ATTI DEL REALE ISTITUTO VENETO DI SCIENZE, LETTERE ED ARTI.
Anno accademico 1913-914 - Tomo LXXIII - Parte seconda.

Sulla misura della concentrazione e della variabilità dei caratteri

CORRADO GINI
prof. ord. di Statistica nella R. Università di Padova

(presentata dal prof. C. F. Ferraris, m. e., nell' adun. del 29 marzo 1914)

1. — Si abbiano n quantità che misurano l' intensità di un certo carattere in n casi. Ordiniamo tali quantità in modo che ognuna di esse sia inferiore od eguale alla successiva, e con a_k ($k = 1, 2 .... n$) indichiamo la quantità che, nella successione così ottenuta, presenta il numero d' ordine k.

Per due qualsiasi valori (i, l) di k, tali che sia $i < l$, sarà

$$a_i \leq a_l \tag{1}$$

$$\frac{1}{i} \sum_{k=1}^{i} a_k \leq \frac{1}{l} \sum_{k=1}^{l} a_k \tag{2}$$

Da questa relazione si ricava l' altra

$$\frac{\sum_{k=1}^{i} a_k}{\sum_{k=1}^{l} a_k} \leq \frac{i}{l} \tag{3}$$

e quindi, come caso particolare, per $l = n$,

$$\frac{\sum_{k=1}^{i} a_k}{\sum_{k=1}^{n} a_k} \leq \frac{i}{n}$$

Poniamo, per semplicità, $\sum_{k=1}^{i} a_k = A_i$; $\sum_{k=1}^{n} a_k = A_n$; $\frac{A_i}{A_n} = q_i$; $\frac{i}{n} = p_i$; $\frac{A_i}{i} = M_i$; $\frac{A_n}{n} = M_n$; $\frac{A_n - A_i}{n - i} = M_{n-i}$, dove è sempre $i < n$.

Diremo che la *concentrazione del carattere è tanto più forte quanto più è forte, per gli* n — 1 *valori di* i, *la disuguaglianza*

$$p_i > q_i \tag{4}$$

In altre parole, diremo che *la concentrazione di un carattere è tanto più forte, quanto più è piccola la parte che, sull' ammontare totale del carattere, spetta a quella parte dei casi, in cui l' intensità del carattere non supera un certo limite.*

La *concentrazione del carattere* si dirà *perfetta* quando l' intensità del carattere sia $= 0$ in $n - 1$ casi ed $= A_n$ in un caso. Sarà allora, per gli $n - 1$ valori di i, $q_i = 0$ e quindi $p_i - q_i = p_i$. Se in tutti i casi il carattere presenterà la stessa intensità, in modo che sia, per tutti gli $n - 1$ valori di i, $p_i = q_i$, diremo che la *concentrazione del carattere è nulla* o in altre parole che vi è *equidistribuzione del carattere.*

Se, con una o più costanti, si riesce ad esprimere, tra i due membri della disuguaglianza (4) (o di una disuguaglianza che da essa si deduce), una relazione valida, con maggiore o minore approssimazione, per tutti i valori di i, potremo assumere tale costante, o l' insieme di tali costanti, come *indice di concentrazione* del carattere.

In luogo della disuguaglianza (4) è talvolta conveniente considerare, a tale scopo, le disuguaglianze

$$1 - p_i < 1 - q_i \tag{5}$$

$$M_{n-i} > M_n \tag{6}$$

che dalla (4) facilmente si deducono.

In un nostro precedente studio, abbiamo distinto gli indici di concentrazione in *semplici* e *complessi*, secondo che risultano da una o più costanti, e, degli uni e degli altri, abbiamo dato qualche esempio.

L' indice di concentrazione dei redditi individuali o famigliari e quello dei valori locativi famigliari possono desumersi quasi sempre con sufficiente approssimazione dalla equazione

$$1 - p_i = (1 - q_i)^{\delta} \; ; \tag{7}$$

quello della prolificità delle coppie matrimoniali dalla equazione

$$q_i = p_i^{\delta} \tag{8}$$

o, con migliori approssimazioni, dall' altra

$$q_i = p_i^{\delta + \varepsilon (a_i - 1)} \tag{9}$$

o anche dalla seguente

$$M_{n-i} = M_n + \delta a_i + \frac{a_i^2}{100} \; . \tag{10}$$

Quello del valore delle successioni nelle provincie italiane e nei dipartimenti francesi si può pure ricavare dall' equazione (8) (1).

(1) Cfr., per tutto questo, la nostra memoria *Indici di concentrazione e di dipendenza* in *Biblioteca dell' Economista*, V. Serie, vol. XX, Torino, Unione Tipografico-editrice, 1910. Un riassunto di questa memoria, presentata alla *Terza riunione della Società italiana per il progresso delle scienze* (ottobre 1909) fu pubblicato negli *Atti* di detta Società (Roma, Bertero 1910). Il concetto di concentrazione e l' indice di concentrazione per i redditi e i valori locativi erano stati però esposti già un anno prima alla *Seconda riunione della Società italiana per il progresso delle scienze* (ottobre 1908), in una comunicazione dal titolo *Il*

Questi varii indici hanno una innegabile utilità, in quanto permettono di confrontare la concentrazione di uno stesso carattere in popolazioni e in tempi diversi (per es. la concentrazione dei redditi individuali nei diversi paesi e nei diversi anni); talora anche di caratteri differenti, che presentano però distribuzioni analoghe e per cui l'indice si può pertanto ricavare dalla medesima formula (per es. la concentrazione dei redditi individuali e quella dei valori locativi). Ma, da una parte, i caratteri, per cui fu trovata una formula, che permetta di desumere un indice di concentrazione, sono finora ben pochi; dall'altra, anche se altre formule si trovassero per altri caratteri, resterebbero sempre impossibili i confronti tra i caratteri per cui l'indice di concentrazione si desume da formule diverse.

La presente nota ha lo scopo di proporre una misura della concentrazione, che sia indipendente dalla curva di distribuzione del carattere e permetta quindi di eseguire paragoni tra la concentrazione dei caratteri più varii.

2. — La disuguaglianza (4) è tanto più forte, in valore assoluto, quanto maggiore è la differenza $p_i - q_i$; tanto più forte in valore relativo quanto più è alto il rapporto

---

diverso accrescimento delle classi sociali e la concentrazione della ricchezza, che vide la luce sul *Giornale degli Economisti* (gennaio 1909).

Per applicazioni degli indici di concentrazione, cfr. G. Del Vecchio, *Ricerche statistiche sui depositi a risparmio*. Udine. Tosolini, 1910; L. V. Furlan, *Neue Literatur zur Einkommensverteilung in Italien*, in *Jahrbücher für Nationalökonomie und Statistik* (agosto 1911); F. Corridore, *Relazioni tra affitto reale e valore locativo fiscale nel Belgio*. Roma, Loescher 1911; E. Porru, *La concentrazione della ricchezza nelle diverse regioni d'Italia*, in *Studi economico-giuridici pubblicati per cura della Facoltà di giurisprudenza della R. Università di Cagliari*. Vol. IV. Parte I^a, 1911-1912; F. Savorgnan, *La distribuzione dei redditi nelle provincie e nelle grandi città dell'Austria*, in *Pubblicazioni del Museo commerciale di Trieste*. 1912; C. Gini, *Variabilità e Mutabilità, contributo allo studio delle distribuzioni e delle relazioni statistiche*. Fasc. I. *Introduzione — Indici di variabilità — Indici di mutabilità*, in *Studi economico-giuridici pubblicati per cura della Facoltà di giurisprudenza della R. Università di Cagliari*. Vol. III. P. II^a. Bologna, Cuppini. 1912.

$$R_i = \frac{p_i - q_i}{p_i}.$$

R_i rappresenta il coefficiente per cui bisogna moltiplicare la frazione p_i dei casi, in cui l'intensità del carattere non supera un dato limite, per ottenere la differenza tra codesta frazione e la frazione dell'ammontare del carattere che spetta a tali casi. Il valore di R_i passerà da un valore $= 1$, nel caso di concentrazione perfetta, a un valore $= 0$, nel caso di equidistribuzione del carattere.

A misura della concentrazione del carattere potremo assumere la media ponderata degli $n - 1$ valori di R_i, in cui ogni R_i entri con un peso proporzionale al valore di p_i.

Tale media sarà

$$R = \frac{\sum_{i=1}^{n-1} (p_i - q_i)}{\sum_{i=1}^{n-1} p_i}. \tag{11}$$

Chiameremo R *rapporto di concentrazione.*

Moltiplicando per R un valore di p_i scelto a caso, si ottiene il valore probabile della corrispondente differenza $p_i - q_i$.

Nel caso di equidistribuzione del carattere, sarà, per tutti i valori di i, $p_i = q_i$ e quindi $R = 0$; nel caso di massima concentrazione, in cui a un solo caso spetti tutto l'ammontare del carattere, sarà, per gli $n - 1$ valori di i, $q_i = 0$ e quindi $R = 1$.

Il rapporto di concentrazione è dunque un indice di concentrazione che verifica le condizioni seguenti:

a) cresce col crescere della concentrazione,

b) diviene $= 1$ nel caso di concentrazione massima,

c) diviene $= 0$ nel caso di concentrazione minima,

d) costituisce il coefficiente, per cui bisogna moltiplicare la frazione (p_i) dei casi, in cui il carattere presenta un'intensità sotto un certo limite, per ottenere il valore probabile della differenza ($p_i - q_i$) fra tale frazione e la frazione (q_i) dell'ammontare del carattere che spetta a tali casi.

3. — Vediamo come si determini praticamente il valore di R.

Sostituendo a p_i, q_i i rispettivi valori $\frac{i}{n}$ e $\frac{A_i}{A_n}$ e tenendo presente che è $\sum_{i=1}^{n-1} i = \frac{n(n-1)}{2}$, la (11) si riduce alla forma seguente, più comoda per il calcolo,

$$R = 1 - \frac{2}{(n-1)A_n} \sum_{i=1}^{n-1} A_i \tag{12}$$

o anche, tenendo presente che è $\sum_{i=1}^{n-1} A_i = \sum_{i=1}^{n-1} (n-i) a_i$, alla forma

$$R = \frac{2\sum_{i=1}^{n} (i-1) a_i}{(n-1)A_n} - 1 \tag{12bis}$$

Essa si può semplificare ulteriormente, nel caso in cui parecchi a_i presentino lo stesso valore. Se con x_l $(l = 1, 2, \ldots, s)$ si indica uno degli s valori, che assumono le n quantità a_i, con f_l la rispettiva frequenza, con i_l il numero delle quantità inferiori od eguali ad x_l e con $i_{l-1} = i_l - f_l$ il numero delle quantità inferiori ad x_l, la (12bis) si riduce facilmente alla forma seguente

$$R = \frac{2\sum_{l=1}^{s} x_l \sum_{i=i_{l-1}+1}^{i=i_l} (i-1)}{(n-1)A_n} - 1$$

o anche, tenendo presente che è

$$\frac{1}{f_l} \sum_{i=i_{l-1}+1}^{i=i_l} (i-1) = \frac{i_{l-1} + i_l - 1}{2},$$

alla seguente [1]

---

(1) In modo analogo si può semplificare la (12). Se si pone $A_l = \sum_{i=1}^{i=i_l} A_i$, $A_{l-1} = \sum_{i=1}^{i=i_{l-1}} A_i$ e si tiene presente che è $\sum_{i=1}^{n-1} A_i = \sum_{i=1}^{n} A_i - A_n$,

$$\frac{1}{f_l} \sum_{i=i_{l-1}+1}^{i=i_l} A_i = \frac{A_l + A_{l-1} + x_l}{2},$$

la (12) si riduce alla seguente

$$R = 1 - \frac{\sum_{l=1}^{s} f_l (A_l + A_{l-1}) - A_n}{(n-1)A_n}. \tag{13bis}$$

$$R = \frac{\sum_{l=1}^{s} (i_{l-1} + i_l - 1) f_l x_l}{(n - 1) A_n} - 1 \tag{13}$$

Questa formula permette di determinare rapidamente il valore di R quando le statistiche dieno i valori del carattere nei singoli casi. L' esempio seguente, relativo alle pulsazioni del cuore di 263 Pelli Rosse (maschi, adulti) mostra come si impostano e come si eseguono le operazioni.

Determinazione del rapporto di concentrazione secondo la formula (13)
Pulsazioni del cuore di 263 Pelli Rosse secondo Hrdlicka (1).

x_i	f_i	$f_i\ x_i$	i_i	$i_{i-1}-1$	$i_i + i_{i-1}-1$	$(i_i + i_{i-1}-1) f_i\ x_i$
44	1	44	1	— 1	0	0
45	1	45	2	0	2	90
48	3	144	5	1	6	864
49	1	49	6	4	10	490
50	4	200	10	5	15	3,000
51	3	153	13	9	22	1,166
52	5	260	18	12	30	7,800
53	2	106	20	17	37	3,922
54	12	648	32	19	51	33,048
55	4	220	36	31	67	14,740
56	19	1064	55	35	90	95,760
57	7	399	62	54	116	46.284
58	24	1392	86	61	147	204,624
59	7	413	93	85	178	73,514
60	23	1380	116	92	208	287,040
61	2	122	118	115	233	28,426
62	19	1178	137	117	254	299.212
63	11	693	148	136	284	196,812
64	19	1216	167	147	314	381,824
65	3	195	170	166	336	65,520
66	32	2112	202	169	371	783,552
67	5	335	207	201	408	136,680
68	18	1224	225	206	431	527,544
69	1	69	226	224	450	31,050
70	12	840	238	225	463	388,920
71	2	142	240	237	477	67,734
72	12	864	252	239	491	424,224
73	1	73	253	251	504	36,792
74	3	222	256	252	508	112,776
75	1	75	257	255	512	38,400
76	2	152	259	256	515	78,280
78	3	234	262	258	520	121,680
80	1	80	263	261	524	41,920
Totali	263	16343	—	—	—	4,533,688

(1) A. Hrdlicka. *Physiological and medical observations among the Indians of Southwestern United States and Northern Mexico.* Smithsonian Institution Bureau of American Ethnology. Washington. Government printing office. 1908, pagg. 348-371.

È $n - 1 = 262$

$A_n = 16,343$

$$R = \frac{4,533,688}{262 \times 16,343} - 1 = 5,88\ \%$$

In modo analogo ho calcolato il valore di R per le pulsazioni del cuore di 94 Egiziani (maschi, adulti) dell' oasi di Kharga (1). Risulta $R = 6,73\ \%$.

Molte volte però le statistiche danno solo le intensità del carattere raggruppate in classi più o meno ampie, talvolta aggiungendo, tale altra no, la somma delle intensità dei casi che rientrano nelle singole classi.

Consideriamo dapprima il caso, in cui, per ognuna delle classi, sia nota la somma delle intensità. Sieno r le classi in cui si dividono le n quantità e con k ($k = 1, 2 \ldots. r$) indichiamo il numero d' ordine di una di esse cominciando dalla più bassa. Indichiamo con f_k il numero delle quantità che rientrano nella classe k^{ma}, con S_k la somma di tali quantità, con l_k il limite superiore e con l_{k-1} il limite inferiore della classe, con i_k il numero delle quantità inferiori ad l_k e con i_{k-1} il numero delle quantità inferiori ad l_{k-1}.

Sia $\delta_{kl} = a_i - \frac{S_k}{f_k}$ (dove $l = 1, 2 \ldots. f_k$) lo scostamento di un valore a_i dalla media aritmetica degli f_k valori analoghi che rientrano nella stessa classe, e sia

$$\varepsilon_{kl} = (i - 1) - \frac{i_k + i_{k-1} - 1}{2}$$

lo scostamento di un valore $(i - 1)$ della media aritmetica degli f_k valori analoghi che rientrano nella stessa classe. Sarà

$$\sum_{l=1}^{f_k} \delta_{kl} = 0 \ , \quad \sum_{l=1}^{f_k} \varepsilon_{kl} = 0 \ ,$$

e quindi

$$2 \sum_{i=i_{k-1}+1}^{i=i_k} (i - 1)\, a_i = (i_k + i_{k-1} - 1)\, S_k + 2 \sum_{l=1}^{f_k} \varepsilon_{kl}\, \delta_{kl} \ .$$

(1) A. Hrdlicka, *The natives of Kharga Oasis Egypt*. Smithsonian Miscellaneous Collections. Vol. 59. Number 1, City of Washington, 1912. Pagg. 112-115.

Sostituendo nella (12^{bis}) in base a questa uguaglianza, si ottiene

$$R = \frac{\sum_{k=1}^{r} (i_k + i_{k-1} - 1) S_k + 2 \sum_{k=1}^{r} \sum_{l=1}^{f_k} \varepsilon_{kl} \delta_{kl}}{(n-1) A_n} - 1 \qquad (14)$$

La quantità $\sum_{k=1}^{r} \sum_{l=1}^{f_k} \varepsilon_{kl} \delta_{kl}$ è sempre positiva (1). Posto

---

(1) Si tenga presente che i valori di a_i crescono coi valori di $(i-1)$ e che i valori di $(i-1)$ costituiscono una progressione aritmetica a ragione 1.

Per ciò, in ogni classe, la mediana degli $(i-1)$ coinciderà con la loro media aritmetica $= \frac{i_{k-1} + i_k - 2}{2}$.

a) Nel caso che anche la mediana degli a_i della classe k^{ma} coincida con la loro media aritmetica $\frac{S_k}{f_k}$, tutti gli f_k prodotti $\varepsilon_{kl} \delta_{kl}$ saranno positivi, perchè ε_{kl} e δ_{kl} avranno sempre il medesimo segno.

b) Nel caso che la mediana degli a_i della classe k^{ma} non coincida con la loro media aritmetica $\frac{S_k}{f_k}$, sia v il numero dei termini che intercedono tra le due medie. Degli f_k prodotti $\varepsilon_{kl} \delta_{kl}$, v avranno allora segno negativo, perchè in v casi gli scostamenti ε_{kl} e δ_{kl} avranno segno contrario; ed $f_k - v$ avranno segno positivo.

α) Supponiamo che la media aritmetica sia superiore alla mediana degli a_i. Vi saranno allora v prodotti negativi in cui ε_{kl} avrà segno positivo e δ_{kl} avrà segno negativo, $\frac{f_k}{2} - v$ prodotti positivi in cui ε_{kl} e δ_{kl} avranno entrambi segno positivo, $\frac{f_k}{2}$ prodotti positivi in cui ε_{kl} e δ_{kl} avranno entrambi segno negativo. Si avverta ora che tutti i v valori di δ_{kl} nei prodotti negativi sono minori in valore assoluto di qualunque valore di δ_{kl} nei prodotti positivi in cui δ_{kl} ed ε_{kl} sono entrambi negativi, e che i v valori di ε_{kl} nei prodotti negativi sono eguali, ciascuno a ciascuno, ai v più piccoli valori di ε_{kl} nei prodotti positivi in cui δ_{kl} ed ε_{kl} sono entrambi negativi. Perciò il valore assoluto della somma dei v prodotti negativi sarà minore della somma dei v più piccoli prodotti positivi in cui i due fattori sono negativi, e *a fortiori* più piccolo della somma di tutti i prodotti positivi.

β) Analoga è la dimostrazione se si suppone che la mediana degli a_i sia superiore alla media aritmetica.

$$R' = \frac{\sum_{k=1}^{r} (i_k + i_{k-1} - 1) S_k}{(n-1) A_n} - 1 \tag{15}$$

sarà quindi sempre

$$R' < R \tag{16}$$

La differenza $R - R'$ cresce col crescere dei valori $\frac{\varepsilon_{kl}}{n-1}$ e $\frac{\delta_{kl}}{A_n}$; ora i valori di $\frac{\varepsilon_{kl}}{n-1}$ e, a parità di altre condizioni, anche quelli di $\frac{\delta_{kl}}{A_n}$ crescono col crescere di $\frac{f_k}{n}$, i valori di $\frac{\delta_{kl}}{A_n}$, poi, a parità di $\frac{f_k}{n}$, crescono anche col crescere delle differenze tra i valori di a_i. Potremo dunque dire che la differenza $R - R'$ cresce col crescere della concentrazione del carattere e della comprensività delle classi, in cui le intensità del carattere sono raggruppate.

In pratica però, quando si hanno dieci classi, di comprensività non troppo diverse tra loro, R' si può riguardare come un valore sufficientemente approssimato di R.

Nella tavola seguente sono eseguite le operazioni per il calcolo di R', in base alla formula (15), sui dati delle pulsazioni del cuore delle 263 Pelli Rosse, raggruppate in 7 classi [1].

---

(1) Quando n è abbastanza elevato, R' può determinarsi, anzi che in base alla (15), in base alla formula seguente

$$R'_2 = \frac{\sum_{k=1}^{n} (i_k + i_{k-1}) S_k}{n A_n} - 1 \,. \tag{15bis}$$

Ci si persuade facilmente che è $R'_2 = \frac{n-1}{n} R'$.

Determinazione del rapporto di concentrazione secondo la formula (15)
Pulsazioni del cuore di 263 Pelli Rosse secondo HRDLICKA.

l_{k-1}	l_k	f_k	S_k	i_k	$i_{k-1}-1$	$i_k+i_{k-1}-1$	$(i_k+i_{k-1}-1)S_k$
44	49	6	282	6	— 1	5	1410
50	54	26	1367	32	5	37	50,579
55	59	61	3488	93	31	124	432,512
60	64	74	4589	167	92	259	1,188,551
65	69	59	3935	226	166	392	1,542,520
70	74	30	2141	256	225	481	1,029,821
75	80	7	541	263	255	518	280.238
	Totali	263	16,343	—	—	—	4,525,631

$$R' = \frac{4,525,631}{262 \times 16,343} - 1 = 5,69\ \%$$

Il valore del rapporto di concentrazione, così ottenuto, è più basso del 3 % del valore esatto (= 5.88 %) dato dalla formula (13). Per gli Egiziani di Kharga, raggruppando i dati in 9 classi, si ottiene $R' = 6.59\ \%$, un valore più basso del 2 % del valore esatto = 6.73 %.

Per mettere in luce l'influenza del numero delle classi sul valore di R′, il dott. G. Pietra, della Direzione generale della Statistica in Roma, ha calcolato il valore di R′, per la proprietà terriera in mano di privati dello Stato di Victoria nel 1910 (1), considerando dapprima la classificazione minuta in 30 categorie, data dalle Statistiche ufficiali, e restringendo poi via via a 25, 20, 15, 10, 8, 6, 5, 4 il numero delle categorie.

(1) I dati sono ricavati dallo *Statistical Register of the State of Victoria for the year 1911*. Melbourne, Mullet. Part VIII. Production, pag. 107.

Victoria, 1910. Proprietà terriera privata

Numero delle classi	Valore di R'
30	69,0 %
25	68,9 %
20	68,8 %
15	68,4 %
10	67,8 %
8	66,5 %
6	65,8 %
5	64,6 %
4	58,1 %

Come si vede, il valore di R' diminuisce col diminuire del numero delle classi considerate; per 10 classi, la riduzione di R' (del 2 % di fronte al valore di R' per 30 classi) si può ancora ritenere del tutto trascurabile.

In altri casi, naturalmente, la riduzione di R' potrà essere alquanto diversa. Nell'esempio seguente, essa risulta alquanto maggiore.

Belgio, 1911. Stipendi degli impiegati pubblici (1)

Numero delle classi	Valore di R'
32	33,9 %
12	31,5 %

Il valore di R' per gli stipendi degli impiegati pubblici del Belgio diminuisce del 4 % riducendo il numero delle classi da 32 a 12 (2).

Si intende come non sia il numero delle classi di per sè che influisce, ma il fatto che, diminuendo le classi, queste si fanno più e più comprensive. Naturalmente poi i valori di R' possono risultare diversi se le quantità si raggruppano nello stesso numero di classi, ma in modi diversi. E può anche darsi che

---

(1) Cfr. *Ministère des Finances. - Secrétariat général. - Tableau statistique du nombre et des traitements des magistrats, fonctionnaires et employés civils de l'Etat et des Ministres des cultes retribués par l'Etat. 1-1-1911.* Bruxelles, G. Piquart. 1913. Nelle seriazioni che servirono al calcolo di R', non erano compresi gli stipendi dei ministri del culto.

(2) Il calcolo degli R' relativi al Belgio è dovuto al prof A. DE' STEFANI, libero docente di Economia politica nella R. Università di Padova.

con raggruppamenti diversi si ottengano valori di R′ inferiori con un numero di classi superiore, se le classi si fanno più disuguali tra loro. Per la proprietà terriera della Victoria, il dott. Pietra ha calcolato i valori di R′ raggruppando le 30 classi in 22 in tre modi diversi. Diversi risultarono i valori di R′, e in tutti e tre i casi inferiori al valore di R′ ottenuto su 20 classi di estensione meno disuguale.

Valori di R′

su 22 classi ottenute riunendo le classi			su 20 classi ottenute riunendo le classi
dalla 1ª alla 5ª e dalla 26ª alla 30ª	dalla 6ª alla 10ª e dalla 21ª alla 25ª	dalla 11ª alla 15ª e dalla 16ª alla 20ª	8ª-9ª; 11ª-13ª; 14ª-16ª; 17ª-18ª; 19ª-20ª; 23ª-24ª; 25ª-26ª; 29ª-30ª
0,6844	0,6698	0,6875	0,6879

Una seconda approssimazione del valore di R si può ottenere calcolando approssimativamente il valore di

$$\frac{2}{(n-1)\,A_n}\sum_{k=1}^{r}\sum_{l=1}^{f_k}\varepsilon_{kl}\,\delta_{kl}\ ,$$

in base a certe ipotesi.

Supponiamo, per esempio, che gli f_k valori di a_i compresi in una stessa classe costituiscano una progressione aritmetica. Sarà allora, per ogni valore di l e di k, $\delta_{kl}=\varepsilon_{kl}\dfrac{c_k}{f_k}$, dove è $c_k=l_k-l_{k-1}$, e quindi

$$\sum_{k=1}^{r}\sum_{l=1}^{f_k}\varepsilon_{kl}\,\delta_{kl}=\sum_{k=1}^{r}\sum_{l=1}^{f_k}\varepsilon_{kl}^{2}\,\frac{c_k}{f_k}\ .$$

Si dimostra (1) che è

$$\sum_{l=1}^{f_k}\varepsilon_{kl}^{2}=\frac{f_k\,(f_k^2-1)}{12}\ .$$

(1) Cfr. la memoria citata *Variabilità e Mutabilità* etc., pagg. 51-52.

e quindi

$$\frac{2}{(n-1)\,A_n}\sum_{k=1}^{r}\sum_{l=1}^{f_k}\varepsilon_{kl}\,\delta_{kl}=\frac{1}{6\,(n-1)\,A_n}\sum_{k=1}^{r}(f_k^2-1)\,c_k\ .$$

Si giunge così, in base all' ipotesi suesposta, al seguente valore approssimato di R [1]

$$R''=\frac{1}{(n-1)\,A_n}\left(\sum_{k=1}^{r}(i_k+i_{k-1}-1)\,S_k+\frac{1}{6}\,(f_k^2-1)\,c_k\right)-1 \quad (17)$$

Si avverta che, nella determinazione approssimata di $\sum\limits_{k=1}^{r}\sum\limits_{l=1}^{f_k}\varepsilon_{kl}\,\delta_{kl}$, si suppone che tutti questi prodotti sieno positivi, ciò che avviene in realtà soltanto se $\frac{S_k}{f_k}$ coincide con la mediana della classe. Ciò spiega come molte volte R″ dia un valore approssimato per eccesso di R.

Nel caso delle pulsazioni del cuore delle Pelli Rosse e degli Egiziani, si può porre, qualunque sia k, $c_k=5$. Si ha allora, per la seriazione delle Pelli Rosse,

$$\sum_{k=1}^{r}\left(f_k^2-1\right)=14{,}332$$

$$R''=5{,}97\ \%_0\ ;$$

per la seriazione degli Egiziani,

$$\sum_{k=1}^{r}\left(f_k^2-1\right)=1499$$

$$R''=6{,}78\ \%_0\ .$$

I valori esatti sarebbero 5,88 %₀ e 6,73 %₀, inferiori a quelli così ottenuti, ma per differenze del tutto trascurabili.

---

(1) Quando n è abbastanza elevato, R″ può determinarsi, anzi che in base alla (17), in base alla formula seguente

$$R''_2=\frac{1}{n\,A_n}\left(\sum_{k=1}^{r}(i_k+i_{k-1})\,S_k+\frac{1}{6}\,f_k^2\,c_k\right)-1\ . \qquad (17^{bis})$$

Ci si persuade facilmente che è

$$R''_2=\frac{n-1}{n}\,R''+\frac{D}{6\,(n-1)\,A_n}\quad ,$$

dove $D=\sum\limits_{k=1}^{r}c_k$ indica la differenza tra il massimo e il minimo valore di a.

Talvolta le statistiche danno il massimo e il minimo della seriazione: ciò avviene spesso nelle seriazioni di caratteri biologici; altre volte invece, e questo è quasi sempre il caso nelle statistiche economiche relative a redditi, affitti, patrimoni, manca tale indicazione. Non si conosce allora il limite superiore della classe più alta, nè il limite inferiore della classe più bassa, per modo che i valori di c_1 e di c_r non possono determinarsi che in via approssimativa, con qualche espediente.

Per determinare c_1, si può ammettere talvolta che il limite inferiore della classe più bassa sia $=0$. Non ci si allontana certamente molto dal vero, per esempio, supponendo che sia $=0$ l'area minima di una proprietà terriera.

Un espediente, a cui si può ricorrere più spesso, per la determinazione di c_1 o di c_r, è quello di supporre che la intensità media della classe coincida con la semisomma dei suoi limiti. Risolvendo l'uguaglianza

$$\frac{S_k}{f_k} = \frac{1}{2}\,(l_k + l_{k-1})$$

rispetto ad l_k e a l_{k-1}, si trova facilmente il valore ignoto di l_k o di l_{k-1}, essendo noti gli altri termini dell'uguaglianza.

Per esempio, nella classificazione della proprietà terriera della Victoria per il 1910, l'ultima categoria abbraccia 2 proprietà di un'estensione complessiva di 116,486 acri.

Dalla equazione

$$\frac{116{,}486}{2} = \frac{1}{2}\,(50{,}000 + x)$$

si ricava facilmente il valore del limite superiore

$$x = 66{,}486\ .$$

Nella classificazione corrispondente per il 1906, l'ultima categoria abbracciava invece 6 proprietà con 366,766 acri.

Dalla equazione

$$\frac{366{,}766}{6} = \frac{1}{2}\,(50{,}000 + x)$$

si ricava, per limite superiore,

$$x = 72,255$$

Si intende che l' approssimazione ottenuta con questo procedimento varia secondo la distribuzione delle quantità nella prima e nell' ultima categoria.

Per le pulsazioni del cuore delle Pelli Rosse e degli Egizioni di Kharga, le approssimazioni così ottenute sono buone. Operando con questo procedimento sulla prima e sull' ultima categoria della tavola a pag. 12 relativa alle Pelli Rosse, si ottengono per valori estremi 79,6 e 45; i veri valori estremi sono invece 80 e 44. Per gli Egiziani di Kharga, i valori calcolati sono 104,3 e 57 e i valori reali 105 e 54. Anche errori notevoli nella determinazione dei valori di c_1 e c_r non potranno però avere di solito che un effetto molto limitato sul valore di R''.

Per la proprietà terriera della Victoria nel 1910, il dott. Pietra ha anche calcolato il valore di R'' mediante la formula (17^{bis}) variando il numero delle classi. Come si vede, fino a che il numero delle classi non scende a 4, i valori di R'' si mantengono molto uniformi, le differenze essendo sempre dell' ordine del terzo decimale.

Victoria, 1910. Proprietà terriera privata

Numero delle classi	Valore di R''
30	69,1 %
25	69,2 %
22	69,5 - 69,3 - 69,2 % (1)
20	69,2 %
15	69,1 %
10	69,2 %
8	69,1 %
6	69,6 %
5	69,5 %
4	71,1 %

Anche per gli stipendi degli impiegati pubblici del Belgio, il riunire in 12 le 32 classi date dalla statistiche esercita un' influenza soltanto sul terzo decimale di R''. In base alla formula (17^{bis}), il prof. de' Stefani ha ottenuto i seguenti risultati.

(1) Questi valori si sono ottenuti in base a raggruppamenti diversi delle classi. I valori rispettivi di R' sono 67.0, 68.8, 68.4. Cfr. pag. 14.

Belgio, 1911. Stipendi degli impiegati pubblici

Numero delle classi	Valore di R''
32	34,1 %
12	34.3 %

Spesso le statistiche danno, per ogni classe, solo il valore di f_k e non quello di S_k. Si può allora determinare un valore approssimato di R supponendo che, per ogni classe, l' intensità media corrisponda alla semisomma dei limiti estremi (1). Posto

$$S_k = \frac{1}{2} f_k (l_k + l_{k-1})$$

e quindi

$$A_n = \sum_{k=1}^{r} S_k = \frac{1}{2} \sum_{k=1}^{r} f_k (l_k + l_{k-1}) \ ,$$

la (17) diventa (2)

$$R''' = \frac{\sum_{k=1}^{r} (i_k + i_{k-1} - 1) f_k (l_k + l_{k-1}) + {}^1/_3 (f_k^2 - 1) c_k}{(n-1) \sum_{k=1}^{r} f_k (l_k + l_{k-1})} - 1 \qquad (18)$$

In tal caso la conoscenza dei limiti inferiore e superiore delle classi serve, non solo per determinare il termine di correzione $\frac{1}{6} \sum_{k=1}^{r} (f_k^2 - 1) c_k$, ma anche per determinare approssimativamente il valore di S_k. La ignoranza della conoscenza del mas-

(1) Alla stessa ipotesi si può ricorrere per la determinazione di R quando le statistiche danno, per ogni classe, solo il valore di S_k e non quello di f_k. Questo caso è piuttosto raro; esso si presenta nelle Statistiche relative all' estensione delle imprese agricole *(Betriebe)* della Svizzera nel 1905.

(2) Quando n è abbastanza elevato, si può far uso della formula

$$R_2''' = \frac{\sum_{k=1}^{r} (i_k + i_{k-1}) f_k (l_k + l_{k-1}) + {}^1/_3 f_k^2 c_k}{n \sum_{k=1}^{r} f_k (l_k + l_{k-1})} - 1 \qquad (18^{bis})$$

che si deduce dalla (17^{bis}) in modo analogo a quello con cui la (18) si deduce dalla (17).

simo e del minimo della seriazione può costituire naturalmente, in questo caso, un inconveniente molto più grave che quando si tratta di determinare il valore di R″; tanto più grave in quanto che, non essendo dati i valori di S_k, non si può ricorrere, per la determinazione del massimo o del minimo, all' espediente a cui abbiamo accennato. Quando le statistiche non dieno il massimo e il minimo della seriazione, nè i valori di S_k, il limite inferiore della classe più bassa e il limite superiore della classe più alta non si possono determinare senza arbitrio; ma l'errore nella determinazione dei loro valori avrà, sul valore di R‴, un' influenza notevole solo nel caso in cui la prima e l' ultima classe sieno molto numerose.

Quando le classi intermedie presentano tutte la stessa estensione, un' ipotesi plausibile è quella che, anche per le classi estreme, l' estensione non sia diversa. Questo è il caso per la classificazione che abbiamo fatto delle pulsazioni del cuore delle Pelli Rosse e degli Egiziani.

La tavola seguente contiene gli elementi per la determinazione di R‴ per le pulsazioni del cuore delle Pelli Rosse; il massimo (79) e il minimo (45) della seriazione furono determinati in base all' ipotesi suddetta.

Determinazione del rapporto di concentrazione secondo la formula (18)
Pulsazioni del cuore di 263 Pelli Rosse secondo HRDLICKA.

l_{k-1}	l_k	$l_{k-1}+l_k$	f_k	$f_k\,(l_{k-1}+l_k)$	$i_k+i_{k-1}-1$	$f_k\,(l_{k-1}+l_k)\,(i_k+i_{k-1}-1)$
45	49	94	6	564	5	2,820
50	54	104	26	2704	37	100,048
55	59	114	61	6954	124	862,296
60	64	124	74	9176	259	2,376,384
65	69	134	59	7906	392	3,099,152
70	74	144	30	4320	481	2,077,920
75	79	154	7	1078	518	558,404
	Totali	—	263	32,702	—	9,077,024

È
$$\sum_{k=1}^{r} f_k\,(l_{k-1}+l_k) = 32,702$$

$$\sum_{k=1}^{r} f_k (l_{k-1} + l_k)(i_k + i_{k-1} - 1) = 9{,}077{,}024$$

e abbiamo già trovato

$$\sum_{k=1}^{r} (f_k^2 - 1) = 14{,}332 \qquad r_k = 5$$

Se ne ricava

$$R''' = \frac{9{,}077{,}024 + {}^5/_3\, 14{,}332}{262 \times 32702} - 1 = 6{,}22\ \%$$

Per le pulsazioni del cuore degli Egiziani di Kharga, si trova $R''' = 6{,}90\ \%$.

In entrambi i casi, il valore di R''' risulta superiore a quello di R ricavato esattamente dallla formula (13) ($= 5{,}88\ \%$ per le Pelli Rosse e $= 6{,}73\ \%$ per gli Egiziani).

Altre volte converrà ricorrere ad altre ipotesi; e queste dovranno naturalmente variare di caso in caso, in armonia con la distribuzione del carattere esaminato.

Nel caso della distribuzione della proprietà terriera nella Victoria, parve plausibile di fare $= 0$ il limite inferiore della classe più bassa e di calcolare approssimativamente il limite superiore della classe più alta, supponendo che nelle tre ultime classi la curva di frequenza del carattere sia un tratto di iperbole equilatera. Riducendo progressivamente il numero delle classi, e calcolando ogni volta nuovamente il limite superiore della classe più alta, il dott. Pietra ha ottenuto, in base a queste ipotesi, i seguenti valori di R_2'''.

Victoria, 1910. Proprietà terriera privata

Numero delle classi	Valori di R_2'''
30	69,5 %
25	69,7 %
22	70,4 - 71,2 - 67,6 %
20	69,7 %
15	69,5 %
10	70,1 %
8	71,6 %
6	70,2 %
5	68,1 %
4	52,5 %

Salvo per il caso di 4 classi, i valori di R_2''' non discordano essenzialmente da quelli di R″ dati a pag. 17. I valori di R_2''' furono anche determinati calcolando il limite superiore dell'ultima classe in base ad ipotesi differenti dalla precedente, senza giungere a risultati sensibilmente diversi.

Queste applicazioni delle formule (18) e (18^{bis}) hanno un'importanza in quanto mostrano come, di regola, sia possibile determinare, con approssimazione praticamente sufficiente, il rapporto di concentrazione di un carattere anche nei casi, in cui le statistiche danno soltanto una classificazione delle sue intensità e, per ogni classe, indicano solo il numero dei casi che in questa rientrano.

4. — Riportiamo qui sotto, in graduatoria crescente, alcuni rapporti di concentrazione calcolati dal prof. de' Stefani, dal dott. Pietra, dal prof. Savorgnan e da noi per caratteri fisici della specie umana e per caratteri economici.

Per ogni carattere, sono considerati solo gli individui che ne sono forniti (per es., per il numero dei figli, sono considerate solo le famiglie con figli; per i patrimoni e per le successioni, solo i censiti o i deceduti a cui fu accertato un patrimonio). Vedremo più innanzi quali risultati si ottengano considerando anche gli individui in cui il carattere presenta un'intensità nulla.

Il confronto dei varii rapporti di concentrazione ottenuti può suggerire considerazioni interessanti, sia dal punto di vista biologico, sia dal punto di vista economico. Basti qui far notare quanto diversa sia la concentrazione per i caratteri economici e per i fisici: per la gran parte dei caratteri fisici, il valore di R resta inferiore al 10 %; per la gran parte dei caratteri economici, esso supera invece il 50 %. È particolarmente interessante il confronto tra la concentrazione del numero dei figli lasciato allo scioglimento del matrimonio (34 %) e la concentrazione dei patrimoni tra gli abbienti (71-93 %), poichè la concentrazione del numero dei figli sopravviventi allo scioglimento del matrimonio, rappresenta un limite minimo della concentrazione dei patrimoni, che si avvererebbe in una generazione, per il solo fatto che le famiglie sono diversamente prolifiche, qualora ogni famiglia della generazione precedente, al suo sciogliersi, lasciasse alla prole lo stesso patrimonio (1).

Concentrazione di alcuni caratteri fisici dell' uomo (2)

Carattere		Valore di R
Temperatura sublinguale	Egiziani di Kharga (3)	0.4 %
Modulo cefalico	" " "	1.2 %
Diametro antero-posteriore della testa	" " "	1.5 %
Distanza dal meato auditivo al bregma	" " "	1.6 %
Statura	" " "	1.7 %
Diametro trasversale massimo della testa	" " "	1.8 %
Indice cefalico	" " "	2.0 %
Lunghezza del piede sinistro	" " "	2.3 %
Statura.	Coscritti italiani nati nel 1890 (4)	2.5 %*
Indice del piede sinistro	Egiziani di Kharga (3)	2.9 %
Lunghezza dell' orecchio sinistro	" " "	3.0 %
Larghezza del piede sinistro	" " "	3.2 %
Altezza dell' orecchio sinistro	" " "	3.5 %
Pulsazioni del cuore	Pelli Rosse (5)	5.9 %*
" " "	Egiziani di Kharga (3)	6.7 %*
Respirazione	" " "	7.7 %
Forza di pressione della mano destra	" " "	9.7 %
Forza di trazione	" " "	14.8 %

(1) Cfr. in proposito, la citata memoria sugli *Indici di concentrazione*, etc. in *Biblioteca dell' Economista*, V Serie. Vol. XX, pag. 78-79. Ciò che ivi si dice a proposito dell' indice di concentrazione δ, si può ripetere per il rapporto di concentrazione. Anche il rapporto di concentrazione del numero dei figli sopravviventi allo scioglimento del matrimonio si può infatti ritenere equivalente, in un numero abbastanza grande di osservazioni, al rapporto di concentrazione dei patrimoni che si verificherebbe tra i figli qualora tutte le coppie matrimoniali, sciogliendosi, lasciassero ai figli lo stesso patrimonio. Ci si persuade agevolmente di ciò rappresentando graficamente la concentrazione dei due caratteri col metodo indicato al § 6.

(2) I valori di R contrassegnati da un asterisco furono calcolati dall' A.; gli altri dal prof. de' Stefani e dai suoi allievi dell' Istituto tecnico di Vicenza.

(3) Hrdlicka. *The natives*, etc. Op. cit..

(4) Ministero della Guerra. *Della leva di terra sui giovani nati nell' anno 1890*. Roma, Voghera, 1913.

(5) Hrdlicka. *Physiological and medical observations*, etc. Op. cit..

Carattere		Valore di R
Numero dei figli lasciati allo scioglimento del matrimonio.	Budapest 1903 - 1908 (1)	33.5 % *
Numero dei figli viventi al momento del censimento.	Francia 1901 (2)	34.0 % *
Numero dei figli nati fino al momento del censimento.	Francia 1906 (2)	37.4 % *
Numero dei figli nati fino allo scioglimento del matrimonio.	Budapest 1903 - 908 (1)	37.5 % *

Concentrazione di alcuni caratteri economici (3)

Carattere		Valore di R
Estensione delle imprese agricole.	Germania	1907 (4) 23 % **
" " " "	Belgio	1895 (5) 25 % **
" " " "	Francia	1892 (6) 28 % **
Stipendi degli impiegati pubblici.	Belgio	1911 (7) 34 %
Estensione delle imprese agricole gestite da proprietari.	Svizzera	1905 (8) 43 % **
	Danimarca	1905 (9) 43 % **
Redditi da lavoro.	Argovia	1892 (10) 44 %

(1) *Statistisches Jahrbuch der Haupt-und Residenzstadt Budapest.* 1903 - 1908.

(2) STATISTIQUE GÉNÉRALE DE LA FRANCE. *Statistique des familles en 1906*. Paris. Imprimerie nationale, 1906.

(3) I valori di R contrassegnati con un asterisco furono calcolati dall' A.; quelli con due asterischi dal dott. G. Pietra; gli altri dal prof. de' Stefani e dai suoi allievi dell' Istituto tecnico di Vicenza. Per il rapporto di concentrazione del risparmio postale in Austria, cfr. nota (1) a pag. 25.

(4) *Statistisches Jahrbuch für das Deutsche Reich.* 1913.

(5) *Annuaire statistique de la Belgique.* 1912.

(6) *Statistique agricole de la France. Résultats généraux de l' enquête décennale de 1892.*

(7) *Tableau statistique*, etc. Op. cit.

(8) *Résultats du recensement fédéral des entreprises agricoles, industr. et com. du 9 Août 1905.*

(9) *Danmarks jordbrug, 1850-1905, udgivet af statens statistiskt bureau.* Copenaghen, 1907.

(10) J. KISTLER. *Erhebungen über Vermögen, Schulden und Erwerb im Kanton Aargau in den Jahren 1892, 1886 und 1872.* Bern. Stämpfli u. Cie. 1895.

Carattere			Valore di R
Redditi	Danimarca	1909 (1)	46 %
Estensione delle imprese agricole.	Serbia	1905 (2)	50 %
Redditi.	Sassonia	1910 (3)	50 %
„	Norvegia	1906 (4)	52 %
Valori locativi dei locali di abitazione.	Parigi	1911 (5)	63 % *
Estensione delle imprese agricole	Olanda	1887 (6)	63 % **
Valore della proprietà terriera.	Victoria	1911 (7)	64 %
Estensione delle imprese agricole.	Irlanda	1896 (6)	66 % **
„ „ „ „	Grambrettagna	1895 (6)	68 % **
Redditi da “personal exertion„.	Victoria	1911 (7)	69 %
„ da proprietà.	Victoria	1911 (7)	69 %
Estensione della proprietà terriera.	Victoria	1910 (8)	69 % **
„ „ „ „	Australia occ.	1911 (9)	70 % **
Patrimoni.	Argovia	1892 (10)	71 %
Estensione della proprietà terriera.	Australia merid.	1911 (9)	72 % **
	Danimarca	1901 (11)	76 % **

(1) STATISTIQUE DU DANEMARK. *Revenus et fortunes d' après la taxation pour 1909-10.* Kiobenhavn. Bianco Lunos. 1912.

(2) In base a dati manoscritti comunicati all' Istituto Int. di Agricoltura in Roma.

(3) L. DUGÉ DE BERNONVILLE. *Distribution de salaires et de revenus en divers pays. Bull. de la Stat. générale de la France.* Juillet 1913.

(4) A. N. KIAER. *Répartition des revenus en Norvège.* Kristiania. Bjoernstads. 1907-1910.

(5) *Annuaire statistique de la Ville de Paris.* 1911.

(6) *Handwörterbuch der Staatswissenschaften.* Zweite Auflage, Zweiter Band. pag. 437 e segg..

(7) *Statistical Register of the State of Victoria for the year 1912. Part II. Finance.*

(8) *Statistical Register of the State of Victoria for the year 1911. Part VIII, Production.*

(9) *Official Year Book of the Commonwealth of Australia.*

(10) J. KISTLER. *Erhebungen über Vermögen, Schulden und Erwerb im Kanton Aargau in den Jahren 1892, 1886 und 1872.* Bern. Stämpfli u. Cie. 1895.

(11) *Annuaire statistique du Danemark.*

Carattere			Valore di R
Risparmio postale.	Austria	1900	(1) 78 %
Successioni.	Victoria	1908-1910	(2) 81 %
Patrimoni.	Cantone di Zurigo	1909	(3) 82 %
Estensione della proprietà terriera.			
	Tasmania	1911	(4) 83 % **
	N. Galles del Sud.	1911	(4) 85 % **
Successioni.	Francia	1904	(5) 88 %
Patrimoni.	Danimarca	1909	(6) 93 %

5. — Vi sono caratteri che, in un numero maggiore o minore di casi, presentano un' intensità positiva e, negli altri casi, un' intensità nulla; per esempio, il patrimonio individuale. Chiameremo i casi del primo gruppo *casi positivi,* quelli del secondo gruppo *casi nulli*, il loro insieme *casi totali.*

Ci si può proporre di misurare la concentrazione di un carattere tra i casi totali o solo tra i casi positivi. Sieno n i casi positivi e v i casi nulli.

Determinato il rapporto di concentrazione del carattere tra i casi positivi — che indicheremo con R_p — è facile passare al rapporto di concentrazione tra i casi totali — che indicheremo con R_t — mediante l' espressione

$$R_t = R_p\, m + (1 - m) \qquad (19)$$

dove m indica il rapporto $\frac{n-1}{n+v-1}$.

(1) Rapporto di concentrazione tolto dalla memoria del prof. F. SAVORGNAN. *Il risparmio postale in Austria dal 1883 al 1912. Ricerche statistiche.* Trieste, 1914. Balestra.

(2) A. M. LAUGHTON. *Victorian Year Book 1910-11.* Part IV. *Accumulation.*

(3) *Uebersicht der Vermögens- und Einkommenssteuerpflichtigen des Kantons Zürich*, etc. Zürich. G. Meyer. 1910.

(4) *Official Year Book of the Commonwealth of Australia.*

(5) *Bulletin de Statistique et de Législation comparée.* 1905.

(6) STATISTIQUE DU DANEMARK. *Revenus et fortunes d'après la taxation pour 1909-10.* Kiobenhavn. Bianco Lunos. 1912.

È sempre $R_t > R_p$ e la differenza cresce col diminuire di m e di R_p [1].

Qui sotto sono messi a confronto, per alcuni caratteri, i valori di R_p e di R_t.

	m	R_p	R_t
Numero dei figli al momento del censimento.			
Francia:			
Figli sopravviventi (1901)	0.841	34.0 %	44.5 %
Figli nati (1906)	0.885	37.4 %	44.6 %
Numero dei figli allo scioglimento del matrimonio. Budapest (1903 - 1908):			
Figli sopravviventi	0.686	33.5 %	54.4 %
Figli nati (vivi e morti)	0.737	37.5 %	53.9 %
Successioni. Victoria (1908 - 1910) . . .	0.371 [2]	80.5 %	92.8 %
Successioni. Francia (1903 - 1904) . . .	0.487 [2]	87.8 %	91.6 %
Patrimoni. Argovia (1892)	0.860 [3]	71.2 %	75.2 %
Patrimoni. Danimarca (1909)	0.836 [3]	92.7 %	93.9 %

Questi risultati sono atti a fare intendere come possa essere fallace trarre conclusioni sulla concentrazione della ricchezza in paesi diversi fondandosi unicamente sulla concentrazione che essa mostra tra gli abbienti. La concentrazione delle successioni,

(1) Il valore di R_p si ricava dalla (13), che si può mettere sotto la forma

$$R_p = \frac{\sum_{l=1}^{s} (i_{l-1} + i_l - 1) f_l x_l - (n-1) A_n}{(n-1) A_n},$$

il valore corrispondente di R_t sarà

$$R_t = \frac{\sum_{l=1}^{s} (i_{l-1} + v + i_l + v - 1) f_l x_l - (n + v - 1) A_n}{(n + v - 1) A_n}$$

che si può anche mettere sotto la forma

$$R_t = R_p \frac{n-1}{n+v-1} + \frac{2v \sum_{l=1}^{s} f_l x_l - v A_n}{(n+v-1) A_n} .$$

Questa si riduce facilmente alla (19), tenendo presente che è

$$A_n = \sum_{l=1}^{s} f_l x_l , \quad m = \frac{n-1}{n+v-1} .$$

(2) Rapporto delle successioni ai morti adulti.

(3) Rapporto delle persone, che figurano nelle statistiche come possessori di patrimoni, alle persone che vi figurerebbero nel caso di massima diffusione della ricchezza.

tra coloro di cui viene tassata l' eredità, risulta nella Victoria ($R = 80.5\ {}^0/_0$) molto più bassa della concentrazione dei patrimoni tra coloro di cui viene accertato un patrimonio in Danimarca ($R = 92.7\ {}^0/_0$); ma, se si tiene conto anche delle persone che figurano nelle statistiche come nullatenenti, il rapporto di concentrazione risulta di ben poco diverso nei due paesi (Victoria: $= 92.8\ {}^0/_0$; Danimarca $= 93.9\ {}^0/_0$).

6. — Al rapporto, che noi proponiamo in questa nota, come misura appropriata della concentrazione, si giunge anche perfezionando un metodo grafico che alcuni autori, il Lorenz, il

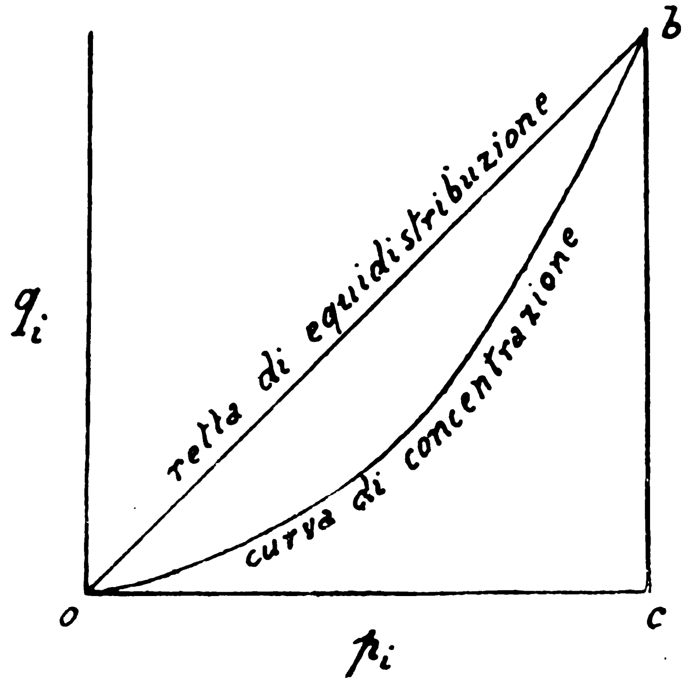

Fig. 1

Chatelain, il Séailles [1], hanno già proposto per giudicare della maggiore o minore disuguaglianza di distribuzione della ricchezza.

(1) Vedi M. O. LORENZ, *Methods of measuring the concentration of wealth*, in *Publications of American Statistical Association*. N. 70, June 1905; e, a proposito della memoria del Lorenz. G. P. WATKINS, *Comment*

Se, in un diagramma a coordinate cartesiane ortogonali, si segnano sull' asse delle ascisse i valori di p_i e sull' asse delle ordinate i valori di q_i, e si fa passare una linea continua per le estremità delle ordinate che misurano i valori di q_i, si ottiene (Fig. 1) una curva (*curva di concentrazione*) ascendente da sinistra a destra, convessa verso l' asse delle ascisse (1).

La curva di concentrazione è tanto meno accentuata quanto meno disuguale è la distribuzione della ricchezza, fino a diventare, nel caso di perfetta uguaglianza di distribuzione, una retta (*retta di equidistribuzione*).

Gli autori sopra nominati trassero partito di questa proprietà della curva di concentrazione per eseguire confronti sulla distribuzione della ricchezza.

Disegnando sullo stesso diagramma più curve relative a tempi o luoghi diversi, essi erano in grado di giudicare in quale tempo o in quale luogo la ricchezza risultava più concentrata.

Questo metodo grafico presentava due inconvenienti, debitamente riconosciuti dal Lorenz e dal King:

a) di non dare una misura precisa della concentrazione;

b) di non permettere neppure, in certi casi, di giudicare dove o quando la concentrazione è più forte. Se, infatti, le curve

on the method of measuring concentration of wealth. Ibidem. N. 72, December 1905; G. P. WATKINS. *An interpretation of certain statistical evidence of concentration of wealth.* Ibidem. N. 81, March 1908; W. M. PERSONS. *The variability in the distribution of wealth and income*, in *The Quarterly Journal of Economics.* Vol. XXIII, N. 3. May 1909; G. P. WATKINS, *The measurement of concentration of wealth.* Ibidem. Vol. XXIV, N. 1. November 1909; W. M. PERSONS. Ibidem; W. J. KING. *The elements of Statistical Methods.* New Jork. The Macmillan Company, 1912.

Vedi pure E. CHATELAIN, *Les successions déclarées en 1905,* in *Revue politique et parlementaire,* 1907; *Le tracé de la courbe des successions en France*, in *Journal de la Soc. de Stat. de Paris*, 1910. Pag. 352 e segg.; *La fortune française d' après les successions en 1909*, in *La Démocratie.* 20 Janvier 1911; J. SÉAILLES, *La répartition des Fortunes en France*, Paris. Alcan, 1910.

(1) Che la curva sia ascendente da sinistra a destra si deduce dalla relazione (1) e che sia convessa verso l' asse delle ascisse si deduce dalla relazione (2).

di concentrazione si accavallano (Fig. 2), non è sempre possibile dire se sia più accentuata l' una o l' altra. Questo incon-

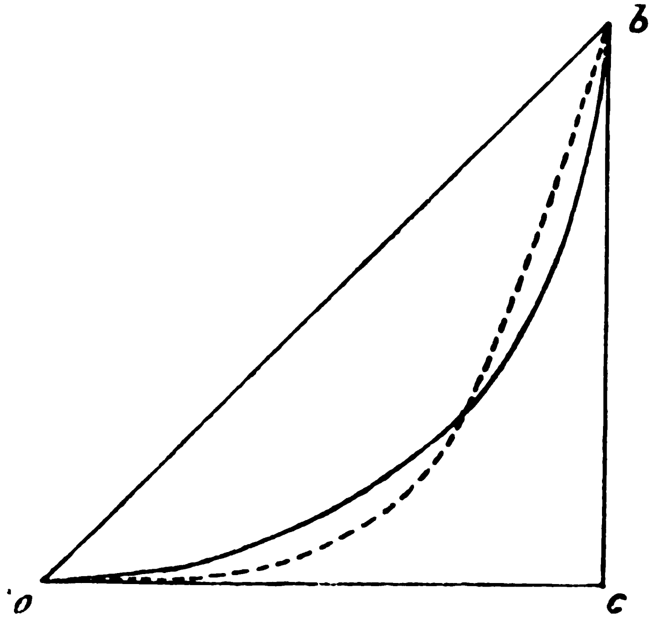

Fig. 2

veniente, che si può ritenere trascurabile nei paragoni tra la concentrazione dei fenomeni dello stesso genere (per es. dei redditi di due anni o di due paesi diversi), assume particolare gravità nei confronti tra la concentrazione di fenomeni differenti, per cui è diversa la forma della distribuzione.

L' uno e l' altro inconveniente scompaiono se si conviene di assumere a misura della concentrazione il rapporto dell' area compresa tra la curva di concentrazione e la retta di equidistribuzione *(area di concentrazione)* all' area del triangolo $o\ b\ c$, che rappresenta l' area di concentrazione nel caso di concentrazione massima.

Ora è facile mostrare che tale rapporto altro non è che il limite, a cui tende il rapporto di concentrazione R, quando cresce il numero n dei casi osservati, mantenendosi però uguale la loro distribuzione.

Si consideri per ciò il diagramma seguente (Fig. 3), in cui è $n = 14$. Gli $n - 1$ piccoli rettangoli limitati dall'asse delle ascisse

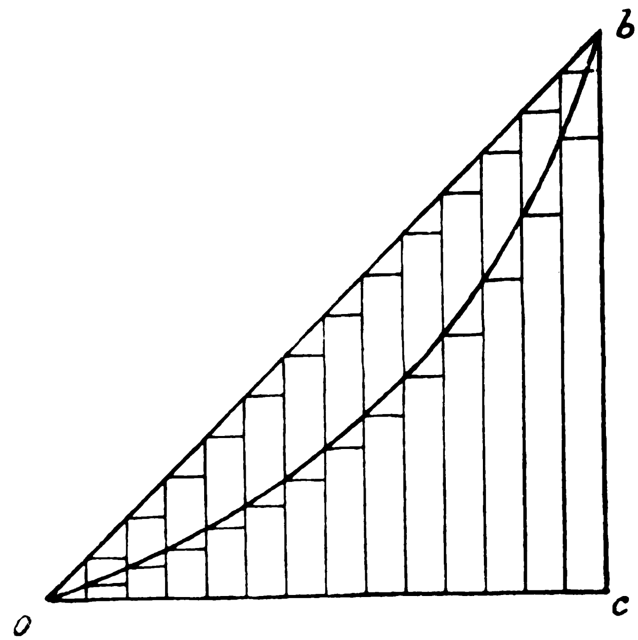

Fig. 3

e dalla curva di concentrazione hanno l'altezza $= q_i$ e la base $= \frac{1}{n}$; la somma delle loro aree sarà $= \frac{1}{n} \sum_{i=1}^{n-1} q_i$. Gli $n - 1$ piccoli rettangoli limitati dall' asse delle ascisse e dalla retta di equidistribuzione hanno la base $= \frac{1}{n}$ e l'altezza $= p_i$, poichè ognuna delle ordinate forma con la retta di equidistribuzione e con l' asse delle ascisse un triangolo isoscele: la somma delle loro aree sarà $= \frac{1}{n} \sum_{i=1}^{n-1} p_i$. La differenza tra le due somme $= \frac{1}{n} \sum_{i=1}^{n-1} (p_i - q_i)$ sarà uguale alla somma delle aree degli $n - 1$ rettangoli limitati dalla retta di equidistribuzione e dalla curva di concentrazione. Ora, col crescere di n, diminuisce l' area delle ipccole superfici comprese tra la retta di equidistribuzione e la sommità dei rettangoli di altezza p_i e diminuisce analogamente l'area delle piccole superfici comprese tra la curva di concentrazione e la sommità dei rettangoli di altezza q_i. Col crescere di n, l'area del triangolo $o\,b\,c$ tende dunque a divenire $= \frac{1}{n} \sum_{i=1}^{n-1} p_i$, l'area di concentrazione a divenire $= \frac{1}{n} \sum_{i=1}^{n-1} (p_i - q_i)$ e il rapporto di questa a quella a divenire eguale al rapporto di concentrazione $R = \frac{\sum_{i=1}^{n-1} (p_i - q_i)}{\sum_{i=1}^{n} p_i}$.

Si intende facilmente che, per descrivere la curva di concentrazione, non è necessario di conoscere tutti i valori di p_i e q_i. Basta la conoscenza di 4 o 5 di questi valori per che la curva possa essere descritta con sufficiente approssimazione.

7. — Queste osservazioni suggeriscono un altro procedimento per determinare praticamente il valore di R.

Si tracci la curva di concentrazione, e si prenda come unità di misura il segmento oc. L'area del triangolo $o\,b\,c$ sarà $= 1/2$; l'area di concentrazione si potrà facilmente misurare mediante un integrafo o un planimetro o anche, più grossolanamente, descrivendo il diagramma con grande accuratezza su carta millimetrata a scala molto ampia e contando il numero dei millimetri quadrati e delle loro frazioni che rientrano nell'area di concentrazione.

Questo metodo, evidentemente molto semplice e rapido, si può teoricamente seguire, non solo tutte le volte che si conosca l'intensità del carattere nei singoli casi, ma anche quando si conosca soltanto il numero dei casi e il loro ammontare totale per categorie, anche se queste sieno poco numerose. Nella pratica però esso è pericoloso quando la curva di concentrazione è poco accentuata e quindi l'area di concentrazione è ristretta, perchè in tal caso piccole inesattezze nella costruzione del diagramma possono avere un'influenza non trascurabile sui risultati. E, anche quando questo inconveniente non si presenti, resta sempre quello, inevitabile, che la misura dell'area, sia mediante l' integrafo o il planimetro, sia mediante l' uso di carta millimetrata, è influenzata dalla tendenza personale del calcolatore e da cause accidentali, che faranno sì che due persone o una stessa persona in due misure otterranno sempre risultati più o meno diversi. Per ciò io penso che, quando sia data l' intensità del carattere per tutti i casi o quando le classi, in cui i casi vengono raggruppati, sieno abbastanza numerose per che si possa ritenere che la formula (17) conduca a buone approssimazioni del valore di R, non ci si debba esimere dal calcolo aritmetico del rapporto di concentrazione, e del calcolo grafico ci si possa servire solo per una prima e rapida ispezione. Il calcolo grafico conserverà naturalmente di fronte al calcolo aritmetico importanza tanto maggiore quanto minore è il numero delle classi, per assumere importanza preponderante quando questo sia così ristretto da non permettere di ritenere che la formula (17) conduca a buone approssimazioni del valore di R.

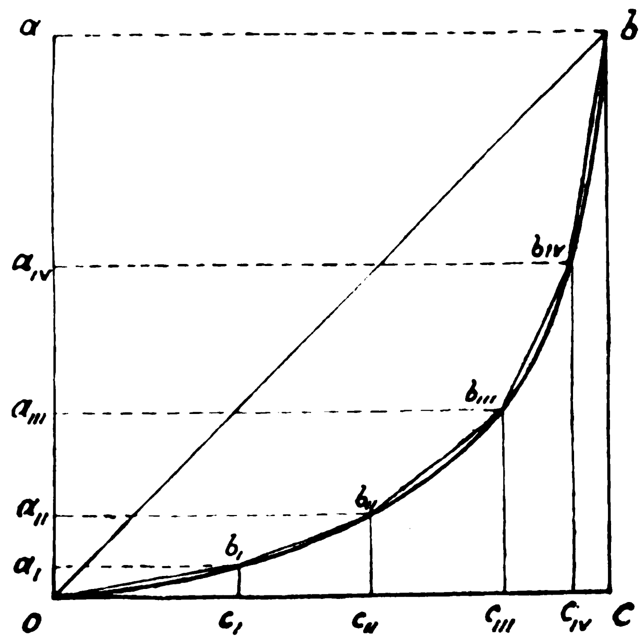

Fig. 4

8. — La rappresentazione grafica della curva di concentrazione è utile per illustrare la relazione che intercede tra il valore esatto R del rapporto di concentrazione secondo la formula (13) e il suo valore approssimato per difetto R′, secondo la formula (15). Si tenga perciò sott'occhio il diagramma a pag. 32 (Fig. 4).

Sieno 5 le classi in cui sono classificate le n intensità del carattere e sieno $\overline{a_\mathrm{I}\, b_\mathrm{I}}$, $\overline{a_\mathrm{II}\, b_\mathrm{II}}$, $\overline{a_\mathrm{III}\, b_\mathrm{III}}$, $\overline{a_\mathrm{IV}\, b_\mathrm{IV}}$, $\overline{ab}$ le ascisse proporzionali ai 5 valori di i_k e $\overline{b_\mathrm{I}\, c_\mathrm{I}}$, $\overline{b_\mathrm{II}\, c_\mathrm{II}}$, $\overline{b_\mathrm{III}\, c_\mathrm{III}}$, $\overline{b_\mathrm{IV}\, c_\mathrm{IV}}$, $\overline{bc}$ le ordinate proporzionali ai corrispondenti valori di A_k. Si dimostra facilmente che il rapporto all'area del triangolo obc dell'area compresa tra la retta ob e la spezzata $o\, b_\mathrm{I}\, b_\mathrm{II}\, b_\mathrm{III}\, b_\mathrm{IV}\, b$ costituisce il limite a cui tende R′ col crescere di n (1); mentre, come si è mostrato al § 6, il rapporto all'area del triangolo obc dell'area compresa tra la retta $o\,b$ e la curva è il limite a cui, col crescere

(1) Indichiamo con a_k uno degli r valori di a e con b_k il corrispondente valore di b.

Le superficie del trapezio $a_k\, a_{k-1}\, b_{k-1}\, b_k$ sarà

$$= \overline{a_k\, a_{k-1}} \times \frac{1}{2} (\overline{a_k\, b_k} + \overline{a_{k-1}\, b_{k-1}}) \ .$$

Si noti ora che è

$$\overline{a_k\, a_{k-1}} = \frac{S_k}{A_n} \ ; \ \overline{a_k\, b_k} = \frac{i_k}{n} \ ; \ \overline{a_{k-1}\, b_{k-1}} = \frac{i_{k-1}}{n} \ .$$

Detta superficie sarà quindi $= \dfrac{(i_k + i_{k-1})\, S_k}{2\, n\, A_n}$; la superficie racchiusa tra i segmenti $\overline{o\, a}$ e $\overline{a\, b}$ e la spezzata sarà $= \dfrac{\sum_{k=1}^{r} (i_k + i_{k-1})\, S_k}{2\, n\ A_n}$; la superficie racchiusa tra la retta di equidistribuzione e la spezzata sarà $= \dfrac{\sum_{k=1}^{r} (i_k + i_{k-1})\, S_k}{2\, n\ A_n} - \dfrac{1}{2}$, e il rapporto di quest'ultima superficie a quella del triangolo obc sarà

$$= \frac{\sum_{k=1}^{r} (i_k + i_{k-1})\, S_k - n\, A_n}{n\, A_n} \ . \tag{20}$$

Dalla (15) si deduce facilmente, tenendo presente che è $\sum_{k=1}^{r} S_k = A_n$, il seguente valore di R′

$$R' = \frac{\sum_{k=1}^{r} (i_k + i_{k-1})\, S_k - n\, A_n}{(n-1)\, A_n} \ ,$$

che coincide a meno di $\dfrac{1}{n}$ col valore dato dalla (20) e tende a questo col crescere di n.

di n, tende R. Poichè la curva di concentrazione è convessa verso l'asse delle ascisse; la spezzata $o\ b_{I}\ b_{II}\ b_{III}\ b_{IV}\ b$ sarà inscritta nella curva di concentrazione e l'area tra detta spezzata e la retta $o\ b$ sarà minore dell'area tra la curva di concentrazione e la retta $o\ b$. Si comprende anche facilmente dal diagramma che tale differenza è tanto maggiore quanto più la curva di concentrazione è accentuata, quanto minore è il numero delle classi e, generalmente, a pari numero di classi, quanto più è diverso il numero dei casi che nelle varie classi rientrano.

9. — Esaminiamo la relazione, in cui il rapporto di concentrazione sta con gli indici di variabilità, che si usano per caratterizzare la distribuzione dei caratteri.

Gli indici di variabilità si possono dividere in due categorie:

a) indici, in cui la variabilità del carattere viene misurata mediante un valore medio degli scostamenti (presi in valore assoluto) da una media delle intensità del carattere. Secondo che, tra i valori degli scostamenti, si sceglie l'uno o l'altro valore medio, si hanno indici di variabilità diversi. Usualmente si sceglie, o la media aritmetica, e si ha lo *scostamento semplice medio*, o la media quadratica (1), e si ha lo *scostamento quadratico medio*, o la mediana, e si ha lo *scostamento probabile*. Ognuno di questi indici di variabilità può a sua volta distinguersi secondo che gli scostamenti sono misurati dall'una o dall'altra media delle intensità del carattere: di solito gli scostamenti si misurano dalla media aritmetica, talvolta dalla mediana.

b) indici, in cui la variabilità del carattere viene desunta da un valore medio delle differenze (prese in valore assoluto) che si possono instituire tra le intensità del carattere. Se, fra tutti i possibili valori medii, si sceglie la media aritmetica, si ottiene l'indice di variabilità detto *differenza media*.

Gli indici della prima categoria rispondono al problema: Quale è la differenza tra le intensità osservate del carattere e una certa loro media? Essi hanno particolare valore nelle ricerche, in cui si ha ragione di ritenere che codesta media tra le intensità osservate corrisponda ad una grandezza effettivamente

(1) Media quadratica fra più quantità chiamo la radice quadrata della media aritmetica fra i quadrati della quantità.

esistente, dalla quale le intensità osservate si discostano per errori di osservazione (come nelle rilevazioni della fisica o dell' astronomia o della geodesia), o a una grandezza che può venire riguardata come tipica, da cui le intensità osservate si possono ritenere allontanarsi per oscillazioni accidentali (come spesso avviene nelle rilevazioni relative ai caratteri di una specie).

Gli indici della seconda categoria rispondono al problema: Quale è la differenza tra le varie intensità del carattere? Essi hanno particolare valore nelle ricerche, in cui non sia soddisfatta una delle due circostanze sopra accennate, per esempio nelle rilevazioni relative ai redditi, ai patrimoni, agli affitti, al numero dei nati, dei matrimoni, dei morti; in generale nelle ricerche di statistica economica e demografica.

Nella nostra memoria *Variabilità e Mutabilità*, abbiamo messo in luce le relazioni che intercedono tra gli indici della prima categoria e quelli della seconda, e abbiamo esaminato partitamente in quali casi rispondano meglio agli scopi della ricerca gli uni e in quali casi gli altri.

Dimostreremo ora che il rapporto di concentrazione coincide col rapporto della differenza media al valore massimo che questa può assumere, o in altre parole, col rapporto della differenza media al doppio della media aritmetica del carattere.

Sappiamo che la differenza media è data dalla formula

$$\Delta = \frac{2}{n(n-1)} \sum_{i=1}^{\frac{n+1}{2}} (n+1-2i)(a_{n-i+1} - a_i)$$

e che il suo massimo valore possibile è $= 2\,M_n = \frac{2\,A_n}{n}$ (1)

Basterà quindi dimostrare che sussiste l' uguaglianza

$$R = \frac{\sum_{i=1}^{\frac{n-1}{2}} (n+1-2i)(a_{n+1-i} - a_i)}{(n-1)\,A_n} \tag{21}$$

o anche, poichè si sa dalla (12) che è

---

(1) Cfr. *Variabilità e Mutabilità*, pag. 22 e pag. 80.

$$R = \frac{(n-1)\,A_n - 2\sum_{i=1}^{n-1} A_i}{(n-1)\,A_n},$$

che sussiste l' uguaglianza

$$\sum_{i=1}^{\frac{n+1}{2}} (n+1-2i)\,(a_{n+1-i} - a_i) = (n-1)\,A_n - 2\sum_{i=1}^{n-1} A_i \tag{22}$$

Cominciamo per ciò col notare che è

$$\begin{gathered}(n-1)\,(a_1 + a_2 + \ldots + a_n) = \\ = (n-1)\,a_1 + (n-2)\,a_2 + (n-3)\,a_3 + \ldots + 2\,a_{n-2} + a_{n-1} + \\ + a_2 + 2\,a_3 + \ldots + (n-3)\,a_{n-2} + (n-2)\,a_{n-1} + (n-1)\,a_n\end{gathered}$$

Sottraendo da ambo i membri la somma

$$2\left\{(n-1)\,a_1 + (n-2)\,a_2 + (n-3)\,a_3 + \ldots + 2\,a_{n-2} + a_{n-1}\right\},$$

si ottiene l' uguaglianza

$$\begin{aligned}(n-1)(a_1+a_2+\ldots+a_n) - 2\left\{(n-1)\,a_1 + (n-2)\,a_2 + \ldots + a_{n-1}\right\} = \\ = a_2 + 2\,a_3 + \ldots + (n-2)\,a_{n-1} + (n-1)\,a_n - \\ - (n-1)\,a_1 - (n-2)\,a_2 - (n-3)\,a_3 - \ldots - a_{n-1}\end{aligned}$$

e da questa, sostituendo in base alle uguaglianze seguenti,

$$\begin{gathered}(n-1)\,A_n = (n-1)\,(a_1 + a_2 + \ldots + a_n) \\ \sum_{i=1}^{n-1} A_i = (n-1)\,a_1 + (n-2)\,a_2 + \ldots + 2\,a_{n-2} + a_{n-1}\end{gathered}$$

$$\begin{aligned}\sum_{i=1}^{\frac{n+1}{2}} (n+1-2i)(a_{n+1-i} - a_i) = (n-1)\,a_n + (n-2)\,a_{n-1} + \ldots + 2a_3 + a_2 - \\ - (n-1)\,a_1 - (n-2)\,a_2 - \ldots - 2\,a_{n-2} - a_{n-1}\end{aligned}$$

si ricava immediatamente la (22).

10. — La relazione trovata fra R e Δ permette di dare un' altra dimostrazione delle relazioni che intercedono tra R e i suoi valori approssimati R′ ed R″.

Se le n quantità, che hanno per differenza media Δ vengono raggruppate in r classi, si dimostra facilmente che la differenza media delle n quantità che si ottengono dalle prime, sostituendo ad ogni quantità la media della classe a cui appartiene, è

$$\Delta' = \Delta - \frac{\sum_{i=1}^{r} f_k (f_k - 1) \Delta_k}{n(n-1)} \tag{23}$$

dove Δ_k $(k = 1, 2 \dots r)$ indica la differenza media tra le f_k quantità della classe k^{ma} (1).

Ma è, per quanto si è dimostrato al paragrafo precedente, $R = \frac{\Delta}{2 M_n}$ e analogamente $R' = \frac{\Delta'}{2 M_n}$.

Sarà quindi

$$R' = R - \frac{\sum_{i=1}^{r} f_k (f_k - 1) \Delta_k}{2n(n-1) M_n} \tag{24}$$

Questa uguaglianza mostra come il valore approssimato R' si avvicini tanto più al valore vero R quanto minore è, relativamente alla media generale M_n, la variabilità Δ_k nelle singole classi e quanto minore è, relativamente al numero totale $n(n-1)$ delle differenze, la somma $\sum_{k=1}^{r} f_k (f_k - 1)$. Ci si persuade facilmente che questa somma è tanto maggiore quanto più è piccolo il numero r delle classi in cui si sono distinte le n osservazioni e

(1) Si osservi per ciò che le $n(n-1)$ differenze, a cui danno luogo le n quantità, che hanno per differenza media Δ, si possono dividere in due categorie: *a)* differenze che passano tra quantità che appartengono alla stessa classe; queste differenze sono in numero di $\sum_{k=1}^{r} f_k (f_k - 1)$; *b)* differenze che passano tra quantità che appartengono a classi diverse.

Se ad ogni quantità si sostituisce la media delle quantità della classe, la somma delle differenze della seconda categoria resta inalterata, mentre le differenze della prima categoria divengono tutte $= 0$. Per ottenere il valore di $n(n-1)\Delta'$ bisognerà dunque sottrarre dal valore di $n(n-1)\Delta$ la somma delle differenze tra quantità appartenenti alla medesima classe, che è $\sum_{k=1}^{r} f_k (f_k - 1) \Delta_k$.

quanto più sono disuguali i numeri f_k dei casi che rientrano nelle singole classi.

Del secondo termine del secondo membro della (24) si può dare un valore approssimato nell' ipotesi che le f_k quantità della classe k^{ma} seguano una progressione aritmetica a ragione $\frac{c_k}{f_k}$, dove c_k indica l' estensione della classe.

In questa ipotesi, è (1)

$$\Delta_k = \frac{(f_k + 1)\, c_k}{3\, f_k}$$

Se nella (24) si sostituiscono, in base a questa uguaglianza, i valori di Δ_k si ottiene, per R, un valore approssimato R″ che è precisamente quello dato dalla (17).

II. — La relazione trovata tra R e Δ è importante in quanto permette di usufruire delle conclusioni a cui, nella nostra memoria *Variabilità e Mutabilità*, siamo giunti relativamente a Δ.

In detta memoria:

a) abbiamo esaminato l' influenza che sul valore probabile di Δ esercita la scelta a caso con ripetizione e la scelta a caso senza ripetizione (2).

b) abbiamo esaminato la relazione tra i valori di Δ ottenuti per seriazioni parallele (3).

c) abbiamo esaminato le relazioni che, per una data seriazione, intercedono tra il valore di Δ e lo scostamento semplice o quadratico medio dalla mediana o dalla media aritmetica. Tali relazioni variano secondo la distribuzione del carattere. Noi abbiamo messo in luce le relazioni che intercedono tra i detti indici di variabilità quando le intensità del carattere si distribuiscono secondo una progressione aritmetica, secondo una progressione geometrica, secondo la curva degli errori accidentali, secondo una curva di tipo iperbolico, nel caso di massima disuguaglianza di distribuzione (4).

(1) Cfr *Variabilità e Mutabilità*, pag. 52.
(2) Pagg. 37-46.
(3) Pagg. 46-49.
(4) Pagg. 49-83.

d) abbiamo riscontrato quali discordanze possono in pratica presentarsi nei risultati secondo che, nei confronti tra la variabilità dei caratteri, ci si fondi sulla differenza media o invece su uno degli altri indici di variabilità (1). Altri, e più numerosi, riscontri furono eseguiti, in questo campo, dal dott. Giovanni Dettori (2).

e) abbiamo messo in luce le relazioni, che intercedono tra il rapporto della differenza media alla media aritmetica, l' indice α di distribuzione dei redditi introdotto dal Pareto e l' indice δ di concentrazione dei redditi secondo la formula (7) (3).

f) abbiamo esaminato quando, nelle ricerche sulla variabilità dei caratteri, sia più conveniente fondarsi sulla differenza media e quando su di un altro indice di variabilità (4).

g) abbiamo esaminato quando, e in quale misura, nel confronto tra la variabilità dei caratteri, si debba tener conto dei loro valori medii.

Tutte codeste conclusioni si possono importare senz' altro nella teoria del rapporto di concentrazione, tenendo presente che il rapporto di concentrazione è uguale alla differenza media divisa per il doppio della media aritmetica del carattere.

12. — Talvolta non si conosce la seriazione intera del carattere, ma solo una seriazione tronca, che contiene i casi in cui il carattere si presenta con un' intensità superiore ad un dato limite. Ciò avviene spesso nelle statistiche economiche, rilevate a scopo fiscale, relative a redditi, patrimoni, stipendi, affitti, per cui è concesso un limite di esenzione. Vediamo quale influenza eserciti sul rapporto di concentrazione il considerare la seriazione tronca in luogo della seriazione intera del carattere.

Sieno k i casi trascurati sugli n casi totali, p_k sarà la frazione dei casi trascurata nella seriazione tronca e q_k la frazione,

---

(1) Pagg. 83-86.

(2) *Contributo allo studio della variabilità dei prezzi* in *Studi economico-giuridici della R. Università di Cagliari.* Vol. IV. Parte 1ª. Cagliari. Dessì, 1912.

(3) Pagg. 70-80.

(4) Pagg. 86-99.

che l'ammontare del carattere spettante a questi casi rappresenta sull'ammontare totale del carattere.

Nel diagramma seguente (Fig. 5), in cui è al solito $\overline{oc} = p_n = 1$; $\overline{cb} = q_n = 1$, sarà $\overline{og} = p_k$; $\overline{dg} = q_k = \overline{ec}$; $gc = 1 - p_k = \overline{de}$; $\overline{be} = 1 - q_k = \overline{ef}$; $\overline{df} = (1 - q_k) - (1 - p_k) = p_k - q_k$; area $obc = \frac{1}{2}$; area $bde = \frac{1}{2}(1 - p_k)(1 - q_k)$; area $fbd = \frac{1}{2}(p_k - q_k)$ $(1 - q_k)$; area $dfo = \frac{1}{2}(p_k - q_k)\, q_k$.

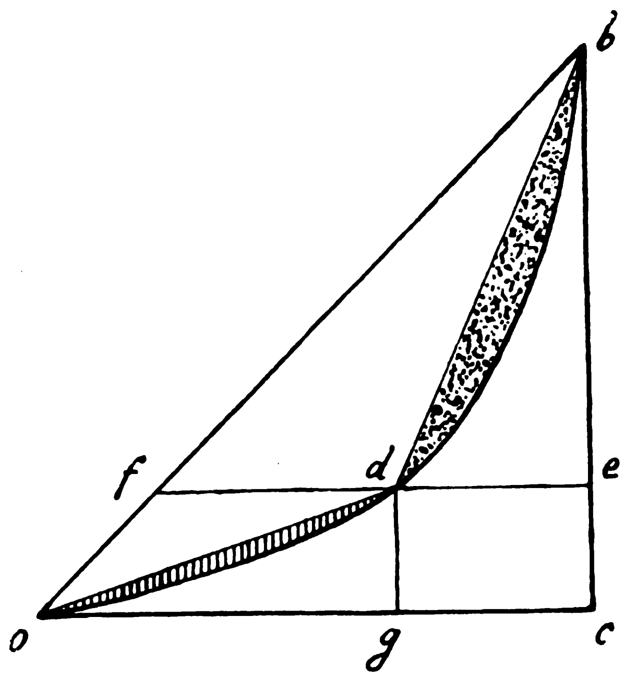

Fig. 5

Indichiamo con C l'area di concentrazione per la seriazione intera, con C′ l'area di concentrazione per la seriazione tronca (area punteggiata del diagramma), con T l'area tratteggiata del diagramma, con R il rapporto di concentrazione per la seriazione intera e con R_{-p_k} il rapporto di concentrazione per la seriazione tronca ottenuta trascurando la frazione p_k dei casi.

Sarà

$$R = \frac{C_1 + \text{area } fdb + \text{area } fdo + T}{\text{area } obc}$$

$$R_{-p_k} = \frac{C_1}{\text{area } bde},$$

da cui, sostituendo in base alle uguaglianze poste e riducendo, si ottiene facilmente

$$R = R_{-p_k}(1 - p_k)(1 - q_k) + p_k - q_k + 2T \qquad (25)$$

Il valore di T si può calcolare approssimativamente nell' ipotesi che le intensità del carattere nei casi trascurati formino una progressione aritmetica. Si intende che questa ipotesi può portare, secondo i casi, ad approssimazioni molto diverse. Quando si possa anche ammettere che il primo termine della progressione sia molto vicino allo 0, si può ritenere (1)

$$\frac{T}{\text{area } odg} = \frac{1}{3}$$

e quindi la (25) si riduce alla forma

$$R = R_{-p_k}(1 - p_k)(1 - q_k) + p_k - q_k + \frac{1}{3} p_k q_k \qquad (26)$$

---

(1) Infatti, se con Δ_k si indica la differenza media dei k casi trascurati e con M_k la loro media aritmetica, sarà

$$\frac{T}{\text{area } odg} = \frac{\Delta_k}{2 M_k}.$$

Ma, se le intensità del carattere nei k casi trascurati costituiscono una progressione aritmetica a ragione H, sarà (cfr. *Variabilità e Mutabilità*, pag. 52

$$\Delta_k = \frac{H(k+1)}{3}$$

e

$$M_k = a_1 + \frac{H(k-1)}{2}$$

dove a_1 indica l'intensità minima del carattere, da cui

$$\frac{T}{\text{area } odg} = \frac{k+1}{3(k-1)} \frac{M_k - \frac{1}{2} a_1}{M_k}$$

che, per k abbastanza grande e a_1 molto vicino allo 0, si può ritenere $= \frac{1}{3}$.

da cui si ricava che è

$$\mathrm{R} \gtrless \mathrm{R}_{-p_k} \quad \text{per} \quad \mathrm{R}_{-p_k} \gtrless 1 - \frac{q_k\left(2 - \frac{4}{3}\,p_k\right)}{p_k + q_k(1-p_k)} \tag{27}$$

La relazione (27) mostra che, secondo i casi, può essere $\mathrm{R} \gtrless \mathrm{R}_{-p_k}$; in altre parole, il rapporto di concentrazione calcolato sulla seriazione tronca può essere diverso e, secondo i casi, maggiore o minore di quello calcolato sulla seriazione intera.

Vi è un caso, che ha notevole importanza pratica, in cui è approssimativamente $\mathrm{R} = \mathrm{R}_{-p_k}$. Se

$$f_l = \mathrm{V}\, x_l^{-h}\, d\, x_l \tag{28}$$

indica il numero delle volte in cui il carattere assume un valore compreso tra x_l e $x_l + d x_l$, e sono soddisfatte le condizioni seguenti: *a*) che il valore minimo di x_l sia molto piccolo di fronte al suo valore massimo, *b*) che $h - 2$ sia positivo e non vicinissimo allo 0, *c*) che il numero $n - k$ dei casi considerati sia abbastanza grande per che possa trascurarsi il fattore $\frac{n-k}{n-k-1}$, la differenza media tra gli $n - k$ valori considerati del carattere può scriversi (1)

$$\Delta = \frac{2}{2h-3}\,\mathrm{M}_{n-k}$$

e il rapporto di concentrazione diviene

$$\mathrm{R} = \frac{1}{2h-3}\,, \tag{29}$$

indipendente dal valore minimo di x_l e dal numero k dei casi trascurati.

Avverandosi le condizioni suddette, il rapporto di concentrazione per la seriazione tronca può ritenersi approssimativamente uguale a quello per la seriazione intera, restando indifferente la frazione dei casi trascurati. Ora le condizioni suddette si verificano quasi sempre per le seriazioni dei redditi globali.

(1) Cfr. *Variabilità e Mutabilità*, pag. 63.

Questi infatti, nella loro distribuzione, seguono, e quasi sempre con buone approssimazioni, la (28), come ha messo in luce il Pareto; i valori di h finora determinati variano da 2,15 a 2,9, il numero dei redditieri dato dalle statistiche di uno Stato è sempre molto alto, e il reddito minimo tassato è sempre molto piccolo in confronto del reddito massimo.

Per i redditi globali, dunque, sarà da attendersi che il rapporto di concentrazione per una seriazione tronca sia approssimativamente rappresentativo del rapporto di concentrazione per la seriazione intera.

Il prof. de' Stefani ha verificato questa previsione relativamente ai redditi della Danimarca (1909). Egli ha calcolato dapprima il valore di R″ (= 45,6 %) e poi i valori di R''_{-r_k}, astraendo successivamente dalla 1ª, dalla 1ª e 2ª, dalla 1ª, 2ª e 3ª classe, e così via. Ecco i risultati ottenuti.

Danimarca, 1909. Redditi

R″ = 45,6 %

Valori di p_k	Valori di R''_{-r_k}
0,63	41,3 %
0,68	40,3 %
0,76	40,1 %
0,83	40,3 %
0,89	40,9 %
0,93	41,8 %
0,97	41,6 %
0,98	41,0 %
0,990	41,1 %
0,994	40,6 %
0,996	40,1 %
0,998	38,9 %

Il rapporto di concentrazione diminuisce alquanto astraendo dalla prima classe, ciò che conferma la circostanza, già osservata dal Pareto, che i redditi minimi non seguono molto bene la (28). Astraendo via via anche dalle classi successive, il rapporto di concentrazione varia ora in un senso, ora nell' altro, ma non fortemente: per che si riscontri un abbassamento sensibile nel suo valore, bisogna considerare solo il 0,2 % dei casi. Si può ritenere confermata, da questo esempio, la previsione che, per i redditi globali, il rapporto di concentrazione calcolato per una seriazione tronca non è essenzialmente diverso da quello per la seriazione intera.

Ben diversa è la conclusione per le seriazioni dei patrimoni e degli stipendi. Il de' Stefani ha calcolato i valori di R''_{-p_k} per i patrimoni della Danimarca (1909) e per gli stipendi degli impiegati pubblici del Belgio (1911).

Danimarca, 1909. Patrimoni

$R'' = 92.7\,\%$

Valori di p_k	Valori di R''_{-p_k}
0,78	64,8 %
0,90	57,4 %
0,93	55,9 %
0,97	54,7 %
0,986	54,1 %
0,995	52,0 %

Belgio, 1911. Stipendi dei pubblici impiegati.

$R'' = 34{,}3\,\%$

Valori di p_k	Valori di R''_{-p_k}
0,009	33,9 %
0,054	37,2 %
0,640	26,0 %

Per i patrimoni danesi, il rapporto di concentrazione diminuisce continuamente e fortemente col crescere della frazione dei casi trascurati; per gli stipendi del Belgio, il suo andamento è più irregolare. In ogni modo, per questi caratteri, il rapporto di concentrazione determinato per una seriazione tronca non può darci un'idea neppure approssimata di quello che sarebbe il suo valore per la seriazione intera.

Vediamo ora con quanta approssimazione, dai valori di p_k, q_k, R_{-p_k}, si possa risalire, in base alla (26), ai valori di R. Diamo qui, i valori di p_k e di q_k e i valori calcolati di R per i patrimoni e i redditi della Danimarca e per gli stipendi del Belgio; nelle tavole precedenti, si possono trovare i corrispondenti valori di R_{-p_k}.

Danimarca, 1909. Redditi

R″ = 45,6 %

Valori di p_k	Valori di q_k	Valori di R calcolati in base alla (26)
0,633	0,302	50,0 %
0,683	0,339	50,6 %
0,762	0,414	50,8 %
0,828	0,491	50,8 %
0,885	0,573	50,1 %
0,933	0,661	48,7 %
0,968	0,749	46,4 %

Danimarca, 1909. Patrimoni

R″ = 92,7 %

Valori di p_k	Valori di q_k	Valori di R calcolati in base alla (26)
0,783	0,052	87,8 %
0,900	0,181	82,1 %
0,972	0,419	69,8 %
0,995	0,644	56,6 %

Belgio, 1911. Stipendi di impiegati pubblici

R″ = 34,3 %

Valori di p_k	Valori di q_k	Valori di R calcolati in base alla (26)
0,009	0,0016	34,3 %
0,054	0,019	38.1 %
0,640	0,388	39,2 %

Per i redditi della Danimarca e per gli stipendi del Belgio, i valori di R calcolati in base alla (26) non sono essenzialmente diversi, per quanto sensibilmente più elevati, dal rispettivo valore R″ determinato sulla seriazione intera; per i patrimoni danesi, invece, essi rimangono molto al di sotto di R″ e tanto più se ne discostano quanto più è elevata la frazione p_k dei casi trascurati.

Varrebbe la pena, io penso, di estendere queste indagini ad altri caratteri e, per ogni carattere, a parecchie seriazioni, per stabilire quale influenza sul rapporto di concentrazione esercita, nei varii casi, l' operare su di una seriazione tronca e con quale approssimazione, in base alla (26) o ad altre formule, che si possono facilmente escogitare fondandosi su ipotesi diverse e meglio rispondenti al vero, si possa risalire dal rapporto di concentrazione per una seriazione tronca al rispettivo rapporto di concentrazione per la seriazione intera.

13. — Riassumiamo i risultati presentati in questa memoria.

1. Prendendo le mosse dal concetto di *indice di concentrazione*, da noi svolto e applicato fin dal settembre 1908 e più completamente nel settembre 1909, abbiamo proposto un indice di concentrazione (detto *rapporto di concentrazione*) valido per tutti i caratteri, qualunque sia la loro distribuzione.
2. Abbiamo dimostrato come al rapporto di concentrazione si pervenga anche perfezionando un metodo grafico proposto da Lorenz (1905), Chatelain (1907), Séailles (1910) per rappresentare la distribuzione della ricchezza.
3. Abbiamo dimostrato pure come il rapporto di concentrazione stia in una relazione molto semplice con la *differenza media* tra le intensità del carattere: esso è precisamente il rapporto della differenza media al valore che questa assumerebbe nel caso di massima disuguaglianza di distribuzione, o, in altre parole, il rapporto della differenza media al doppio della media aritmetica del carattere.
4. Abbiamo indicato i procedimenti aritmetici e grafici, convenienti per il calcolo del rapporto di concentrazione, sia quando le statistiche forniscono gli elementi completi sulle intensità del carattere, sia in molti casi, in cui esse forniscono soltanto elementi più o meno incompleti.
5. Del rapporto di concentrazione abbiamo dato numerose applicazioni, sia a caratteri fisici dell'uomo, sia a caratteri economici.

Le teorie degli indici di concentrazione e degli indici di variabilità vengono così collegate fra loro e resta dimostrato il significato che, rispetto ad esse, hanno i metodi grafici escogitati per rappresentare la distribuzione della ricchezza.

(Licenziate le bozze per la stampa il giorno 25 Giugno 1914)