senso* varia la media quando variano in un certo senso o solo i termini o solo i pesi. Ma è molto interessante d'istituire un confronto fra il *valore numerico* (valore assoluto) della variazione che subisce la media ponderata per effetto dell' *incremento di un termine* a_h e il valore numerico della variazione che la stessa media subisce per effetto d'un uguale *incremento del peso* corrispondente p_h .

Sia Δa_h l'incremento del termine a_h . Allora la media μ_a , come si può facilmente dedurre dalla formola (16) del § 20, subisce la variazione:

$$\Delta' \mu = \frac{p_h \Delta a_h}{\sum p_i}.$$

Chiamiamo ε_a l'incremento $\frac{\Delta a_h}{a_h}$ che tocca a ciascuna *unità* di a_h , per modo che:

$$\Delta a_h = \varepsilon_a a_h.$$

Allora:

$$\Delta' \mu = \frac{\varepsilon_a a_h p_h}{\sum p_h}.$$

Sia Δp_h l'incremento del peso p_h . Allora la media μ_a , come si può facilmente dedurre dalla formula (13) del § 17, subisce la variazione:

$$\Delta'' \mu = \frac{s_h \Delta p_h}{\sum p_h + \Delta p_h}.$$

Chiamiamo ε_p l'incremento $\frac{\Delta p_h}{p_h}$ che tocca a ciascuna *unità* di p_h , per modo che:

$$\Delta p_h = \varepsilon_p p_h.$$

Allora:

$$\Delta'' \mu = \frac{\varepsilon_p p_h s_h}{\sum p_h + \varepsilon_p p_h}.$$

Supposto ora che sia:

$$\varepsilon_a = \varepsilon_p = \eta$$

il rapporto tra i due incrementi diventa:

$$(17) \quad \frac{\Delta' \mu}{\Delta'' \mu} = \frac{a_h}{s_h} \left(1 + \eta \frac{p_h}{\sum p_h} \right).$$

Se $\Delta' \mu$ misura la *sensibilità della media* alle variazioni del termine a_h , ed $\Delta'' \mu$ misura la *sensibilità della media* alle variazioni del peso p_h , possiamo dire che *la media ponderata è assai più sensibile alle variazioni nei termini che non alle variazioni nei pesi.*

Infatti il termine a_h sarà di regola superiore in valore assoluto, e notevolmente superiore, al suo *scostamento* dalla media ponderata μ_a . Quindi il rapporto $\frac{a_h}{s_h}$ sarà assai superiore all'unità. Questo rapporto va moltiplicato ancora per una quantità superiore all'unità; sicchè non vi è dubbio che $\Delta' \mu$ debba essere assai maggiore di $\Delta'' \mu$.

Se η fosse una frazione assai piccola dell'unità, il prodotto η per la quantità $\frac{p_h}{\sum p_h}$ minore dell'unità potrebbe trascurarsi e avremmo:

$$(18) \quad \frac{\Delta' \mu}{\Delta'' \mu} = \frac{a_h}{s_h}.$$

Cioè: la variazione che la media subisce per l'incremento proporzionale di un *termine* è, in valore assoluto, tanto maggiore della variazione dovuta al medesimo incremento proporzionale nel *peso* corrispondente, quanto più grande è il *termine* per rapporto al suo *scostamento* dalla media ponderata. Ciò significa che, se un termine è per esempio 15 volte maggiore del suo scostamento dalla media ponderata, questa media è 15 volte meno sensibile alle variazioni del peso (ridotte all'unità) che non alle variazioni del termine (ridotte anch'esse all'unità). Sicchè, se si ha ragione di ritenere che un errore del 2 % nel termine produrrebbe nella media una variazione trascurabile, si ha uguale ragione di trascurare una variazione del 30 % nel peso del termine (1).

(1) Anche senza dare una dimostrazione rigorosa, come nel testo, s'intuisce che una variazione in un *termine* debba essere risentita dalla media più che una variazione nel *peso*, perchè l'incremento del termine fa variare soltanto il numeratore del rapporto $\frac{\sum ap}{\sum p}$, mentre una variazione nel peso fa variare nello stesso senso tanto il numeratore quanto il denominatore del rapporto medesimo. Vogliamo mostrare i diversi effetti delle due variazioni *in un caso limite*.

1° Sappiamo già che, se tutti i termini d'una serie vengono moltiplicati per una stessa costante k , la media rimane moltiplicata per quella costante (cfr. § 5). Basta porre

$$k = 1 + \epsilon$$

dove ϵ è una frazione (positiva o negativa) dell'unità, e il teorema sopra annunciato si traduce nell'altro: *Se tutti i termini d'una serie subiscono uno stesso aumento (o una stessa diminuzione) proporzionale, anche la media ponderata subisce il medesimo aumento (o la medesima diminuzione) proporzionale.*

2° Se invece tutti i pesi vengono moltiplicati e divisi per una stessa costante k , la media rimane inalterata. Infatti:

$$\frac{\sum kap}{\sum kp} = \frac{\sum ap}{\sum p} = \mu.$$

Ponendo ancora:

$$k = 1 + \epsilon$$

il teorema sopra enunciato si traduce nell'altro: *Se tutti i pesi di una serie subiscono il medesimo aumento (o la medesima diminuzione) proporzionale, la media ponderata dei termini non cambia.*

Nel primo caso abbiamo una sensibilità massima, nel secondo caso una sensibilità nulla.

SEZIONE VII. — Paradosso della media ponderata.

§ 22. — Enunciazione del paradosso nel caso di una serie di fenomeni a due dimensioni. — Abbiamo visto che, se aumentano di valore i *termini* di una serie, la media *aumenta* sempre, mentre, se aumentano di valore i *pesi*, la media può *crescere* o *diminuire*, secondo che i pesi cresciuti si riferiscono a termini superiori o inferiori alla media.

Questa conseguenza è molto importante, perchè dimostra che l'accrescimento d'un termine può essere compensato (e anche più che compensato) dalla diminuzione di un peso di braccio superiore al medio o dall'aumento di un peso di braccio inferiore al medio.

Cosicchè è pensabile un caso addirittura paradossale, in cui l'aumento generale dei termini produce una diminuzione della loro media, o in cui la diminuzione generale dei termini abbia come conseguenza un accrescimento della loro media.

Chiamiamo come al solito Δa_i gli incrementi dei termini, Δp_i gli incrementi dei pesi, e chiamiamo μ_a^{IV} la media che si ottiene quando variano contemporaneamente i termini e i pesi. Abbiamo:

$$\mu_a^{IV} = \frac{\sum (a_i + \Delta a_i) (p_i + \Delta p_i)}{\sum (p_i + \Delta p_i)}.$$

Sviluppando e ricordando la (2) del § 1 e la (5) del § 6 abbiamo:

$$\begin{aligned} \mu_a^{IV} &= \frac{\mu_a \sum p_i + \sum (\mu_a + s_i) \Delta p_i + \sum p_i \Delta a_i + \sum \Delta a_i \Delta p_i}{\sum (p_i + \Delta p_i)} \\ &= \mu_a + \frac{\sum s_i \Delta p_i + \sum p_i \Delta a_i + \sum \Delta a_i \Delta p_i}{\sum (p_i + \Delta p_i)}. \end{aligned}$$

Infine, ponendo:

$$p'_i = p_i + \Delta p_i$$

otteniamo:

$$(19) \quad \mu_a^{IV} = \mu_a + \frac{\sum s_i \Delta p_i + \sum p'_i \Delta a_i}{\sum p'_i}.$$

Sicchè il segno della differenza:

$$\mu_a^{IV} - \mu_a$$

è il medesimo del segno della somma algebrica:

$$(20) \quad \zeta = \sum s_i \Delta p_i + \sum p_i' \Delta a_i$$

1° caso. — Supponiamo che *tutti i* Δa_i *siano positivi*. Allora il termine $\sum p_i' \Delta a_i$ è positivo. Se vogliamo che sia:

$$\mu_a^{IV} < \mu_a$$

ossia che ζ sia negativo, dobbiamo quindi fare in modo che l'altro termine $\sum s_i \Delta p_i$ sia negativo, e *che superi in valore assoluto* il valore del primo termine.

Per essere assolutamente sicuri che $\sum s_i \Delta p_i$ sia negativo, basta dare ai vari Δp_i il *segno opposto* al segno del corrispondente s_i , ossia basta *accrescere* i pesi dei termini *inferiori* alla media e *diminuire* i pesi dei termini *superiori* alla media.

Per fare poi in modo che il valore numerico di $\sum s_i \Delta p_i$ superi il valore numerico di $\sum p_i' \Delta a_i$, basta scegliere opportunamente i valori dei vari Δa_i o dei vari Δp_i .

2° caso. — Supponiamo che *tutti i* Δa_i *siano negativi*. È facile dimostrare che dando ai vari Δp_i lo *stesso segno* dei corrispondenti s_i e scegliendo opportunamente i valori numerici dei vari Δa_i o dei vari Δp_i si può fare in modo che ζ sia *positivo*, ossia che si abbia:

$$\mu_a^{IV} > \mu_a.$$

Questo principio che: *tutti i termini di una serie possono variare in un senso eppure la loro media ponderata può variare nel senso opposto* si potrebbe chiamare il *paradosso della media ponderata*.

Per esempio è teoricamente possibile che tutti i rendimenti per ettaro di una determinata cultura in più paesi, o in più regioni di un paese, *aumentino* da un anno all'altro e che la media generale del rendimento per ettaro *diminuisca*: oppure che tutti i prezzi a cui si sono vendute determinate partite di una stessa merce in diversi mercati *crescano* da un momento all'altro e che il prezzo medio *diminuisca*; e via dicendo.

§ 23. — Esempio numerico. — Diamo un esempio numerico. Sia:

$$\begin{array}{ll} a_1 = 10 & p_1 = 7.000 \\ a_2 = 14 & p_2 = 9.000 \\ a_3 = 16 & p_3 = 4.000 \\ a_4 = 9 & p_4 = 20.000. \end{array}$$

La media ponderata risulta uguale a:

$$\mu_a = 11.$$

Supponiamo che tutti i termini *a* diminuiscono di 1; ossia poniamo:

$$\Delta a_1 = \Delta a_2 = \Delta a_3 = \Delta a_4 = -1.$$

Vediamo quali valori debbano avere i Δp_i perchè la nuova media risulti *superiore* all'antica.

Chiamiamo Δ_i i *valori numerici* dei Δp_i . A questi Δ_i daremo lo stesso *segno* che hanno gli s_i corrispondenti. I calcoli richiesti dalla (20) sono eseguiti nel tabellino qui appresso:

s_i	Δp_i	Δa_i	$s_i \Delta p_i$	$\frac{p'_i \Delta a_i}{(p_i + \Delta p_i) \Delta a_i}$	$s_i \Delta p_i + p'_i \Delta a_i$
-1	$-\Delta_1$	-1	$+\Delta_1$	$\Delta_1 - 7.000$	$2\Delta_1 - 7.000$
+3	$+\Delta_2$	-1	$+3\Delta_2$	$-\Delta_2 - 9.000$	$2\Delta_2 - 9.000$
+5	$+\Delta_3$	-1	$+5\Delta_3$	$-\Delta_3 - 4.000$	$4\Delta_3 - 4.000$
-2	$-\Delta_4$	-1	$+2\Delta_4$	$\Delta_4 - 20.000$	$3\Delta_4 - 20.000$

$$\zeta = 2\Delta_1 + 2\Delta_2 + 4\Delta_3 + 3\Delta_4 - 40.000.$$

Perchè dunque la nuova media risulti *superiore* all'antica occorre che ζ sia *positivo*, ossia che

$$2\Delta_1 + 2\Delta_2 + 4\Delta_3 + 3\Delta_4 > 40.000.$$

Vi sono infiniti modi di soddisfare questa disuguaglianza. Uno per es. consiste nel porre:

$$\begin{array}{l} \Delta_1 = 2000 \\ \Delta_2 = 4000 \\ \Delta_3 = 4000 \\ \Delta_4 = 5000. \end{array}$$

Allora abbiamo:

$$p'_1 = 7000 - 2000 = 5000$$

$$p'_2 = 9000 + 4000 = 13000$$

$$p'_3 = 4000 + 4000 = 8000$$

$$p'_4 = 20000 - 5000 = 15000$$

e siccome

$$a'_1 = 9$$

$$a'_2 = 13$$

$$a'_3 = 15$$

$$a'_4 = 8$$

la nuova media risulta uguale a:

$$\mu_a^{IV} = \frac{454}{41} = 11,1.$$

Sebbene tutti i termini della serie siano *diminuiti*, la nuova media è *cresciuta* rispetto alla precedente.

§ 24. — Il paradosso nelle medie di indici. — Il principio dimostrato nel paragrafo precedente si traduce subito in linguaggio di indici.

Dire infatti che tutti i termini di una serie vengono accresciuti equivale a dire che, mettendo uguali a 100 i valori iniziali dei vari termini, *gli indici* dei nuovi valori rispetto ai precedenti *risultano tutti superiori a 100*. Dire che la media ponderata può diminuire, equivale a dire che, mettendo uguale a 100 il valore iniziale della media, *l'indice della nuova media* rispetto all'antica può risultare *inferiore a 100*.

Il paradosso enunciato nel paragrafo precedente si traduce allora in quest'altra forma che: *sebbene gli indici della nuova serie di termini siano tutti superiori (o inferiori) a 100, l'indice della nuova media può riuscire inferiore (o superiore) a 100*, qualora i pesi non rimangano inalterati (1).

Ciò basta per poter affermare che l'indice di una media ponderata *non merita il nome di media* degli indici dei vari termini. Infatti il paradosso dimostra che non è soddisfatto il secondo dei requisiti essenziali di una media (vedi § 4).

(1) La proposizione dimostrata nel § 8 [cfr. la (10)], si riferisce al caso in cui i pesi rimangano costanti e varino soltanto i termini e diventando *e'*.

§ 25. — Il paradosso nelle medie di valori proporzionali. —
Date le due serie (vedi § 8):

$$\begin{array}{c} c_1, c_2 \dots c_k \\ d_1, d_2 \dots d_k \end{array}$$

e calcolati i valori proporzionali:

$$\rho_1, \rho_2 \dots \rho_k$$

supponiamo che tanto i termini della prima serie quanto quelli della seconda subiscano tali variazioni da far variare i vari ρ tutti nello stesso senso. Si domanda se il valore proporzionale medio muterà nello stesso senso o in senso opposto ai vari ρ .

Abbiamo dimostrato nel § 8 che il valore proporzionale medio non è altro che la media ponderata dei vari ρ , e quindi la risposta è che sono possibili tutti i casi. Cioè tanto nell'ipotesi di un aumento generale dei vari ρ , quanto nell'ipotesi di una loro generale diminuzione, può darsi che la loro media cresca, diminuisca o rimanga costante.

Possiamo citare un esempio tolto dalla statistica del bestiame. In un primo periodo abbiamo lo specchietto seguente (1):

Paese	Anno della rilevazione	Popolazione (b')	Numero assoluto dei bovini (c')	Numero dei bovini per 1000 abitanti (p')
Giappone	1880	35.929.060	1.124.564	31,3
Russia Asiatica	1872	12.807.124	5.365.000	418,9
Totali e media		48.736.184	6.489.564	133,2

In un secondo periodo abbiamo quest'altro specchietto (1):

Paese	Anno della rilevazione	Popolazione (b'')	Numero assoluto dei bovini (c'')	Numero dei bovini per 1000 abitanti (p'')
Giappone	1908	52.340.220	1.297.974	24,8
Russia Asiatica	1909	27.262.800	9.637.330	353,5
Totali e media		79.603.020	10.935.304	137,4

Come si vede, entrambi i rapporti per 1000, o valori proporzionali, sono *diminuiti* dal 1° al 2° periodo, e ciò non di meno il rapporto medio è *cresciuto*.

(1) Questi dati si possono trovare in una pubblicazione dell'Istituto Internazionale d'Agricoltura « *Statistique des superficies cultivées, de la production végétale et du bétail dans les Pays adhérents* » p. 156 e 149, 154 e 148.

SEZIONE VIII. — **Medie ponderate doppie, triple, ennuple.**

§ 26. — **Fenomeni a più di 2 dimensioni.** — Nel § 7 abbiamo parlato di fenomeni a due dimensioni. Per analogia potremo parlare di fenomeni a tre dimensioni. Diremo che un certo fenomeno G è a tre dimensioni, quando esso si può decomporre in 3 altri fenomeni P, Q, F collegati tra loro in tal modo che sia:

$$g_i = p_i q_i f_i$$

dove:

g_i indica genericamente una misura del fenomeno G ,

p_i indica genericamente una misura del fenomeno P ,

q_i indica genericamente la misura di un certo requisito posseduto da ciascuna unità di p_i ,

f_i indica genericamente la misura di un certo requisito posseduto da ciascuna unità di q_i .

Chiamiamo R il fenomeno a 2 dimensioni che risulta dalla composizione dei due fenomeni P e Q ed r_i indichi genericamente una misura del fenomeno R . Avremo:

$$r_i = p_i q_i.$$

Allora possiamo dire che il fenomeno G risulta dalla composizione dei due fenomeni F ed R e potremo scrivere:

$$g_i = f_i r_i.$$

L'intensità media del fenomeno G riferita all'unità di misura del fenomeno R non è altro che la media ponderata:

$$\mu_r = \frac{\sum f_i r_i}{\sum r_i}.$$

Ma essa può scriversi anche:

$$(21) \quad \mu_r = \frac{\sum f_i p_i q_i}{\sum p_i q_i}$$

ed esprime allora l'intensità media del fenomeno G riferita all'unità di misura del fenomeno Q (che è la stessa unità di misura del fenomeno R).

La quantità μ_r così calcolata si può chiamare una media *ponderata doppia* dei termini f_i in quanto ogni termine f_i entra nella media in ragione composta dei due pesi p_i e q_i .

Sarebbe un errore dire per esempio che la intensità media del fenomeno G sia misurata dalla media:

$$\frac{\sum f_i p_i}{\sum p_i}$$

È facile vedere che una media ponderata doppia gode di tutte le proprietà della media ponderata ordinaria (cfr. § 6) purchè si abbia l'accortezza di considerare il prodotto dei due pesi p_i e q_i come un peso unico.

Per dare un esempio di fenomeni a 3 dimensioni diciamo che:

P rappresenta il fenomeno « superficie coltivata a frumento in diversi paesi ». Esso ha per unità di misura l'ettaro.

Q rappresenta il fenomeno « rendimento per ettaro della coltura del frumento in diversi paesi ». Esso ha per unità di misura l'ettolitro.

F rappresenta il fenomeno « peso (1) di un ettolitro di frumento in vari paesi ». Esso ha per unità di misura il quintale.

Allora G può rappresentare il fenomeno « produzione complessiva della coltura del frumento in vari paesi » e può considerarsi come un fenomeno a tre dimensioni. Esso ha per unità di misura la stessa unità di misura del fenomeno F, cioè il quintale.

Se ora cerchiamo l'intensità media del fenomeno G espressa in unità di misura del fenomeno F (che è la stessa unità di misura di G) e riferita all'unità di misura del fenomeno Q, dobbiamo per l'appunto fare una media ponderata doppia.

Una media ponderata *tripla* dovremmo fare per calcolare il prezzo medio di un quintale qualora fossero dati anche i *prezzi* di ciascun

(1) Si tratta di peso *fisico* e non di peso *statistico*.

quintale. Infatti il valore (di scambio) complessivo del raccolto si ottiene moltiplicando ogni superficie

per il rendimento per ettaro (espresso in ettolitri),
 per il peso di un ettolitro (espresso in quintali),
 per il prezzo di un quintale (espresso in lire),

e addizionando insieme tutti i prodotti così ottenuti. La somma si dovrebbe poi dividere per il numero complessivo dei quintali, il quale si ottiene moltiplicando ogni superficie

per il rendimento per ettaro (espresso in ettolitri),
 per il peso di un ettolitro (espresso in quintali),

e addizionando i prodotti così ottenuti. Il quoziente dei due prodotti indica il prezzo medio di un quintale.

In generale, dato un fenomeno ad n dimensioni — e quindi data una serie di termini ed $n - 1$ serie di pesi — la media ponderata di grado n^{mo} si trova moltiplicando ciascun termine per il prodotto dei rispettivi $n - 1$ pesi, addizionando i prodotti, e dividendo la somma così ottenuta per la somma dei prodotti degli $n - 1$ pesi. Il quoziente, o media ponderata di grado n^{mo} , indica l'intensità media del fenomeno espressa nella stessa unità di misura dei termini e riferita all'unità di misura dell'ultima serie di pesi, ossia della serie dei pesi di ordine $n - 1$.

§ 27. — **Indici di grado superiore al primo.** — Se la quantità e_1 , che esprime una certa misura del fenomeno E in un certo momento assume alla fine di un primo, di un secondo, di un n^{o} intervallo i valori:

$$e'_1, e''_1, \dots, e^n_1$$

noi chiameremo rispettivamente *indici di 1^o, di 2^o, di n^{mo} grado* delle quantità $e'_1, e''_1, \dots, e^n_1$ rispetto al termine di riferimento e_1 , i rapporti:

$$(22) \quad \frac{100e'_1}{e_1} = i'_1, \quad \frac{100e''_1}{e_1} = i''_1, \dots, \frac{100e^n_1}{e_1^{n-1}} = i^n_1.$$

Supponiamo ora di avere una serie di indici di *secondo* grado:

$$i''_1, i''_2, \dots, i''_k$$

la quale non può essere evidentemente che una serie nel tempo. Proponiamoci di calcolare la media di questi indici. Essa è (cfr. § 8):

$$i_{\mu}'' = \frac{100 \sum e_i''}{\sum e_i''}.$$

Siccome in generale:

$$e_i' = \frac{e_i i_i'}{100}$$

$$e_i'' = \frac{e_i i_i' i_i''}{100^2}$$

abbiamo:

$$i_{\mu}'' = \frac{\sum (e_i i_i') i_i''}{\sum (e_i i_i')}$$

la quale è una *media ponderata doppia* dei vari i_i'' , ossia una media fatta in ragione composta di due ordini di pesi costituiti dai *primi termini di riferimento* e_i e dagli *indici di primo grado* i_i' . In generale la *media di indici di grado n^{mo}* è una *media ponderata n^{pla}* ossia una media degli i^n fatta in ragione composta dei primi termini di riferimento e di tutti gli indici di grado inferiore a n :

$$(23) \quad i_{\mu}^n = \frac{\sum (e_i i_i' i_i'' \dots i_i^{n-1}) i_i^n}{\sum (e_i i_i' i_i'' \dots i_i^{n-1})}.$$

Esempi di indici di grado superiore al primo sarebbero, per rimanere nella cerchia degli esempi che abbiamo finora dati, indici che esprimessero la previsione dei raccolti in percentuali di un rendimento anteriormente previsto, il quale a sua volta fosse già espresso in percentuale di un rendimento determinato o di una percentuale di questo rendimento; oppure indici che esprimessero il rapporto fra i prezzi attuali e i prezzi di un periodo anteriore, essendo questi ultimi già espressi a loro volta in percentuali di prezzi antecedenti, e così via.

Se ciò avvenisse, e si volessero poi calcolare le medie degli indici di 2°, di 3°, di n^{mo} grado, si dovrebbero fare medie ponderate doppie, triple, ennuple.

