METRON

RIVISTA INTERNAZIONALE DI STATISTICA — REVUE INTERNATIONALE DE STATISTIQUE
INTERNATIONAL REVIEW OF STATISTICS — INTERNATIONALE STATISTISCHE ZEITSCHRIFT

DIRETTORE PROPRIETARIO — DIRECTEUR ET PROPRIÉTAIRE
EDITOR AND PROPRIETOR — HERAUSGEBER UND EIGENTHÜMER

Prof. Dott. Corrado Gini, *direttore dell'Istituto di Statistica della R. Università di Roma, presidente dell'Istituto Centrale di Statistica del Regno d'Italia.*

COMITATO DIRETTIVO — COMITÉ DE DIRECTION
EDITORIAL COMMITTEE — DIREKTION-KOMITEE

Prof. A. Andréadès (*Athènes*) - **Prof. A. E. Bunge** (*Buenos Aires*) - **Prof. F. P. Cantelli** (*Roma*)
Prof. C. V. L. Charlier (*Lund*) - **Prof. F. v. Fellner** (*Budapest*) - **Prof. A. Flores de Lemus** (*Madrid*)
Prof. M. Greenwood (*London*) - **Ing. L. March** (*Paris*) - **Prof. A. W. Methorst** (*La Haye*)
Prof. A. Julin (*Bruxelles*) - **Prof. R. Pearl** (*Baltimore*) - **Prof. H. Westergaard** (*Copenhagen*)

AMMINISTRATORE — ADMINISTRATEUR — MANAGER — VERWALTER

Dott. Silvio Orlandi, *Istituto di Statistica e Politica Economica della R. Università di Roma*

SEGRETARI DI REDAZIONE — SECRÉTAIRES DE RÉDACTION
EDITORIAL SECRETARIES — REDACTIONSSECRETAERE

Prof. Luigi Galvani - Dott. Mario Saibante

ISTITUTO CENTRALE DI STATISTICA BIBLIOTECA

Vol. VIII - N. 3. **28-II-1930.**

SOMMARIO — SOMMAIRE — CONTENTS — INHALT

C. Gini. *Sul massimo degli indici di variabilità assoluta e sulle sue applicazioni agli indici di variabilità relativa e al rapporto di concentrazione* . . . . *Pag.* 3

G. A. Baker. *Random Sampling from non-homogeneous populations.* » 67

B. de Finetti e U. Paciello. *Calcolo della differenza media* . . . » 89

V. A. Nekrassoff. *Nomography in applications of Statistics.* . . » 95

R. Roy. *La demande dans ses rapports avec la répartition des revenus* . . . . » 101

ROMA
AMMINISTRAZIONE DEL « METRON »
R. UNIVERSITÀ — ISTITUTO DI STATISTICA

La Rivista internazionale di Statistica METRON esce in fascicoli. Quattro fascicoli consecutivi costituiscono complessivamente un volume di 700-800 pagine.

METRON accoglie articoli originali di metodologia statistica e di applicazioni statistiche alle varie discipline, e rassegne o discussioni di risultati raggiunti col metodo statistico in diversi campi della scienza o tali da poter interessare il cultore della statistica. Pubblica altresì una bibliografia di tutte le opere e riviste ricevute in omaggio od in cambio.

Articoli e rassegne potranno essere scritti in italiano, francese, inglese o tedesco. I manoscritti in lingua francese, inglese o tedesca dovranno essere dattilografati.

La collaborazione non è retribuita. Gli autori riceveranno gratuitamente 25 estratti dei lavori pubblicati.

I manoscritti per la pubblicazione dovranno essere indirizzati al *Prof. Corrado Gini, R. Università di Roma — Istituto di Statistica*, oppure al membro del Comitato direttivo che rappresenta lo Stato a cui l'autore appartiene. Gli autori sono pregati di conservare copia del manoscritto inviato, poichè, nel caso che questo non venga pubblicato, la Direzione non ne garantisce la restituzione.

Al Prof. Corrado Gini dovranno pure essere indirizzate le richieste di cambi da parte di riviste o di altri periodici e ogni pubblicazione inviata in cambio od in omaggio.

Le richieste di abbonamenti, del pari che i versamenti, dovranno invece essere indirizzati alla *Amministrazione del « Metron »* presso l'*Istituto di Statistica della R. Università di Roma — Via delle Terme di Diocleziano*, 10.

Il prezzo di abbonamento per ciascun Volume è di **100 Lire italiane** e quello del fascicolo di **30 Lire italiane**, porto compreso.

La Revue Internationale de Statistique METRON paraît par livraisons. Quatre livraisons consécutives forment un volume de 700-800 pages.

METRON publie des articles originaux de méthodologie statistique et d'applications statistiques aux différentes disciplines, ainsi que des revues ou des discussions des résultats obtenus par la méthode statistique dans toutes les sciences ou bien intéressant les savants qui s'occupent de statistique.

METRON publie aussi une bibliographie de tous les ouvrages et revues reçues en hommage ou en échange.

Les articles et les revues pourront être écrites en français, en italien, en anglais ou en allemand. Les manuscrits en français, en anglais ou en allemand doivent être envoyés dactylographiés.

On enverra gratis aux auteurs 25 copies tirées à part de leurs travaux après publication.

On adressera les manuscrits pour la publication à *M. le Prof. Corrado Gini, Istituto di Statistica, R. Università di Roma (Italie)*, ou bien au membre du comité de direction représentant le pays de l'auteur. On prie les auteurs de garder une copie du manuscrit qu'ils adressent à la Revue, car, en cas de non publication, la rédaction ne garantit pas de pouvoir le renvoyer.

Les demandes d'échange de la part des Revues et des autres périodiques, ainsi que toutes les publications envoyées en échange ou en hommage doivent aussi être adressées au Prof. Corrado Gini.

Les demandes de nouveaux abonnements, ainsi que tout payement, devront être adressés à l'*Administration du « Metron »* auprès de l'*Institut de Statistique de l'Université Royale de Rome — Italie.*

Le prix d'abonnement par volume est fixé à **100 Lires it.** et le prix par fascicule est de **30 Lires it.** frais d'envoi compris.

METRON

RIVISTA INTERNAZIONALE DI STATISTICA — REVUE INTERNATIONALE DE STATISTIQUE
INTERNATIONAL REVIEW OF STATISTICS — INTERNATIONALE STATISTISCHE ZEITSCHRIFT

DIRETTORE PROPRIETARIO — DIRECTEUR ET PROPRIÉTAIRE
EDITOR AND PROPRIETOR — HERAUSGEBER UND EIGENTHÜMER

Prof. Dott. Corrado Gini, *direttore dell'Istituto di Statistica della R. Università di Roma, presidente dell'Istituto Centrale di Statistica del Regno d'Italia.*

COMITATO DIRETTIVO — COMITÉ DE DIRECTION
EDITORIAL COMMITTEE — DIREKTION-KOMITEE

Prof. A. Andréadès, *de Science des finances à l'Université d'Athènes* (Grèce).
Prof. A. E. Bunge, *director general de Estadística de la Nación, Buenos Aires* (Argentina).
Prof. F. P. Cantelli, *professore di Matematica Attuariale nel R. Istituto Superiore di Scienze Economiche e Commerciali di Napoli* (Italia).
Prof. C. V. L. Charlier, *professor der Astronomie an der Universität Lund* (Schweden).
Prof. F. von Fellner, *o. off. Universitäts-Professor in Budapest* (Ungarn).
Prof. A. Flores de Lemus, *jefe de Estadística del Ministerio de Hacienda, Madrid* (España).
Prof. M. Greenvood, *professor of Epidemiology and Vital Statistics in the University of London* (England).
Ing. L. March, *directeur honoraire de la Statistique générale de la France, Paris* (France).
Prof. H. W. Methorst, *directeur de l'Office permanent de l'Institut International de Statistique et du Bureau central de Statistique, La Haye* (Pays-Bas).
Prof. A. Julin, *secrétaire général du Ministère de l'Industrie et du Travail, Bruxelles* (Belgique).
Prof. R. Pearl, *director of the Institute for Biological Research at the J. Hopkins University, Baltimore* (U. S. A.).
Prof. H. Westergaard, *professor in the University of Copenhagen* (Denmark).

AMMINISTRATORE — ADMINISTRATEUR — MANAGER — VERWALTER

Dott. Silvio Orlandi, *Istituto di Statistica e Politica Economica della R. Università di Roma.*

SEGRETARI DI REDAZIONE — SECRÉTAIRES DE RÉDACTION
EDITORIAL SECRETARIES — REDACTIONSSECRETAERE

Prof. Luigi Galvani — Dott. Mario Saibante

Vol. VIII - N. 3. **28-II-1930.**

SOMMARIO — SOMMAIRE — CONTENTS — INHALT

C. Gini. *Sul massimo degli indici di variabilità assoluta e sulle sue applicazioni agli indici di variabilità relativa e al rapporto di concentrazione* Pag. 3
G. A. Baker. *Random Sampling from non homogeneous populations. . .* » 67
B. de Finetti e U. Paciello. *Calcolo della differenza media . . . . . .* » 89
V. A. Nekrassoff. *Nomography in applications of Statistics. . . . . . .* » 95
R. Roy. *La demande dans ses rapports avec la répartition des revenus. .* » 101

ROMA

AMMINISTRAZIONE DEL « METRON »
R. UNIVERSITÀ — ISTITUTO DI STATISTICA

ARTICOLI GIUNTI ALLA RIVISTA CHE VERRANNO PUBBLICATI NEI PROSSIMI NUMERI.

(*Secondo l'ordine d'arrivo*)

ARTIKEL DIE AN DIE ZEITSCHRIFT ANGELANGT SIND UND WELCHE IN DEN NACHFOLGENDEN NUMMERN ERSCHEINEN WERDEN.

(*Nach der Reihenfolge des Eingangs*)

ARTICLES REÇUS PAR LA REVUE ET À PARAÎTRE PROCHAINEMENT.

(*D'après la date de réception*)

ARTICLES RECEIVED BY THE REVIEW WHICH WILL BE PUBLISHEED IN FUTURES ISSUES.

(*According to date of receipt*)

M. Fréchet. *Sur la convergence en probabilité.*

J. O. Irwin, M. A., M. Sc. *On the frequency distribution of the means of samples from populations of certain of Pearson's types.*

L. I. Dublin and A. J. Lotka. *The True Rate of Natural Increase of the Population of the United States. Revision on Basis of Recent Data.*

P. D. Rediadis. *The Greek National Income and Wealth.*

F. Savorgnan. *La Fecondità delle Aristocrazie.*

Gli Autori degli articoli inviati per la pubblicazione nella Rivista, rinunciano in favore della medesima alla proprietà letteraria degli articoli stessi, qualora vengano pubblicati.

The Authors of papers sent for publication in the Review are supposed to give up their copyright in favour of the Review if the papers are published.

Les Auteurs des articles envoyés à la Revue pour y être publiés, renoncent, en faveur de celle-ci, à la propriété littéraire de leurs articles, s'ils sont acceptés.

Die Verfasser der zur Veröffentlichung in der Zeitschrift zugesandten Aufsätze, werden, falls selbige veröffentlicht werden, auf ihre Verfasserrechte zu Gunsten der Zeitschrift verzichten müssen.

CORRADO GINI

Sul massimo degli indici di variabilità assoluta e sulle sue applicazioni agli indici di variabilità relativa e al rapporto di concentrazione.

1. S'intende per *massimo di un indice di variabilità* il valore che detto indice assume nel caso di disuguaglianza massima fra le quantità della seriazione. Chiameremo *seriazione massimante dell'indice di variabilità* la seriazione fittizia per cui si verifica tale circostanza.

Sia una seriazione di n quantità, A sia la media aritmetica di queste e quindi $T = n\,A$ la somma di tutte le quantità.

Il massimo dei seguenti indici di variabilità assoluta:

a) scostamento semplice medio (${}^1S_A$) dalla media aritmetica,
b) scostamento semplice medio (${}^1S_M$) dalla mediana,
c) scostamento quadratico medio (${}^2S_A$) dalla media aritmetica,
d) scostamento quadratico medio (${}^2S_M$) dalla mediana,
e) differenza media (Δ),
f) differenza media con ripetizione (Δ_R),

venne determinato nell'ipotesi che una quantità sia uguale a T e le altre quantità siano uguali a zero. Si trova allora (*)

$$\left.\begin{array}{ll} \text{Max } {}^1S_A = 2\dfrac{n-1}{n}A & \text{Max } {}^1S_M = A \\[2ex] \text{Max } {}^2S_A = \sqrt{n-1}\,A & \text{Max } {}^2S_M = \sqrt{n}\,A \\[2ex] \text{Max } \Delta = 2\,A & \text{Max } \Delta_R = 2\dfrac{n-1}{n}A \end{array}\right\} \quad (1)$$

2. Se non che, non sempre si può ammettere che una quantità sia uguale a T. Ciò si può ammettere, invero, tutte le volte che per l'in-

(*) Cfr. C. GINI, *Variabilità e Mutabilità*, in « Studi economico-giuridici della R. Università di Cagliari », Anno III, Parte 2ª, 1912, pag. 80.

tensità del carattere non si può porre *a priori* un limite superiore, come, per esempio, avviene per il patrimonio, per il reddito, per l'altitudine, per la densità della popolazione e, nella maggior parte dei paesi, per la proprietà terriera. Noi vediamo — è vero — che nel fatto questi caratteri non oltrepassano certi valori, ma non possiamo escludere che essi possano oltrepassarli in altro tempo e in altri paesi. Vi sono invece molti altri caratteri i quali, per definizione, non possono oltrepassare un dato limite : tali sono, in generale, quelli che si esprimono con frazioni proprie e con percentuali, come, per esempio, la natalità, la mortalità, la nuzialità, e tali sono ancora parecchi altri, come la latitudine, la longitudine. Ora per le serie relative a tali caratteri, può avvenire, e spesso effettivamente avviene, che non si possa ammettere che un termine raggiunga il valore T. Se noi esaminiamo, ad esempio, la variabilità dei coefficienti di mortalità durante il secolo XIX in Isvezia (media annuale $= 21{,}4\ {}^0/_{00}$) e vogliamo determinare il massimo che può assumere, per esempio, la differenza media tra i 100 coefficienti di mortalità, evidentemente non possiamo supporre che in un anno si sia verificata la mortalità $= 21{,}4 \times 100 = 2140\ {}^0/_{00}$; tutt'al più potremo supporre che in un anno tutti gli abitanti sieno morti, ciò che darebbe una mortalità del $1000\ {}^0/_{00}$. Ora, per serie siffatte, è pure importante determinare il massimo degli indici di variabilità.

Noi lo determineremo dapprima facendo, per semplicità, le ipotesi :

a) che a tutti i termini delle seriazioni si attribuisca lo stesso peso ; per esempio, nel caso della mortalità della Svezia nel secolo XIX, che si attribuisca lo stesso peso ai 100 coefficienti annuali di mortalità ;

b) che, se L indica il limite superiore, $k = \frac{nA}{L}$ sia un numero intero ≥ 1, in altre parole che l'ammontare totale del carattere $T = nA$ sia multiplo esatto del limite superiore in modo che esso possa distribuirsi esattamente su un certo numero di termini, qualora questi presentino un valore corrispondente a tale limite.

Vedremo più innanzi quando e come si possa prescindere da tali ipotesi e quale ne sia la portata.

3. La relazione $k = \frac{nA}{L}$ esprime anche che il limite superiore L, dei valori del carattere è $\frac{nA}{k}$ e che, nell'ipotesi di disugua-

glianza massima, k saranno i casi che presentano il limite superiore L, ed $n - k$ i casi che presentano il valore 0.

Sarà quindi:

$$\left.\begin{aligned}
\text{Max } {}^{1}S_A &= \frac{1}{n}\left\{(n-k)(A-0) + k\left(\frac{nA}{k} - A\right)\right\} = 2\,\frac{n-k}{n}\,A \\
\text{Max } {}^{2}S_A &= \sqrt{\frac{1}{n}\left\{(n-k)(A-0)^2 + k\left(\frac{nA}{k} - A\right)^2\right\}} = \sqrt{\frac{n-k}{k}}\,A \\
\text{Max } \Delta &= \frac{1}{n(n-1)}\Big\{k(k-1)\,0 + (n-k)(n-k-1)\,0 + \\
&\quad + 2k(n-k)\frac{nA}{k}\Big\} = 2\,\frac{n-k}{n-1}\,A \\
\text{Max } \Delta_R &= \frac{1}{n^2}\left\{k^2\,0 + (n-k)^2\,0 + 2k(n-k)\frac{nA}{k}\right\} = 2\,\frac{n-k}{n}\,A
\end{aligned}\right\}\quad(2)$$

S'intende che, in conformità all'ipotesi *b*), k è un numero intero ≥ 1. Per $k = 1$, queste formule si riducono alle (1).

Per gli scostamenti dalla mediana, conviene distinguere due casi, a seconda che sia $k \gtrless n - k$.

Se è $k < n - k$, la mediana sarà uguale a 0, e si avrà:

$$\left.\begin{aligned}
\text{Max } {}^{1}S_M &= \frac{1}{n}\left\{k\left(\frac{nA}{k} - 0\right) + (n-k)(0-0)\right\} = A \\
\text{Max } {}^{2}S_M &= \sqrt{\frac{1}{n}\left\{k\left(\frac{nA}{k} - 0\right)^2 + (n-k)(0-0)^2\right\}} = \sqrt{\frac{n}{k}}\,A
\end{aligned}\right\}\text{a)}$$

Se è $k > n - k$, la mediana sarà $= \dfrac{nA}{k}$ e si avrà

$$\left.\begin{aligned}
\text{Max } {}^{1}S_M &= \frac{1}{n}\left\{k\left(\frac{nA}{k} - \frac{nA}{k}\right) + (n-k)\left(\frac{nA}{k} - 0\right)\right\} = \frac{n-k}{k}\,A \\
\text{Max } {}^{2}S_M &= \sqrt{\frac{1}{n}\left\{k\left(\frac{nA}{k} - \frac{nA}{k}\right)^2 + (n-k)\left(\frac{nA}{k} - 0\right)^2\right\}} = \\
&= \sqrt{\frac{n}{k}\,\frac{n-k}{k}}\,A
\end{aligned}\right\}\text{b)}\quad(2)$$

Si noti che, per $k = 1$, è $k < n - k$, e quindi sono applicabili le (2^a) e non le (2^b). Ora per $k = 1$ le (2^a) si riducono alle (1).

Noi siamo così pervenuti ad una generalizzazione delle formule (1) che altra volta avevamo dato per i massimi degli indici di variabilità. Le formule (1) valgono nel caso particolare che l'ammontare totale del carattere si possa concepire concentrato in una sola osservazione; le formule generalizzate (2) valgono in ogni caso, sempre che — ben inteso — si verifichino le due ipotesi suesposte.

Ci si persuade facilmente che le (2), essendo $k > 1$, danno sempre valori inferiori ai valori rispettivi ottenuti dalle (1), (*) salvo per ${}^1S_M$ nel caso che sia $k < n - k$, in cui le due formule coincidono.

Il massimo degli indici di variabilità fornito dalle (1) viene dunque ad essere esagerato tutte le volte che sia $k > 1$, salvo per lo scostamento semplice medio dalla mediana, nel caso di $k < n - k$.

4. Le formule (2) si possono mettere sotto la forma seguente che è spesso più conveniente, anche perchè, sotto questa forma, esse non contengono esplicitamente k.

$$\left.\begin{array}{ll}
\text{Max } {}^1S_A = \dfrac{2A(L-A)}{L} & \text{Max } {}^1S_M = \begin{cases} A \\ L-A \end{cases} \\
& \text{a seconda che è } L \gtreqless 2A \\
\text{Max } {}^2S_A = \sqrt{A(L-A)} & \text{Max } {}^2S_M = \begin{cases} \sqrt{AL} \\ \sqrt{(L-A)L} \end{cases} \\
& \text{a seconda che è } L \gtreqless 2A \\
\text{Max } \Delta = 2\dfrac{n}{n-1}\dfrac{A(L-A)}{L} & \text{Max } \Delta_R = 2\dfrac{A(L-A)}{L}
\end{array}\right\} \quad (2 \text{ bis})$$

Per $L = nA$, le (2 bis) si riducono alle (1). Il caso $L > nA$, che corrisponde all'ipotesi $k < 1$, resta escluso dalla premessa *b*) fatta al § 2. In ogni modo s'intende che nella pratica è come se in tal caso il limite non esistesse e quindi valgono le formule (1).

5. I massimi degli indici di variabilità sono utili per la determinazione degli indici di variabilità relativa.

Gli indici di variabilità assoluta spesso non sono infatti comparabili tra di loro, sia perchè i caratteri sono misurati con unità diverse (es. lunghezza, peso, capacità, ecc.), sia anche perchè è diverso il massimo che essi possono raggiungere. Si superano entrambe codeste difficoltà, ragguagliando gli indici di variabilità assoluta al massimo che essi possono raggiungere. Gli indici di variabilità relativa, che così si ottengono, esprimono a quale percentuale della massima

(*) Nel confronto fra le (2^b) e le (1) rispettive, si tenga presente che, per $k > n - k$, è $\dfrac{n-k}{k} < 1$.

variabilità concepibile corrisponde la variabilità effettivamente osservata nella serie.

Nel caso che il massimo dell'indice di variabilità assoluta sia dato dalle formule (1), questo punto di vista veniva a giustificare la consuetudine di ragguagliare l'indice di variabilità assoluta alla media aritmetica, ai fini di ottenere l'indice di variabilità relativa, fin tanto che l'indice di variabilità assoluta si desume dallo scostamento semplice medio dalla media aritmetica o dalla mediana o dalla differenza media semplice o con ripetizione. Il massimo dello scostamento semplice medio dalla mediana e quello della differenza media semplice sono infatti esattamente proporzionali alla media aritmetica; e il massimo dello scostamento semplice medio dalla media aritmetica e quello della differenza media con ripetizione sono approssimativamente (a meno di $1/n$) proporzionali a tale media (*).

Le formule (2 bis) fanno ora comprendere che, se tale procedimento è giustificato quando un limite superiore L dell'intensità del carattere non esiste (o quanto meno quando non esiste un limite superiore L, tale che sia $L < nA$), esso non lo è più quando esiste un tale limite.

In tal caso, anzi che ragguagliare lo scostamento semplice medio o la differenza media semplice o con ripetizione alla media aritmetica A converrà ragguagliarli alla quantità $\frac{A\ (L - A)}{L}$. Ciò significa, in altre parole, introdurre nel denominatore un coefficiente di correzione $\frac{L - A}{L}$, che è sempre una frazione propria. Ragguagliando i detti indici di variabilità assoluta alla media aritmetica, il valore dell'indice di variabilità relativa risulta dunque inferiore al reale quando esiste un limite superiore dell'intensità del carattere, e tanto più inferiore quanto più il detto limite risulta vicino alla media aritmetica.

Per ciò che concerne lo scostamento semplice medio dalla mediana, converrà ragguagliarlo alla media aritmetica o al suo valore complementare secondo che è minore l'uno o l'altro. In questo caso, il procedimento usuale di ragguagliarlo ogni volta alla media aritmetica poteva dunque talora, ma non sempre, portare a valori dell'indice di variabilità relativo inferiori al vero.

Per ciò che concerne infine lo scostamento quadratico medio dalla

(*) Cfr. *Variabilità e Mutabilità*, pag. 102.

media o dalla mediana, il valore massimo dell'indice di variabilità assoluta (e quindi anche il valore dell'indice di variabilità relativa) mentre risulta dipendente dal numero delle osservazioni n quando l'intensità del carattere non ammette un limite superiore (Cfr. formule (1)) (*), ne risulta indipendente quando il limite esiste ed è inferiore a T (Cfr. formule (2 bis)).

Ora si noti che i valori massimi di tali indici, e cioè $\sqrt{A(L-A)}$ per lo scostamento quadratico medio dalla media, e $\sqrt{AL}$ o $\sqrt{(L-A)L}$ per lo scostamento quadratico medio dalla mediana, possono essere, secondo i casi, superiori o inferiori ad A, per modo che, ragguagliando tali indici di variabilità assoluta alla media aritmetica, come si faceva usualmente, anzichè al loro valore massimo, si ottengono valori ora inferiori ora superiori ai veri, ciò che dà luogo a inconvenienti manifesti nei paragoni.

6. Se A indica la probabilità elementare p di un fenomeno, l'espressione $\sqrt{A(L-A)}$ assume la forma $\sqrt{p(1-p)}$ ben nota nel calcolo delle probabilità e nella teoria della dispersione. Questa teoria viene così ad essere direttamente collegata con la teoria degli indici di variabilità relativa.

7. Si avverta come i massimi di tutti gli indici di variabilità assoluta considerati (e quindi gli indici di variabilità relativa che se ne ricavano), risultano simmetrici rispetto ad A e a $L-A$. Ciò significa che gli indici di variabilità relativa risultano identici per due caratteri complementari (per es. mortalità e sopravvivenza, percentuali degli immobili e dei mobili nell'annualità ereditaria, ecc.), come appare naturale.

Resta così eliminato un inconveniente degli indici di variabilità relativa desunti dal rapporto degli indici di variabilità assoluta alla media aritmetica, che io avevo altra volta messo in luce (**).

8. Deve notarsi che non sempre il limite superiore L dei valori del carattere è un limite raggiungibile. Ciò si verifica anzi spesso

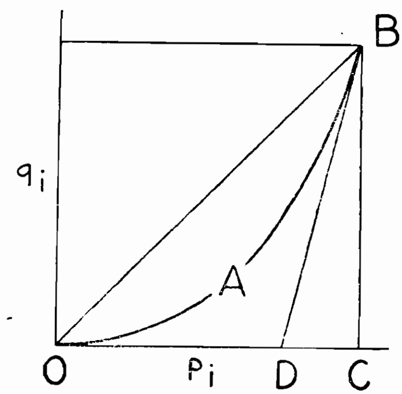
(*) In questo caso i paragoni tra gli indici di variabilità relativa desunti dal rapporto dello scostamento quadratico medio dalla media o dalla mediana al loro valore massimo sono legittimi solo tra serie aventi lo stesso numero di termini.

(**) Cfr. *Variabilità e Mutabilità*, pagg. 100-101.

in quei rapporti statistici che si sogliono chiamare *rapporti di derivazione generica*. Per esempio, è possibile che tutti i viventi muoiano e quindi che il coefficiente di mortalità generale per l'insieme della popolazione raggiunga il 1000 $^0/_{00}$, ma non è possibile invece che tutti i viventi partoriscano in un anno o che tutti si sposino, e quindi non sarà mai possibile che il coefficiente di natalità generica dato dal rapporto dei nati agli abitanti o il coefficiente di nuzialità generica dato dal numero dei matrimoni agli abitanti raggiunga il 1000$^0/_{00}$. D'altra parte riesce difficile, per non dire praticamente impossibile, nella maggior parte dei casi, di determinare il limite superiore che siffatti rapporti di derivazione potrebbero effettivamente raggiungere, nè è detto (anzi per la maggior parte delle serie può dirsi escluso) che il limite superiore raggiungibile sarebbe lo stesso per tutti i termini della serie. Quindi in tali casi è forza procedere considerando un limite superiore che nel fatto non è raggiungibile. S'intende che, qualora esistesse un limite superiore ugualmente raggiungibile per tutti i termini e questo si potesse determinare, gli indici di variabilità relativa risulterebbero più elevati e diverso risulterebbe il loro massimo. Si verifica in sostanza, in tal caso, per gli indici di variabilità relativa dei rapporti di derivazione generica, un inconveniente analogo a quello che gli statistici hanno da tanto tempo segnalato per la stessa intensità di tali rapporti e per evitare il quale, ricorrono, ogni qualvolta sia possibile, ai rapporti di *derivazione specifica*. Il rapporto di derivazione è, invero, diretto a ragguagliare l'intensità di un fenomeno o carattere, all'intensità di un altro fenomeno o carattere che del primo costituisca il presupposto, per modo che il denominatore dovrebbe a rigore essere tale che da tutti i suoi elementi potessero derivare gli elementi del numeratore. Così è per il coefficiente di mortalità generale, da ciascun abitante potendo derivare un decesso; così non è per i coefficienti di natalità generica e di nuzialità generica, non potendosi da tutti gli abitanti attendere una nascita o un matrimonio. Perciò appunto si fa ricorso, ogni qualvolta riesce possibile, ai coefficienti di natalità specifica ragguagliando i nati alla popolazione in età riproduttiva, o più specificamente alla popolazione femminile in età riproduttiva, e analogamente ai coefficienti di nuzialità specifica, ragguagliando i matrimonî alle persone in età atta al matrimonio o ai non coniugati in età atta al matrimonio.

In questo, come in tanti altri campi, la metodologia statistica deve accontentarsi d'indicare i metodi esatti e di segnalare, senza avere la pretesa di eliminarle, le difficoltà di applicazione dipendenti dal materiale disponibile.

9. Strettamente collegata con la teoria degli indici di variabilità relativa è la teoria degli indici di concentrazione, e in particolare del rapporto di concentrazione (*). Se, ordinati i casi in cui si presenta un carattere secondo l'intensità crescente di questo, si portano sull'asse delle ascisse le percentuali p_i che sul numero totale rappresentano quei casi che presentano un'intensità inferiore a un dato limite, e sull'asse delle ordinate le percentuali q_i che, sull'ammontare totale T, rappresentano in complesso i valori del carattere che spettano ai detti casi, e si considera il luogo dei punti determinati dalle coppie di coordinate corrispondenti p_i e q_i, tale luogo costituisce ciò che si dice una *curva di concentrazione*, la quale è convessa (**) verso l'asse delle ascisse e analoga alla $O\ A\ B$ del diagramma qui accanto.



Ora il rapporto dell'*area di concentrazione*, compresa fra detta curva e il segmento OB, all'area rappresentata dal triangolo OBC, dà il valore del *rapporto di concentrazione*. Si dimostra che l'area di concentrazione è proporzionale alla differenza media con ripetizione e l'area del triangolo OBC è proporzionale al suo valore massimo, che si verifica nell'ipotesi che tutto l'ammontare del

(*) Cfr., per questa teoria, la memoria *Sulla misura della concentrazione e della variabilità dei caratteri*, in « Atti del R. Istituto Veneto di Scienze, Lettere ed Arti », Anno accademico 1913-14, Tomo LXXIII, Parte seconda.

(**) Infatti è facile vedere che le prime due differenze sono positive, ovvero, ragionando sul continuo, che le prime due derivate sono positive.

carattere T, corrispondente all'ordinata CB, si concentri in un solo caso, così che il rapporto di concentrazione è esprimibile col rapporto della differenza media con ripetizione al suo massimo, nell'ipotesi che l'intensità del carattere non abbia limite o abbia un limite superiore a T.

Questa ipotesi però non era stata enunciata, per modo che si sarebbe potuto pensare a calcolare nello stesso modo il rapporto di concentrazione anche quando l'ipotesi non si avverava e l'intensità del carattere ammetteva un limite inferiore a T. Quanto si è detto nei paragrafi precedenti fa comprendere come il procedimento grafico per la determinazione del rapporto di concentrazione possa facilmente generalizzarsi in modo da adattarsi a tutte le ipotesi.

Se sul segmento OC, proporzionale al numero n dei casi della serie, si prende, a partire da C, il segmento CD proporzionale a k, che rappresenta il numero minimo dei casi in cui potrebbe concentrarsi tutto l'ammontare T del carattere effettivamente distribuito sugli n casi, il rapporto di concentrazione sarà dato dal rapporto dell'area di concentrazione all'area del triangolo OBD. Nel caso di $k = 1$, si rientra nel caso precedente.

Nella pratica le differenze tra i risultati ottenuti col metodo originario e col metodo generalizzato possono essere notevoli e portare a conclusioni diverse nei confronti tra la concentrazione di più caratteri.

Nella tavola seguente sono dati, oltre le medie aritmetiche (A), i rapporti di concentrazione calcolati col metodo originario (R_o) e col metodo generalizzato (R_g) per 9 caratteri di cui 3 (quelli ai numeri 7, 8 e 9) non ammettono, e gli altri 6 ammettono un limite, essendo per tutti questi 6 caratteri $L < nA$. A fianco dei valori R_o e R_g sono indicati tra parentesi i posti nelle rispettive graduatorie decrescenti (*).

(*) I dati della tavola sono ricavati dalla comunicazione : *Une application de la méthode représentative aux données du recensement italien en 1921*, presentata alla XVII Sessione dell'Istituto Internazionale di Statistica (Cairo 1927-28) e pubblicata nel « Bulletin » di detto Istituto, Tome XXIII, 2ème Livraison, p. 198-216. In detta memoria sono state date le prime applicazioni del metodo generalizzato per il calcolo degli indici di variabilità esposto in questa memoria. Un'estensione di detta comunicazione, a cura dello scrivente e del Prof. Galvani, è stata pubblicata negli « Annali di Statistica », Serie VI, Vol. IV.

CARATTERI	Intensità media A	Limite superiore L	Rapporto di concentrazione secondo il metodo originario $R_o = \Delta_R : 2A$	metodo generalizzato $R_g = \Delta_R : 2A \frac{L-A}{L}$
1. Natalità	30.23 ‰	1000	11.8 % (7)	12.2 % (7)
2. Mortalità	17.54 ‰	1000	8.4 % (8)	8.6 % (8)
3. Accrescimento naturale.	12.70 ‰	1000	22.2 % (4)	22.5 % (6)
4. Nuzialità	10.45 ‰	1000	6.2 % (9)	6.3 % (9)
5. Percentuale della popolazione agglomerata.	73.25 %	100	14.6 % (6)	54.6 % (2)
6. Percentuale della popolazione agricola maschile (*).	47,36 %	100	21,1 % (5)	40.1 % (3)
7. Altitudine sul livello del mare (**).	171.35 metri	—	59.0 % (1)	59.0 % (1)
8. Reddito medio (***).	4152.62 lire	—	27.2 % (3)	27.2 % (5)
9. Densità della popolazione.	129.59 ab. per km.^{2}	—	39.2 % (2)	39.2 % (4)

(*) Su 100 maschi in età superiore ai 10 anni.

(**) Media ponderata delle altitudini dei capoluoghi di circondario secondo la popolazione dei circondari.

(***) Redditi di ricchezza mobile, Categorie B e C.

Come si vede, i valori del rapporto di concentrazione per i 4 primi caratteri subiscono lievi spostamenti, ma per i due successivi risultano invece modificati essenzialmente.

Nelle due graduatorie, solo 4 caratteri su 9 occupano lo stesso posto. La modificazione più forte, sia nel valore del rapporto di concentrazione, sia nel posto della graduatoria, viene naturalmente presentata dalla percentuale della popolazione agglomerata, la cui media (73.25 %) più si avvicina al limite superiore (100).

10. Possiamo ora riprendere in esame le ipotesi *a*) e *b*) considerate al n. 2 e vedere di quanto differirebbero i risultati qualora dette ipotesi non si verificassero.

Consideriamo anzitutto il caso in cui non si verifichi l'ipotesi *b*) vale a dire il caso in cui l'ammontare totale del carattere $T = nA$ non sia un multiplo esatto del limite superiore, in modo che esso non possa distribuirsi esattamente su un numero di termini che presentino il valore massimo.

Se k sono i termini in cui il valore è eguale al limite superiore, anzi che essere $nA = kL$, come abbiamo supposto per giungere alle formule (2) e (2 bis), sarà pertanto

$$nA = kL + R.$$

dove è $R < L$.

È facile mostrare che gli indici di variabilità assumono in tal caso, nell'ipotesi di disuguaglianza massima, i valori seguenti:

$$\text{Max } {}^{1}S_A = \left\{ \begin{array}{l} 2\,A\,\dfrac{n-k}{n} - \dfrac{2\,A}{n} \\[2ex] 2\,A\,\dfrac{n-k}{n} - \dfrac{2\,R}{n} \end{array} \right\} \text{ a seconda che è } R \gtreqless A$$

$$\text{Max } {}^{1}S_M = \left\{ \begin{array}{l} A\,\dfrac{n-k}{k} - \dfrac{R}{k} \\[2ex] A \\[2ex] A - \dfrac{R}{n} \end{array} \right\} \text{ a seconda che è } \left\{ \begin{array}{l} M = L \\[2ex] M = 0 \\[2ex] M = R \text{ (*)} \end{array} \right.$$

$$\text{Max } {}^{2}S_A = \sqrt{A^2\,\frac{n-k}{k} - \frac{R}{n}\,(L-R) - A\,\frac{R}{k}}$$

$$\text{Max } {}^{2}S_M = \left\{ \begin{array}{l} \sqrt{A^2\,\dfrac{n}{k}\;\dfrac{n-k}{k} - \dfrac{R}{n}(L-R) - \dfrac{R}{k}\,\dfrac{A\,(2n-k)-R}{k}} \\[2ex] \sqrt{A^2\,\dfrac{n}{k} - \dfrac{R}{n}\,(L-R) - A\,\dfrac{R}{k}} \\[2ex] \sqrt{A^2\,\dfrac{n}{k} - \dfrac{R}{n}\,(L-R) - R\,(2\,A-R) - A\,\dfrac{R}{k}} \end{array} \right\} \text{ a seconda che è } \left\{ \begin{array}{l} M = L \\[2ex] M = 0 \\[2ex] M = R \end{array} \right. \qquad (3)$$

$$\text{Max } \Delta_R \text{ (**)} = 2\,A\,\frac{n-k}{n} - \frac{2\,R}{n^2}\,(k+1)$$

(*) Si tenga presente che in questo caso è $n = 2k + 1$.

(**) Ritengo inutile dare la formula per Max Δ che si deduce immediatamente da quella per Max Δ_R tenendo presente che è

$$\text{Max } \Delta = \frac{n}{n-1}\,\text{Max } \Delta_R$$

che si possono anche mettere sotto la forma seguente:

$$\left.\begin{array}{l}
\text{Max } {}^{1}S_A = \left\{\begin{array}{l} 2\dfrac{A\,(L-A)}{L} - \dfrac{2\,A}{n}\,\dfrac{L-R}{L} \\[2ex] 2\dfrac{A\,(L-A)}{L} - \dfrac{2\,R}{n}\,\dfrac{L-A}{L} \end{array}\right\} \begin{array}{l}\text{a seconda} \\ \text{che è } R \gtreqless A\end{array} \\[4ex]
\text{Max } {}^{1}S_M = \left\{\begin{array}{l} L-A \\ A \\ A-\dfrac{R}{n} \end{array}\right\} \text{ a seconda che è } \left\{\begin{array}{l} M=L \\ M=0 \\ M=R \end{array}\right. \\[4ex]
\text{Max } {}^{2}S_A = \sqrt{A\,(L-A) - \dfrac{R}{n}\,(L-R)} \\[4ex]
\text{Max } {}^{2}S_M = \left\{\begin{array}{l} \sqrt{L\,(L-A) - \dfrac{R}{n}\,(L-R)} \\[2ex] \sqrt{A\,L - \dfrac{R}{n}\,(L-R)} \\[2ex] \sqrt{A\,L - \dfrac{R}{n}\,(L-R) - R\,(2\,A-R)} \end{array}\right\} \text{a seconda che è} \left\{\begin{array}{l} M=L \\[2ex] M=0 \\[2ex] M=R \end{array}\right. \\[4ex]
\text{Max } \Delta_R\ (*) = 2\,\dfrac{A\,(L-A)}{L} - 2\,\dfrac{R}{n^2}\,\dfrac{L-R}{L}
\end{array}\right\} (3\ \text{bis})$$

Confrontando queste formule con le (2) e le (2 bis) ci si persuade facilmente che, quando non vi è coincidenza, le (3) e (3 bis) si possono sempre scindere in due o più termini, di cui il primo, positivo, coincide coi valori dati dalle (2) e dalle (2 bis) e gli altri sono sempre negativi (**).

(*) Vedi nota (**) pag. precedente.

(**) Ciò può non riuscire evidente a tutta prima per quel che riguarda il valore di ${}^{2}S_M$, nel caso particolare in cui sia $M = R$. Conviene per ciò dimostrare che il termine $-R\,(2\,A - R)$ è sempre negativo, cioè che è $2\,A - R > 0$.

Si avverta perciò che è, nel caso particolare in esame, $n = 2k + 1$. Ora, posto $\varepsilon = \dfrac{L-R}{L}$ si trova immediatamente

$$\frac{R}{2} = \frac{L}{2} - \frac{\varepsilon\,L}{2} \qquad \text{(a)}$$

e d'altra parte, dalle

$$A\,n = k\,L + R$$
$$A\,(2\,k + 1) = L\,(k + 1 - \varepsilon)$$

È anche evidente che il valore di questi termini negativi tende a zero quando n tende all'∞, ossia che il valore dei termini stessi diviene praticamente trascurabile per n abbastanza grande.

È infine degno di nota che le differenze fra le (3) e le (2) sono in ogni caso maggiori o non minori che fra le (3 bis) e le (2 bis).

Possiamo dunque concludere che il fare l'ipotesi b) quando essa non risponde al vero, non porta nella pratica, quando il numero dei casi sia sufficientemente grande, a inesattezze notevoli. Quando i risultati ottenuti non sono esatti l'approssimazione ottenuta è in tal caso maggiore, o quanto meno non minore, con le formule (2 bis) che con le formule (2) (*). Sia con le (2 bis) che con le (2), l'approssimazione è sempre in eccesso.

---

si ricava facilmente

$$A = \frac{L}{2} + \frac{L}{2(2k+1)} - \frac{\varepsilon L}{2k+1} \qquad \text{(b)}$$

Essendo sempre $k \geq 1$, è pure sempre

$$\frac{L}{2(2k+1)} > 0 \qquad \frac{\varepsilon L}{2k+1} < \frac{\varepsilon L}{2}$$

e quindi

$$A > \frac{R}{2} \qquad \text{e} \qquad 2A - R > 0 \qquad \text{c. d. d.}$$

(*) Applicare, quando si verifica un resto R, le formule (2), in cui k indica il numero dei termini il cui valore raggiunge, nella seriazione massimante, il limite superiore L, equivale a supporre che sia $R = 0$.

Se invece si supponesse $R = L$, ciò porterebbe a sostituire, nelle (2), al valore di k il valore di $k + 1$, ove $k + 1$ indica il numero dei termini a cui, nella seriazione massimante, non viene attribuito il valore 0.

In quest'ultimo caso, i valori ottenuti dalle (2) risulterebbero uguali a quelli ottenuti dalle (3) per ${}^{1}S_A$, quando è $R > A$, oltre che per ${}^{1}S_M$ quando è $M = 0$ essi risulterebbero invece diversi da quelli ottenuti dalle (3) sia per ${}^{1}S_A$ e ${}^{1}S_M$ negli altri casi, sia per gli altri indici semplici o quadratici, la diversità dipendendo essenzialmente dalla differenza fra R ed A.

Se a k, anzichè un valore intero, si attribuisce il valore, di solito frazionario, $\frac{A\,n}{L}$ le (2) vengono a coincidere con le (2 bis). Le differenze fra i valori delle (2 bis) e quelli delle (2) saranno quindi tanto più sensibili quanto più è forte la disuguaglianza $\frac{A\,n}{L} > k$, ossia quanto più è prossimo all'unità il rapporto $\frac{R}{L}$; e viceversa le differenze fra i valori delle (2 bis) e quelli che si ottengono dalle (2) sostituendo, a k, $k + 1$, saranno tanto più sensibili quanto più è forte la disuguaglianza $k + 1 > \frac{A\,n}{L}$ ossia quanto più è prossimo a 0 il rapporto $\frac{R}{L}$.

Si noti come, nel caso che si verifichi l'ipotesi *b*) e sia quindi $R = 0$, le (3) e (3 bis) si riducono rispettivamente alle (2) e (2 bis).

Nel caso, invece, in cui la intensità del carattere ammette un limite superiore L, ma questo non impedisce che l'intensità tutta del carattere $T = n\,A$ possa essere concentrata in un solo termine, essendo $T < L$, sarà $k = 0$ e $R = T = n\,A$; e a questo caso, ponendo $L = \infty$, può ricondursi il caso in cui le intensità del carattere non ammettano un limite superiore. Ora, in tal caso, le (3 bis) si riducono alle (1), e alle (1) si riducono pure le (3) relative agli indici lineari ${}^1S_A$, ${}^1S_M$..... Δ_R mentre, per gli indici quadratici ${}^2S_A$, ${}^2S_M$ le (3) perdono significato.

Si conclude quindi che le formule (3 bis) debbono riguardarsi come assolutamente generali per il calcolo dei massimi degli indici di variabilità. Esse valgono infatti, per gli indici di variabilità considerati, sia che esista, sia che non esista un limite superiore dell'intensità, sia che la seriazione massimante dia luogo, sia che non dia luogo a un resto. Altrettanto si può dire per le formule (3), ma limitatamente agli indici lineari considerati, mentre esse non servono per determinare gli indici quadratici quando non esista un limite superiore o questo sia più elevato dell'ammontare totale del carattere.

11. Consideriamo ora il caso in cui non si faccia ricorso neppure all'ipotesi *a*), vale a dire il caso in cui si tenga conto del diverso peso dei termini nella determinazione del massimo degli indici di variabilità.

Si noti, a questo proposito, che per lo più gli indici di variabilità vengono calcolati attribuendo ai varî termini uno stesso peso, anche se da un certo punto di vista il peso possa essere riguardato diverso. Ciò può essere fatto per due considerazioni, o perchè in realtà il peso può essere considerato uguale dallo speciale punto di vista che interessa la ricerca; o perchè, pure dovendo esso a rigore essere considerato come diverso, il tener conto della differenza di peso riuscirebbe malagevole per la maggiore lunghezza dei calcoli, o addirittura impossibile.

Si consideri, per esempio, la variabilità dei coefficienti annuali di mortalità in uno Stato. La popolazione dello Stato è diversa negli anni successivi e in questo senso i coefficienti annuali di mortalità hanno un peso differente, ma, se ciascun coefficiente si considera come indice delle condizioni igieniche presentate in un dato intervallo di tempo, i varî coefficienti devono considerarsi come di uguale peso in quanto si riferiscono a intervalli di tempo uguali.

Si voglia invece determinare la variabilità dei prezzi di una data merce da regione a regione ; qui in realtà al prezzo di una regione dovrebbe attribuirsi un peso proporzionale al volume delle contrattazioni eseguite ; ma, se questo volume non si conosce, forza è attribuire ad ogni termine lo stesso peso.

Ancora, si voglia misurare la variabilità della ricchezza media degli abitanti da provincia a provincia. In tal caso converrebbe a rigore attribuire al dato di ogni provincia un peso proporzionale alla popolazione rispettiva ; ma, molte volte, per brevità, si attribuisce a tutti i dati provinciali lo stesso peso.

Quando ai termini della serie si attribuisca, nella determinazone dell'indice di variabilità, lo stesso peso, si intende che lo stesso peso deve pure attribuirsi ad essi nel determinare il massimo valore che l'indice può raggiungere.

L'eliminazione dell'ipotesi *a*) nella determinazione dei valori massimi dovrà pertanto essere presa in considerazione nei soli casi, che costituiscono per vero l'eccezione, in cui l'indice di variabilità venga determinato tenendo conto del diverso peso dei termini.

In tal caso, anzichè misurare gli scostamenti dalla media semplice A o dalla mediana semplice M, conviene misurarli dalla media ponderata o dalla mediana ponderata, che indicheremo rispettivamente con A' ed M'.

Se rappresentiamo, al solito, con L il limite superiore del valore dei termini e indichiamo con N la somma dei pesi degli n termini, con N_k la somma dei pesi dei k termini che, nell'ipotesi di massimo, raggiungerebbero il limite superiore, con N_h la somma dei pesi degli h termini, che in tale ipotesi presenterebbero il valore 0, si avrà, nel caso che si verifichi l'ipotesi *b*),

$$A' N = L N_k \quad ; \quad h = n - k$$

e nel caso che l'ipotesi *b*) non si verifichi,

$$A' N = L N_k + R \quad ; \quad h = n - k - 1$$

Indichiamo, in questo secondo caso, con D il valore e con N_1 il peso del termine residuo : sarà $D = \frac{R}{N_1}$.

Le formule (4), o sotto altra forma le (4 bis), dànno i valori massimi che assumono gli indici di variabilità nel caso che l'ipotesi *b*) non si verifichi. Facendo $D = 0$, (il che implica che sia $D < A'$ e che M' possa acquistare i soli valori L o 0) si ottengono dalle (4) o (4 bis) le rispettive formule, che diremo (5) e (5 bis), nel caso che l'ipotesi *b*) sia verificata. Le (5) e (5 bis) si riducono rispettivamente alle (2) e (2 bis), e le (4), (4 bis) rispettivamente alle (3), (3 bis), sostituendo ad A' e M' i simboli A ed M e a D, N, N_1, ed N_k i simboli R, $n N_1$, $\frac{N}{n}$ e $k N_1$ e ricordando che, per $M = R$, è $n = 2k + 1$.

$$
\left.
\begin{aligned}
&\text{Max } {}^1S_{A'} = \begin{cases} 2A' \dfrac{N - N_k}{N} - 2A' \dfrac{N_1}{N} \\[2ex] 2A' \dfrac{N - N_k}{N} - 2D \dfrac{N_1}{N} \end{cases} \text{a seconda che è } D \gtreqless A' \\[3ex]
&\text{Max } {}^1S_{M'} = \begin{cases} A' \dfrac{N - N_k}{N_k} - D \dfrac{N_1}{N_k} \\[2ex] A' \\[2ex] A' - D \dfrac{2N_k + 2N_1 - N}{N} \end{cases} \begin{matrix} \text{a seconda} \\ \text{che è} \end{matrix} \begin{cases} M' = L \\[2ex] M' = 0 \\[2ex] M' = D \end{cases} \\[3ex]
&\text{Max } {}^2S_{A'} = \sqrt{A'^2 \frac{N - N_k}{N_k} - D(L - D)\frac{N_1}{N} - A' D \frac{N_1}{N_k}} \\[3ex]
&\text{Max } {}^2S_{M'} = \\
&\begin{cases} \sqrt{A'^2 \dfrac{N}{N_k} \dfrac{N - N_k}{N_k} - \dfrac{N_1}{N} D(L - D) - \dfrac{N_1}{N_k} D \dfrac{A'(2N - N_k) - D N_1}{N_k}} \\[3ex] \sqrt{A'^2 \dfrac{N}{N_k} - \dfrac{N_1}{N} D(L - D) - A' D \dfrac{N_1}{N_k}} \\[3ex] \sqrt{A'^2 \dfrac{N}{N_k} - \dfrac{N_1}{N} D(L - D) - D(2A' - D) - A' D \dfrac{N_1}{N_k}} \end{cases} \text{a seconda che è} \begin{cases} M' = L \\[3ex] M' = 0 \\[3ex] M' = D \end{cases} \\[3ex]
&\text{Max } \Delta_R \text{ (*)} = 2A' \frac{N - N_k}{N} - 2D \frac{N_1}{N^2}(N_k + N_1)
\end{aligned}
\right\} (4)
$$

(*) Per il valore di **Max** Δ, vedi nota (**) pag. 13.

$$
\left.\begin{array}{l}
\text{Max } {}^{1}S_{A'} = \begin{cases} 2\,\dfrac{A'\,(L-A')}{L} - 2\,\dfrac{N_1}{N}\,\dfrac{A'\,(L-D)}{L} \\[2ex] 2\,\dfrac{A'\,(L-A')}{L} - 2\,\dfrac{N_1}{N}\,\dfrac{D\,(L-A')}{L} \end{cases} \text{a seconda che è } D \gtreqless A' \\[6ex]
\text{Max } {}^{1}S_{M'} = \left.\begin{cases} L - A' \\ A' \\ A' - D\left(\dfrac{2\,A'-L}{L} - 2\,\dfrac{N_1}{N}\,\dfrac{L-D}{L}\right) \end{cases}\right\} \text{a seconda che è } \begin{cases} M' = L \\ M' = 0 \\ M' = D \end{cases} \\[6ex]
\text{Max } {}^{2}S_{A'} = \sqrt{A'\,(L-A') - \dfrac{N_1}{N}\,D\,(L-D)} \\[4ex]
\text{Max } {}^{2}S_{M'} = \\
\left.\begin{cases} \sqrt{L\,(L-A') - \dfrac{N_1}{N}\,D\,(L-D)} \\[2ex] \sqrt{A'\,L - \dfrac{N_1}{N}\,D\,(L-D)} \\[2ex] \sqrt{A'\,L - \dfrac{N_1}{N}\,D\,(L-D) - D\,(2\,A'-D)} \end{cases}\right\} \text{a seconda che è } \begin{cases} M' = L \\ M' = 0 \\ M' = D \end{cases} \\[8ex]
\text{Max } \Delta_R\,(*) = 2\,\dfrac{A'\,(L-A')}{L} - 2\,\dfrac{N_1^2}{N^2}\,\dfrac{D\,(L-D)}{L}
\end{array}\right\} \quad (4 \text{ bis})
$$

E per $D = 0$:

$$
\left.\begin{array}{l}
\text{Max } {}^{1}S_{A'} = 2\,A'\,\dfrac{N-N_k}{N} \\[4ex]
\text{Max } {}^{1}S_{M'} = \left.\begin{cases} A'\,\dfrac{N-N_k}{N_k} \\[2ex] A' \end{cases}\right\} \text{a seconda che è } \begin{cases} M' = L \\ M' = 0 \end{cases} \\[6ex]
\text{Max } {}^{2}S_{A'} = A'\,\sqrt{\dfrac{N-N_k}{N_k}}
\end{array}\right\} \quad (5)
$$

(*) Per il valore di Max Δ, vedi nota (**) pag. 13.

$$\left.\begin{array}{l} \text{Max } {}^2S_{M'} = \begin{cases} \dfrac{A'}{N_k}\sqrt{N\,(N-N_k)} \\[2ex] A'\sqrt{\dfrac{N}{N_k}} \end{cases} \quad \text{a seconda che è} \quad \begin{cases} M' = L \\[2ex] M' = 0 \end{cases} \\[6ex] \text{Max } \Delta_R = 2\,A'\,\dfrac{N-N_k}{N} \end{array}\right\} (5)$$

$$\left.\begin{array}{l} \text{Max } {}^1S_{A'} = \dfrac{2\,A'\,(L-A')}{L} \\[3ex] \text{Max } {}^1S_{M'} = \begin{cases} L - A' \\ A' \end{cases} \quad \text{a seconda che è} \quad \begin{cases} M' = L \\ M' = 0 \end{cases} \\[3ex] \text{Max } {}^2S_{A'} = \sqrt{A'\,(L-A')} \\[3ex] \text{Max } {}^2S_{M'} = \begin{cases} \sqrt{L\,(L-A')} \\ \sqrt{A'\,L} \end{cases} \quad \text{a seconda che è} \quad \begin{cases} M' = L \\ M' = 0 \end{cases} \\[3ex] \text{Max } \Delta_R = 2\,\dfrac{A'\,(L-A')}{L} \end{array}\right\} (5 \text{ bis})$$

Si noti che le formule (5) e (5 bis), quando non coincidono con le (4) e (4 bis), ne differiscono per dei termini in cui entra come fattore D. Ora ci si persuade facilmente che il valore di tali termini è sempre negativo (*). Anche in questo caso il supporre verificata l'ipotesi *b*),

---

(*) Anche in questo caso è opportuno dimostrare che, per $M' = D$, è $2\,A' - D > 0$.

Si consideri per ciò che, posto $\varepsilon = 1 - \dfrac{D}{L}$, è

$$\frac{D}{2} = \frac{L}{2} - \frac{\varepsilon L}{2}$$

Ricordando, d'altra parte, che, per $M' = D$, è $N_k + N_1 > \dfrac{N}{2}$, quando nel fatto non lo è, conduce quindi ad innalzare il valore massimo degli indici di variabilità.

Resta eccettuato naturalmente, per ${}^{1}S_{M'}$, il caso in cui è $M' = 0$, nell'applicazione delle (5), e quelli in cui è $M' = L$ e $M' = 0$, nell'applicazione delle (5 bis), restando in tali casi inalterato il valore massimo degli indici.

Ci si persuade pure facilmente che, quando le (5) differiscono dalle (4), le differenze sono sempre maggiori, o non minori, delle differenze corrispondenti fra le (5 bis) e le (4 bis). L'applicazione delle (5 bis), in luogo delle (4) o (4 bis), quando non si verifica l'ipotesi *b*), porta dunque a risultati più approssimati dell'applicazione delle (5) (*).

---

si trova
$$A' = \frac{L\,N_k + D\,N_1}{N} > \frac{L\,N_k + D\,N_1}{2(N_k + N_1)}$$

e poichè è

$$\frac{L\,N_k + D\,N_1}{2(N_k + N_1)} = \frac{L}{2} - \frac{\varepsilon\,L\,N_1}{2(N_k + N_1)} > \frac{L}{2} - \frac{\varepsilon\,L}{2},$$

sarà, a più forte ragione,

$$A' > \frac{D}{2} \text{ e quindi } 2A' - D > 0 \qquad \text{c. d. d.}$$

(*) Se, nelle (5), al valore di $\frac{N_k}{N}$ si sostituisce il valore $\frac{A'}{L}$, i risultati ottenuti dalle (5) coincidono con quelli ricavati dalle (5 bis).

Se, infine, nelle (5), al valore di $\frac{N_k}{N}$ si sostituisce il valore di $\frac{N_k + N_1}{N}$, i risultati ottenuti con le (5) vengono a coincidere con quelli ottenuti con le (4) o (4 bis) per ${}^{1}S_{A'}$ quando sia $D > A'$ (oltre che per ${}^{1}S_{M'}$ quando sia $M' = 0$), mentre ne divergono in ogni altro caso, per differenze che dipendono essenzialmente dalla differenza fra D e A'.

Se ne deduce che le differenze fra i risultati ottenuti con le (5 bis) e le (5) sono tanto più sensibili quanto più è forte la disuguaglianza $\frac{A'}{L} > \frac{N_k}{N}$, ossia quanto più è alto $\frac{N_1}{N}$ e, a parità di $\frac{N_1}{N}$, quanto più $\frac{D}{L}$ si avvicina all'unità; mentre le differenze fra i risultati ottenuti con le (5 bis) e quelli ottenuti con le (5) sostituendo a $\frac{N_k}{N}$ il valore $\frac{N_k + N_1}{N}$ sono tanto più sensibili quanto

12. Una difficoltà si presenta nell'applicazione pratica delle formule suesposte. I risultati, a cui esse conducono, differiscono a seconda dei valori di N_1 e di D, e questi, a loro volta, variano a seconda dei termini nei quali si suppone sia raggiunto il limite superiore L e del termine che si considera come residuo. Fra le varie distribuzioni possibili, è da intendersi che sia scelta, per l'applicazione delle formule, quella che fa acquistare alle formule stesse i valori massimi (*distribuzione massimante*). Il prof. Galvani, che ha fatto di questo argomento speciale studio, conclude che è difficile dare criterî generali per la determinazione a priori della distribuzione massimante, salvo che in alcuni casi più semplici, che sono — del resto — quelli che più frequentemente si presentano di fatto.

In generale per ottenere la distribuzione massimante bisogna — secondo le sue conclusioni — seguire il criterio di attribuire il valore L ad un termine o a un complesso di termini di peso complessivo massimo e di attribuire l'ammontare residuo R del carattere al termine di peso minimo. Perciò :

a) Se non è possibile attribuire il valore L neanche ad un solo termine, l'ammontare totale del carattere si dovrà attribuire al termine di peso minimo fra tutti i dati, affinchè il corrispondente quoziente D abbia il valore massimo.

Per es., fra le applicazioni che seguono poco oltre, per la nuzialità generica la distribuzione $401.580.000 = 1000 \times 0 + 857{,}057 \times 468.557$ (468.557 peso minimo) è la massimante.

b) Quando è possibile attribuire il valore L ad un termine, ma a non più di uno, possono darsi due casi : o il termine di maggior peso tra quelli a cui può attribuirsi L non è il termine che ha, fra tutti, peso minimo, oppure è quello che ha il peso minimo fra tutti i termini.

Nella prima eventualità, attribuito L a quel certo termine, ed attribuito l'ammontare residuo del carattere al termine di peso minimo fra tutti i dati si avrà senz'altro la distribuzione massimante. Per es., fra le stesse applicazioni, per la nuzialità specifica, la distribuzione : $401.582.000 = 1000 \times 334.856 + 957{,}867 \times 69.661$ in cui 69.661 è il peso minimo, è la massimante.

---

più è forte la disuguaglianza $\frac{N_k + N_1}{N} > \frac{A'}{L}$, ossia quanto più è alto $\frac{N_1}{N}$ e, a parità di $\frac{N_1}{N}$, quanto più $\frac{D}{L}$ si avvicina allo *0*.

Nell'altra eventualità, attribuito il valore L al termine di peso minimo fra tutti i dati ed il residuo al termine di peso N_1, immediatamente superiore, potrà darsi che la distribuzione ottenuta non sia la massimante : per decidere si dovrà anche considerare la distribuzione ottenuta coll'attribuire l'ammontare totale del carattere $A' N$ a quel solo termine di peso N_1 tale che il corrispondente quoziente D abbia il valore massimo compatibile con la condizione di non superare L. Una fra queste due distribuzioni è la massimante, e potrà facilmente essere individuata.

Per esempio per la mortalità generale, considerata nelle applicazioni che seguono, la distribuzione :

$$652.712.000 = 1000 \times 468.557 + 249,509 \times 738.070$$

(468.557 peso minimo) non è la massimante, e lo è, invece, la

$$652.712.000 = 1000 \times 0 + 884,350 \times 738.070.$$

c) Infine, se il valore L si può attribuire ad un complesso di termini, è necessario prescegliere fra i diversi complessi quello in cui la somma N_k risulta massima, a condizione che rimanga ancora disponibile il termine di peso minimo fra tutti i dati per attribuire a tale termine il resto R.

Così per la natalità specifica, pure considerata fra le applicazioni che seguono, la distribuzione $1.127.834.000 = 1000\ (896.490 + 160.523) + 712,008 \times 103.688$ (103.688 peso minimo) è la massimante.

13. Le formule (4) e (4 bis) del pari che le (5) e (5 bis) sono valide solo per $k \geqq 1$. Nel caso che sia $k = 0$ e quindi $N_k = 0$ e $A'N = R$, in luogo delle (4) e (4 bis) devono adoperarsi le formule (6).

$$\left.\begin{aligned} &\text{Max } {}^1S_{A'} = 2\,A' \frac{N - N_1}{N} \text{ essendo sempre } D > A' \\ &\text{Max } {}^1S_{M'} = \begin{cases} A' \\ D - A' = A' \dfrac{N - N_1}{N_1} = D \dfrac{N - N_1}{N} \end{cases} \begin{matrix} \text{a se-} \\ \text{conda} \\ \text{che sia} \\ (*) \end{matrix} \begin{cases} M' = 0 \\ M' = D \end{cases} \end{aligned}\right\} (6)$$

(*) L'ipotesi $M' = L$ è esclusa a priori, essendo $k = 0$. L'ipotesi $M' = D$ corrisponde al caso in cui sia $N_1 > N - N_1$ e $2\,A' > L$. Nel caso in cui sia

$$\left.\begin{array}{l}
\text{Max } {}^2S_{A'} = \sqrt{A'(D-A')} = A'\sqrt{\dfrac{N-N_1}{N_1}} \\[3ex]
\text{Max } {}^2S_{M'} = \left\{\begin{array}{l} \sqrt{A'D} = A'\sqrt{\dfrac{N}{N_1}} \\[2ex] \sqrt{D(D-A')} = A'\sqrt{\dfrac{N}{N_1}\,\dfrac{N-N_1}{N_1}} \end{array}\right\} \begin{array}{c}\text{a se-}\\ \text{conda}\\ \text{che sia}\\ (*)\end{array} \left\{\begin{array}{l} M' = 0 \\[2ex] M' = D \end{array}\right. \\[3ex]
\text{Max } \Delta_R = 2A'\,\dfrac{N-N_1}{N}
\end{array}\right\} (6)$$

Queste formule vengono a coincidere con le (1) facendo $A' = A$, $N_1 = \frac{N}{n}$. Si noti anche che le formule (6), quando sia $N_k = 0$ e quindi $A'N = DN_1$ vengono a coincidere sempre con le (4 bis) e per Max ^1S_A, Max ^1S_M e Max Δ_R, anche con le (4). Le formule (4 bis) in ogni caso e le formule (4) per i massimi degli indici lineari hanno dunque valore assolutamente generale, mentre altrettanto non si può dire delle stesse formule (4) per i massimi degli indici quadratici di variabilità, in quanto dette formule (4) non sono applicabili a questi indici quando sia $N_k = 0$.

14. Il prof. Galvani ha fatto applicare le formule sopra indicate alla determinazione dei massimi degli indici di variabilità, non ponderati e ponderati, per alcuni caratteri: natalità generica, nuzialità generica, mortalità generale, natalità specifica, nuzialità specifica, inerenti ai diversi Compartimenti del Regno d'Italia entro i vecchi confini, in base ai numeri delle nascite, dei matrimoni e delle morti del quadriennio 1920-23 ed alla popolazione censita al 1° dicembre 1921.

I valori della natalità e della nuzialità generica e quelli della mortalità generale per i singoli Compartimenti, nonchè le medie

---

$N_1 > N - N_1$, ma $2A' < L$, la distribuzione massimante non si ottiene riversando l'ammontare totale del carattere sul termine a peso massimo che comprende N_1 casi, ma sopra altri termini, applicando, secondo i casi, o le stesse formule (6), o le formule (5), (5 bis) o le (4), (4 bis).

(*) Vedi nota pag. precedente.

aritmetiche e le mediane, semplici e ponderate, risultano dal seguente prospetto :

Prospetto I.

Compartimenti	Popolazione	Natalità generica ‰	Nuzialità generica ‰	Mortalità generale ‰
Piemonte	3.383.646	19,936	11,575	15,853
Liguria	1.326.904	20,331	9,459	14,947
Lombardia	5.103.329	27,681	11,146	17,719
Veneto	3.957.335	34,849	10,845	16,188
Emilia	2.954.687	30,234	10,034	16,188
Toscana	2.830.184	27,483	11,725	15,770
Marche	1.148.296	33,016	11,467	17,843
Umbria	738.070	32,878	11,918	17,737
Lazio	1.517.292	30,903	10,566	16,889
Abruzzi	1.432.940	35,415	11,723	19,680
Campania	3.546.641	34,266	10,684	19,381
Puglie	2.297.061	37,971	10,632	21.046
Basilicata	468.557	39,139	12,649	22,228
Calabrie	1.512.318	36,748	11,308	18,892
Sicilia	4.061.452	28,704	9,551	17,663
Sardegna	864.174	32,466	9,617	19,772
Regno (vecchi confini) .	37.142.886	30,365	10,812	17,573
A = Medie aritmetiche (non ponderate) :		31,376	10,931	17,987
M = Mediane (non ponderate) . .		32,672	10,996	17,728
A' = Medie aritmetiche ponderate .		30,365	10,812	17,573
M' = Mediane ponderate		30,324	10,845	17,663

I valori della natalità specifica sono ricavati riferendo la media dei nati nel quadriennio 1920-23 alle femmine in età feconda (da 15 a 45 anni), il cui numero è ottenuto aggiungendo a quello delle donne di età nota entro questi limiti una quota proporzionale delle donne di età ignota. Si ha così:

Prospetto II.

Compartimenti	Femmine in età di 15-45 anni	Natalità specifica ‰
Piemonte	824.610	81,806
Liguria	329.944	81,762
Lombardia	1.245.003	113,466
Veneto	896.490	153,832
Emilia	658.119	135,740
Toscana	661.749	117,540
Marche	260.977	145,270
Umbria	160.523	151,168
Lazio	354.295	132,345
Abruzzi	323.325	156,954
Campania	799.489	152,007
Puglie	511.129	170,644
Basilicata	103.688	176,867
Calabrie	340.171	163,371
Sicilia	921.210	126,552
Sardegna	196.857	142,520
Regno (vecchi confini). .	8.587.579	131,333
A = Media aritmetica (non ponderata)		137,615
M = Mediana (non ponderata) . . .		143,895
A' = Media aritmetica ponderata .		131,333
M' = Mediana ponderata		132,345

Infine i valori della nuzialità specifica si hanno mediante riferimento della media dei matrimonî nel solito quadriennio 1920-23 alle donne coniugabili, cioè al numero delle donne nubili aumentato di quello delle vedove e di una quota proporzionale delle donne di stato civile ignoto :

Prospetto III.

Compartimenti	Donne coniugabili	Nuzialità specifica ‰
Piemonte	667.132	58,710
Liguria	260.520	48,177
Lombardia.	920.852	61,771
Veneto	675.235	63,560
Emilia	480.708	61,299
Toscana	467.696	70,950
Marche	187.225	70,332
Umbria	106.625	82,495
Lazio	253.230	63,306
Abruzzi	219.286	76,603
Campania	593.558	63,837
Puglie	334.856	72,933
Basilicata	69.661	85,083
Calabrie	241.398	70,842
Sicilia.	627.233	61,841
Sardegna	149.522	55,584
Regno (vecchi confini) .	6.254.737	64,204
A = Media aritmetica (non ponderata)		66,708
M = Mediana (non ponderata) . .		63,699
A' = Media aritmetica ponderata . .		64,204
M' = Mediana ponderata		63,306

Gli indici di variabilità assoluti, semplici e ponderati, relativi alle cinque serie considerate, hanno i seguenti valori :

Prospetto IV.

Caratteri	Non ponderati					Ponderati				
	${}^{1}S_{A}$	${}^{2}S_{A}$	${}^{1}S_{M}$	${}^{2}S_{M}$	Δ_{R}	${}^{1}S_{A'}$	${}^{2}S_{A'}$	${}^{1}S_{M'}$	${}^{2}S_{M'}$	Δ_{R}
Natalità generica	4,295	5,396	4,159	5,519	5,928	4,178	5,165	4,170	5,166	5,708
Nuzialità generica	0,758	0,900	0,758	0,902	1,019	0,619	0,762	0,615	0,763	0,856
Mortalità generale	1,634	1,981	1,585	1,998	2,229	1,347	1.683	1,335	1,685	1,858
Natalità specifica	21,762	27,134	21,149	27,851	29,871	21,653	26,192	21,580	26,211	29,104
Nuzialità specifica	7,786	9,387	7,427	8,858	10,434	4,[illegible]09	6,642	4,534	6,699	6,790

Volendo ora determinare i massimi degli indici di variabilità per ciascuno dei cinque caratteri esaminati, converrà anzitutto decidere quali siano, caso per caso, le formule da applicare, correttamente o in via approssimata, e ciò, separatamente, per i caratteri non ponderati e per i ponderati.

A tal fine si dovranno in primo luogo individuare le distribuzioni massimanti degli indici di variabilità.

Pei caratteri non ponderati, tenendo conto che il limite superiore è per definizione, sia pure astrattamente, uguale a 1000 per ciascuno di essi, le distribuzioni massimanti saranno:

Prospetto V.

Caratteri non ponderati	Distribuzione massimante
Natalità generica . . .	$502{,}020 = 1000 \times 0 + 502{,}020$
Nuzialità generica . . .	$174{,}899 = 1000 \times 0 + 174{,}899$
Mortalità generale . . .	$287{,}796 = 1000 \times 0 + 287{,}796$
Natalità specifica . . .	$2.201{,}844 = 1000 \times 2 + 201{,}844$
Nuzialità specifica . . .	$1.067{,}323 = 1000 \times 1 + 67{,}323$

Difatti, per la natalità generica, nuzialità generica e mortalità generale, l'ammontare totale T del carattere, risultando minore di 1000, può (astrattamente) concentrarsi in un solo elemento della serie, cioè in un solo compartimento. Ciò non è, invece, possibile nè per la natalità specifica, nè per la nuzialità specifica : per questi caratteri l'ammontare totale T consente il raggiungimento del limite superiore 1000 in 2 e rispettivamente in un circondario, restando però, in entrambi i casi, un resto da attribuirsi ad altro circondario. Pertanto, in via rigorosa, i massimi cercati si avranno per la natalità generica, nuzialità generica e mortalità generale con applicazione delle formule (1) — o delle (3 bis) che, come si è osservato, danno sempre risultati esatti — mentre per la natalità specifica, e per la nuzialità specifica le cui distribuzioni comportano un resto, si dovranno applicare le (3) o (3-bis). Nelle tabelle VII-XI tali massimi sono indicati in carattere grassetto.

Qualora si applichino formule diverse da quelle indicate, i relativi risultati costituiranno in generale dei valori approssimati dei massimi cercati, ma, in taluni casi, li rappresenteranno esattamente. Così, per la natalità generica, la nuzialità generica e la mortalità generale, se si applicano le (2) per $k = 1$ si avranno gli stessi risultati forniti dalle (1), cioè i valori esatti, poichè appunto le (2) vengono allora a coincidere con le (1). Applicando, invece, le (2) per $k = \frac{nA}{L}$, o, ciò che è lo stesso, le (2-bis) si otterranno valori approssimati dei massimi cercati. Similmente, per la natalità specifica e per la nuzialità specifica, l'applicazione delle (1) fornirà risultati generalmente approssimati per eccesso; le (2) per $k = \frac{nA}{L}$ e le (2-bis) forniranno altri valori approssimati e coincidenti fra di loro ; infine le (2), quando si attribuiscano a k i valori approssimati a meno di una unità per difetto e per eccesso del quoziente $\frac{nA}{L}$, forniranno ancora altri valori approssimati dei massimi che si vogliono determinare, a meno che — come avviene per la nuzialità specifica — il quoziente per difetto non sia uguale ad 1, nel qual caso si ha coincidenza colle formule (1).

Pei caratteri ponderati le distribuzioni massimanti degli indici di variabilità sono :

Prospetto VI.

Caratteri ponderati	Distribuzione massimante
Natalità generica	$1.127.834.000 = 1000 \times 864.174 + 562{,}706 \times 468.557$
Nuzialità generica	$401.580.000 = 1000 \times 0 + 857{,}057 \times 468.557$
Mortalità generale	$652.712.000 = 1000 \times 0 + 884{,}350 \times 738.070$
Natalità specifica	$1.127.834.000 = 1000 \times (896.490 + 160.523) + 712{,}008 \times 103.688$
Nuzialità specifica	$401.580.000 = 1000 \times 334.856 + 957{,}838 \times 69.661$

Si constata che, per la nuzialità generica e per la mortalità generale, l'ammontare totale $A'N$ del carattere può concentrarsi in un solo elemento della serie, e precisamente in quello di peso minimo: i massimi degli indici di variabilità saranno dunque forniti esattamente pei caratteri stessi dalle formule (6). Per la natalità generica e per la nuzialità specifica, il limite superiore L viene attribuito ad un termine (che non è quello di peso minimo), ed il resto al termine di peso minimo; sono pertanto da applicarsi, in via rigorosa, le formule (4-bis). Infine, per la natalità specifica, in cui il valore L si attribuisce a un complesso di termini, ed il resto al termine di peso minimo, sono pure da applicarsi le formule (4-bis). I massimi, correttamente calcolati, sono scritti nelle Tabelle VII-XI in carattere grassetto.

Anche qui l'applicazione di formule diverse da quelle specificatamente idonee in ogni caso, fornisce valori più o meno approssimati dei massimi cercati e talora anche dei valori esatti. Così, come si è già osservato, la applicazione della (4-bis) fornisce in tutti i casi i valori esatti dei massimi.

Per la natalità generica, per la natalità specifica e per la nuzialità specifica, le (5) e (5-bis) forniscono valori più o meno approssimati dei massimi che si vogliono determinare. Per la nuzialità generica e la mortalità generale, le (6) forniscono gli stessi valori delle (4-bis) mentre le (5) e (5-bis) non sono applicabili.

Ciò posto, ecco i prospetti che danno, per ciascun carattere, i massimi calcolati, sia correttamente che approssimativamente. I massimi calcolati correttamente sono indicati in grassetto.

Prospetto VII. — Natalità generica.

		Max ^1S_A	Max ^2S_A	Max ^1S_M	Max ^2S_M	Max Δ_R
Non ponder.	Applicaz. formule (1) o (3 bis). Applic. form. (2) per k = 1	**58,830**	**121,519**	**31,376**	**125,504**	**58,830**
	Applic. form. (2) per k = 0,502; Applicaz. form. (2 bis)	60,783	174,332	31,376	177,133	60,783
Ponder.	Applicaz. formule (4 bis)	**58,551**	**162,292**	**30,365**	**165,109**	**58,808**
	Applic. form. (5) per N_k = 864.174	59,317	196,743	30,365	199,072	59,317
	Applic. form. (5) per N_k = 1.332.731	58,551	157,400	30,365	160,302	58,551
	Applic. form. (5) per N_k = 1.127.834; Applicaz. formule (5 bis)	58,886	171,590	30,365	174,256	58,886
	Applicazione formule (6)	59,964	268,639	30,365	270,340	59,964
		Max $^1S_{A'}$	Max $^2S_{A'}$	Max $^1S_{M'}$	Max $^2S_{M'}$	Max Δ_R

Prospetto VIII. — Nuzialità generica.

		Max ^1S_A	Max ^2S_A	Max ^1S_M	Max ^2S_M	Max Δ_R
Non ponder.	Applicaz. formule (1) o (3 bis). Applic. form. (2) per k = 1	**20,496**	**42,336**	**10,931**	**43,724**	**20,496**
	Applic. form. (2) per k = 0,175; Applicaz. formule (2 bis)	21,623	103,978	10,931	104,551	21,623
Ponder.	Applicazione formule (6); Applicaz. formule (4 bis)	**21,351**	**95,654**	**10,812**	**96,263**	**21,351**
		Max $^1S_{A'}$	Max $^2S_{A'}$	Max $^1S_{M'}$	Max $^2S_{M'}$	Max Δ_R

Prospetto IX. — Mortalità generale.

		Max ^1S_A	Max ^2S_A	Max ^1S_M	Max ^2S_M	Max Δ_R
Non pond.	Applicaz. formule (1) o (3 bis).	**33,726**	**69,664**	**17,987**	**71,948**	**33,726**
	Applic. form. (2) per k = 1					
	Applic. form. (2) per k = 2.288 / Applicaz. form. (2 bis)	35,327	132,903	17,987	134,116	35,327
Ponder.	Applicazione formule (6) / Applicaz. formule (4 bis)	**34,447**	**123,418**	**17,573**	**124,662**	**34,447**
		Max $^1S_{A'}$	Max $^2S_{A'}$	Max $^1S_{M'}$	Max $^2S_{M'}$	Max Δ_R

Prospetto X. — Natalità specifica.

		Max ^1S_A	Max ^2S_A	Max ^1S_M	Max ^2S_M	Max Δ_R
Non ponder.	Applicazione formule (1)	258,028	532,983	137,615	550,460	258,028
	Applic. form. (2) per k = 2	240,826	364,095	137,615	389,234	240,826
	Applic. form. (2) per k = 3	223,624	286,468	137,615	317,088	223,624
	Applic. form. (2) per k = 2,202 / Applicaz. form. (2 bis)	237,354	344,495	137,615	370,965	237,354
	Applicazione formule (3 = 3 bis)	**223,624**	**329,557**	**137,615**	**357,135**	**236,096**
Ponder.	Applicaz. formule (4) o (4 bis)	**227,164**	**334,079**	**131,333**	**358,967**	**228,109**
	Applic. form. (5) per N_k = 1.057,013	230,335	350,549	131,333	374,342	230,335
	Applic. form. (5) per N_k = 1.160.701	227,164	332,213	131,333	357,231	227,164
	Applic. form. (5) per N_k = 1.127.834 / Applicaz. formule (5 bis)	228,169	337,764	131,333	362,399	228,169
	Applicazione formule (6)	259,495	1188,038	131,333	1195,262	259,495
		Max $^1S_{A'}$	Max $^2S_{A'}$	Max $^1S_{M'}$	Max $^2S_{M'}$	Max Δ_R

Prospetto XI. — Nuzialità specifica.

		Max $^{1}S_A$	Max $^{2}S_A$	Max $^{1}S_M$	Max $^{2}S_M$	Max Δ_R
Non ponder.	Applicazione formule (1) — Applic. form. (2) per $k =$ 1	125,078	258,360	66,708	266,832	125,078
	Applic. form. (2) per $k =$ 2	116,739	176,493	66,708	188,689	116,739
	Applic. form. (2) per $k =$ 1,067 — Applicaz. formule (2 bis)	124,516	249,516	66,708	258,279	124,516
	Applicazione formule (3) o (3 bis)	**116,739**	**241,523**	**66,708**	**250,567**	**124,026**
Ponder.	Applicazione formule (4) o (4 bis)	**120,104**	**244,197**	**64,204**	**252,497**	**120,154**
	Applic. formule (5) per $N_k =$ 334.856	121,533	269,954	64,204	277,483	121,533
	Applic. formule (5) per $N_k =$ 404.517	120,104	244,163	64,204	252,463	120,104
	Applic. formule (5) per $N_k =$ 401.580 — Applicaz. formule (5 bis)	120,164	245,116	64,204	253,385	120,164
	Applicaz. delle form. (6)	126,978	604,994	64,204	608,397	126,978
		Max $^{1}S_{A'}$	Max $^{2}S_{A'}$	Max $^{1}S_{M'}$	Max $^{2}S_{M'}$	Max Δ_R

Riuniamo, infine, in un solo prospetto, sia per gli indici di variabilità non ponderati che per i ponderati, tanto i valori relativi ottenuti per riferimento dei valori assoluti degli indici ai «massimi» calcolati con applicazione delle formule (1), (2) e (2 bis) [e rispettivamente (6), (5) e (5 bis)]; come anche quelli ottenuti per riferimento ai massimi calcolati correttamente con eventuale applicazione di formule diverse dalle (1) e (6), a seconda dei casi. I valori degli indici relativi corrispondenti ai massimi calcolati correttamente vengono indicati in carattere grassetto.

Prospetto XII. — Indici relativi non ponderati.

Caratteri			^1S_A	^2S_A	^1S_M	^2S_M	Δ_R
Natalità generica	proced. consueto (corretto) dalle form. (1), o dalle (2) per $k = 1$, o dalle (3 bis).		**0,073**	**0,044**	**0,133**	**0,044**	**0,101**
Nuzialità generica	proced. consueto (corretto) dalle form. (1), o dalle (2) per $k = 1$, o dalle (3 bis).		**0,037**	**0,021**	**0,069**	**0,021**	**0,050**
Mortalità generale	proced. consueto (corretto) dalle formule (1), o dalle (2) per $k = 1$, o dalle (3 bis).		**0,048**	**0,028**	**0,088**	**0,028**	**0,066**
Natalità specifica	proced. consueto dalle form. (1)		0,084	0,051	0,154	0,051	0,116
Natalità specifica	proced. dalle form. (2) per $k =$	2	0,090	0,075	0,154	0,072	0,124
Natalità specifica	proced. dalle form. (2) per $k =$	3	0,097	0,095	0,154	0,088	0,134
Natalità specifica	proced. dalle form. (2) per $k =$ proced. dalle formule (2 bis)	2,202	0,092	0,079	0,154	0,075	0,126
Natalità specifica	proc. (corretto) dalle form. (3) o (3 bis).		**0,097**	**0,082**	**0,154**	**0,078**	**0,127**
Nuzialità specifica	proc. consueto dalle form. (1) prced. dalle form. (2) per $k =$	1	0,062	0,036	0,111	0,037	0,083
Nuzialità specifica	prced. dalle form. (2) per $k =$	2	0,067	0,053	0,111	0,047	0,089
Nuzialità specifica	prced. dalle form. (2) per $k =$ proced. dalle formule (2 bis)	1,067	0,063	0,038	0,111	0,034	0,084
Nuzialità specifica	proc. (corretto) dalle form. (3) o (3 bis).		**0,067**	**0,039**	**0,111**	**0,039**	**0,084**

Prospetto XIII. — Indici relativi ponderati.

Caratteri			$^1S_{A'}$	$^2S_{A'}$	$^1S_{M'}$	$^2S_{M'}$	Δ_R
Natalità generica	dalle formule (6)		0,071	0,019	0,137	0,019	0,095
	dalle formule (5) per N_k =	864.174	0,070	0,026	0,137	0,026	0,096
		1.332.731	0,071	0,033	0,137	0,032	0,097
	dalle formule (5) per N_k = 1.127.834; dalle formule (5 bis)	1.127.834	0,071	0,030	0,137	0,030	0,097
	dalle form. (4) o (4 bis) (proced. corretto).		**0,071**	**0,032**	**0,137**	**0,031**	**0,097**
Nuzialità generica	dalle form. (6) o dalle (4 bis) (proced. corretto).		**0,029**	**0,008**	**0,057**	**0,008**	**0,040**
Mortalità generale	dalle form. (6) o dalle (4 bis) (proced. corretto).		**0,039**	**0,014**	**0,076**	**0,014**	**0,054**
Natalità specifica	dalle formule (6)		0,083	0,022	0,164	0,022	0,112
	dalle formule (5) per N_k =	1,057.013	0,094	0,075	0,164	0,070	0,126
		1.160.701	0,095	0,079	0,164	0,073	0,128
	dalle formule (5) per N_k = 1.127.834; dalle formule (5 bis)	1.127.834	0,095	0,078	0,164	0,072	0,128
	dalle form. (4) o (4 bis) (proc. corretto).		**0,095**	**0,078**	**0,164**	**0,073**	**0,128**
Nuzialità specifica	dalle formule (6)		0,038	0,011	0,071	0,011	0,053
	dalle formule (5) per N_k =	334.856	0,040	0,025	0,071	0,024	0,056
		404.517	0,040	0,027	0,071	0,027	0,057
	dalle formule (5) per N_k = 401.580; dalle formule (5 bis)	401.580	0,040	0,027	0,071	0,026	0,057
	dalle form. (4) o (4 bis) (proc. corretto).		**0,040**	**0,027**	**0,071**	**0,027**	**0,057**

15. Le applicazioni fatte al numero precedente non sono fra le più significative ai fini di mettere in evidenza le analogie o le differenze fra i risultati che, nella determinazione degli indici di variabilità, si ottengono a seconda che si fa uso dell'uno o dell'altra formula, e ciò perchè il numero dei termini delle serie considerate è basso (16) e perchè pochi sono, rispetto alla totalità, i termini sui quali si può riversare l'ammontare totale del carattere.

Nel caso, invero, degli indici non ponderati, per tre dei caratteri (natalità generica, nuzialità generica, mortalità generale), l'ammontare totale del carattere si può effettivamente attribuire a un solo termine (come appare chiaro dalle distribuzioni massimanti), mentre per gli ultimi due caratteri (natalità specifica e nuzialità specifica) sono soltanto due, e rispettivamente uno, i termini ai quali può attribuirsi il valore massimo L, mentre a un altro termine va attribuito il resto.

Per ciò che concerne gli indici ponderati, sono due i caratteri (nuzialità generica e mortalità generale), per i quali l'ammontare totale del carattere può riversarsi su un solo termine, due (natalità generica e nuzialità specifica) quelli per cui un termine ottiene il valore massimo L e un altro il resto, mentre, per il quinto carattere (natalità specifica), il massimo L viene raggiunto da due termini, mentre a un terzo spetta il resto.

Quando l'ammontare totale del carattere può essere attribuito a un solo termine, come si verifica per tre caratteri negli indici non ponderati e per due negli indici ponderati, le formule usuali risultano naturalmente corrette e i loro risultati coincidono con quelli delle formule (3 bis) o (4 bis). Ed è, d'altra parte, chiaro che, più è piccolo il numero dei termini sui quali si può distribuire l'ammontare totale del carattere, e minore risulta il vantaggio nell'adoperare le formule corrette (3) o (3 bis), per gli indici non ponderati, e (4) o (4 bis) per i ponderati, in confronto delle rispettive formule usuali (1) e (6).

Ciò non ostante, differenze non trascurabili, e talvolta anzi notevoli, si osservano fra i risultati delle formule corrette e quelli delle formule usuali; ciò soprattutto per i valori degli indici quadratici, in minor grado per i valori degli indici lineari. Le differenze si verificano sempre in un senso; gli indici secondo le formule usuali risultano, cioè più bassi che secondo le corrette.

Per ciò che concerne gli indici relativi non ponderati (prospetto XII), le differenze fra i risultati delle formule (1) e delle formule (3) o (3 bis) sono, della nuzialità specifica, lievi, per quanto sempre avvertibili, sia per ciò che concerne gli indici lineari ${}^1S_A$ e Δ_R, sia per ciò che concerne gli indici quadratici. Per la natalità specifica, le differenze sono abbastanza notevoli per gli indici lineari ${}^{1}S_A$ e Δ_R e notevoli per gli indici quadratici.

Per ciò che concerne gli indici relativi ponderati (Cfr. prospetto XIII), le differenze sono, per la natalità generica e per la nuzialità specifica, lievi per gli indici lineari ${}^{1}S_A$ e Δ_R e notevoli per gli indici quadratici. Esse sono abbastanza notevoli per gli indici lineari ${}^{1}S_A$ e Δ_R e molto notevoli per gli indici quadratici per la natalità specifica.

Si avverta come, sia per gli indici non ponderati che per i ponderati, le differenze — conformemente alle previsioni — risultano, per la natalità specifica, per la quale l'ammontare totale del carattere si distribuisce su tre termini (due con valore massimo e uno col resto), molto maggiori che per gli altri caratteri, per cui l'ammontare totale dei caratteri si distribuisce su due termini (uno con valore massimo e uno col resto).

Nessuna differenza risulta per ciò che riguarda lo scostamento semplice della mediana, il cui valore massimo, secondo le formule usuali e secondo le formule corrette, coincide tutte le volte che la mediana è uguale a zero; come sempre avviene nelle applicazioni del paragrafo precedente.

Lo scarso numero dei termini fa sì che il trascurare o il considerare il resto e l'usare quindi le formule (2) o (2 bis) in luogo delle (3) o (3 bis) per gli indici non ponderati, e rispettivamente le formule (5) e (5 bis) in luogo delle (6) o (6 bis) per gli indici ponderati, porti a differenze più sensibili che se il numero dei termini fosse maggiore.

Malgrado ciò, le differenze fra i risultati ottenuti con le dette formule non sono mai forti, e talvolta addirittura non si avvertono fino al terzo decimale, e ciò sia per gli indici relativi non ponderati, sia per i ponderati.

Gli indici relativi ottenuti con le (2 bis) e (5 bis) risultano sempre approssimati per difetto rispetto ai valori esatti ottenuti con le (3 bis) e (4 bis), ciò che è conforme alle conclusioni raggiunte ai numeri 10 e 11.

In confomità alle dette conclusioni, risulta pure che le formule (2) e (5) — quando con k, o con N_k, si indica, in dette formule, il numero, e rispettivamente il peso, dei termini a cui può attribuirsi il limite superiore L — danno sempre valori approssimati per difetto e sensibilmente più lontani da quelli esatti, dei valori rispettivi ottenuti mediante le (2 bis) e (5 bis). Talvolta l'approssimazione delle formule (2) e (5), senza potersi dire cattiva, non può ritenersi però sufficiente, ciò che invece non si verifica mai per le formule (2 bis) e (5 bis).

Se, invece, nelle (2) e nelle (5) si indica con k ed N_k il numero, e rispettivamente il peso, dei termini a cui non può attribuirsi il valore 0, dette formule danno, oltre che per lo scostamento semplice medio della mediana anche per lo scostamento semplice medio della media aritmetica valori coincidenti con quelli esatti forniti dalle (3) o (3 bis) e dalle (4) o (4 bis), e ciò pure in confomità alle previsioni teoriche, essendo sempre $R > A$, $D > A$; per gli altri indici, dette formule danno valori talora praticamente uguali, spesso più elevati, di quelli esatti ottenuti dalle (3) o (3 bis) e dalle (4) o (4 bis) e, in questo secondo caso, per lo più, meno vicini ai valori esatti di quelli forniti dalle (2 bis) e (5 bis).

Questi risultati suggeriscono la conclusione che le formule (2) o meglio le (2 bis) e le formule (5) o meglio ancora le (5 bis) possono usarsi tranquillamente anche in applicazioni come le precedenti, in cui è piccolo il numero dei termini delle serie. Essi mostrano, d'altra parte, che, anche in applicazioni come queste (in cui tuttavia le divergenze fra le formule corrette e le formule usuali sono relativamente poco forti, in quanto l'ammontare totale del carattere può venire riversato sopra un numero di termini piccolo rispetto alla totalità), le formule usuali possono portare a risultati lontani dal vero, soprattutto per ciò che concerne gli indici quadratici.

Se, per uno stesso carattere, consideriamo l'altezza dei valori relativi ottenuti per i vari indici, riscontriamo che, per tutti i caratteri e per tutte le formule, e sia per gli indici ponderati che per i non ponderati, i valori più elevati sono forniti dagli indici lineari e i meno elevati dagli indici quadratici. Fra gli indici lineari, si succedono nell'ordine, ben distanziati l'uno dall'altro, lo scostamento semplice dalla mediana, la differenza media e lo scostamento semplice della media aritmetica. Gli indici quadratici hanno valori che si avvicinano molto tra di loro : nelle formule corrette (3) o (3 bis), (4) o (4 bis) come nelle approssimate (2) o (2 bis), (5) o (5 bis), quando una differenza c'è, risulta più alto lo scostamento medio dalla media aritmetica.

Confrontando i diversi caratteri per ciò che concerne l'intensità della loro variabilità, si riscontra che essi si succedono per lo più in questo ordine decrescente : natalità specifica, natalità generica, nuzialità specifica, mortalità generale, nuzialità generica. Quest'ordine è conservato con tutti gli indici non ponderati qualunque sia la formula adoperata ; esso è conservato altresì quando si fa uso delle formule (4) o (4 bis) o delle (5) e (5 bis), con tutti gli indici ponderati, salvo per ${}^{1}S_M$, che dà una variabilità per la nuzialità specifica inferiore che per la mortalità generale. Non è conservato invece con le formule usuali, in quanto la variabilità della nuzialità specifica risulta, con queste, inferiore per tutti gli indici a quella della mortalità generale.

Se infine confrontiamo il valore di un indice relativo ponderato ottenuto con una certa formula, col valore rispettivo dello stesso indice non ponderato otteuto con la formula corispondente (per es. ${}^{1}S_M$ con ${}^{1}S_{M'}$, ${}^{1}S_A$ con ${}^{1}S_{A'}$, ecc.), notiamo, che per due caratteri (natalità generica e mortalità specifica), le differenze sono piccole, ma non si verificano sempre nello stesso senso ; per gli altri tre, sono più o meno forti. Le differenze sono più accentuate per gli indici quadratici che per i lineari ; e più accentuate per le formule usuali (6) che per le corrette (4) e (4 bis) o per le approssimate (5) (5 bis).

Nelle applicazioni eseguite al numero precedente, le formule corrette, del pari che le approssimate, risultano pertanto meno perturbate dalla mancanza di ponderazione delle formule usuali, e meglio di queste conservano l'ordine dei caratteri secondo l'altezza della variabilità.

16. Allo scopo di mettere meglio in luce la differenza o le analogie tra i risultati che si ottengono con le formule usuali e quelli a cui si perviene con le formule esatte o approssimate proposte in questa memoria, un'altra applicazione è stata fatta, relativa ai seguenti caratteri: natalità generica, nuzialità generica, mortalità generale, accrescimento naturale, percentuale della popolazione agglomerata, percentuale della popolazione agricola maschile nell'insieme dei maschi di età superiore ai 10 anni. L'intensità di tali caratteri fu determinata per i 214 Circondari che costituivano il Regno d'Italia, nel territorio compreso nei vecchi confini, al 1º dicembre 1921, data dell'ultimo censimento generale della popolazione. Come denominatori dei primi cinque quozienti demografici considerati si assunsero le popolazioni presenti, censite alla data predetta, nei singoli circondari, e come denominatori dell'ultimo quoziente i numeri dei maschi in età superiore ai 10 anni presenti alla stessa data. I numeratori furono poi determinati come medie aritmetiche dei nati, dei matrimoni, dei morti, etc., nel biennio 1921–1922.

Si omette la trascrizione dei valori dei 6 caratteri esaminati in ciascuno dei 214 Circondari, valori che potranno trovarsi altrove (*), e ci si limita a indicare per ogni carattere (Prosp. XIV) il valore medio aritmetico, non ponderato e ponderato, nella totalità dei Circondari.

Prospetto XIV.

	Natalità ‰	Nuzialità ‰	Mortalità ‰	Accrescimento naturale ‰	Popolazione agglomerata %	Popolazione agricola %
A = medie aritmetiche non ponderate	31,12	10,61	17,93	13,19	72,66	54,36
M = mediane non ponderate . . .	32,38	10,70	17,50	14,18	77,15	57,75
A' = medie aritmetiche ponderate.	30,23	10,45	17,54	12,70	73,25	47,36
M' = mediane ponderate	31,69	10,56	16,89	13,34	78,60	51,50

Relativamente alle sei serie considerate si trova che gli indici assoluti di variabilità, semplici e ponderati, hanno i valori risultanti dal seguente prospetto:

Prospetto XV.

Caratteri	Non ponderati				
	${}^1S_A$	${}^2S_A$	${}^1S_M$	${}^2S_M$	Δ_R
Natalità	5,147	6,185	5,028	6,311	6,978
Nuzialità	0,924	1,160	0,919	1,164	1,309
Mortalità	2,196	2,667	2,179	2,702	3,021
Accrescimento naturale . .	4,020	4,919	3,956	5,018	5,528
Popolazione agglomerata. .	16,507	19,297	16,086	19,813	21,827
Popolazione agricola	11,414	14,630	11,026	15,018	15,722

(*segue*)

(*) V. Tav. I nel Vol. IV degli « Annali di Statistica », serie VI : C. GINI e L. GALVANI, *Di un'applicazione del metodo rappresentativo all'ultimo censimento italiano della popolazione*. Roma, 1929-VII. Nella stessa tavola si troveranno anche indicate le popolazioni dei singoli circondari.

Seguito *Prospetto XV*.

Caratteri	Ponderati				
	${}^1S_{A'}$	${}^2S_{A'}$	${}^1S_{M'}$	${}^2S_{M'}$	Δ_R
Natalità	5,339	6,348	4,716	6,514	7,204
Nuzialità	0,897	1,102	0,893	1,108	1,269
Mortalità	2,158	1,516	1,956	1,649	2,945
Accrescimento naturale	4,126	4,975	4,106	5,016	5,633
Popolazione agglomerata	16,839	20,816	16,300	21,493	21,282
Popolazione agricola	14,657	18,005	14,304	18,475	19,995

Come sappiamo, la determinazione dei massimi dei diversi indici di variabilità è subordinata alla conoscenza della distribuzione massimante corrispondente a ciascuna serie.

Per i caratteri non ponderati basta, per individuare tale distribuzione, eseguire una divisione avente come dividendo l'ammontare totale del carattere e come divisore il limite superiore L del carattere stesso. Dal quoziente ottenuto e dalla eventuale presenza di un resto, si deduce immediatamente quali siano le formule da applicare in via rigorosa, e quali possano essere applicate in via approssimata.

Per i caratteri non ponderati, dei quali i primi quattro hanno come limite superiore teorico 1000 e gli ultimi due 100, le distribuzioni massimanti sono:

Prospetto XVI.

Caratteri non ponderati	Distribuzione massimante
Natalità	$6.660,12 = 1000 \times 6 + 660,12$
Nuzialità	$2.270,21 = 1000 \times 2 + 270,21$
Mortalità	$3.836,91 = 1000 \times 3 + 836,91$
Accrescimento naturale	$2.823,67 = 1000 \times 2 + 823,67$
Popolazione agglomerata	$15.548,5 = 100 \times 155 + 48,5$
Popolazione agricola	$11.632,3 = 100 \times 116 + 32,3$

Si osserva che tutte queste distribuzioni comportano un resto, e che il limite superiore di ciascun carattere può essere attribuito a più di un elemento (circondario). Perciò, in via rigorosa, saranno per tutti i sei caratteri da applicare le formule (3) o (3 bis), mentre in via approssimata potranno essere anche applicate le (2) e (2 bis); naturalmente poi l'applicazione delle (2) per il valore esatto di k fornirà lo stesso valore dato dalle (2 bis). Sono anche state applicate le (1), allo scopo di paragonarne i risultati con quelli ottenuti dalle (3) e dalle (2) e (2 bis).

In relazione alle osservazioni fatte al numero precedente, notiamo che, per i primi quattro caratteri, la distribuzione massimante concentra in pochi termini la totalità del carattere, mentre per gli ultimi due la distribuzione massimante effettua la concentrazione dall'ammontare totale del carattere in un numero di termini molto rilevante (155 e 116). Perciò è da prevedere che le differenze più cospicue fra gli indici di variabilità calcolati correttamente e quelli ottenuti con le formule usuali si riscontreranno appunto per questi due caratteri: popolazione agglomerata e popolazione agricola; ciò che risulta confermato dall'ispezione del prospetto XXV.

Per quanto riguarda i caratteri ponderati, la determinazione delle distribuzioni massimanti non può essere fatta secondo una regola fissa, ma viene agevolata dall'impiego delle norme esposte al n. 12, le quali sono di facile applicazione quando i termini delle serie siano poco numerosi (come negli esempi del numero precedente), poichè allora la scelta dell'aggregato di termini ai quali conviene attribuire il valore massimo L del carattere è limitata ai soli aggregati che contengono un piccolo numero di termini. Se invece, come negli esempi attuali, le serie comprendono molti termini, e l'ammontare totale del carattere è grande, cosicchè esso si debba distribuire, attribuendo il valore massimo L ad un rilevante numero di termini, allora, pur tenendo presenti le norme dette, è generalmente necessario eseguire molte prove prima di ottenere quell'aggregato di termini, ai quali viene conferito il valore L tale che $L\ N_k$ sia il massimo a non superare l'ammontare totale $A'\ N$ del carattere.

Tuttavia, se invece di scegliere quel gruppo di termini (o uno di quei gruppi di termini) che dà la massima somma N_k, se ne fosse — per insufficienza del numero di prove eseguite o per altra causa — scelto un altro avente una somma N_k non massima, ma prossima alla massima, le formule fornirebbero valori prossimi ai veri e praticamente accettabili come tali. In altri termini, le formule (4) o (4 bis) (che tengono conto del resto) e le formule (5) o (5 bis) (che tale resto trascurano), fornirebbero praticamente gli stessi valori.

Ciò posto, relativamente ai caratteri ponderati, le distribuzioni massimanti, o assunte come tali, sono:

Prospetto XVII.

Caratteri ponderati	Distribuzione massimante
Natalità	1.122.829.141,48 = 1000 × 1.122.824 + 0,178 × 28.914
Nuzialità	388.143.054,20 = 1000 × 388.134 + 0,313 × 28.914
Mortalità	651.486.045,04 = 1000 × 651.473 + 0,451 × 28.914
Accrescimento naturale	471.714.525,20 = 1000 × 471.714 + 0,018 × 28.914
Popolazione agglomerata	2.720.715.667,00 = 100 × 27.207.142 + 0,051 × 28.914
Popolazione agricola	686.499.444,48 = 100 × 6.864.985 + 0,086 × 10.969

Il gruppo di elementi, a cui si è attribuito il valore 1000 o il valore 100 a seconda dei casi, è costituito, per ciascuno dei caratteri indicati, dai circondari seguenti, per ognuno dei quali viene indicato il peso (popolazione per ciascuno dei primi cinque caratteri; popolazione maschile in età superiore ai 10 anni per l'ultimo carattere).

Prospetto XVIII.

Natalità	
Milano	1.002.788
Orvieto	60.109
Rocca S. Casciano	59.927
N_k =	1.122.824

Nuzialità	
Gerace Marina .	150.092
Vallo della Lucania	98.090
Montepulciano .	79.674
San Bartolomeo in Galdo. . .	60.278
N_k =	338.134

Mortalità	
Belluno	111.200
Chiavari	102.716
Forlì	94.132
Spoleto	84.401
Porto Maurizio .	60.201
Cittaducale . .	60.086
Bovino	52.588
Civitavecchia .	46.348
Pieve di Cadore	39.801
N_k =	651.473

Accrescimento naturale	
Cittaducale . .	60.086
Camerino . . .	50.324
Civitavecchia .	46.348
Savona	145.378
Crema	109.300
San Bartolomeo in Galdo. . .	60.278
N_k =	471.714

Popolazione agglomerata		Popolazione agglomerata		Popolazione agglomerata	
Campagna . . .	99.111	S. Severo	171.660	Piacenza	202.389
Abbiategrasso . .	134.578	Pieve di Cadore .	39.801	Parma	207.680
Altamura	136.571	Porto Maurizio. .	60.201	Spezia	195.925
Treviglio	140.248	S. Bartolomeo in Galdo	60.278	Foggia	234.254
Monteleone Cal. .	150.858	Mortara	155.409	Modena	238.758
Gerace Marina . .	150.092	Varese	174.029	Ferrara	244.891
Massa Carrara . .	142.459	Reggio Calabria .	178.862	Modica	252.546
Oristano	137.010	Alessandria . . .	159.994	Gallarate	267.414
Urbino	144.148	Pavia	159.663	Reggio Emilia . .	266.611
Pistoia	140.375	Grosseto	164.990	Pisa	259.973
Pordenone . . .	156.945	Cuneo	171.944	Taranto	274.907
Ivrea	148.074	Teramo	189.114	Novara	261.206
Potenza	151.436	Avellino	184.258	Girgenti	272.880
Voghera	138.991	Siena	168.168	Messina	285.948
Piazza Armerina .	138.287	Rovigo	187.270	Perugia	291.982
Saluzzo	147.873	Siracusa	172.424	Como	299.286
Volterra	100.814	Lodi	179.419	Salerno	314.008
Orvieto	60.109	Cassino	192.794	Arezzo	298.519
S. Miniato . . .	143.284	Brindisi	188.039	Bergamo	331.075
Acireale	151.200	Cosenza	205.760	Monza	322.032
Savona	145.378	Asti	182.121	Ancona	334.654
Biella	156.732	Frosinone	211.281	Caserta	331.025
Lecco	157.662	Trapani	198.473	Lucca	346.602
Alba	151.039	Gallipoli	207.562	Brescia	334.948
Gaeta	167.177	Viterbo	196.842	Mantova	376.901
Vercelli	146.553	Macerata	217.436	Barletta	385.697
Caltagirone . .	154.329	Castell. Stab. . .	216.622	Udine	420.938
Mondovì	152.742	Cagliari	211.245	Bari delle Puglie.	430.243
Sora	170.967	Lecce	215.540	Venezia	451.323
Casale Monferr. .	148.789	Cremona	203.155	Catania	448.880
Palmi	173.357			Bologna	484.191

Popolazione agglomerata		Popolazione agglomerata		Popolazione agglomerata	
Vicenza	517.119	Palermo	617.662	Napoli	952.488
Verona	518.256	Genova	657.657	Roma	958.436
Treviso	548.487	Firenze	698.191	Milano.	1.002.788
Padova	588.043	Torino	818.787		

$N_k = 27.207,142$

Popolazione agricola		Popolazione agricola		Popolazione agricola	
Portoferraio	12.135	Pisa	106.336	Udine	156.280
Borgotaro	14.085	Novara	107.235	Bari delle Puglie	160.595
Piedimonte d'Alife	17.153	Girgenti	107.453	Venezia	171.930
Camerino	18.744	Messina	110.582	Catania	175.749
S. Bartolomeo in Galdo	22.068	Perugia	114.287	Bologna	194.114
Rocca S. Casciano	23.656	Como	117.489	Vicenza	197.361
Terranova Sicilia.	34.952	Salerno	118.682	Verona	204.026
Cremona.	81.975	Arezzo.	119.643	Treviso	206.396
Parma.	82.924	Bergamo.	123.072	Padova	221.485
Spezia	83.527	Monza	125.646	Palermo	232.351
Foggia.	89.523	Ancona	125.836	Genova	276.115
Ferrara	93.795	Caserta	128.634	Firenze	284.393
Modena	93.494	Lucca	130.541	Torino.	350.265
Gallarate	103.873	Brescia	132.223	Napoli.	375.544
Reggio Emilia	103.979	Mantova.	150.346	Roma	402.314
		Barletta	146.938	Milano	415.241

$N_k = 6.864.985$

Tutte le distribuzioni massimanti considerate ammettono un certo resto, attribuito al termine di peso minimo fra tutti i dati. Tale termine è costituito, per i primi cinque caratteri, dal Circondario di Portoferraio, il cui peso (popolazione) è 28.914; e, per l'ultimo carattere, dal Circondario di Varallo, il cui peso (popolazione maschile in età superiore ai 10 anni) è 10.969.

Perciò, la determinazione dei massimi degli indici di variabilità si dovrà fare in modo corretto applicando le (4) o (4 bis) e in modo approssimato applicando le (5) o (5 bis).

Naturalmente, si applicheranno anche le (6) supponendo che tutto l'ammontare del carattere possa essere riversato sopra un solo termine e precisamente su quello di peso minimo (*), allo scopo di paragonare i risultati da esse forniti con quelli delle (4) o (4 bis), (5) o (5 bis).

Richiamandoci, anche qui, alle considerazioni del precedente numero, costatiamo che le distribuzioni massimanti relative ai primi quattro caratteri: natalità, nuzialità, mortalità, accrescimento naturale, effettuano la concentrazione dell'ammontare totale del carattere fra tanti circondari aventi all'incirca un peso complessivo dell'ordine di 1/30-1/100 del peso complessivo totale, mentre le ultime due realizzano una distribuzione fra tanti circondari che hanno un peso complessivo molto maggiore, e precisamente dell'ordine di grandezza di oltre ½, del peso complessivo totale. Fra la distribuzione non consentita, che risulterebbe dal concentrare l'ammontare totale di un carattere in un solo elemento della serie, e quella massimante da noi giustamente considerata, corre dunque una diversità molto maggiore per gli ultimi due caratteri che non per i primi quattro: e ciò fa prevedere che gli indici relativi di variabilità da noi calcolati correttamente differiranno da quelli usuali in molto maggior misura per gli ultimi due caratteri che per i primi quattro: anche questa previsione è pienamente confermata dall'esame del prospetto XXVI.

Ecco, qui di seguito, le tabelle contenenti i massimi cercati per ciascuno dei sei caratteri considerati.

(*) In tal caso sarebbe $M' = 0$ e perciò per ${}^1S_M$ e ${}^2S_M$ sono state applicate le formule (6) ottenute per $M' = 0$.

Prospetto XIX. — Natalità.

		Max $^{1}S_{A}$	Max $^{2}S_{A}$	Max $^{1}S_{M}$	Max $^{2}S_{M}$	Max Δ_{R}
Non ponderati	Applicazione formula (1)	61,949	454,181	31,12	455,245	61,949
	Applicaz. form. (2) per k = 6	60,494	183,228	31,12	185,855	60,494
	Applicaz. form. (2) per k = 7	60,204	169,231	31,12	172.066	60,205
	Applicaz. form. (2) per k = 6,660 / Applicaz. form. (2 bis)	60,303	173,642	31,12	176,409	60,303
	Applicaz. formule (3) o (3 bis).	**60,204**	**170,596**	**31,12**	**173,412**	**60,293**
Ponderati	Applicaz. formule (4) o (4 bis). / Applicaz. formule (5) o (5 bis).	**58,632**	**171,219**	**30,23**	**173,867**	**58,632**
	Applicazione formula (6)	60,413	1083,060	30,23	1083,482	60,413
		Max $^{1}S_{A'}$	Max $^{2}S_{A'}$	Max $^{1}S_{M'}$	Max $^{2}S_{M'}$	Max Δ_{R}

Prospetto XX. — Nuzialità.

		Max $^{1}S_{A}$	Max $^{2}S_{A}$	Max $^{1}S_{M}$	Max $^{2}S_{M}$	Max Δ_{R}
Non ponderati	Applicazione formula (1)	21,121	154,848	10,61	155,201	21,121
	Applicaz. form. (2) per k = 2	21,022	109,241	10,61	109,750	21,022
	Applicaz. form. (2) per k = 3	20,923	88,975	10,61	89,612	20,923
	Applicaz. form. (2) per k = 2,270 / Applicaz. form. (2 bis)	20,995	102,457	10,61	103,005	20,995
	Applicaz. formule (3) o (3 bis).	**20,923**	**97,857**	**10,61**	**98,430**	**20,986**
Ponderati	Applicaz. formule (4) o (4 bis). / Applicaz. formule (5) o (5 bis).	**20,681**	**101,689**	**10,45**	**102,224**	**20,682**
	Applicazione formula (6)	20,884	374,396	10,45	374,541	20,884
		Max $^{1}S_{A'}$	Max $^{2}S_{A'}$	Max $^{1}S_{M'}$	Max $^{2}S_{M'}$	Max Δ_{R}

Prospetto XXI. — Mortalità.

		Max ${}^{1}S_{A}$	Max ${}^{2}S_{A}$	Max ${}^{1}S_{M}$	Max ${}^{2}S_{M}$	Max Δ_{R}
Non ponderati	Applicaz. formule (1) .	35,692	261,679	17,93	262,293	35,692
	Applicaz. form. (2) per $k =$ 3	35,357	150,361	17,93	151,437	35,357
	Applicaz. form. (2) per $k =$ 4	35,190	129,921	17,93	131,140	35,190
	Applicaz. form. (2) per $k =$ 3,857 / Applicaz. form. (2 bis)	35,217	132,697	17,93	133,903	35,217
	Applic. form. (3) o (3 bis)	**35,190**	**130,272**	**17,93**	**131,500**	**35,211**
Ponderati	Applic. form. (4) o (4 bis) / Applic. (form. (5) o (5 bis)	**34,465**	**131,271**	**17,54**	**132,437**	**34,466**
	Applicaz. formule (6) .	35,053	628,412	17,54	628,656	35,053
		Maz ${}^{1}S_{A'}$	Max ${}^{2}S_{A'}$	Max ${}^{1}S_{M'}$	Max $S^{1}_{M'}$	Max Δ_{R}

Prospetto XXII. — Accrescimento naturale.

		Max ${}^{1}S_{A}$	Max ${}^{2}S_{A}$	Max ${}^{1}S_{M}$	Max ${}^{2}S_{M}$	Max Δ_{R}
Non ponderati	Applicaz. formule (1) .	26,257	192,501	13,19	192,953	26,257
	Applicaz. form. (2) per $k =$ 2	26,133	135,799	13,19	136,439	26,133
	Applicaz. form. (2) per $k =$ 3	26,010	110,611	13,19	111,403	26,010
	Applicaz. form. (2) per $k =$ 2,824 / Applicaz. form. (2 bis)	26,032	114,088	13,19	114,848	26,032
	Applic. form. (3) o (3 bis)	**26,010**	**111,074**	**13,19**	**111,854**	**26,026**
Ponderati	Applic. form. (4) o (4 bis) / Applic. form. (5) o (5 bis)	**25,077**	**111,976**	**12,70**	**112,694**	**25,077**
	Applicaz. formule (6) .	25,380	455,007	12,70	455,184	25,380
		Max ${}^{1}S_{A'}$	Max ${}^{2}S_{A'}$	Max ${}^{1}S_{M'}$	Max ${}^{2}S_{M'}$	Max Δ_{R}

Prospetto XXIII. — Popolazione agglomerata.

		Max ${}^{1}S_A$	Max ${}^{2}S_A$	Max ${}^{1}S_M$	Max ${}^{2}S_M$	Max Δ_R
Non ponderati	Applicazione formule (1)	144,641	1060,436	72,66	1062,921	144,641
	Applic. form. (2) per $k =$ 155	40,065	44,829	27,658	52,674	40,065
	Applic. form. (2) per $k =$ 156	39,386	44,304	27,015	51,891	39,386
	Applic. form. (2) per $k =$ 155,485 Applicaz. form. (2 bis)	39,736	44,574	27,345	52,293	39,736
	Applicaz. formule (3) o (3 bis).	**39,612**	**44,443**	**27,345**	**52,182**	**39,735**
Ponderati	Applicaz. formule (4) o (4 bis). Applicaz. formule (5) o (5 bis).	**39,189**	**44,265**	**26,75**	**51,720**	**39,189**
	Applicazione formule (6)	146,386	2624,353	73,25	2625,374	146,386
		Max ${}^{1}S_{A'}$	Max ${}^{2}S_{A'}$	Max ${}^{1}S_{M'}$	Max ${}^{2}S_{M'}$	Max Δ_R

Prospetto XXIV. — Popolazione agricola.

		Max ${}^{1}S_A$	Max ${}^{2}S_A$	Max ${}^{1}S_M$	Max ${}^{2}S_M$	Max Δ_R
Non ponderati	Applicazione formule (1)	108,212	793,357	54,36	795,216	108,212
	Applic. form. (2) per $k =$ 116	49,788	49,964	45,925	67,864	49,788
	Applic. form. (2) per $k =$ 117	49,280	49,496	45,068	66,940	49,280
	Applic. form. (2) per $k =$ 116,323 Applicaz. form. (2 bis)	49,624	49,736	45,646	67,564	49,624
	Applicaz. formule (3) o (3 bis).	**49,486**	**49,710**	**45,646**	**67,482**	**49,623**
Ponderati	Applicaz. formule (4) o (4 bis). Applicaz. formule (5) o (5 bis).	**49,861**	**49,930**	**47,36**	**68,819**	**49,861**
	Applicazione formule (6)	94,649	1720,989	47,36	1721,640	94,649
		Max ${}^{1}S_{A'}$	Max ${}^{2}S_{A'}$	Max ${}^{1}S_{M'}$	Max ${}^{1}S_{M'}$	Max Δ_R

Nei prospetti XXV e XXVI, che seguono, sono raccolti i valori degli indici relativi di variabilità calcolati in rapporto ai massimi, testè determinati, in via rigorosa od approssimata.

Prospetto XXV. — Indici relativi non ponderati.

Caratteri	${}^1S_A$	${}^2S_A$	${}^1S_M$	${}^2S_M$	Δ_R
Natalità:					
dalle formule (1)	0,08308	0,01362	0,16157	0,01386	0,11264
dalle form. (2) per $k =$ 6	0,08508	0,03376	0,16157	0,03401	0,11535
dalle form. (2) per $k =$ 7	0,08549	0,03655	0,16157	0,03668	0,11590
dalle form. (2) per $k =$ 6,66012 / dalle formule (2 bis). . .	0,08535	0,03562	0,16157	0,03578	0,11571
dalle form. (3) o (3 bis) .	**0,08549**	**0,03626**	**0,16157**	**0,03639**	**0,11573**
Nuzialità:					
dalle formule (1)	0,04375	0,00749	0,08662	0,00750	0.06198
dalle form. (2) per $k =$ 2	0,04395	0,01062	0,08662	0,01061	0,06227
dalle form. (2) per $k =$ 3	0,04416	0,01304	0,08662	0,01299	0,06256
dalle form. (2) per $k =$ 2,27021 / dalle formule (2 bis). . .	0,04401	0,01132	0,08662	0,01130	0,06235
dalle form. (3) o (3 bis) .	**0,04416**	**0,01185**	**0,08662**	**0,01182**	**0,06237**
Mortalità:					
dalle formule (1)	0,06153	0,01019	0,12153	0,01030	0,08464
dalle form. (2) per $k =$ 3	0,06211	0,01774	0,12153	0,01784	0,08544
dalle form. (2) per $k =$ 4	0,06240	0,02064	0,12153	0,02060	0,08585
dalle form. (2) per $k =$ 3,83691 / dalle formule (2 bis) . .	0,06236	0,02010	0,12153	0,02018	0,08579
dalle form. (3) o (3 bis) .	**0,06240**	**0,02047**	**0,12153**	**0,02055**	**0,08580**

Segue *Prospetto XXV.*

Caratteri	$^{1}S_{A}$	$^{2}S_{A}$	$^{1}S_{M}$	$^{2}S_{M}$	Δ_{R}
Accrescimento naturale:					
dalle formule (1)	0,15154	0,02555	0,29992	0,02601	0,21053
dalle form. (2) per k = 2	0,15383	0,03623	0,29992	0,03678	0,21153
dalle form. (2) per k = 3	0,15456	0,04447	0,29992	0,04504	0,21253
dalle form. (2) per k = 2,82367 / dalle formule (2 bis) . .	0,15443	0,04312	0,29992	0,04370	0,21235
delle form. (3) o (3 bis) .	**0,15456**	**0,04430**	**0,29992**	**0.04487**	**0,21239**
Popolazione agglomerata:					
dalle formule (1)	0,11412	0,01820	0,22139	0,01864	0,15090
dalle form. (2) per k = 155	0,41201	0,43046	0,58160	0,37614	0,54479
dalle form. (2) per k = 156	0.41911	0,43556	0,59545	0,38182	0,55418
dalle form. (2) per k = 155,485 / dalle formule (2 bis) . .	0,41542	0,43292	0,58826	0,37888	0,54930
dalle form. (3) o (3 bis) .	**0,41672**	**0,43420**	**0,58826**	**0,37969**	**0,54931**
Popolazione agricola:					
dalle formule (1)	0,10548	0,01844	0,20283	0,01889	0,14529
dalle form. (2) per k = 116	0,22925	0,29281	0,24009	0,22130	0,31578
dalle form. (2) per k = 117	0,23162	0,29558	0,24465	0,22435	0,31903
dalle form. (2) per k = 116,323 / dalle formule (2 bis) . .	0,23001	0,29415	0,24155	0,22228	0,31682
dalle form. (3) o (3 bis) .	**0,23065**	**0,29431**	**0,24155**	**0,22254**	**0,31683**

Prospetto XXVI. — Indici relativi ponderati.

Caratteri	$^1S_{A'}$	$^2S_{A'}$	$^1S_{M'}$	$^2S_{M'}$	Δ_R
Natalità :					
dalle form. (4) o (4 bis).					
» » (5)	**0,09106**	**0,03708**	**0,15600**	**0,03747**	**0,12287**
» » (5 bis) . . .					
dalle form. (6)	0,08838	0,00586	0.15600	0,00601	0,11925
Nuzialità:					
dalle form. (4) o (4 bis)					
» » (5)	**0,04337**	**0,01084**	**0,08545**	**0,01084**	**0,06136**
» » (5 bis) . . .					
dalle form. (6)	0,04295	0,00294	0,08545	0,00296	0,06076
Mortalità:					
dalle form. (4) o (4 bis)					
» » (5)	**0,06261**	**0,01155**	**0,11152**	**0,01245**	**0,08545**
» « (5 bis) . . .					
dalle form. (6)	0,06156	0,00241	0,11152	0,00262	0,08402
Accrescimento naturale:					
dalle form. (4) o (4 bis)					
» » (5)	**0,16453**	**0,04443**	**0,32331**	**0,04451**	**0,22463**
» » (5 bis) . . .					
dalle form. (6)	0,16257	0,01093	0,32331	0,01102	0,22195
Popolazione agglomerata :					
dalle form. (4) o (4 bis)					
» » (5)	**0,42969**	**0,47026**	**0,60935**	**0,41556**	**0,54306**
» » (5 bis) . . .					
dalle form. (6)	0,11503	0,00793	0,22252	0,00819	0,14538
Popolazione agricola :					
dalle form. (4) o (4 bis)					
» » (5)	**0,29396**	**0,36060**	**0,30203**	**0,26846**	**0,40102**
» » (5 bis) . . .					
dalle form. (6)	0,15486	0,01046	0,30203	0,01073	0,21126

17. Se esaminiamo i risultati ottenuti nel numero precedente ed esposti nei prospetti XXV e XXVI, riscontrimo che le differenze fra i risultati forniti dalle formule usuali (1) e (6) e quelli ottenuti con le formule corrette (3) o (3 bis) e (4) o (4 bis) sono sempre in uno stesso senso, come era da prevedersi, queste ultime portando a valori superiori.

Tali differenze sono, sia per gli indici non ponderati che per i ponderati, molto maggiori per gli ultimi due caratteri che per i primi quattro: avevamo già previsto tale risultato (Cfr. pagg. 42 e 46) considerando il numero e il peso dei termini — piccolo, relativamente al totale, per i primi quattro caratteri, notevole per gli ultimi due — che nella distribuzione massimante assumevano il valore massimo. Ma, anche per i primi quattro caratteri, le differenze risultano piccole solo per gli indici lineari, mentre risultano forti per gli indici quadratici e particolarmente per i ponderati. Per i due ultimi caratteri, poi, le differenze, forti per gli indici lineari, si accentuano per i quadratici, e ciò soprattutto per i ponderati, per i quali i valori degli indici calcolati secondo le formule usuali sono dell'ordine da 1/30 a 1/60 di quelli calcolati secondo le formule corrette.

Per ciò che concerne lo scostamento semplice dalla mediana nessuna differenza si osserva tra i valori delle formule usuali e quelli delle formule corrette per gli indici non ponderati di 4 sui 6 caratteri e nessuna altresì per gli indici ponderati di 5 caratteri; ma per gli indici non ponderati della popolazione agricola e della popolazione agglomerata e per gli indici ponderati della popolazione agglomerata, una differenza, invece, si verifica, ed è forte. Ciò dipende dal fatto che, in questi casi, la mediana semplice o, rispettivamente, ponderata corrisponde, nella seriazione massimante, al limite superiore, mentre in tutti gli altri casi, essa assume il valore *0*.

Passiamo a considerare le differenze tra i risultati delle formule esatte e quelli delle approssimate (2) o (2 bis), (5) o (5 bis).

Per ciò che riguarda gli indici non ponderati, si riscontra che le formule (2 bis) danno sempre approssimazioni per difetto, come era previsto. Le differenze sono sempre piccole, così che praticamente si può dire che non vi sarebbe inconveniente a sostituire i valori esatti con gli approssimati. Le differenze sono, per i primi quattro caratteri, più sensibili per gli indici quadratici che per i lineari.

Nello stesso senso, ma più forti, sono sempre le differenze che si verificano tra i risultati esatti ottenuti dalle (3) e quelli ottenuti dalle (2) quando per k si prende il valore più basso, che indica il numero dei termini a cui nella distribuzione massimante si attribuisce il limite superiore ; anche qui le differenze sono, per i primi quattro caratteri, più sensibili per gli indici quadratici, per i quali anzi le approssimazioni non si potrebbero giudicare sufficienti agli scopi pratici.

Se, invece, nelle formule (2) si attribuisce a k il valore (più elevato del precedente di un'unità) che indica il numero dei termini cui nella distribuzione massimante non si attribuisce il valore 0, si ottengono risultati che, oltre che per lo scostamento semplice medio delle mediane, coincidono coi corretti per lo scostamento semplice medio della media aritmetica, come era da prevedersi, essendo sempre $R > A$; per gli altri indici i risultati così ottenuti sono sempre più elevati dei corretti, non di rado con approssimazioni migliori di quelle che si ottengono, in senso inverso, con le (2 bis), ma più spesso con approssimazioni meno buone.

Relativamente poi ai caratteri ponderati, l'applicazione delle formule (5) o (5 bis) fornisce, come si era preveduto, risultati praticamente identici con quelli esatti forniti dalle (4) e (4 bis) : per constatare una differenza fra gli uni e gli altri, sarebbe necessario tener conto di un maggior numero di cifre decimali. Questa pratica coincidenza conferma appunto che, anche nel caso in cui non si riesca a determinare con certezza la distribuzione massimante (per l'elevato numero di termini della data serie), basterà praticamente trovarne una della quale sia lecito presumere che sia molto prossima alla massimante : il simultaneo impiego delle (4) e (5) fornirà, a posteriori, nel caso che i loro risultati praticamente coincidano, una prova che la distribuzione assunta è molto prossima alla massimante.

Questi risultati confermano quelli raggiunti nell'applicazione esposta al numero 15 e suggeriscono la conclusione che per lo più le formule (2) e (5) e sempre le (2 bis) e (5 bis) danno approssimazioni praticamente soddisfacenti ; le formule usuali (1) e (6) al contrario portano spesso a risultati molto lontani dal vero.

Il confronto fra l'altezza dei valori ottenuti per i vari indici relativi, sia ponderati, sia non ponderati, concernenti un dato carattere, ci mostra che i valori ottenuti con le formule usuali presentano un andamento perfettamente uguale per tutti i caratteri, risultando sempre gli indici lineari più elevati dei quadratici e tra gli uni come fra gli altri l'ordine essendo sempre lo stesso.

Il più elevato fra gli indici lineari risulta lo scostamento semplice medio dalla mediana seguito, sempre a notevole distanza, dalla differenza media e dallo scostamento semplice medio dalla media aritmetica, mentre fra gli indici quadratici, che presentano valori spesso vicinissimi, lo scostamento quadratico medio dalla mediana supera lo scostamento quadratico medio dalla media aritmetica.

Sostanzialmente lo stesso è l'ordine dei vari indici ottenuti con le formule corrette o con le approssimate per ciò che concerne i primi 4 caratteri, con la sola differenza che, nella nuzialità, l'ordine dei due indici quadratici è invertito, lo scostamento dalla media aritmetica superando lievemente quello dalla mediana. Ma, per la popolazione agglomerata e per la popolazione agricola, l'ordine subisce una radicale variazione: per il primo carattere, lo scostamento semplice medio dalla mediana e la differenza media conservano i primi posti, seguìti poi dallo scostamento quadratico dalla media aritmetica, dallo scostamento semplice da detta media, e infine dallo scostamento quadratico dalla mediana. Per la popolazione agricola, invece, il valore massimo è fornito dalla differenza media seguìto dallo scostamento quadratico dalla media aritmetica e quindi dagli scostamenti lineari dalla mediana e dalla media aritmetica, e infine dallo scostamento quadratico dalla mediana. Quest'ordine risulta identico per gli indici relativi ponderati e per i non ponderati.

Più importante è il confronto fra i diversi caratteri per ciò che concerne l'intensità della loro variabilità ; e qui si palesa un netto vantaggio degli indici calcolati secondo le formule corrette o approssimate in confronto di quelli calcolati con le formule usuali.

Secondo le formule corrette o approssimate, la popolazione agglomerata risulta il carattere più variabile fra i 6 considerati ; ad essa tiene dietro la popolazione agricola e quindi l'accrescimento naturale, la natalità, la mortalità e, da ultimo, la nuzialità. Quest'ordine si conserva per tutti gli indici di variabilità con un solo piccolo spostamento per lo scostamento semplice medio dalla mediana, secondo il quale l'accrescimento naturale risulta più variabile della popolazione agricola. L'ordine risulta identico sia per gli indici ponderati sia per gli indici non ponderati.

In base alle formule usuali, il carattere più variabile risulta, invece, secondo gli indici ponderati, l'accrescimento naturale, a cui tiene dietro la popolazione agricola, la popolazione agglomerata, la natalità, la mortalità e infine la nuzialità. Questo è l'ordine risultante da tutti gli indici, con due piccole variazioni per ciò che concerne i due indici quadratici secondo i quali la nuzialità risulterebbe più variabile della mortalità. Ma questo ordine non viene conservato dagli indici non ponderati, non solo perchè le due eccezioni accennate scompaiono, ma anche perchè secondo la maggior parte degli indici, e precisamente secondo tutti e tre gli indici lineari, la popolazione agricola risulta più variabile della popolazione agglomerata.

Altro vantaggio degli indici calcolati secondo le formule corrette o approssimate, in confronto di quelli calcolati secondo le formule usuali, si manifesta nel confronto fra i valori relativi ponderati e quelli corrispondenti non ponderati, che riguardano lo stesso carattere e lo stesso indice. Con le formule corrette o approssimate, una differenza notevole si verifica unicamente per i due indici quadratici relativi alla mortalità, e una differenza sensibile per gli indici relativi alla popolazione agricola. Con le formule usuali, una differenza notevole si manifesta, invece, per gli indici quadratici relativi a tutti i caratteri, mentre anche per gli altri indici della popolazione agricola le differenze sono più elevate di quelle che si riscontrano con le formule corrette o approssimate.

18. Molte volte avviene di dover paragonare le differenze tra due coppie di quantità, per esempio tra la natalità di due paesi in uno stesso intervallo di tempo o di uno stesso paese in due intervalli diversi. Se le due quantità possono riguardarsi come due termini di una stessa serie, si è in un caso particolare del paragone tra la variabilità di due serie, in cui il numero dei termini è $n = 2$; tale è, per esempio, il caso quando si paragona la natalità del Piemonte nel 1928 con quella contemporanea della Liguria, o la natalità verificatasi in Italia nel 1928 con quella verificatasi nel 1927.

Altre volte, invece, delle due quantità, una rappresenta il valore medio o normale e l'altra un valore singolo od eccezionale; così avviene, ad esempio, se il secondo rappresenta il prodotto per ettaro durante un anno e il primo durante un decennio; oppure se il secondo rappresenta la mortalità durante la guerra e il primo la mortalità nel periodo antebellico.

In tal caso la differenza della seconda quantità dalla prima si suole considerare alla stregua di uno scostamento di un valore singolo da un valore medio. Ciò, per vero, non è sempre rigorosamente giustificato nell'ipotesi che la prima quantità sia un valore normale e la seconda un valore eccezionale, perchè nella determinazione del valore normale possono non intervenire i valori eccezionali del tipo di cui la seconda quantità fornisce un esempio: è questo il caso nel confronto tra la mortalità normale e la mortalità bellica. Tuttavia, per analogia, appare conveniente di assimilare anche questi casi a quello di uno scostamento dalla media.

Ora, molte volte, ciò che interessa di misurare è la differenza *assoluta* tra le due quantità : tale è ad esempio il caso quando si voglia misurare la differenza tra l'altezza di una montagna e l'altitudine di una pianura o altipiano su cui la montagna sorge. Ma per lo più interessa, invece, di misurare la differenza *relativa*, o, in altre parole, l'importanza della differenza, tenuto conto dell'altezza del valore normale o medio. È ovvio, per esempio, che una differenza di 10 centimetri nella statura non ha la stessa importanza per un bambino o per un adulto ; che una differenza di 10 punti nel coefficiente di mortalità non ha la stessa importanza per un paese in cui il coefficiente di mortalità è molto alto e per uno in cui esso è molto basso, e via dicendo.

Per tener conto dell'altezza del valore normale o medio, si suole generalmente ragguagliare ad esso la differenza assoluta. Se con A s'indica il valore normale o medio, che diremo *valore fondamentale*, e con a il valore singolo che con esso si confronta, la misura della differenza tra a ed A è data in tal caso dall'espressione :

$$\frac{a - A}{A} \tag{1}$$

Quando i valori di a ed A si riferiscono ad un carattere che non ammette un limite superiore, nulla vi è da eccepire a tale procedimento. Ma, se il carattere ammette un limite superiore L, e quindi un carattere complementare misurato dall'intensità $L - a$, si presenta l'obiezione che, a parte il segno contrario, l'importanza dello scostamento di a dal valore fondamentale A, misurata mediante la (1), non è equivalente all'importanza dello scostamento dal rispettivo valore fondamentale del corrispondente valore $L - a$ del carattere complementare. È infatti

$$\frac{|a - A|}{A} \gtreqless \frac{|(L - a) - (L - A)|}{L - A} = \frac{|a - A|}{L - A}$$

per $A \lesseqgtr L - A$.

Così, per esempio, l'importanza dell'aumento della mortalità verificatosi durante il periodo di guerra risulterebbe diversa dall'importanza della contemporanea diminuzione della sopravvivenza (*).

Per ovviare a tale inconveniente è necessario, come abbiamo già accennato al numero 7, che lo scostamento venga riferito ad una funzione simmetrica rispetto ad A e a $L - A$.

Rispondono a tale condizione le due formule seguenti, che ho altra volta avuto occasione di proporre (**).

$$\frac{a - A}{\sqrt{A\,(L - A)}} \tag{2}$$

$$\frac{a - A}{\dfrac{A\,(L - A)}{L}} \tag{3}$$

La prima formula risponde al criterio di riguardare come maggiore o minore la differenza — considerata come significativa — tra due quantità, a seconda che, a pari numero di osservazioni, essa avrebbe minore o maggiore probabilità di verificarsi per puro effetto del caso, qualora essa fosse non significativa, ma accidentale (***).

(*) Cfr. *Sull'aumento di mortalità determinato dalla guerra* in « Rivista italiana di Sociologia », settembre-dicembre 1916, riprodotta in *Problemi sociologici della guerra*, Zanichelli, Bologna, 1921 e, per il caso generale, il passo già citato di *Variabilità e Mutabilità*, pagg. 100-101.

(**) La formula (2) fu proposta nell'articolo citato *Sull'aumento di mortalità*, ecc. (1916) : la formula (3) fu comunicata al Prof. F. VINCI in occasione del suo articolo *La concentrazione dei capitali nelle nostre società ordinarie per azioni* (« Rivista delle Società Commerciali », marzo 1918), nel quale egli ne fece applicazione.

(***) Secondo la consueta enunciazione del teorema di BERNOULLI, la probabilità che un evento E, il quale ha una probabilità p, costante in ogni prova, si presenti in n prove, con una frequenza f, che differisca in più o in meno da p per meno di l, in modo che sia $|f - p| < l$, è data dalla formula :

$$P(\gamma) = \frac{2}{\sqrt{\pi}} \int_0^{\gamma} e^{-x^2}\, dx + \frac{e^{-\gamma^2}}{\sqrt{2\pi n p (1 - p)}} \tag{1}$$

dove è $\gamma = l \sqrt{\dfrac{n}{2 p (1 - p)}}$.

Poichè il secondo termine della somma è piccolo rispetto al primo per n sufficientemente grande, si suole limitare la formula al primo termine. È il

Era dunque, questa che si proponeva, una *convenzione*, come d'altra parte su delle convenzioni si basano tutti i sistemi di misura (*).

---

caso però di osservare che il secondo termine della somma corrisponde alla probabilità che lo scarto sia compreso entro i limiti l ed $l - \frac{1}{2}$, per modo che la formula (1), arrestata al primo termine, dà precisamente la probabilità che sia $|f - p| < l - \frac{1}{2}$.

Ora, i numeri che si incontrano nelle statistiche sono di solito arrotondati, in modo che per il numero l si intendono nel fatto tutte le infinite quantità comprese tra $l + \frac{1}{2}$ ed $l - \frac{1}{2}$. Ne viene che, nella maggior parte delle applicazioni alla statistica, volendo determinare la probabilità che uno scarto inferiore ad l si avveri per puro effetto del caso, sarà perfettamente rispondente allo scopo di applicare la formula (1) arrestata al primo termine.

Tale formula ci dice che, se si hanno due masse, ciascuna di n osservazioni, relative a due eventi E' ed E'' dotati di probabilità costanti p' e p'', ed f' ed f'' rappresentano le frequenze con cui rispettivamente nelle n osservazioni si è verificato l'evento E' e l'evento E'', vi è la stessa probabilità che nella massa degli E' si verifichi uno scarto $|f' - p'|$ tale che sia

$$\gamma > |f' - p'| \sqrt{\frac{n}{2\,p'(1 - p')}}$$

e nella massa degli E'' si verifichi uno scarto $|f'' - p''|$ tale che sia

$$\gamma > |f'' - p''| \sqrt{\frac{n}{2\,p''(1 - p'')}}.$$

Se ne deduce facilmente che vi è la stessa probabilità che si verifichi nella massa degli E' uno scarto $|f' - p'|$ e nella massa degli E'' uno scarto $|f'' - p''|$, se è

$$\frac{|f' - p'|}{\sqrt{p'(1 - p')}} = \frac{|f'' - p''|}{\sqrt{p''(1 - p'')}};$$

e quindi che vi è maggiore probabilità che si verifichi nella massa degli E' uno scarto $|f' - p'|$ o nella massa degli E' uno scarto $|f'' - p''|$ a seconda che sia

$$\frac{|f' - p'|}{\sqrt{p'(1 - p')}} \lesseqgtr \frac{|f'' - p''|}{\sqrt{p''(1 - p'')}}.$$

(*) La proposta fu oggetto di critica da parte del CANTELLI, al quale, secondo a me pare, è sfuggita la sua precisa natura. L'obbiezione del CANTELLI consiste in sostanza nel contestare la legittimità di far riferimento a un teorema del calcolo della probabilità, relativo a scarti accidentali, per inferirne un criterio di misura per scarti sistematici o riguardati come tali. Applicate a scarti sistematici, le formule del teorema di Bernoulli — egli dice — non

La seconda formula poteva venire giustificata in base alle considerazioni seguenti.

Si era osservato che il valore massimo, che può assumere la differenza $a - A$ per $a > A$, è $L - A$ e che perciò la differenza stessa, ragguagliata al suo massimo, dava

$$\frac{a - A}{L - A}$$

Si era proposto appunto di ricorrere a tale formula per misurare l'aumento di mortalità dovuto alla guerra (*). Esprimerebbe, infatti, tale formula la parte che, sul totale delle persone sopravvissute alle cause di morte normali, e quindi esposte a morire solo per le cause eccezionali di guerra, rappresentano le persone morte per cause di guerra.

In altre parole, la frazione $\frac{A}{L}$ rappresenterebbe la probabilità di morte per cause normali e la frazione $\frac{a - A}{L - A}$ la probabilità di morte per cause eccezionali dovute alla guerra (**).

---

hanno senso; tutta la questione gli appare pertanto trattata in modo arbitrario (Cfr. *Sull'aumento di mortalità dovuto alla guerra*, « Giornale degli Economisti e Rivista di Statistica » novembre 1917; per la letteratura delle discussioni che ne seguirono, cfr. la nota in *Problemi sociologici della guerra*, op. cit., pagina 94-95). Se, in luogo di usare la parola arbitrario, egli avesse usato la parola « convenzionale », saremmo d'accordo. Ora una convenzione si può accettare o non accettare. Si può accettare quando abbia un'utilità e non dia luogo a contraddizioni, come questo mi pare appunto il caso. Non si comprende una discussione sulla sua « legittimità ». L'obbiezione del CANTELLI rassomiglia, da un certo punto di vista, a quella di chi avesse contrastato la legittimità di misurare la potenza dei motori meccanici in cavalli-vapore, facendo così riferimento al lavoro compiuto da un motore animale. Evidentemente trattasi di due categorie di motori essenzialmente diversi, e in certo modo non ha senso parlare di cavalli per una macchina, ciò che non impedisce che si sia potuto per convenzione utilmente misurare la potenza dell'una in funzione del lavoro degli altri. Il carattere di « convenzione » della proposta fu bene richiamato dal BOLDRINI (*I figli di guerra*, in « Giornale degli Economisti » giugno 1919, nota (2) a pag. 293) che applicava la formula alla misura della diminuzione della natalità.

(*) Dal CANTELLI, nell'articolo ricordato. Essa veniva pure applicata dal MORTARA alla misura del grado di eccedenza della mortalità di un paese rispetto a quella di un altro (cfr. *Lezioni di statistica metodologica*, 1922, pag. 78-79).

(**) Il ragionamento del CANTELLI è stato per questa parte sottoposto a critica da parte di F. INSOLERA (*Sulla misura dell'aumento di mortalità per effetto indiretto della guerra* « Rivista Italiana di Sociologia » gennaio-giugno

Se non che, che cosa in realtà si vuole misurare? Si vuol misurare la mortalità eccezionale dovuta alla guerra, oppure l'importanza di questa rispetto alla mortalità normale? Se questo secondo è lo scopo, la misura di tale importanza relativa pare possa opportunamente ricavarsi dal rapporto della mortalità eccezionale alla mortalità normale che è dato dalla espressione:

$$\frac{\dfrac{a - A}{L - A}}{\dfrac{A}{L}}$$

la quale si riduce immediatamente alla (3).

Sia la formula (2) che la (3) ebbero parecchie applicazioni (*).

1918) e da A. T. TRAVERSI (« Journal of the Institute of Actuaries », Vol. LI, pag. 304 e segg., 1919). Per ciò che concerne la critica dell'INSOLERA, vedi la risposta del CANTELLI nella nota *Sulle applicazioni delle probabilità parziali alla statistica*, in « Giornale di Matematica finanziaria », giugno-settembre 1919; per ciò che concerne la critica del TRAVERSI, vedi la lettera di G. BALDUCCI in argomento sul « Journal of the Institute of Actuaries » (vol. LII, 1921, pagina 154) e il recente articolo del prof. C. E. BONFERRONI sullo stesso periodico (vol. LIX, Parte I, 1928, pag. 62, *On the measure of the variations of a Collectivity*).

(*) La formula (2) fu applicata dallo scrivente nel citato articolo *Sull'aumento della mortalità determinato dalla guerra* (1916), dal Prof. F. VINCI nel citato articolo *La concentrazione dei capitali*, ecc. (1917), dal Dott. A. MANCINI, in « Bollettino Mensile dell'Ufficio Municipale del Lavoro di Roma », marzo 1919, pag. 91, dal Prof. L. LIVI nell'articolo *La tubercolosi in Italia*, in « La riforma sociale », maggio-giugno 1919, dal Prof. M. BOLDRINI nel citato articolo *I figli di guerra* (giugno 1919), dal Dott. M. SAIBANTE nell'articolo *I profitti delle Società per azioni e la concentrazione dei capitali industriali*, in « Metron » Vol. VI, 1, 1926.

La formula (3) fu applicata dal Prof. F. VINCI e dal Dott. M. SAIBANTE negli articoli citati. In queste applicazioni, che si riferivano ai rapporti di concentrazione dei capitali delle società per azioni, entrambi gli autori non tennero però conto della circostanza che detti rapporti sono espressi in percentuali, per modo che anzi che applicare la formula

$$\frac{a - A}{A\,\dfrac{L - A}{L}}$$

applicarono, in sostanza, l'altra

$$\frac{a - A}{A\,(L - A)}$$

giungendo a risultati che divengono esatti solo quando si moltiplicano per il fattore costante 100.

Ora, i risultati esposti al numero 4 di questa memoria, permettono di dare delle formule (2) e (3) anche un'altra giustificazione.

La formula (3) corrisponde infatti — a parte un fattore costante uguale a 2 — al rapporto dello scostamento dalla media al valore massimo che può essere raggiunto dallo scostamento semplice medio dalla media ; la formula (2) corrisponde al rapporto dello scostamento dalla media al valore massimo che può essere raggiunto dallo scostamento quadratico medio dalla media.

Riassunto.

Il calcolo dei valori massimi degli indici assoluti di variabilità di una seriazione (scostamenti medî semplice e quadratico ${}^1S_A$, ${}^2S_A$ dalla media aritmetica, scostamenti medî semplice e quadratico ${}^1S_M$, ${}^2S_M$ dalla mediana, differenza media semplice Δ e con ripetizione Δ_R) è, nella teoria delle seriazioni statistiche, di importanza essenziale per tre ordini di considerazioni : *a*) per la migliore conoscenza del comportamento intrinseco dei detti indici ; *b*) perchè in molti casi conviene calcolare gli indici relativi di variabilità riferendo gli indici assoluti ai rispettivi valori massimi ; *c*) perchè attraverso la conoscenza di tali massimi si riesce a giustificare, per gli indici lineari di variabilità, la consuetudine di riferire gli indici stessi alla media aritmetica del carattere che ha dato origine alla seriazione.

Nella Memoria *Variabilità e Mutabilità*, la determinazione di quei massimi era stata fatta nell'ipotesi che l'ammontare totale del carattere potesse concentrarsi in un solo termine della seriazione : ma tale circostanza non può verificarsi in molti casi, e precisamente per quei caratteri la cui intensità non può oltrepassare un certo limite (come sono i quozienti di natalità, di mortalità, di nuzialità, la latitudine, la longitudine e in generale le ampiezze angolari, ecc.). Di qui l'opportunità di riprendere il problema di determinazione dei massimi in tutta la sua generalità. A questo scopo è appunto rivolta la presente memoria, che esamina, via via, tutti i casi possibili : per serie di termini aventi peso uguale o invece peso diverso ; sia nella eventualità che l'ammontare totale del carattere possa distribuirsi esattamente fra un certo numero di termini, raggiungendo in ciascuno di questi il limite superiore che gli è consentito ; sia nell'altra che, saturati alcuni termini, si abbia un resto del carattere da attribuire a un altro termine. A ciascuno dei casi possibili corrisponde un particolare gruppo di formule atte a dare rigorosamente i valori massimi dei vari indici. Le formule

più generali (3 bis) pei caratteri non ponderati (V. pag. 14) e (4 bis) pei caratteri ponderati (V. pag. 19), comprendono come caso particolare le formule valide per il caso più semplice di concentrazione della totalità dei caratteri in un solo termine della seriazione. È anche mostrato se e quando, ai fini dell'approssimazione che si vuole raggiungere, sia lecito, per il calcolo dei massimi, sostituire alle formule corrispondenti a un certo tipo di distribuzione, quelle corrispondenti a un altro tipo.

Il lavoro è corredato da due serie di applicazioni di cui la seconda — riferendosi a seriazioni più numerose e nelle quali è relativamente maggiore il numero dei termini su cui può riversarsi la totalità del carattere — è meglio atta a lumeggiare le analogie e le differenze che, in base ai vari tipi di formule, risultano sia pei massimi degli indici assoluti sia per gli indici relativi ottenuti come quozienti degli indici assoluti per quei massimi. Da queste applicazioni risulta confermata la previsione teorica che, specie per gli indici quadratici, l'impiego delle formule usuali, che presuppongono la concentrazione di tutti i caratteri in un termine, può condurre a errori veramente enormi ; e risulta altresì confermato che, nella maggioranza dei casi, e cioè quando il carattere possa riversarsi su diversi termini, si possono praticamente impiegare le formule (2 bis) pei caratteri non ponderati (Vedi pag. 6) e le (5 bis) per i ponderati (V. pag. 20), le quali forniscono ottime approssimazioni, ed hanno, presso a poco, lo stesso grado di semplicità delle formule usuali.

Passando, infine, a considerare il caso di una particolare seriazione composta di due soli termini e quello, più generale, di due quantità tra cui si vuol misurare la disuguaglianza, vengono nuovamente giustificate, con le risultanze teoriche della Memoria, le formule già proposte dall'A. in altra occasione per la misura dello scostamento relativo dell'intensità di un carattere dal suo valore fondamentale, quando il carattere stesso ammetta un limite superiore.

Summary.

The calculation of the maximum values of the absolute indices of variability of a seriation (simple mean deviation and quadratic mean deviation from arithmetic mean S_A, ${}^2S_A$; simple mean deviation and quadratic mean deviation from median S_M, ${}^2S_M$; simple mean difference Δ and with repetition Δ_R) is of essential importance in the theory of statistical series, for three reasons : first, for

a better knowledge of the intrinsic behaviour of mentioned indices; second, for the convenience, in many cases, of calculating the relative indices of variability referring the absolute indices to the respective maximum values; third, for the possibility to justify, through the knowledge of such maximums, in relation to the lineal indices of variability, the use of referring the same indices to the arithmetic mean of the character which has given origin to the seriation.

In the work *Variabilità e Mutabilità* the determination of those maximums, has been made in the hypothesis that the total amount of the character might concentrate in one term of seriation only; but such a circumstance, might not be verified in many cases, and precisely, for those characters whose intensity cannot go further than a certain limit, as for instance the quotients of natality, mortality, nuptiality; the latitude, longitude, and in general, the angular extensions, etc. Hence the opportunity to take up again the problem of determination of maximums on the whole.

The present work, examining by and by all possible cases, by series of terms having equal weigh or different weigh, just suits to this purpose: as in the case of exact distribution of total amount of character among a certain number of terms, reaching in each of them its due upper limit; as in the case, when some terms filled up, there is a rest of character left to be applied to another term. To each of possible cases, corresponds a particular group of formules suitable to give strictly the maximum values of the various indices.

The most general formules (3bis) for unweighted characters (see page 14) and (4bis) for weighted characters (see page 19) include, as a particular case, the formules which refer to the simplest case of concentration of the totality of characters in one term of seriation only.

It is also shown, whether and when, aiming at the reachable approximation wanted, it be convenient, for the calculation of the maximums, to substitute to the formules corresponding to a certain type of distribution, those corresponding to another type.

The work is completed by two series of applications. The second, which refers to more numerous seriations, where the number of terms on which it is possible to apply the totality of character, is relatively greater, is more able to lighten the analogies and the differences which on the basis of the various types of formules, are resulting for the maximums of asbsolute indices, and for relative indices obtained as a quotient of absolute indices for these maximums.

Through these applications the theoretical prevision is confirmed that, mainly for the quadratic indices, the employment of the usual formules, which presuppose the concentration of all characters on one term only, can really lead to enormous mistakes. It is also confirmed that in most cases, namely when the character might apply to various terms, it is practically possible to employ the formules (2bis) for the unweighted (see page 4) and the (5bis) for the weighted characters (see page 20), which afford best approximation and have, nearly, the same degree of simplicity of the usual formules.

Passing, at last, to consider the case of a particular seriation formed with two terms only, and the more general case of two quantities, between whom the inequality ought to be measured, the theoretical results of the work, justify the formules already suggested by the A. in another opportunity for the measure of relative deviation of intensity of a character from its fundamental value, when the character itself is susceptible to have an upper limit.

G. A. BAKER

Random sampling from non-homogeneous populations.

It is the purpose of this paper to develop the theory of the distributions of estimates of statistical constants by means of random samples from non-homogeneous populations of one variate to a basis that is comparable to that already developed for homogeneous populations.

I. — Combinations of normal populations.

We shall first consider combinations of normal populations, as that affords a convenient point of departure.

The sum of two normal populations can obviously be represented as

$$f(x) = \frac{N_1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} + \frac{N_2}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(x-m)^2}{\sigma^2}} \tag{a}$$

1. *Distribution of the Means of Samples of* n *Drawn from the Combination of Normal Populations.*

We shall prove at this point the following lemmas.

Lemma I. If x and y are independent variables $0 \leq x \leq \infty$, $0 \leq y \leq \infty$ and the probability of an x in dx is $f(x)\,dx$ and the probability of a y in dy is $\varphi(y)\,dy$, then the probability of a value of their sum $v = x + y$ being between v and $v + dv$ is proportional to

$$\left[\int_0^v f(x)\,\varphi(v-x)\,dx\right] dv$$

and

$$z = f(x)\,\varphi(y)$$

represent a surface in three dimensional space. Then

$$x + y = v$$

considering v as constant, is a plane parallel to the z-axis. Now the probability of a v in dv is proportional to the content of the space enclosed between $z = o$, $z = f(x)\,\varphi(y)$, $x + y = v$, and $x + y = v + dv$ which is

$$\left[\int_{o}^{v} \sqrt{2}\, f(x)\, \varphi(v - x)\, d\,x\right] d\,v$$

which proves the lemma.

An obvious extension of Lemma I is

Lemma II. If x and y are independent variables, $-\infty \leqq x \leqq \infty$, $-\infty \leqq y \leqq \infty$ and the probability of an x in dx is $f(x)\,d\,x$ and the probability of a y in dy is $\varphi(y)\,dy$ then the probability of a $v = x + y$ in dv is proportional to

$$\left[\int_{-\infty}^{\infty} f(x)\, \varphi(v - x)\, d\,x\right] d\,v$$

Similar lemmas could be stated for other ranges of the independent variables by making proper changes in the limits of integration but these two are the ones of greatest importance in what follows.

Suppose we consider a population represented by

$$f(x) = \frac{N_1}{\sigma_1 \sqrt{2\pi}}\, e^{-\frac{1}{2}\frac{(x+m_1)^2}{\sigma_1^2}} + \frac{N_2}{\sigma_2 \sqrt{2\pi}}\, e^{-\frac{1}{2}\frac{(x-m_2)^2}{\sigma_2^2}} \tag{b}$$

The first two moments of this population about its mean will be

$$u_0 = N_1 + N_2 \tag{1}$$

$$u_1 = -N_1 m_1 + N_2 m_2 \tag{2}$$

Now we wish to estimate u_1, by means of samples of n drawn from (b) We may consider separately the $n + 1$ cases in which we get $n - s$ individuals from the first component and s individuals from the second. The probability of getting $n - s$ individuals from the first component and s individuals from the second is proportional to

$$\binom{n}{s} k^s \tag{3}$$

where

$$k = \frac{N_2}{N_1}$$

Now consider that we are estimating u_1 by means of $n - s$ individuals drawn from the first component and s from the second. Then (2) becomes.

$$\overline{u}_1 = -(n - s)\,\overline{m}_1 + s\,\overline{m}_2 \tag{4}$$

where the bar over a letter denotes an estimate of this parameter.

Taking the representation of the population in the form (a) we have by Lemma II

$$\varphi(y) = c_1\, e^{-\frac{1}{2}\frac{y^2}{n-s}}$$

$$f(x) = c_2\, e^{-\frac{1}{2}\frac{(x - s\,m)^2}{s\,\sigma^2}}$$

Whence

$$\int_{-\infty}^{\infty} c_1\, c_2\, e^{-\frac{1}{2}\frac{(x - s\,m)^2}{s\,\sigma^2}}\, e^{-\frac{1}{2}\frac{(v - x)^2}{n-s}}\, d\,x$$

is proportional to

$$l^{-\frac{1}{2}\frac{(v - \cdot\, m)^2}{[s\,\sigma^2 + n - s]}}$$

Where v is n times the mean of samples of n. This gives the distribution of the totals of samples of n drawn from a population represented by (a) on the hypothesis that $n - s$ come from the first component and s from the second. To get the total probability of a value of a total of a sample of n we must sum on s and multiply by a suitable constant to make to $s + 1^{th}$ term total up to $\binom{n}{s} k^s$ as the mean ranges from $-\infty$ *to* $+\infty$. That is, we will have that the distribution of the totals of samples of n drawn from a population represented by (a) will be proportional to

$$\sum_{s=0}^{s=n} \binom{n}{s} \frac{k^s}{\sqrt{s\,\sigma^2 + n - s}}\, e^{-\frac{1}{2}\frac{(v - s\,m)^2}{[s\,\sigma^2 + n - s]}} \tag{5}$$

or that the means will be distributed as proportional to

$$\sum_{s=0}^{s=n} \binom{n}{s} \frac{k^s}{\sqrt{s\,\sigma^2 + n - s}}\, e^{-\frac{(n\,x - s\,m)^2}{2\,[s\,\sigma^2 + n - s]}} \tag{6}$$

In case we have the difference of two normal populations which can be represented by (a) regarding k replaced by $-k$ and subject to the condition that there is alwyas a positive frequency (6) becomes

$$\sum_{s=0}^{s=n} \binom{n}{s} \frac{(-1)^s k^s}{\sqrt{s\sigma^2 + n - s}} e^{-\frac{(nx - sm)^2}{2[s\sigma^2 + n - s]}} \tag{7}$$

PEARSON has suggested an interpretation for this state of affairs (see 9).

Suppose that we have the following combination of p normal populations

$$f(x) = \frac{\pm 1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} + \frac{k_2}{\sigma_2\sqrt{2\pi}} e^{-\frac{(x-m_2)^2}{2\sigma_2^2}} + - - - - +$$

$$+ \frac{k_p}{\sigma_p\sqrt{2\pi}} e^{-\frac{(x-m_p)^2}{2\sigma_p^2}} \tag{8}$$

Where the k_i' s may be either plus or minus but, nevertheless, subject to the restriction that $f(x)$ shall be positive throughout its entire range. Then we find that the means of samples of n drawn from a population represented by (8) will be distributed as proportional to

$$\sum \frac{n!}{v_1! \ldots v_p!} \frac{(\pm 1)^{v_1} k_2^{v_2} \ldots k_p^{v_p}}{\sqrt{v_1 + v_2\sigma_2^2 + \ldots + v_p\sigma_p^2}} e^{-\frac{(nx - v_2 m_2 - \ldots - v_p m_p)^2}{2[v_1 + v_2\sigma_2^2 + \ldots + v_p\sigma_p^2]}}$$

summed for all integral non-negative values of the v_i s such that

$$v_1 + v_2 + \ldots + v_p = n$$

where v_i represents the number of individuals obtained from the i^{th} component.

2. The Distribution of the Standard Deviations of Samples of n *Drawn from Combinations of Normal Populations.*

Consider a population represented by (b). Its second moment about its mean will be

$$u_2 = N_1(\sigma_1^2 + m_1^2) + N_2(\sigma_2^2 + m_2^2) \tag{1}$$

Let us form the expression

$$w = N_1(\sigma_1^2 + m_1^2) + N_2(\sigma_2^2 + m_2^2) - \frac{(-N_1 m_1 + N_2 m_2)^2}{N_1 + N_2} =$$

$$= N_1\sigma_1^2 + N_2\sigma_2^2 + \frac{N_1 N_2}{N_1 + N_2}(m_1 + m_2)^2 \tag{2}$$

Now estimates of this quantity w will be estimates of n times the standard deviation squared of the population by means of samples of n. Let us consider, as before, that we draw $n - s$ individuals from the first component and s from the second. Then (2) becomes

$$\bar{w} = (n - s)\,\bar{\sigma}_1^2 + s\,\bar{\sigma}_2^2 + \frac{(n-s)\,s}{n}\,(\bar{m}_1 + \bar{m}^2)^2 \tag{3}$$

where the bar means that these quantities are estimates of the true parameters of the populations. If $n - s = 0$ we have only the second term of (3) left, if $s = 0$ we have only the first term of (3) left, if $n - s = 1$ we have the second and last terms of (3) left, if $s = 1$ we have the first and last terms of (3) left. Further, if $n = 2$, $n - s = s = 1$, (3) reduces to

$$\bar{w} = \frac{1}{2}\,(\bar{m}_1 + \bar{m}_2)^2$$

which is exceptional.

Bearing these facts in mind and making use of Lemma I, we may state the following theorem.

Theorem I. — If samples of n, $n > 2$ are drawn from the population represented by (a) then the standard deviations of these samples are distributed as proportional to

$$c_0\, x^{n-2}\, e^{-\frac{n x^2}{2}} + c_1\, n\, k\, x\, e^{-\frac{1}{2} n x^2} \int_0^{n x^2} \frac{(n x^2 - z)^{\frac{n-1-3}{2}}}{\sqrt{z}}\, e^{\sqrt{\frac{n-1}{n}} m z}\, dz +$$

$$+ \sum_{s=2}^{s=n-2} \binom{n}{s} k^s\, c_s\, x\, \psi_s(n x^2) +$$

$$+ n\, k^{n-1}\, c_{n-1}\, x \int_0^{n x^2} \frac{(n x^2 - z)^{\frac{n-1-3}{2}}}{\sqrt{z}}\, e^{-\frac{1}{2}\frac{(n x^2 - z)}{\sigma^2}}\, e^{-\frac{1}{2}\left(\sqrt{z} - \sqrt{\frac{n-1}{n}} m\right)^2} dz +$$

$$+ k^n\, c_n\, x^{n-2}\, e^{-\frac{n x^2}{2 \sigma^2}}$$

where

$$\psi_s(n x^2) =$$

$$\int_0^{n x^2} \left[\frac{e^{-\frac{1}{2}\left(\sqrt{n x^2 - v} - \sqrt{\frac{s(n-s)}{n}} m\right)^2}}{\sqrt{n x^2 - v}} \int_0^{v} y^{\frac{n-s-3}{2}} (v - y)^{\frac{s-3}{2}}\, e^{-\frac{y}{2}}\, e^{-\frac{v-y}{2\sigma^2}}\, dy\right] dv$$

and the $c_i's$ are constants such that the integral from 0 to ∞ of the $s + 1^{th}$ term is $\cdot \binom{n}{s} k^s$.

In case $n = 2$ both the results for the means and standard deviations can be checked by changing to the mean and standard deviation as variables following the method used by R. A. FISHER and KARL PEARSON in the case the sampled population was normal. (See 3 and 4).

Any other combination of normal populations may be handled in the same way but the results are rather complicated to writer out in full.

3. *Distributions of the Mean Divided by the Standard Deviation and of the Higher Moments about the Mean of the Sample.*

Consider a population represented by (b) and that we get $n - s$ individuals from the first component and s from the second in drawing samples of n. Then if we denote the means of the samples by $\bar{m}$ the standard deviations of the samples by $\bar{\sigma}$ the m^{th} moment of the sample about its mean by $\bar{u}_m$, we will have the following expressions for the estimation of the abve mentioned quantities.

$$\frac{\bar{m}}{\bar{\sigma}} = \frac{-(n-s)\,\bar{m}_1 + s\,\bar{m}_2}{\left[(n-s)\,\bar{\sigma}_1{}^2 + s\,\bar{\sigma}_2{}^2 + \frac{(n-s)\,s}{n}\,(\bar{m}_1 + \bar{m}_2)^2\right]^{\frac{1}{2}}} \tag{1}$$

$$\bar{u}_3 = \frac{1}{n^2}\left[-3\,n\,(n-s)\,s\,\bar{\sigma}_1{}^2 + 3\,n\,(n-s)\,s\,\bar{\sigma}_2{}^2 + \right.$$
$$\left. + (n-s)\,s\,(n-2\,s)\,(\bar{m} + \bar{m}_2)^3\right] \tag{2}$$

$$\bar{u}_4 = \frac{1}{n^3}\left[3\,n^3\left\{(n-s)\,\bar{\sigma}_1{}^4 + s\,\bar{\sigma}_2{}^4\right\} + 6\,(n-s)\,s^2\,n\,(\bar{m}_1 + \bar{m}_2)^2\,\bar{\sigma}_1{}^2 + \right.$$
$$\left. + 6\,(n-s)^2\,s\,n\,(\bar{m}_1 + \bar{m}_2)^2\,\bar{\sigma}_2{}^2 + (n-s)\,s\,(n^2 - 3\,n\,s + s^2)\,(\bar{m}_1 + \bar{m}_2)^4\right] \tag{3}$$

$$\bar{u}_5 = \frac{1}{n^4}\left[15\,n^3\,(n-s)\,s\,(\bar{m}_1 + \bar{m}_2)\,(\bar{\sigma}_2{}^4 - \bar{\sigma}_1{}^4) - \right.$$
$$- \left\{\bar{m}_1{}^3\left[20\,(n-s)^5 + 80\,(n-s)^4\,s + 120\,(n-s)^3\,s^2 + 90\,(n-s)^2\,s^3 + \right.\right.$$
$$\left. + 30\,(n-s)\,s^4\right] + 10\left[(n-s)^2\,s^3 + (n-s)\,s^4\right]\,(3\,\bar{m}_1{}^2\,\bar{m}_2 + 3\,\bar{m}_1\,\bar{m}_2{}^2 +$$
$$\left. + \bar{m}_2{}^3\right\}\,\sigma_1{}^2 + \left\{\bar{m}_2{}^3\left[20\,s^5 + 30\,(n-s)^4\,s + 90\,(n-s)^3\,s^2 + 120\,(n-s)^2\,s^3 + \right.\right.$$
$$\left. + 80\,(n-s)\,s^4\right] + 10\left[(n-s)^4\,s + (n-s)^3\,s^2\right]\,(3\,\bar{m}_2{}^2\,\bar{m}_1 + 3\,\bar{m}_2\,\bar{m}_1{}^2 +$$
$$\left. + \bar{m}_1{}^3)\right\}\,\bar{\sigma}_2{}^2 + \left[(n-s)^4\,s - (n-s)^3\,s^2 + (n-s)^2\,s^3 - \right.$$
$$\left.\left. - (n-s)\,s^4\right]\,(\bar{m}_1 + \bar{m}_2)^5\right] \tag{4}$$

Now $\bar{m}_1$, $\bar{m}_2$, $\bar{\sigma}_1^2$, $\bar{\sigma}_2^2$ are independent variables and we know the distribution of each of them. In the case of (1), (2), and (3), we can regard $\bar{m}_1 + \bar{m}_2$, $\bar{\sigma}_1^2$, $\bar{\sigma}_2^2$ as our independent variables. Let the frequency surface of these three variables be denoted by

$$z = f\left[(\bar{m}_1 + \bar{m}_2), \bar{\sigma}_1^2, \bar{\sigma}_2^2\right] \tag{5}$$

Then in a space of four dimensions (1), (2), (3) equated to a quantity v, say, will be hypercylinders, and the probability of a value of v between v and $v + dv$ will be proportional to the content of the space enclosed between $z = 0$, $z = f\left[(\bar{m}_1 + \bar{m}_2), (\bar{\sigma}_1^2\ \bar{\sigma}_2^2)\right]$, (1), (2), or (3) $= v$ and (1) or (2) or (3) $= v + dv$. For small values of n particularly and for certain values of $n - s$ and s, in general, we will have degenerate cases as some of the variables will drop out.

In case of (4) we have similar considerations but we must now use $\bar{m}_1$, $\bar{m}_2$, $\bar{\sigma}_1^2$, $\bar{\sigma}_2^2$ as independent variables making necessary the use of five dimensional concepts.

II. — Combinations of other populations.

The foregoing theory enables us, in a simple manner, to specify the distribution of the means of samples of n drawn from any population which is the combination of populations of which we know the distribution of the means of samples of n whether the component populations are of the same type or not. Further, if we know the distribution of the higher moments about a fixed point of each of the components we can deduce the distribution of the higher moments of samples from the resulting population about the same fixed point. This is so because the moments of samples of n drawn from the resulting population are the sums of the moments of samples drawn from the component populations.

A method will be given presently for the calculation of the distributions of the moments of any order about a fixed point of samples of n drawn from any continuous infinite population with finite moments.

We shall now discuss in some detail two other combinations of populations.

1. *Combinations of Rectangular Distributions.*

Suppose we have the following combination of two rectangular distributions

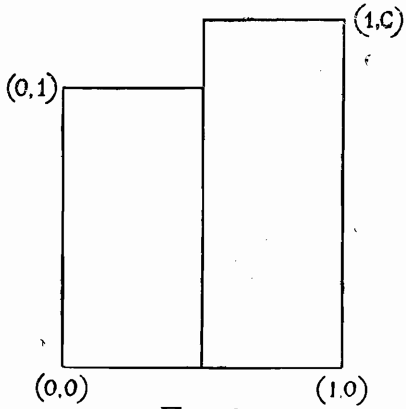


Fig. 1

The distribution of the means of samples drawn from (1) may be deduced by the method of the previous sections but the sampling from (1) is very much simpler and more instructive considered as an extension of HALL's article on sampling from a rectangular distribution (See 5).

We will discuss the case in which samples of two are drawn from (1) and it will then be apparent, from this discussion and HALL's results, what will be true in the general case.

Samples of two drawn from (1) will fill the following space with the designated densities

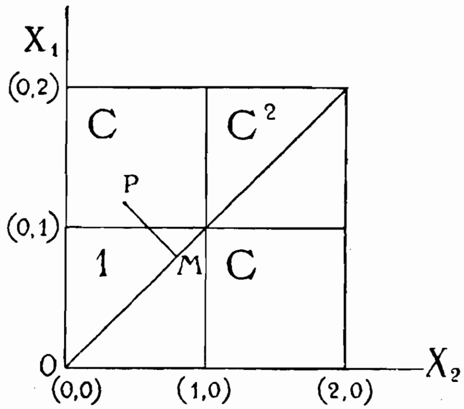


Fig. 2

$$\text{Mean of the sample} = \frac{1}{2}(x_1 + x_2) = \frac{OM}{\sqrt{2}}$$

$$\text{Standard Deviation of the sample} = \frac{1}{2}(x_1 - \bar{x})^2 + \frac{1}{2}(x_2 - \bar{x})^2 =$$

$$= \frac{MP}{\sqrt{2}}.$$

Thus the probability of a value of the mean will be given as proportional to the content of the lines

$$x_1 + x_2 = 2\bar{x}$$

That is the means of samples of two drawn from (1) will be distributed as proportional to

$$\begin{aligned}
f(\bar{x}) &= \bar{x} & , 0 \leq \bar{x} \leq \frac{1}{2} \\
&= -\bar{x} + 1 + 2c\,[2\bar{x} - 1] & , \frac{1}{2} \leq \bar{x} \leq 1 \\
&= c^2\,[\bar{x} - 1] + 2c\,[-2\bar{x} + 3], & 1 \leq \bar{x} \leq 1\frac{1}{2} \\
&= c^2\,[-\bar{x} + 2] & , 1\frac{1}{2} \leq \bar{x} \leq 2
\end{aligned}$$

The probability of a value of the standard deviation will be given as proportional to the content of the two lines parallel to $x_1 = x_2$ and at a distance $\sqrt{2}\,r$ from it, r being the value of the standard deviation. Thus the distribution of the standard deviations will be proportional to

$$\begin{aligned}
f(r) &= 2\,[-4r + 2\sqrt{2} + c^2(-4r + 2\sqrt{2}) + 2c\,2r] & , 0 \leq r \leq \frac{1}{\sqrt{2}} \\
&= 2c\sqrt{2}\,(2 - \sqrt{2}\,r) & , \frac{1}{\sqrt{2}} \leq r \leq \sqrt{2}
\end{aligned}$$

In general if samples of n are drawn from (1) they will fill a space consisting of 2^n n-dimensional cubes $\binom{n}{s}$ of which will have a density c^s. Thus we see that the distribution of the means will be proportional to $n+1$ curves, the area under the $s+1^{th}$ being $\binom{n}{s} c^s$, each consisting of n arcs of degree $n-1$. The distribution of standard deviations will be proportional to the content of the surfaces of hypercylinders inside the cube made up of the 2^n small cubes, the axis of these hypercylinders would be the main diagonal of the large cube.

If the ranges of our two rectangular distributions are unequal we would have rectangles, and rectangular n-dimensional spaces to deal with instead of cubes.

If the ranges were overlapping the n-dimensional cubes or rectangular spaces would also overlap.

2. *Combinations of Type III Populations.*

Suppose that we have a population represented by

$$F(x) = y_0 \left(1 + \frac{x}{a}\right)^{\gamma a} e^{-\gamma x} + y_1 \left(1 + \frac{x}{b}\right)^{\lambda b} e^{-\lambda x} \qquad (1)$$

Now the distribution of the totals of samples of $n-s$ drawn from the first component (See (2) and (7)) is proportional to

$$\left(n - s + \frac{x}{a}\right)^{(n-s)(\gamma a + 1) - 1} e^{-\gamma x} \qquad (2)$$

and the distribution of totals of s drawn from the second component is proportional to

$$\left(s + \frac{y}{b}\right)^{s(\lambda b + 1) - 1} l^{-\lambda x} \qquad (3)$$

Hence by considerations similar to Lemma I we find that the means of samples of n drawn from (1) are distributed as proportional to

$$c_0 \left(1 + \frac{x}{a}\right)^{n(\Upsilon a + 1) - 1} e^{-\Upsilon n x} +$$

$$\sum_{s=1}^{s=n-1} c_s \binom{n}{s} k^s \left[e^{-\lambda n x} \int_{-a}^{nx+b} \left(n - s + \frac{y}{a}\right)^{(n-s)(\Upsilon a + 1) - 1} \right.$$

$$\left. \left(s + \frac{y - nx}{b}\right)^{s(\lambda b + 1) - 1} e^{-(\Upsilon - \lambda)} dy \right]$$

$$+ c_n k^n \left(1 + \frac{x}{b}\right)^{n(\lambda b + 1) - 1} e^{-\lambda n x} \qquad (4)$$

Where k is the ratio of the area of the second component to the area of the first and c_s is such that the integral from α_s to β_s of the $s + 1^{th}$ term is $\binom{n}{s} k^s$ and (α_s, β_s) denotes the range of x for the $s + 1^{th}$ term.

III. — Distributions of the moments of samples of n about a fixed point.

This is an obvious extension of J. O. Irwin's general method for finding the distributions of the totals of samples n drawn from a continuous infinite population with finite moments. (See 7).

If we consider a discrete population in which a value x_i of a character measured from the mean of the population is replaced by x_i^m, the probability of an x_i^m being the same as the probability of an x_i in the original population we are led at once to the following theorem.

Theorem II. — If $f(x)$ is continuous, $a \leq x \leq b$, with finite moments and with its origin at its means, then the distribution of the m_{th} moment of samples of n drawn from the population represented by $f(x)$ and measured from its mean is given as proportional to the solution of the integal equation

$$F_m(z) = \int_{\alpha}^{\beta} \psi(x)\, e^{zx}\, dx$$

where if $m = 2k$

$$\alpha = 0$$

$$\beta = \text{larger of } n a^m,\ n b^m$$

and if $m = 2k + 1$

$$\alpha = n a^m$$

$$\beta = n b^m$$

and where

$$F_m(z) = \left(\int_a^b f(x)\, e^{x^m z}\, dx\right)^n$$

if such solution exists.

For example, if $m = 2$

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2},\ a = -\infty,\ b = \infty \tag{1}$$

by the method used by IRWIN we find that $\psi(x)$ is proportional to

$$\frac{x^{\frac{n-2}{2}} e^{-\frac{1}{2}x}}{\left(\frac{n-2}{2}\right)!}$$

if n is even.

In case the sampled population is normal we may find the distribution of the second moments of samples of n about the mean of the original population very simply in another way.

Suppose we have a population represented by (1). Now an estimate of the second moment in terms of the estimated mean and standard deviation of the sample is given by

$$\bar{u}_2 = n\,\bar{\sigma}^2 + n\,\bar{x}^2 \tag{3}$$

Whence by Lemma I $\bar{u}_2$ will be distributed as proportional to

$$\bar{u}_2^{\frac{n-2}{2}} e^{-\frac{1}{2}\bar{u}_2} \tag{4}$$

if n is odd.

The distribution for the second moment of samples drawn from the combination of normal populations can also be written down at once by means of the theory developed in the preceding sections even though it is very difficult to obtain this distribution as a solution of the integral equation of Theorem II.

IV. — Experimental test of the theory.

To test whether the developed theory actually represented the facts or not we drew samples of four from populatons which were the sum of two normal populations, or were the difference of two normal populations, by means of throwing dice.

1. *Test of means.*

The distributions of the means were tested in the case of five populations, 1038 samples of four being drawn from each population and their means calculated. The results were as follows:

Case I. — The population was represented by

$$f(x) = \frac{648}{5\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(x-15.5)^2}{25}} + \frac{648}{5\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(x-32.5)^2}{25}}$$

Thus, theoretically, the means were distributed as

$$y = \frac{519}{8\sqrt{2\pi}} \left[e^{-\frac{1}{2}x^2} + 4e^{-\frac{1}{2}\left(x-\frac{17}{10}\right)^2} + \right.$$

$$\left. + 6e^{-\frac{1}{2}\left(x-\frac{17}{5}\right)^2} + 4e^{-\frac{1}{2}\left(x-\frac{51}{10}\right)^2} + e^{-\frac{1}{2}\left(x-\frac{34}{5}\right)^2} \right]$$

For agreement between actual observations and theory see Table I. In this table and in all subsequent tables the class interval labelled by a whole number, say, a includes all individuals with a value of a character x such that $a \leqq x \angle a + 1$.

TABLE I.

Size	Actual number	Theoretical number	Size	Actual number	Theoretical number
8	0	.30198	25	83	84.10123
9	1	.59876	26	69	77.31134
10	3	1.42701	27	54	65.73126
11	1	2.90408	28	61	54.65654
12	8	5.04666	29	49	42.96048
13	6	9.09819	30	44	32.29471
14	8	12.51037	31	31	24.01954
15	19	17.18900	32	31	17.18900
16	28	24.01954	33	15	12.51037
17	33	32.29471	34	12	9.09819
18	34	42.96048	35	6	5.04666
19	49	54.65654	36	2	2.90408
20	67	65.73126	37	2	1.42701
21	69	77.31134	38	0	.59876
22	75	84.10123	39	0	.215234
23	90	88.12886	40	1	.087744
24	87	88.12886			

$$X^2 = 47.6093$$

$$P = .037$$

Case II. — In this case the population was

$$f(x) = \frac{864}{10\sqrt{2\pi}} e^{-\frac{(x-30.5)^2}{100}} + \frac{432}{5\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(x-41.5)}{25}}$$

The distribution of the means is given by

$$y = \frac{328}{\sqrt{2\pi}}\left[\frac{1}{2}e^{-\frac{1}{2}\frac{x^2}{\frac{1}{4}}} + \frac{4}{\sqrt{13}}e^{-\frac{1}{2}\frac{\left(x-\frac{11}{40}\right)^2}{\frac{13}{64}}} + \right.$$

$$+ \frac{3\sqrt{2}}{2\sqrt{5}}e^{-\frac{1}{2}\frac{\left(x-\frac{11}{20}\right)^2}{\frac{5}{32}}} + \frac{1}{\sqrt{7}}e^{-\frac{1}{2}\frac{\left(x-\frac{33}{40}\right)^2}{\frac{7}{64}}} +$$

$$\left. + \frac{1}{16}e^{-\frac{1}{2}\frac{\left(x-\frac{11}{10}\right)^2}{\frac{1}{16}}}\right]$$

For agreement between actual observations and theory see Table II.

TABLE II.

Size	Actual number	Theoretical number	Size	Actual number	Theoretical number
18	1	2.5218	33	75	78.9410
19	5	1.8132	34	93	82.6969
20	3	2.8901	35	94	81.5593
21	3	4.4567	36	83	77.5671
22	7	6.6481	37	77	72.0152
23	7	9.8527	38	64	61.8085
24	13	13.8432	39	47	51.3804
25	13	18.9288	40	40	41.1012
26	29	25.1300	41	31	30.5722
27	28	32.5217	42	26	22.1918
28	26	40.7728	43	12	13.2239
29	42	49.4037	44	8	8.0433
30	58	58.2457	45	5	4.4384
31	55	67.2530	46	6	2.3238
32	82	73.7898	47	5	1.9596

$X^2 = 36.5453$
$P = .17$

Case III. — In this case the population is

$$f(x) = \frac{972}{5\sqrt{2\pi}}e^{-\frac{1}{2}\frac{(x-15.5)^2}{25}} + \frac{324}{5\sqrt{2\pi}}e^{-\frac{1}{2}\frac{(x-31.5)^2}{25}}$$

Whence the distribution of the means of samples of four is

$$y = \frac{657}{\sqrt{2\pi}}\left[\frac{1}{2}e^{-\frac{1}{2}x^2} + \frac{2}{3}e^{-\frac{1}{2}\left(x-\frac{8}{5}\right)^2} + \frac{1}{3}e^{-\frac{1}{2}\left(x-\frac{16}{5}\right)^2} + \right.$$
$$\left. + \frac{2}{27}e^{-\frac{1}{2}\left(x-\frac{24}{5}\right)^2} + \frac{1}{162}e^{-\frac{1}{2}\left(x-\frac{32}{5}\right)^2}\right]$$

For agreement between actual observations and theory see Table III.

TABLE III.

Size	Actual number	Theoretical number	Size	Actual number	Theoretical number
7	1	.4428	21	76	79.4227
8	1	1.0861	22	64	68.3437
9	2	3.0314	23	66	56.9442
10	8	7.2248	24	46	45.4676
11	12	15.2942	25	40	34.9127
12	21	27.0010	26	24	25.7212
13	45	41.9642	27	25	18.2412
14	52	57.6956	28	15	12.2448
15	63	71.9240	29	7	8.1507
16	91	82.8926	30	5	5.1967
17	98	90.5719	31	3	3.2399
18	93	94.5360	32	2	1.7897
19	94	94.0163	33	1	1.0216
20	80	88.6049	34	3	1.0170

$X^2 = 17.3511$

$P = .92$

Case IV. — In this case the population was

$$f(x) = \frac{1728}{10\sqrt{2\pi}}e^{-\frac{1}{2}\frac{(x-30.5)^2}{100}} - \frac{432}{5\sqrt{2\pi}}e^{-\frac{1}{2}\frac{(x-35.5)^2}{25}}$$

Whence the distribution of the means of samples of four is

$$y = \frac{32811}{\sqrt{3\pi}}\left[\frac{1}{2}e^{-\frac{1}{2}\frac{x^2}{\frac{1}{4}}} - \frac{2}{\sqrt{13}}e^{-\frac{1}{2}\frac{\left(x-\frac{1}{8}\right)^2}{\frac{13}{64}}} + \right.$$
$$+ \frac{3\sqrt{2}}{8\sqrt{5}}e^{-\frac{1}{2}\frac{\left(x-\frac{1}{4}\right)^2}{\frac{5}{32}}} - \frac{1}{8\sqrt{7}}l^{-\frac{1}{2}\frac{\left(x-\frac{3}{8}\right)^2}{\frac{7}{64}}} +$$
$$\left. + \frac{1}{256}l^{-\frac{1}{2}\frac{\left(x-\frac{1}{2}\right)^2}{\frac{1}{16}}}\right]$$

For agreement between actual observations and theory see Table IV.

TABLE IV.

Size	Actual number	Theoretical number	Size	Actual number	Theoretical number
14	1	2.7924	30	91	70.3738
15	1	2.4432	31	71	70.0778
16	7	4.2107	32	61	60.2234
17	7	6.8703	33	49	53.2194
18	8	10.7324	34	38	50.0532
19	17	15.9725	35	31	37.5582
20	27	22.6731	36	19	30.1810
21	26	29.4506	37	19	25.3298
22	46	38.9797	38	13	17.2839
23	48	44.8890	39	12	12.6019
24	64	55.5311	40	11	8.9075
25	72	62.5850	41	5	6.4679
26	74	68.0261	42	0	3.8504
27	75	71.5158	43	2	2.5067
28	74	75.8491	44	1	3.6774
29	64	73.0864			

$X^2 = 36.0159$
$P = .23$

Case V. — In this case the population was

$$f(x) = \frac{648}{5\sqrt{2\pi}}e^{-\frac{1}{2}\frac{(x-15.5)^2}{25}} + \frac{648}{10\sqrt{2\pi}}e^{-\frac{1}{2}\frac{(x-43.5)^2}{100}}$$

Whence the distribution of the means of samples of four is given by

$$y = \frac{260}{\sqrt{2\pi}}\left[\frac{1}{2}e^{-\frac{x^2}{\frac{1}{4}}} + \frac{4}{\sqrt{7}}e^{-\frac{1}{2}\frac{\left(x-\frac{1}{5}\right)^2}{\frac{7}{16}}} + \right.$$

$$+ \frac{6}{\sqrt{11}}e^{-\frac{1}{2}\frac{\left(x-\frac{14}{5}\right)}{\frac{11}{16}}} + \frac{4}{\sqrt{15}}e^{-\frac{1}{2}\frac{\left(x-\frac{21}{5}\right)^2}{\frac{15}{16}}} +$$

$$\left. + \frac{1}{4}e^{-\frac{1}{2}\left(x-\frac{28}{5}\right)^2}\right]$$

For agreement between actual observations and theory see Table V.

TABLE V.

Size	Actual number	Theoretical number	Size	Actual number	Theoretical number
10	1	2.3355	32	35	45.0972
11	2	2.9137	33	43	41.1296
12	2	5.6318	34	47	38.9627
13	5	8.3458	35	43	35.8226
14	11	11.2997	36	29	31.9529
15	14	13.9880	37	31	29.5427
16	10	15.9986	38	22	26.6005
17	16	18.2464	39	29	22.8845
18	18	21.3363	40	28	20.9188
19	25	25.6743	41	21	18.1148
20	29	30.8612	42	18	15.3877
21	33	36.3240	43	10	12.6509
22	43	40.9314	44	12	10.6307
23	40	43.2945	45	9	8.6248
24	46	44.9101	46	9	6.7753
25	43	45.7143	47	6	5.4139
26	50	46.1941	48	4	4.1485
27	40	48.5057	49	1	3.0880
28	50	47.9640	50	1	2.2837
29	54	47.9897	51	4	1.7967
30	46	48.0221	52	2	1.0301
31	55	48.7921	53	1	1.8665

$X^2 = 31.5024$

$P = .90$

2. *Experimental Test of the Distribution of the Standard Deviations.*

We considered a population represented by

$$f(x) = \frac{648}{5\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(x-15.5)^2}{25}} + \frac{648}{5\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(x-32.5)^2}{25}} \qquad (1)$$

We drew samples of four from (1) and calculated their standard deviations. Theoretically, these standard deviations were distributed as

$$y = y_0 \Big[c_1 \Sigma^2 e^{-2\Sigma^2} + c_2 4 \Sigma e^{-2\Sigma^2} (e^{\sqrt{3}(3.4)\Sigma} - 1) + c_3 3 \Sigma e^{-2\Sigma^2} (e^{2(3.4)\Sigma} - 1)\Big]$$

Where the c_s's are constants so determined that the integral from *0* to ∞ of the first, second, third terms shall be respectively 1, 4, 3.

For agreement between actual observations and theory see Tables VI and VII. The theoretical values are mid-ordinate values of y and not actual areas. Table VI considers one sample of 1,038

TABLE VI.

Size	Actual number	Theoretical number	Size	Actual number	Theoretical number
0-1	0	1.7986	11	129	109.5280
2	11	13.5572	12	54	74.2968
3	27	30.2397	13	36	44.8084
4	58	45.9168	14	15	22.3433
5	66	62.6243	15	8	9.4225
6	79	87.1586	16	3	3.0328
7	109	113.5740	17	0	1.0000
8	141	141.8728	18	0	.2500
9	134	152.2519	19	—	—
10	168	124.7886			

$X^2 = 38.6835$
$P = .002$

and shows an unusual central distribution. Because of the labor of calculating the standard deviations the first, second, etc. hundreds of those already found were taken as samples, except the last sample which used part of the first of the series over. The agreement of these samples with the theoretical distribution is given in Table VII.

TABLE VII.

ELEVEN SAMPLES OF 100 EACH

Size	Theoretical number	I	II	III	IV	V	VI	VII	VIII	IX	X	XI
0	.1732	0	0	0	0	0	0	0	0	0	0	0
1	1.3059	1	1	0	3	1	1	0	3	1	0	0
2	2.9139	1	6	5	3	3	4	0	2	1	3	1
3	4.4232	5	5	7	7	6	5	3	5	8	7	3
4	6.0330	6	6	7	5	5	7	7	12	7	5	4
5	8.3964	8	4	8	9	8	4	9	9	2	10	7
6	10.9203	10	6	10	12	5	19	11	13	12	10	8
7	13.6544	17	16	8	9	14	12	12	10	15	21	15
8	14.6674	13	14	19	17	13	9	15	14	12	8	13
9	12.0212	14	19	12	14	21	12	18	15	14	19	17
10	10.5514	13	13	14	9	11	12	14	7	17	8	20
11	7.1575	4	6	6	5	7	4	9	6	3	4	2
12	4.3163	5	1	3	6	2	7	1	3	4	4	4
13	2.1438	2	1	1	0	1	2	1	1	4	1	3
14	.90779	1	1	0	1	1	2	—	—	—	—	3
15	.2922	—	1	0	—	2	—	—	—	—	—	—
16	.0953	—	—	—	—	—	—	—	—	—	—	—
17	.024	—	—	—	—	—	—	—	—	—	—	—
	$X^2 =$	5.96	20.80	11.99	11.04	23.85	17.27	16.53	14.43	20.42	17.56	25.08
	$P =$	.99	.23	.80	.86	.12	.42	.48	.62	.25	.41	.09

BIBLIOGRAPHY.

(1) CHURCH, A. E. R. *On the Moments of the Distribution of the Squared Standard Deviation for Samples of* n *drawn from an Indefinitely Large Population.* « Biometrika », vol. 17 (1925), p. 79.

(2) CHURCH, A. E. R. *On the Means and Squared Standard Deviations of Small Samples from any Population.* « Biometrika », vol. 18 (1926), pp. 321-394.

(3) Editorial. *Appendix to Papers by « Student »* and R. A. FISHER. « Biometrika », vol. 19 (1925), p. 522.

(4) FISHER, R. A. *Frequency Distribution of the Values of the Correlation Coefficient in Samples from an Indefinitely Large Population.* « Biometrika », vol. 10 (1915), p. 507.

(5) HALL, Philip, B. A. *The distribution of Means for Samples of size* N *drawn from a Population in wich the Variate takes the values between 0 and 1, all such values being equally probable.* « Biometrika », vol. 19 (1927), p. 240.

(6) HELGUERO, Dr. Fernando De. *Sui Massimi Delle Curve Dimorfiche.* « Biometrika », vol. 3 (1904), pp. 84-89.

(7) IRWIN I. O., M. A., M. S. *On the Frequency Distribution of the Means of Samples from a Population having any Law of Frequency with Finite Moments with Special reference to* PEARSON'S *Type II.* « Biometrika », vol. 19 (1927), pp. 225.

(8) PEARSON, Egon S. *Editorial Note on two Preceding Papers.* « Biometrika », vol. 19 (1927), p. 244.

(9) PEARSON, Karl. *Contributions to the Mathematical Theory of Evolution* II. « Philosophical Transactions of the Royal Society of London, Series A, vol. 185, part I. (1894), p. 71.

(10) SPLAWA-NEYMAN, J., Ph. D. *Contributions to the Theory of Small Samples drawn from a Finite Population.* « Biometrika », vol. 17 (1925), pp. 472-479.

(11) STUDENT. *On the Probable Error of a Mean.* « Biometrika », vol. 6 (1908), p. 1-25.

B. DE FINETTI E U. PACIELLO

Calcolo della differenza media.

Uno dei metodi introdotti nella metodologia statistica dal GINI (*) per misurare la variabilità di una seriazione consiste nel calcolo della *differenza media* (con o senza ripetizione, Δ_R e Δ).

Vari metodi sono stati proposti per eseguirne il calcolo.

1) Il GINI stesso, nel lavoro citato, diede un metodo che si basa sulla considerazione delle *distanze graduali* di due elementi simmetrici (che occupano cioè lo stesso posto nella graduatoria rispettivamente in ordine crescente e decrescente). La sua formula è

$$\Delta_R = \frac{n-1}{n} \Delta = \frac{1}{n^2} \sum_{1}^{n}{}_i \, d_{i,\, n-i+1} \left| a_i - a_{n-i+1} \right|$$

ove $d_{i,\, n-i+1} = n + 1 - 2\, i$ è la distanza graduale.

2) Lo CZUBER ha proposto un metodo che utilizza le successive somme parziali dei termini disposti in ordine crescente e decrescente. È da questo metodo che si deduce la disposizione di calcolo data dal PIETRA (**), per il caso cosiddetto della « differenza media ponderata », che qui in particolar modo c'interessa. Un'abbreviazione al calcolo è stata portata recentemente dal DE GLERIA (***), grazie

(*) CORRADO GINI, *Variabilità e Mutabilità,* « Studi economico-giuridici pubblicati per cura della facoltà di Giurisprudenza della R. Università di Cagliari », Anno III, Parte 2ª, 1912.

(**) GAETANO PIETRA, *The theory of statistical relations with special reference to cyclical series,* « Metron », Vol. IV, N. 3-4, 1925.

(***) AMADIO DE GLERIA, *Una abbreviazione nel calcolo della differenza media,* « Rivista Italiana di Statistica », Anno I, N. 4, 1929.

all'osservazione, per se stessa ovvia, che le due serie di somme sono complementari.

La formula di CZUBER è

$$\Delta_R = \frac{n-1}{n}\,\Delta = \frac{2}{n^2}\,(S' - S)$$

dove

$$S = s_1 + s_2 + \ldots + s_n \qquad s_i = a_1 + a_2 + \ldots + a_i$$
$$S' = s'_1 + s'_2 + \ldots + s'_n \qquad s'_i = a_n + a_{n-1} + \ldots + a_{n-i+1}$$

3) Il GINI ha introdotto poi il *rapporto di concentrazione* R (*), e ha dimostrato che è collegato alla differenza media dalla relazione

$$\Delta_R = \frac{n-1}{n}\,\Delta = 2\,A\,R$$

dove A è la media aritmetica delle a_i. Un metodo per il calcolo della differenza media consiste dunque nel dedurla dal rapporto di concentrazione, che si può calcolare con metodi aritmetici od anche grafici, illustrati nel lavoro citato di GINI.

Dovendo però eseguire spesso tali calcoli, specialmente nel caso della « differenza media ponderata », per cui il procedimento preferibile era quello (nominato) del PIETRA, abbiamo veduto che essi riescono molto lunghi e faticosi, e abbiamo cercato quindi di procedere in modo più conveniente e che si presti meglio all'impostazione meccanica. Ne è risultato il metodo che qui esponiamo, e che, impiegato molte volte, ha mostrato di consentire un risparmio enorme di tempo e di lavoro.

1. — Detta

$$a_1 \quad a_2 \quad \ldots \quad a_n \qquad\qquad (a_1 \leq a_2 \leq \ldots \leq a_n)$$

una seriazione statistica, si dicono differenza media con ripetizione, Δ_R, e differenza media senza ripetizione, Δ, le espressioni

$$\Delta_R = \frac{\sum_{1}^{n}{}_{ij}\,|a_i - a_j|}{n^2} \qquad\qquad \Delta = \frac{\sum_{1}^{n}{}_{ij}\,|a_i - a_j|}{n\,(n-1)}$$

(*) CORRADO GINI, *Sulla misura della concentrazione e della variabilità dei caratteri*, « Atti del R. Ist. Veneto di S. L. A. », Anno 1913-14, Tomo LXXV, Parte II.

per il cui calcolo, ovviamente, non interessa in sostanza che valutare la sommatoria

$$\sum_{1}^{n}{}_{ij} \left| a_i - a_j \right|.$$

Possiamo scrivere intanto

$$\sum_{1}^{n}{}_{ij} \left| a_i - a_j \right| = 2 \sum_{1}^{n}{}'_{ij} (a_i - a_j)$$

ove Σ' è esteso alle sole coppie con $i > j$. Ogni termine di tale somma si può scomporre:

$$a_i - a_j = (a_i - a_{i-1}) + (a_{i-1} - a_{i-2}) + \dots \\ + (a_{j+2} - a_{j+1}) + (a_{j+1} - a_j)$$

ed è ovvio quindi che si può scrivere

$$\sum_{1}^{n}{}'_{ij} (a_i - a_j) = \sum_{1}^{n-1}{}_{h} C_h (a_{h+1} - a_h)$$

ove C_h è il numero dei termini del tipo $(a_i - a_j)$ che contengono l'addendo $(a_{h+1} - a_h)$, ossia per cui $i \geq h + 1$, $j \leq h$. Essendo $(n - h)$ i valori di i maggiori o uguali ad $(h + 1)$, e h i valori di j minori o uguali ad h, risulta

$$C_h = h\ (n - h),$$

$$\sum_{1}^{n}{}'_{ij} (a_i - a_j) = \sum_{1}^{n-1}{}_{h} h\ (n - h)\ (a_{h+1} - a_h).$$

2. — Il metodo risulta particolarmente vantaggioso quando molte delle quantità a_i hanno lo stesso valore (o si attribuisce loro lo stesso valore raggruppandole in classi più o meno estese), perchè allora la più parte dei termini è nulla.

Consideriamo infatti il caso in cui la seriazione è costituita degli n valori

$$a_1 \quad a_2 \quad \dots \quad a_n \qquad\qquad (a_1 < a_2 < \dots < a_n)$$

assunti rispettivamente in

$$p_1 \quad p_2 \quad \dots \quad p_n$$

casi. Oppure, ciò che non cambia nulla, supponiamo che p_1 casi siano raggruppati in una classe cui compete il valore a_1, p_2 in una classe cui compete il valore a_2, . . . e si voglia calcolare la differenza media come se nei $p_1, p_2, \ldots, p_n$ casi il carattere a assumesse effettivamente il valore $a_1, a_2, \ldots, a_n$.

Posto allora

$$q_h = p_1 + p_2 + \ldots + p_h \qquad\qquad q = q_n$$

la formula precedente diviene

$$\sum_{1}^{n}{}'_{ij} (a_i - a_j) = \sum_{1}^{n-1}{}_h q_h (q - q_h) (a_{h+1} - a_h)$$

e si presta al calcolo in modo facilissimo, come mostra l'esempio che segue.

3. — Vogliamo calcolare la differenza media della natalità nel 1921 per i 16 Compartimenti, tenendo conto dell'ammontare delle rispettive popolazioni.

Poniamo allora :

h = numero d'ordine dei Compartimenti disposti per natalità crescente,

a_h = natalità nell'h-esimo Compartimento (anno 1921),

p_h = popolazione dell'h-esimo Compartimento (cens. 1° dic. 1921),

e, al solito

$q_h = p_1 + p_2 + \ldots + p_h =$ popolazione dei Compartimenti con natalità non superiore ad a_h,

$q = q_{16} =$ popolazione censita totale,

e quindi

$q - q_h =$ popolazione dei Compartimenti con natalità non inferiore ad a_{h+1}.

Nella tabella seguente le somme S_h sono riportate in migliaia ; il calcolo è stato eseguito tenendo conto di tutte le cifre, e ciò spiega delle piccole differenze che si incontrerebbero ripetendolo. Va notato che, pur tenendo conto di tutte le cifre, il calcolo si è potuto eseguire sulla Nova Brunsviga a (10, 10, 15) cifre, mentre col metodo precedente sarebbe stato necessario ricorrere a macchine con un numero di cifre maggiore.

Ecco la tabella dei dati e dei calcoli che danno la sommatoria

$$\sum_{1}^{n}{}'_{ij} (a_i - a_j).$$

h	Compartimenti	a_h	p_h	q_h	$q - q_h$	$a_{h+1} - a_h$	$q_h (q - q_h) \times$ $\times (a_{h+1} - a_h)$
1	Piemonte	20,22	3.383.646	3.384	33.801	0,82	93.783.718
2	Liguria	21,04	1.335.466	4.719	32.465	6,31	966.743.552
3	Sicilia	27,35	4.061.452	8.781	28.404	0,81	202.016.640
4	Lombardia . . .	28,16	5.086.338	13.867	23.318	0,40	129.337.566
5	Toscana	28,56	2.759.767	16.627	20.558	1,14	389.662.881
6	Lazio	29,70	1.956.908	18.584	18.601	0,62	214.317.350
7	Sardegna	30,32	864.174	19.447	17.737	0,28	96.583.586
8	Emilia	30,60	3.027.009	22.474	14.710	1,97	651.285.855
9	Marche	32,57	1.145.685	23.620	13.564	1,05	336.410.397
10	Umbria	33,62	645.515	24.266	12.919	0,43	134.797.550
11	Campania . . .	34,05	3.243.739	27.510	9.675	0,36	95.815.077
12	Veneto	34,41	3.999.027	31.509	5.676	0,91	162.743.370
13	Abruzzi e Molise	35,32	1.387.215	32.896	4.289	0,57	80.414.886
14	Puglie	35,89	2.307.762	35.204	1.981	0,34	23.709.606
15	Calabrie	36,23	1.512.318	36.716	469	1,65	28.385.855
16	Basilicata	37,88	468.557	37.184	—	—	—
			37.184.578				3.606.007.889

Si ricava

$$\sum_{1}^{n}{}'_{ij} (a_i - a_j) = 3.606.007.889.000.000:$$

vanno aggiunti 6 zeri perchè la popolazione si è valutata in migliaia (mancano 3 cifre a q_h e $q - q_h$), La popolazione totale è 37.184.578, e quindi

$$\Delta_R = 2 \cdot \frac{3.606.007.889.000.000}{(37.184.578)^2} = 5{,}216$$

$$\Delta = 2 \cdot \frac{3.606.007.889.000.000}{37.184.578 \times 37.184.577} = 5{,}216.$$

È ovvio che su un numero di termini così grande la differenza di un'unità non influisce sulle cifre decimali che possono interessare, e si ha quindi praticamente $\Delta = \Delta_R$.

4. — Una semplificazione notevole si ha quando tutte le differenze $(a_{h+1} - a_h)$ sono uguali, come avviene in molti casi. Sono infatti molto usate delle classificazioni in gradi di uguale ampiezza: le età, le stature, i pesi di un gruppo d'individui si ripartiranno ad es. per classi ciascuna d'un anno, d'un centimetro, d'un chilogrammo.

Allora si ha

$$\sum_{1}^{n-1}{}_h\, q_h\,(q - q_h)\,(a_{h+1} - a_h) = d \,.\, \sum_{1}^{n-1}{}_h\, q_h\,(q - q_h)$$

ove

$$d = a_n - a_{n-1} = a_{n-1} - a_{n-2} = \ldots = a_2 - a_1$$

è l'ampiezza d'ogni singola classe.

V. A. NEKRASSOFF

Nomography in applications of statistics

INTRODUCTION.

A successful and effective use of statistical theory often calls for all possible simplifications in the method of applying complicated formulae and tables in carrying out requisite computations. Naturally, one turns to nomography for assistance and customarily finds it if he will but search for what we shall term an efficient nomogram. Typical nomograms of this character found useful in our own work are those designed to meet the following needs : calculation of standard deviation in number of defects or in percentage defective in samples of a given size drawn from a given product ; setting sampling limits on percentage defective ; correction of the estimate of standard deviation to allow for sample size ; setting limit lines on quality control charts (*), and calculation of sampling limits in general.

It is firmly believed that more general application of many important modern theoretical developments will be made when satisfactory nomograms are developed to simplify the use of these results. Of course, such graphical presentations serve only to supplement tables where the accurancy secured by such means is all that is required.

The purpose of the present note is to illustrate how a simple and efficient alignment nomogram can be found for a very complicated, awe-inspiring expression of almost universal importance in

(*) SHEWHART, W. A., *Quality Control*, « Bell System Technical Journal », Vol. VI, October 1927, pp. 722-735.

the modern theory of statistics. The particular case chosen is that of the representation of the expansion of " STUDENT'S " integral as given by R. A. FISHER in this Journal (*).

EFFICIENT NOMOGRAM DEFINED.

It is, in general, well known that the functional relationship between a certain number of variables, say three, may often be represented by several nomograms. Three criteria are of importance in judging the usefulness or efficiency of a nomographic presentation. :

1. The simplicity of reading and ease of understanding the nomographic scales.
2. The accuracy attained.
3. The size, shape, appearance and, in general, the ease with which nomographs may be used.

An efficient nomogram will, therefore, be considered to be the one which satisfies these three requirements.

NOMOGRAPHIC PRESENTATION OF " STUDENT'S " INTEGRAL.

It will be recalled that " Student's " integral gives the probability P_∞ -P_n that the error of an average X of a sample of size n measured from the expected value X' of the universe will fall within the range $-\infty$ to $+ t\,\sigma$, where σ is the observed standard deviation of the average. The extact form of the rather complicated equation is given at the bottom of Fig. 1. For our present purpose, we can simply represent this functional relationship f as

$$P_n = f\,(n, t). \tag{1}$$

The first three nomograms shown in Fig. 1 are based upon assumed approximate relationships between the three variables represented in Eq. (1). The next seven charts were obtained through the use of the exact tabular values given in the paper already referred to, the different forms of the nomographic presentation being obtained through the application of certain mathematical ttricks outlined

(*) FISHER , R. A., *Expansion of* STUDENT'S *Integral in Powers of* n^{-1}, « Metron », Vol. II, No. 3, 1925.

under each figure. All of these charts are somewhat complicated because of the plurality of scales used and hence do not meet the requirements of optimum simplicity.

The search, therefore, had to be continued. The next step taken was to express the relationship between P_n, n and t in the form (*):

$$f_1 (P_n) = f_2 (n) + f_3 (t), \qquad (2)$$

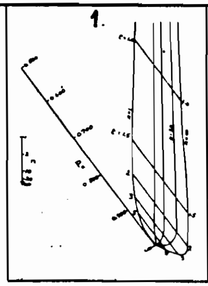
where f_1, f_2 anc f_3 are functions such that any set of values of P_n, n and t satisfying Eq. (2) also satisfy Eq. (1).

Through the use of this approximate relationship it is possible to obtain the simplest form of alignment nomogram having only three linear scales, as is indicated in Fig. 2. However, to avoid more refined and laborious interpolation and possible illegibility of the more minute graduations for a certain range of t values larger than two, a curved portion of the t scale was introduced as indicated in the figure. This nomographic presentation of the very complicated expression shown at the bottom of Fig. 1 meets the above stated three requirements of an efficient nomogram. This chart takes the place in many practical problems of nearly four pages of tables.

(*) I am indebted to Col. E. E. SCHAITANOFF, Russian Ordnance Engineer (Ceskomoravska Kolben Ack. Sp. Praha, Tchechoslovaquie) for his most ingenious solution of this functional equation.

TABLE OF VARIOUS NOMOGRAMS REPRESENTING STUDENT'S INTEGRAL AS GIVEN BY R. A. FISHER'S TABLE.

1.

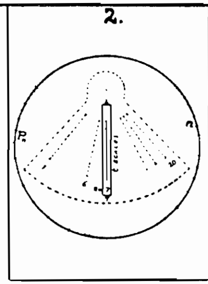


$$(P_n-0.5)^2=\frac{e_0^2t^2f_2(n,t)}{(4e_0^2-.01)t^2+1}$$

$$\begin{vmatrix} (P_n-0.5)^2 & 0 & +1 \\ 0 & e_0^2 & +1 \\ \frac{f_2(n,t)}{4} & \frac{\frac{1}{t^2}-.01}{4} & +1 \end{vmatrix}=0$$

$x=(P_n-0.5)^2$, $y=0$ } P_n-SCALE; $x=0$, $y=e_0^2$ } n-SCALE; $x=\frac{f_2(n,t)}{4}$, $y=\frac{\frac{1}{t^2}-.01}{4}$ } t-SCALES

2.



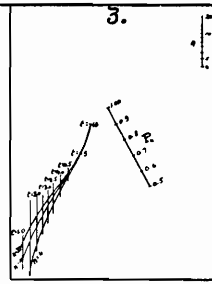
$$\frac{(P_n-0.5)^2}{1-4(P_n-0.5)^2}=\frac{e_0^2F(n,t)}{1-.01t^2}$$

$$\begin{vmatrix} 1 & f_1 & f_1^2 \\ 1 & f_2 & f_2^2 \\ 1 & 0 & -f_3 \end{vmatrix}=0$$

$f_1=\frac{(P_n-0.5)^2}{1-4(P_n-0.5)^2}$, $f_2=\frac{1}{e_0^2}$, $f_3=\frac{F(n,t)}{1-.01t^2}$

$x=\frac{(P_n-0.5)^2}{1-4(P_n-0.5)^2}$, $y=x^2$ } P_n-SCALE; $x=\frac{1}{e_0^2}$, $y=x^2$ } n-SCALE; $x=0$, $y=-\frac{F(n,t)}{1-.01t^2}$ } t-SCALES

3.

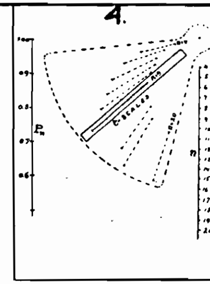


$$(P_n-0.5)=\frac{0.575\,t\,f_3(n,t)}{t+e(n)}$$

$$\begin{vmatrix} 174(P_n-0.5) & 1 & 0 \\ 0 & 1 & \frac{100}{e(n)} \\ -100f_3(n,t) & 1 & -\frac{100}{t} \end{vmatrix}=0$$

$x=174(P_n-0.5)$, $y=0$ } P_n-SCALE; $x=0$, $y=\frac{100}{e(n)}$ } n-SCALE; $x=-100f_3(n,t)$, $y=-\frac{100}{t}$ } t-SCALES

4.



$P_n=0.5+P_\infty f(n,t)$ or $P_n=0.5+f_2(n,t)$

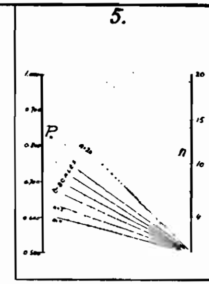
$$2P_n-n\frac{2P_\infty f(n,t)}{n}-1=0$$

$$\frac{u}{\ell_1}\ell_1\frac{L-x}{2Ly}+\frac{v}{\ell_2}\ell_2\frac{L+x}{2Ly}-1=0$$

$u=\ell_1P_n$; P_n-SCALE; $v=-\ell_2 n$; n-SCALE

$$x=L\frac{\frac{P_n-0.5}{n}\frac{\ell_1}{\ell_2}-1}{\frac{P_n-0.5}{n}\frac{\ell_1}{\ell_2}+1},\quad y=\frac{\frac{\ell_1}{2}}{\frac{P_n-0.5}{n}\frac{\ell_1}{\ell_2}+1}$$ } t & n SCALES

5.



$P_n=P_\infty-f_1(n,t)$

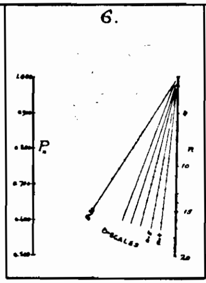
$P_n+f_1(n,t)-P_\infty=0$

$$\frac{u}{\ell_1}\ell_1\frac{L-x}{2Ly}+\frac{v}{\ell_2}\ell_2\frac{L+x}{2Ly}-1=0$$

$u=\ell_1P_n$; P_n-SCALE; $v=\ell_2 n$; n-SCALE

$$x=L\frac{\frac{\ell_1}{n\ell_2}f_1(n,t)-1}{\frac{\ell_1}{n\ell_2}f_1(n,t)+1},\quad y=\frac{\ell_1P_\infty}{\frac{\ell_1}{n\ell_2}f_1(n,t)+1}$$ } t & n SCALES

6.



$P_n=P_\infty-f_1(n,t)$

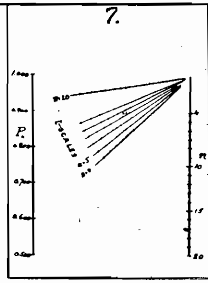
$$-P_n\frac{1}{f_1(n,t)}+\frac{nP_\infty}{nf_1(n,t)}-1=0$$

$$\frac{u}{\ell_1}\ell_1\frac{L-x}{2Ly}+\frac{v}{\ell_2}\ell_2\frac{L+x}{2Ly}-1=0$$

$u=\ell_1P_n$; P_n-SCALE; $v=-\ell_2 n$; n-SCALE

$$x=L\frac{\frac{\ell_1}{\ell_2}\frac{P_\infty}{n}-1}{\frac{\ell_1}{\ell_2}\frac{P_\infty}{n}+1},\quad y=-\frac{\ell_1 f_1(n,t)}{\frac{\ell_1}{\ell_2}\frac{P_\infty}{n}+1}$$ } t & n SCALES

7.



$P_n=P_\infty-f_1(n,t)$

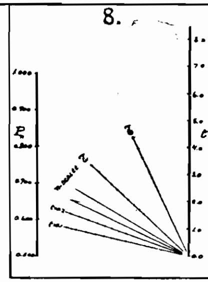
$$P_n\frac{1}{P_\infty}+\frac{-nf_1(n,t)}{-nP_\infty}-1=0$$

$$\frac{u}{\ell_1}\ell_1\frac{L-x}{2Ly}+\frac{v}{\ell_2}\ell_2\frac{L+x}{2Ly}-1=0$$

$u=\ell_1P_n$; P_n-SCALE; $v=-\ell_2 n$; n-SCALE

$$x=L\frac{\frac{\ell_1}{\ell_2}\frac{f_1(n,t)}{n}+1}{\frac{\ell_1}{\ell_2}\frac{f_1(n,t)}{n}-1},\quad y=-\frac{\ell_1P_\infty}{\frac{\ell_1}{\ell_2}\frac{f_1(n,t)}{n}-1}$$ } t & n SCALES

8.



$P_n=P_\infty-f_1(n,t)$

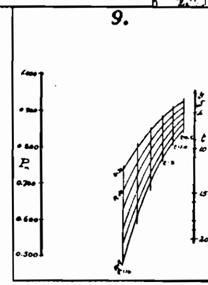
$$P_n\frac{1}{P_\infty}+\frac{tf_1(n,t)}{tP_\infty}-1=0$$

$$\frac{u}{\ell_1}\ell_1\frac{L-x}{2Ly}+\frac{v}{\ell_2}\ell_2\frac{L+x}{2Ly}-1=0$$

$u=\ell_1P_n$; P_n-SCALE; $v=\ell_2 t$; n-SCALE

$$x=L\frac{\frac{\ell_1}{\ell_2}\frac{f_1(n,t)}{t}-1}{\frac{\ell_1}{\ell_2}\frac{f_1(n,t)}{t}+1},\quad y=\frac{\ell_1P_\infty}{\frac{\ell_1}{\ell_2}\frac{f_1(n,t)}{t}+1}$$ } t & n SCALES

9.



$P_n=0.5+f_2(n,t)$

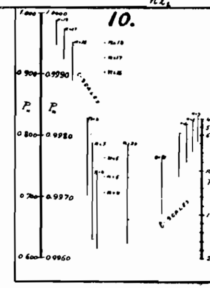
$$P_n\frac{1}{f_2(n,t)}-\frac{0.5t}{f_2(n,t)t}-1=0$$

$$\frac{u}{\ell_1}\ell_1\frac{L-x}{2Ly}+\frac{v}{\ell_2}\ell_2\frac{L+x}{2Ly}-1=0$$

$u=\ell_1P_n$; P_n-SCALE; $v=-\ell_2 t$; t-SCALE

$$x=L\frac{\frac{\ell_1}{\ell_2}-2t}{\frac{\ell_1}{\ell_2}+2t},\quad y=\frac{2\ell_1tf_2(n,t)}{\frac{\ell_1}{\ell_2}+2t}$$ } t & n SCALES

10.



$P_n=0.5+f_2(n,t)$

$$P_n\frac{1}{f_2(n,t)}-\frac{n\,0.5}{nf_2(n,t)}-1=0$$

$$\frac{u}{\ell_1}\ell_1\frac{L-x}{2Ly}+\frac{v}{\ell_2}\ell_2\frac{L+x}{2Ly}-1=0.$$

$u=\ell_1P_n$; P_n-SCALE; $v=-\ell_2 n$; n-SCALE

$$x=L\frac{\frac{\ell_1}{\ell_2}-2n}{\frac{\ell_1}{\ell_2}+2n},\quad y=\frac{2\ell_1nf_2(n,t)}{\frac{\ell_1}{\ell_2}+2n}$$ } t & n SCALES

ORIGINAL FORM OF EQUATION REPRESENTED BY FISHER'S TABLE IS:

$$P_\infty-P_n=\frac{te^{-\frac{t^2}{2}}}{\sqrt{2\pi}}\left\{\frac{t^2+1}{4n}+\frac{3t^6-7t^4-5t^2-3}{96n^2}+\frac{t^{10}-11t^8+14t^6+6t^4-3t^2-15}{384n^3}+\frac{15t^{14}-375t^{12}+2225t^{10}-2141t^8-939t^6-213t^4-915t^2+345}{92160n^4}+\frac{3t^{18}-133t^{16}+1764t^{14}-7516t^{12}+5994t^{10}+2490t^8+1140t^6+180t^4+5355t^2+17955}{368640n^5}\right\}$$

V. A. N. 1-24-1928.

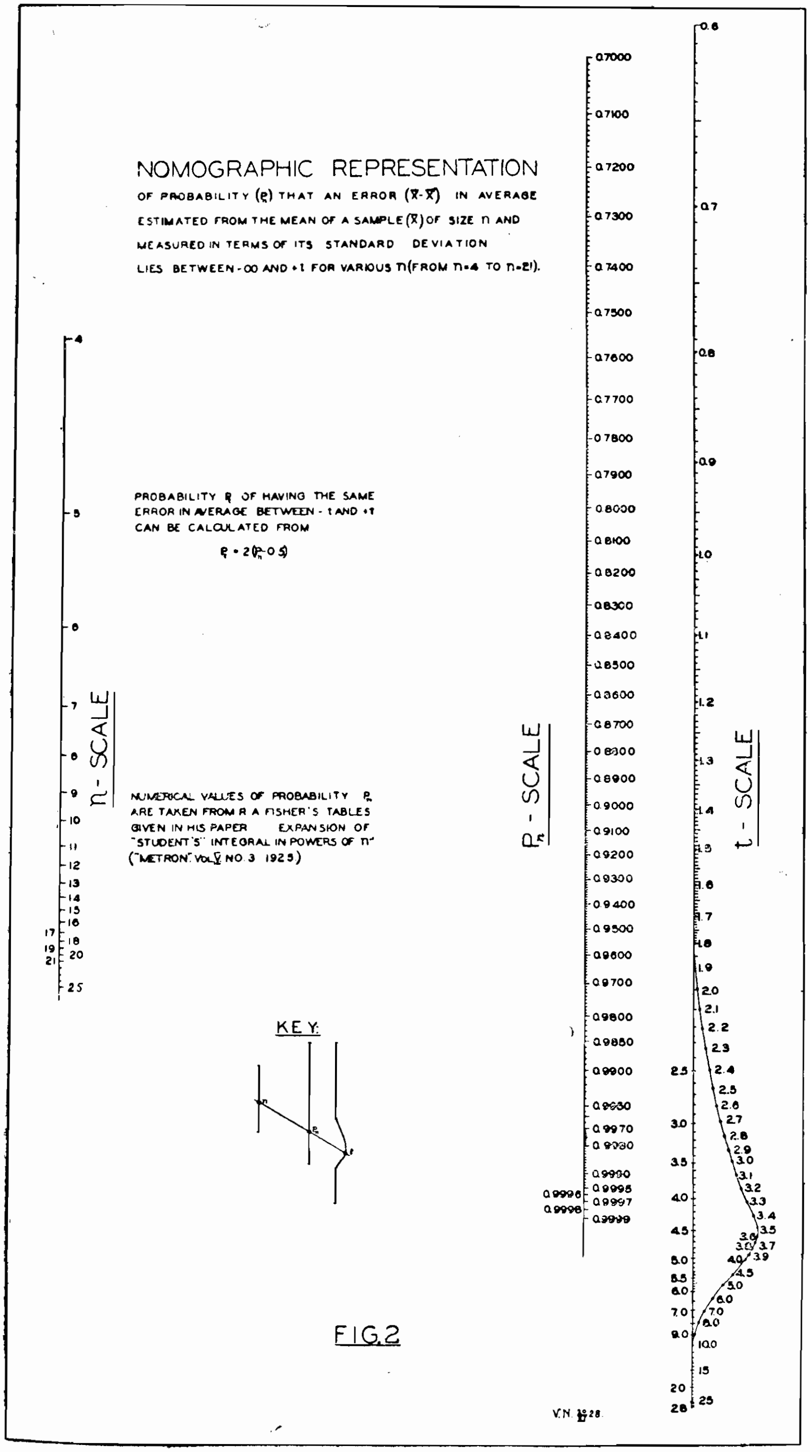

NOMOGRAPHIC REPRESENTATION
OF PROBABILITY (P_n) THAT AN ERROR ($\bar{X}-\bar{\bar{X}}$) IN AVERAGE ESTIMATED FROM THE MEAN OF A SAMPLE ($\bar{X}$) OF SIZE n AND MEASURED IN TERMS OF ITS STANDARD DEVIATION LIES BETWEEN -∞ AND +t FOR VARIOUS n (FROM n=4 TO n=21).
PROBABILITY P OF HAVING THE SAME ERROR IN AVERAGE BETWEEN -t AND +t CAN BE CALCULATED FROM
$P = 2(P_n - 0.5)$
NUMERICAL VALUES OF PROBABILITY P_n ARE TAKEN FROM R. A. FISHER'S TABLES GIVEN IN HIS PAPER EXPANSION OF "STUDENT'S" INTEGRAL IN POWERS OF n^{-1} ("METRON" VOL. V NO. 3 1925)
n - SCALE
P_n - SCALE
t - SCALE
KEY:
V.N. 1928.

FIG. 2

R. ROY

La demande dans ses rapports avec la répartition des revenus.

IMPORTANCE DE LA LOI DE LA DEMANDE.

Dans ses *Eléments d'Economie politique pure* WALRAS étudiant l'échange de deux marchandises, observe que la Demande constitue le fait principal et l'Offre le fait accessoire, car, dit-il « On n'offre pas pour offrir ; on n'offre que parce que l'on ne peut pas demander sans offrir ; l'offre n'est qu'une conséquence de la demande ».

Cette opinion avait été déjà émise par COURNOT qui, dans ses *Principes de la théorie des richesses* s'exprime de la manière suivante :

« Le prix règle la consommation, ou, comme on dit, la demande, et à son tour la demande règle la production ».

MARSHALL au contraire dans ses *Principes d'Economie politique* estime que l'esprit d'entreprise est en dernière analyse le ressort de l'activité économique ; mais il reconnaît cependant l'importance de la loi de la Demande et considère comme regrettable l'oubli dans lequel sont tombés au cours du dix-neuvième siècle les travaux de COURNOT et de DUPUIT, travaux effacés par ceux de RICARDO et de son école sur le coût de production.

Les auteurs qui se sont occupés de la théorie de la valeur sont donc tous d'accord pour reconnaître l'importance de la loi de la demande, et, sans prétendre vouloir trancher la question de prépondérance entre l'Offre et la Demande, nous constatons que la Demande constitue l'un des deux termes du problème de la valeur, et que toute application de la théorie des prix s'appuie sur cette notion.

Il y a donc intérêt à en préciser la signification et, si possible, l'expression mathématique.

Ce qu'il faut entendre par « Loi de la Demande ».

Pour savoir ce qu'il faut entendre exactement par « Loi de la Demande », il nous faut encore avoir recours aux travaux de Cournot qui, dans ce domaine comme en tant d'autres, fait vraiment figure de précurseur.

Sa loi du débit ou de la demande exprime le rapport existant entre la demande effective totale d'un produit et son prix ; en d'autres termes, cette loi définit la consommation ou le débit d'un article en fonction de son prix ; cette notion a été reprise et même développée par Dupuit, qui en a déduit sa définition de l'Utilité, et c'est en somme à ses travaux que l'on doit faire appel en matière de monopole et de tarif.

Cependant, la théorie de Dupuit sur l'Utilité a été fortement combattue par Walras et nous devons insister un peu sur cette divergence, car elle est d'importance capitale, en ce qui touche la conception d'une loi de la Demande.

Dans l'esprit de Cournot et de Dupuit, la relation existante entre le prix et le débit d'un article, relation exprimée par la loi de la Demande, suppose la permanence de tous les autres éléments du problème de l'équilibre économique, et notamment du prix et du débit des autres articles, du chiffre de la population considérée, du revenu total et de sa répartition, des influences du goût et de la mode, etc.

Cette conception s'oppose donc à celle de Walras qui établit sa théorie de la valeur sur la notion d'équilibre général des prix et quantités offertes ou demandées ; dans un tel système, en effet, il suffit de faire varier le prix et le débit d'un article pour que varient immédiatement les éléments relatifs aux autres articles, de manière à rétablir l'équilibre. D'ailleurs, les fonctions de Demande utilisées par Walras ne font intervenir qu'une seule variable dans le cas de l'échange de deux marchandises, mais deviennent immédiatement fonctions de $(n - 1)$ variables dans le cas d'échange de n marchandises entre elles.

Cette conception est également celle de Pareto qui se refuse même à considérer toute loi de la Demande et construit son système économique en partant de la notion de « Lignes d'indifférence » ; cependant, il en arrive lui-même à considérer la courbe individuelle de la Demande et, dans l'appendice à son *Manuel d'Economie politique*, il fournit d'intéressantes précisions sur les caractéristiques de cette courbe.

De son côté WALRAS est bien forcé de reconnaître que l'idée de COURNOT peut recevoir des applications importantes et il aboutit lui aussi aux courbes d'achat qui ne diffèrent en aucune manière des courbes de débit, imaginées par COURNOT ; WALRAS énumère d'ailleurs les conditions auxquelles est subordonné l'emploi des courbes d'achat ; ces conditions peuvent être ainsi résumées :

Pour que l'on puisse admettre que la variation du prix et du débit d'un article n'altèrent pas sensiblement l'équilibre existant antérieurement, il est nécessaire de se trouver en présence d'un système constitué par un très grand nombre d'articles et de faire en outre abstraction de la loi de substitution.

La première de ces conditions nous paraît suffisamment bien établie en pratique, pour que nous soyons en droit de faire appel à la notion de loi de la Demande ; quant à la seconde réserve, nous y reviendrons plus loin.

En ce qui concerne la permanence des autres éléments du problème de l'équilibre économique, nous devons reconnaître avec MARSHALL que cette permanence est loin d'être réalisée en fait, et que la mobilité du milieu économique constitue un des obstacles les plus sérieux pour l'étude statistique des courbes de Demande, et pour l'établissement des conclusions auxquelles cette étude permet d'aboutir ; il est, par exemple, un facteur dont les variations ont grandement influé sur la consommation de certains articles : nous voulons parler du revenu total et de sa répartition, mais nous aurons l'occasion de nous étendre longuement sur cette question.

CARACTÈRES GÉNÉRAUX DE LA DEMANDE.

Ainsi présentée, la loi de la Demande exprime, comme nous l'avons dit, la quantité demandée en fonction du prix de l'article envisagé, tous autres facteurs de l'équilibre étant supposés invariables. Nous considérons ici la Demande totale et non la Demande individuelle.

Tous les auteurs sont d'accord sur les caractères généraux de cette fonction ; ils peuvent se résumer ainsi :

1) La loi de la Demande s'exprime par une fonction définie, c'est-à-dire qu'à tout prix situé dans l'intervalle de variation étudié correspond une demande et une seule.

Cette propriété paraît évidente en raison de la permanence des autres facteurs de l'équilibre économique, mais il semble qu'on doive cependant faire quelques réserves, pour tenir compte de l'irréversibilité du phénomène étudié.

Nous pensons en effet que si le prix d'un article passe d'une valeur à une autre qui en diffère sensiblement, pour revenir à la valeur primitive, la quantité demandée ne repasse pas exactement par les mêmes valeurs ; par suite de la modification des habitudes et des goûts résultant d'une variation de prix, il doit se produire un phénomène identique à celui qu'on observe en magnétisme, phénomène désigné sous le nom d'Hystérésis.

Mais comme il nous est impossible de vérifier l'exactitude de cette hypothèse et d'apprécier l'importance de ses effets, nous négligerons le caractère d'irréversibilité et admettrons qu'a tout instant la fonction exprimant la Demande est définie.

2) En second lieu, la Demande est, en général, considérée comme une fonction continue du prix ; cette propriété est due au fait qu'un grand nombre de consommateurs se trouvent en présence et, pour beaucoup d'articles, la Demande individuelle doit être au contraire tenue pour une fonction discontinue.

Cependant, même dans l'hypothèse où la loi des grands nombres est applicable, il convient de faire état du principe de substitution en vertu duquel une marchandise est remplacée par une autre dès que son prix atteint une certaine valeur.

La notion de continuité ne doit donc être admise que dans certaines limites de prix, variables avec la nature des articles étudiés.

3) La Demande est toujours considérée comme une fonction décroissante du prix, et certains auteurs, comme MARSHALL ont même fait appel à la notion d'Elasticité de la Demande, cette grandeur étant en rapport avec l'inclinaison de la courbe représentative sur l'axe des prix.

Dans l'appendice à son *Manuel d'Economie politique,* PARETO démontre par un procédé rigoureux que la Demande individuelle est fonction décroissante du prix, en s'appuyant simplement sur le fait que l'ophélimité décroît lorsqu'augmente la quantité du bien envisagé.

Nous établirons au cours de cette étude que la Demande totale est décroissante, en nous référant à la loi de distribution des revenus à l'intérieur d'une collectivité, et nous ferons appel à la notion de satiété, qui se confond à la limite avec le caractère décroissant de l'ophélimité.

On admet aussi en général l'existence d'un débit maximum correspondant à un prix nul, et d'un prix-limite à partir duquel la Demande est nulle, mais sur ce dernier point WALRAS, en étudiant l'échange de deux marchandises, admet la possibilité d'une courbe asymptotique à l'axe des prix, et il se contente d'affirmer que la courbe représentative de la Demande est comprise entre les axes de coordonnées et l'hyperbole de la quantité existante.

Nous aurons ultérieurement l'occasion de préciser nous vues sur ces divers points.

4) Enfin, on donne généralement à la courbe représentative de la Demande en fonction du prix, une concavité telle que la tangente à la courbe soit d'autant plus inclinée sur l'axe des abscisses, que le prix est plus bas.

Les justifications que l'on donne à l'appui de cette manière de faire sont à la vérité bien vagues ; on se contente le plus souvent d'affirmer que, plus le prix d'un article baisse, plus il est accessible aux couches les plus nombreuses de la population ; la forme admise pour la courbe de la Demande serait donc liée à la répartition des revenus, ce qui ne paraît pas douteux ; mais les arguments invoqués à cet effet sont loin de constituer une véritable démonstration.

COURNOT lui-même estime qu'une baisse de prix doit déterminer une augmentation de débit supérieure à celle qui résulterait de la loi de proportionnalité inverse, mais ce n'est là qu'une pure affirmation. C'est également COURNOT qui fait observer que lorsqu'on envisage des variations restreintes de prix, la courbe de la Demande peut être assimilée à une droite, et cette manière de voir ne constitue qu'une application d'un principe général exprimé par la formule des accroissements finis.

Cependant, dans une étude publiée par les « Annales des Ponts-et-Chaussées », M. DUMAS a envisagé l'application générale d'une loi linéaire dans tout le champ de variation des prix, en s'appuyant sur la notion un peu vague de « marché saturé », et n'invoquant en somme que le principe de raison suffisante, fort contestable à la vérité.

Nous verrons au cours de cette étude qu'il est possible d'aboutir à des conclusions précises en ce qui touche la courbe de la Demande, moyennant certaines hypothèses sur la légitimité desquelles nous nous étendrons longuement.

5) En dernier lieu, tous les auteurs auxquels nous nous sommes référés jusqu'à présent ont longuement insisté sur le caractère empirique de la loi de la Demande.

Ce caractère est pour COURNOT de même nature que celui des lois observées en démographie, ou plus généralement en « Arithmétique sociale ».

WALRAS a tenté de pousser l'analyse plus loin en faisant appel à l'« utilité d'intensité » et aux courbes de besoin ou de rareté ; mais, au terme de cette analyse, il est bien obligé de reconnaître le caractère empirique des courbes de besoin et celui des courbes de Demande qui s'en déduisent.

L'empirisme de la loi de la Demande semble donc bien un fait acquis, mais il ne s'oppose pas à la recherche des liens pouvant exister entre cette loi et d'autres lois empiriques accessibles à l'observation statistique. Les problèmes qui se posent en cette matière sont de l'essence même de ceux qui se posent en démographie, pour en revenir au parallélisme établi par COURNOT.

BUT DE LA PRÉSENTE ÉTUDE.

Bien que les auteurs aux travaux desquels nous avons fait allusion aient eu tous recours aux notations mathématiques, ils ne se sont pas attachés à la recherche d'une expression quantitative de la loi de la Demande et se sont bornés à souligner son caractère empirique et à mentionner ses caractéristiques générales.

WALRAS présente d'ailleurs son appareil mathématique plutôt comme un mode rationnel d'exposition et d'analyse, que comme une expression quantitative des faits, et il ne songe nullement à nous éclairer sur la forme précise des fonctions ou des courbes représentatives de la Demande.

Pour ces auteurs, la forme précise dont il s'agit ne pourrait être obtenue qu'au moyen de données statistiques dont l'établissement soulèverait d'ailleurs de sérieuses difficultés d'ordre pratique.

Ce n'est que tout-à-fait accessoirement et sous une forme fragmentaire qu'on découvre parfois quelques précisions sur la nature des fonctions exprimant la Demande ; nous avons déjà fait allusion à une opinion de COURNOT, relative à la loi de proportionalité inverse ; nous trouvons un autre exemple de cette nature dans les *Principes d'Economie politique* de MARSHALL, au sujet du blé.

« Les cas de disette sont très exceptionnels en Angleterre depuis la suppression des lois sur les céréales. Mais, d'après l'expérience de temps moins heureux, nous pouvons dire que des déficits de 1, 2, 3, 4, ou 5 dixièmes dans l'offre, amènent une hausse de prix de 3, 8, 16, 28, ou 45 dixièmes ».

Cette estimation précise aurait été établie par GREGORY KING et MARSHALL ajoute :

« Sa portée pour la loi de la Demande est admirablement étudiée par Lord LANDERDALE ; elle est représentée dans la figure I par la courbe *DD'*, le point *A* correspondant aux prix ordinaires ; si nous tenons compte du fait que lorsque le prix du blé est très bas, il peut être employé comme on le fit par exemple, en 1834, pour nourrir le bétail, pour la brasserie et la distillation, la partie inférieure de la courbe prendrait une forme assez semblable à celle de la ligne pointillée sur la figure et, si nous supposons que lorsque le prix est très élevé, des substituts meilleur marché puissent remplacer le blé, la partie supérieure de la courbe prendrait la forme indiquée par la partie supérieure de la ligne pointillée ».

Nous reproduisons à page 153 la figure I indiquée par MARSHALL, sur laquelle nous n'avons pas d'autre renseignement ; elle constitue une donnée tout-à-fait précise, mais il semble difficile d'en pouvoir induire une expresson générale de la loi de la Demande.

JEVONS a bien donné une expression mathématique, également applicable à la demande du blé, en partant de considérations analogues à celles de MARSHALL sur les phénomènes de substitution.

La formule indiquée par JEVONS est la suivante :

$$y = \frac{a}{(x - b)^n}$$

dans laquelle y représente le prix du blé, et x la quantité de blé récoltée.

L'examen des données numériques montre que n est à peu près égal à 2 et, en supposant $n = 2$, on trouve comme valeurs les plus probables de a et de b :

$$\begin{cases} a = 0{,}824 \\ b = 0{,}12 \end{cases}$$

ce qui donne :

$$y = \frac{0,824}{(x - 0,12)^2}$$

ou approximativement :

$$y = \frac{5}{6\left(x - \frac{1}{8}\right)^2}$$

L'approximation de cette formule est appréciée par la comparaison des chiffres qu'elle donne avec ceux de DAVENANT, lesquels ne diffèrent pas de ceux attribués par MARSHALL à GREGORY KING.

La comparaison s'établit comme il suit :

récolte	1	0,9	0,8	0,7	0,6	0,5
prix (DAVENANT) . .	1	1,3	1,8	2,6	3,8	5,55
prix calculés	1,06	1,36	1,78	2,65	3,58	5,71

Sans nous étendre sur les propriétés de cette loi précise, qui ne paraît pas à première vue susceptible de généralisation, nous observerons qu'elle exprime le prix en fonction de la quantité, et non la quantité en fonction du prix, comme nous l'avions supposé jusqu'à présent ; cette circonstance est tout-à-fait accessoire car, si l'on veut faire abstraction du lien de causalité, la loi de la Demande est exprimée par une relation de la forme :

$$f(p, q) = 0$$

permettant de calculer les prix en fonction de la quantité ou la quantité en fonction du prix, suivant la nature des problèmes à traiter.

Des recherches récentes ont permis de se faire une opinion sur la forme précise de la loi de la Demande et sur la valeur de l'élasticité.

Parmi ces travaux, nous devons citer ceux du professeur HENRY MOORE de Columbia University, complétés par ceux du professeur HENRY SCHULTZ, de l'Université de Chicago.

Leurs travaux font appel aux procédés statistiques les plus modernes et portent sur un grand nombre d'observations poursuivies avec méthode ; M. HENRY SCHULTZ observe par exemple que, en ce qui concerne le sucre, la Demande n'apparaît pas très élastique pour les années considérées : une diminution de 1 pour 100 dans le prix du sucre n'entraîne qu'une augmentation de 0,6 pour 100 dans la consommation de la même année.

Il est hors de doute que c'est dans cette voie que doivent résolument s'engager les économistes soucieux de conduire ces études avec toute la rigueur désirable, mais il ne faut pas se dissimuler les difficultés de semblables observations.

A propos de ces difficultés, nous pensons qu'il serait particulièrement indiqué d'appliquer ici la méthode préconisée par HENRI POINCARÉ, dans *La Science et l'Hypothèse*, au sujet des sciences physiques : l'auteur compare le savant à un bibliothécaire qui disposerait de crédits extrêmement limités pour enrichir sa collection ; afin d'éviter des acquisitions inutiles, et obtenir l'utilisation optima de ses crédits, il aurait évidemment intérêt à dresser un catalogue soigneusement tenu à jour, et ce catalogue ne serait autre que les conséquences des lois déduites de ses hypothèses ; il suffirait alors de faire appel à l'observation ou à l'expérience, pour apprécier l'exactitude des hypothèses et des lois qui en découlent.

Le principe auquel nous comptons faire appel ici est le lien qui doit exister entre la loi de répartition des revenus et l'expression générale de la loi de la Demande, lien auquel nous avons déjà fait allusion, et pour l'utilisation duquel nous serons conduits à faire certaines hypothèses sur l'affectation des revenus.

Ces lois établies, à partir de nos hypothèses, il resterait à en contrôler l'exactitude par une étude statistique de la Demande ; mais comme nous ne nous trouvons pas en mesure de procéder à de telles vérifications, la présente étude doit être plutôt considérée comme l'indication d'une méthode que comme la recherche de conclusions définitives.

Cette méthode nous permettra non seulement d'aboutir à une forme précise mais encore de démontrer le caractère rationnel des propriétés générales de la loi de la Demande.

LA RÉPARTITION DES REVENUS.

La loi de répartition des revenus à l'intérieur d'une collectivité peut être ainsi conçue :

Si l'on considère les personnes dont le revenu est compris entre r et $r + dr$, leur nombre peut être représenté par une expression de la forme $n\ (r)\ .\ dr$; $n\ (r)$ désignant une fonction du revenu r.

Le nombre des personnes dont le revenu est compris entre r_1 est r_2 donné par l'expression :

$$\int_{r_1}^{r_2} n(r) \,.\, dr$$

En particulier, le nombre de personnes dont le revenu est rupérieur à x résulte de l'expression :

$$N(x) = \int_x^{\infty} n(r)\, dr$$

Le revenu total de l'ensemble étudié serait donné par la formule

$$R_0 = \int_0^{\infty} r \,.\, n(r) \,.\, dr$$

VILFREDO PARETO s'est particulièrement attaché à l'étude de cette question ; il a pu aboutir à des conclusions précises et fort intéressantes :

En utilisant les recherches statistiques antérieures et notamment celles de GIFFEN pour l'Angleterre en 1893 et 1879-1880, il a pu tracer les graphiques de la fonction N ci-dessus définie, représentant le nombre des personnes ayant un revenu supérieur à x, et parvenir à la relation :

$$\log N = \log A - \alpha \log x$$

D'où :

$$N = \frac{A}{x^{\alpha}}$$

Pour obtenir les constantes A et α PARETO interpole les logarithmes de N suivant la méthode de CAUCHY et parvient aux valeurs ci-après relatives à divers pays et à diverses époques :

Angleterre	1843	$\alpha = 1{,}5$
	1879-80	$\alpha = 1{,}35$
Prusse	1876	$\alpha = 1{,}72$
	1881	$\alpha = 1{,}73$
	1886	$\alpha = 1{,}68$
	1894	$\alpha = 1{,}60$
Saxe	1880	$\alpha = 1{,}58$
	1886	$\alpha = 1{,}51$
Villes italiennes		$\alpha = 1{,}45$
Paris (Statistique des loyers)		$\alpha = 1{,}57$

Pour apprécier l'exactitude de sa formule, PARETO calcule l'expression

$$\Delta (\log N - \log N')$$

N représentant les nombres observés et N' les nombres calculés.

Le maximum de Δ pour les pays observés est 3,6 pour 100 ; on obtient la plupart du temps un écart voisin de 1 pour 100 et souvent de quelques millièmes seulement.

PARETO en conclut que cette concordance n'est pas imputable au seul hasard et qu'il existe une cause provoquant la tendance des revenus à se disposer selon une certaine loi.

La formule à laquelle est parvenu PARETO donne naissance à un tracé rectiligne, si l'on représente N en fonction de x, en utilisant des échelles logarithmiques ; certaines droites accusent une légère concavité, mais c'est pour le grand-duché d'Oldenbourg 1880, que la courbure est le plus accusée ; l'interpolation donne pour ce cas particulier :

$$\log N = \log A - \alpha \log (x + a) - \beta x$$

ce qui peut encore s'écrire :

$$N = \frac{A}{(x + a)^{\alpha}} \cdot 10^{-\beta x}$$

C'est probablement la forme générale des courbes de distribution, mais dans la plupart des cas, on peut se contenter de l'expression simplifiée :

$$N = \frac{A}{x^{\alpha}}$$

Cette manière de faire, est d'autant plus justifiée que la valeur de β se trouve être extrêmement petite vis-à-vis de α ; dans le cas du grand-duché d'Oldenbourg on trouve ainsi :

$$\begin{cases} \alpha = 1{,}465 \\ \beta = 0{,}0000274 \end{cases}$$

Dans une thèse publiée à Paris en 1910, M. J. SÉAILLES observe, en prenant pour base les statistiques successorales que, pour les fortunes supérieures à 2.000 F, le diagramme logarithmique de N est pratiquement rectiligne, et plus encore si l'on considère seulement les revenus supérieurs à 10.000 F. Mais si l'on trace la courbe pour toutes les fortunes, y compris les plus petites, la courbe affecte une allure nettement parabolique, susceptible d'être représentée par une expression telle que :

$$(\log x)^2 = 2p \log N + q$$

PARETO avait d'ailleurs lui-même signalé l'incertitude de la portion de courbe correspondant aux revenus les plus bas ; en ce qui nous concerne, nous négligerons les éléments relatifs aux revenus inférieurs à un valeur minima r, telle que l'on puisse faire abstraction des revenus inférieurs à ce minimum, et du nombre des personnes auquel ils correspondent.

Si l'on observe d'autre part que les valeurs indiquées précédemment pour α sont distribuées autour d'une valeur moyenne égale à 1,5, nous serons fondés à nous servir, au cours de la présente étude, de la formule simplifiée

$$N = \frac{A}{x^{\frac{3}{2}}} = \frac{A}{\sqrt{x^3}}$$

Cette approximation nous paraît d'autant plus légitime que notre but est uniquement l'ordre de grandeur des variations de la Demande en fonction de celles du prix.

Des travaux plus récents que ceux de PARETO ont permis d'aboutir, pour la distribution des revenus, à des expressions mathématiques plus complexes et mieux ajustées aux données de l'observation ; si nous croyons devoir malgré tout nos référer à la formule de PARETO, c'est uniquement en raison de sa forme très simple, traduisant la tendance commune des revenus à se grouper d'une certaine manière.

Cependant, pour présenter un aspect des nouvelles conceptions, nous exposerons les données admises par M. GINI, Professeur à l'Université de Rome et Président de l'Institut Central de Statistique du Royaume d'Italie, dans les deux notes qu'il a publiées au sujet de l'indice de concentration : *Indici di concentrazione e di dipendenza*, « Biblioteca dell'Economista », 5ª Serie, vol. 20°, Torino, 1910) et *Sulla misura della concentrazione e della variabilità dei caratteri* (« Atti del R. Istit. Ven. di S. L. A. » t. LXXIII, 1913-1914).

Dans le premier mémoire, M. GINI définit en particulier un « indice de concentration » que l'on peut aisément rattacher aux grandeurs N (Nombre de revenus supérieurs à x) et R (total des revenus supérieurs à x), considérées jusqu'ici.

L'indice de concentration δ résulte de l'équation :

$$\frac{N}{N_0} = \left(\frac{R}{R_0}\right)^{\delta}$$

N et R désignant l'ensemble de tous les revenus et de leur montant total pour la collectivité étudiée.

M. Gini considère d'ailleurs d'autres formules permettant de définir un indice de concentration relatif à des caractères d'une autre espèce que la valeur des revenus ; mais nous devons nous borner à l'examen de l'indice qui a été plus particulièrement envisagé pour la distribution des revenus.

En admettant que la formule simplifiée de Pareto ($N = Ax^{-\alpha}$) soit applicable, on constate aisément que cet indice de concentration se trouve lié à l'exposant α par la relation :

$$\delta = \frac{\alpha}{\alpha - 1}$$

En particulier, les valeurs de α qui ont été indiquées précédemment, et qui se groupent autour d'une valeur moyenne égale à $\frac{3}{2}$ conduiraient à un indice de concentration égal à 3.

Si la formule de Pareto, dont l'approximation n'est qu'assez grossière, donne un indice de concentration constant, il n'en est pas ainsi pour la plupart des autres lois de distribution ; ainsi, M. Gini se trouve-t-il amené à définir un « rapport de concentration » applicable à l'ensemble d'une répartition ; ces considérations font l'objet de son second mémoire, qui déborde d'ailleurs largement le cadre de la répartition des revenus, et comporte l'étude de certaines caractéristiques physiques comme la température sublinguale, les pulsations du cœur et la taille des individus ; d'autres grandeurs économiques sont également envisagées, notamment la répartition des patrimoines, des valeurs locatives, et pour chacune de ces grandeurs, M. Gini détermine la valeur du rapport de concentration.

Ce qu'il convient de retenir ici, c'est le caractère approximatif de la formule simplifiée de Pareto, ainsi que la diversité des valeurs de l'exposant α, suivant la nature des ensembles étudiés, suivant les régions géographiques et suivant les époques auxquelles sont faites les observations.

A côté des travaux de M. le Prof. Gini, nous devons citer des recherches plus récentes de M. Maurice Fréchet, Professeur à l'Institut Poincaré (*) ; elles s'inspirent des considérations suivantes :

(*) Communication au Congrès de l'Institut International de Statistique (Rome 1925) et « Revue Générale des Sciences » (1926).

La formule simplifiée de PARETO présente l'avantage de se prêter au calcul explicite du total des revenus R, supérieurs à x, par une simple intégration ; mais dans bien des cas, l'approximation donnée par cette formule peut sembler insuffisante. La formule complexe de PARETO donne au contraire une approximation bien supérieure mais elle n'est pas intégrable et ne fournit pas l'expression du revenu R ; or, en fait, les observations à partir desquelles sont déterminées les constantes figurant dans les formules portent fréquemment sur le revenu R, aussi bien que sur N.

Pour échapper à ces inconvénients, M. FRÉCHET remarque tout d'abord que lorsqu'on calcule le quotient $\frac{R}{Nx}$ au moyen de la formule simple de PARETO, on obtient une constante susceptible de caractériser l'inégalité de la répartition des revenus. En fait, l'observation numérique montre que ce quotient varie assez peu quand x et N varient.

Il est important de noter ici que ce quotient se confond alors précisément avec l'indice de concentration défini par M. GINI.

Un calcul simple montre en effet qu'en appliquant la formule de PARETO, le rapport $\frac{R}{Nx}$ se trouve exprimé par la fraction $\frac{\alpha}{\alpha - 1}$ que nous avons trouvée pour la valeur de l'indice de concentration δ.

Il y a là une coïncidence tout-à-fait remarquable, tendant à prouver que l'expression choisie par ces deux auteurs pour caractériser la concentration ou l'inégalité de la répartition des revenus, se trouve pour ainsi dire imposée par la nature même du sujet.

L'intérêt qui doit s'attacher à ces recherches se trouve renforcé de ce fait pour tous ceux qui s'efforcent d'attribuer une signification précise à la notion jusqu'alors un peu vague de la concentration d'une distribution.

Pour en revenir aux travaux de M. FRÉCHET, sa méthode consiste à trouver une formule traduisant les variations de l'inverse $\frac{Nx}{R}$ du quotient envisagé et qui, considérée comme une fonction de la variable auxiliaire $u = L(N)$ soit aisément intégrable.

Cette méthode est justifiée aux yeux de M. FRÉCHET, en premier lieu parce que le quotient $\frac{Nx}{R}$ étant une fonction à variations restreintes, il sera possible de la représenter avec une approximation satisfaisante par des fonctions très diverses de la variable u ; en second lieu, un raisonnement mathématique fournit immédiatement la valeur explicite de R en fonction de N, ainsi qu'une relation explicite entre N et x, à condition que la fonction choisie soit intégrable.

A titre d'indication, on pourrait envisager comme fonction $f(u)$ dans la formule préliminaire à ajuster :

$$\frac{Nx}{R} = f\,[L\,(N)]$$

soit un polynôme, soit une exponentielle : $e^{-\frac{h}{u - u_1}}$ ou encore : $e^{-h^2 (u - u_1)}$.

Le fait d'utiliser des fonctions très diverses permet d'appliquer la méthode à la représentation d'autres phénomènes que la distribution des revenus, par exemple la répartition des fortunes, la répartition des villes suivant leur population, celle du nombre des pièces dans les maisons d'une ville, etc.

Le cas échéant, rien ne s'opposerait pour le problème que nous étudions à l'adoption de telles formules ou même de diagrammes de fréquence et de procédés d'intégration graphique.

Dans un autre ordre d'idées, nous devons également observer que les sommes utilisées par les consommateurs proviennent non seulement de leurs revenus mais encore des prêts qui peuvent leur être consentis ; ce dernier élément pouvant être éliminé, si l'on considère une période de temps suffisament étendue pour permettre le remboursement des sommes prêtées, nous nous bornerons à la seule considération des revenus effectifs.

L'affectation des revenus.

Quelle que soit l'importance du revenu, il existe pour chaque individu un certain ordre de priorité relativement à l'affectation de ses ressources à chaque nature d'article. Ce classement dans les besoins individuels pourrait être mis en évidence par une méthode analogue à celle qui fut utilisée par Dupuit pour l'appréciation de l'utilité des divers biens ; il suffirait, en admettant que tous les prix restent constants, d'imaginer que le revenu de l'individu en cause diminue progressivement, de telle manière qu'ils soit obligé de supprimer peu à peu les articles qu'il considère comme les moins indispensables.

Pour l'ensemble des personnes appartenant à la collectivité étudiée, le classement opéré comme il vient d'être dit ne serait évidemment pas le même pour tous les individus, en raison des divergences de goût, mais on conçoit la possibilité de constituer des groupes d'articles tels que, pour l'ensemble des individus, l'ordre de priorité soit observable, non d'un article à l'autre, mais d'un groupe à l'autre.

La mise en œuvre d'un tel procédé ne reviendrait en somme qu'à préciser la notion courante d'objets de première nécessité, d'articles de demi-luxe, de luxe ou de grand luxe. Une telle classification ne peut être envisagée qu'avec bien des précautions et moyennant sans doute une certaine part d'arbitraire, mais il paraît cependant possible de l'établir de telle sorte que tout bien d'une catégorie déterminée soit choisi par tout individu, dans un ordre invariable, par rapport à tout bien d'une catégorie voisine.

Les catégories dont il s'agit traduiraient en définitive la hiérarchie des besoins, tandis qu'à l'intérieur de ces catégories, la sélection des articles serait effectuée d'après les goûts individuels ; nous établissons donc une distinction entre les besoins et les goûts.

Cette classification étant supposée établie, nous étudierons la manière dont varie la quantité totale des biens entrant dans chaque catégorie, en fonction du prix de ces biens.

Envisagée de cette manière, l'étude de la Demande exige la détermination préalable d'index économiques de prix et de quantités, tels que l'ensemble des sommes consacrées à chaque groupe d'articles puisse être considéré comme le produit d'un seul prix par une seule quantité, la loi de la Demande exprimant alors les variations de l'index de quantité en fonction des variations de l'index de prix.

Or, nous disposons d'index économiques jouissant de cette propriété, et, il nous suffira de faire correspondre à chaque groupe d'articles donnant lieu de la part des consommateurs à un ensemble de payements représenté par le symbole $\Sigma\,(q\,p)$ où q désigne les quantités et p les prix, un index de prix P et un index de quantité Q, respectivement définis par les deux relations différentielles :

$$\left\{\begin{aligned} \frac{dP}{P} &= \frac{\Sigma\,(q\,dp)}{\Sigma\,(q\,p)} \\ \frac{dQ}{Q} &= \frac{\Sigma\,(p\,dq)}{\Sigma\,(p\,q)} \end{aligned}\right.$$

Nous savons en effet que les index ainsi définis sont tels qu'on ait identiquement :

$$\frac{P}{P_0} \cdot \frac{Q}{Q_0} = \frac{\Sigma\,(q\,p)}{\Sigma\,(q_0\,p_0)}$$

P_0, Q_0, p_0, q_0, designant les valeurs prises par les index P, Q et les termes en $p\ q$ à l'instant choisi comme point de départ des observations.

Les équations par lesquelles nous définissons les index de prix et de quantité ne doivent pas être considérés comme déterminées d'une manière arbitraire. La forme de ces index est due à M. Divisia, qui en a étudié les propriétés dans l'ouvrage qu'il a consacré à l'Indice monétaire et à la théorie de la monnaie.

Le principe de tels index peut être établi comme il suit : $\Sigma\,(q\,p)$ désignant un ensemble de payements portant sur un genre de transactions bien déterminées, les index P et Q ci-dessus considérés doivent constamment satisfaire à la relation :

$$P\,Q = k\,\Sigma\,(q\,p)$$

k désignant un coefficient constant.

La différentiation de cette équation permet d'aboutir à la relation :

$$\frac{d\,P}{P} + \frac{d\,Q}{Q} = \frac{\Sigma\,(q\,d\,p)}{\Sigma\,(q\,p)} + \frac{\Sigma\,(p\,d\,q)}{\Sigma\,(p\,q)}$$

Cette identité permet d'établir une distinction entre les variations imputables aux changements de prix et celles dues aux modifications des quantités échangées dans l'unité de temps.

En identifiant, dans les deux membres de cette équation, chaque espèce de variation, on en dédut les relations définissant les index P et Q.

Si l'on considère en particulier l'index de prix P, défini par l'équation différentielle :

$$\frac{d\,P}{P} = \frac{\Sigma\,(q\,d\,p)}{\Sigma\,(q\,p)}$$

on constate que le dit index peut être regardé comme une moyenne pondérée, où les coefficients de pondération sont égaux aux quantités échangées q ; on peut également observer que cet index n'est assimilable à un prix moyen, au sens usuel du mot, que si la variation des termes en q est négligeable ; si l'on suppose, en effet, que le prix de chaque article ne varie pas au cours de l'intervalle étudié, tandis que les quantités varient, l'index P conserve la même valeur, en vertu de l'équation :

$$d\,p = 0$$

tandis que le prix moyen change, par suite des modifications survenues dans la distribution de termes en q.

Dans cette même hypothèse, où les prix p sont invariables, l'index de quantité Q reste constamment proportionnel au montant des payements effectués, par suite de l'équation :

$$\frac{d\,Q}{Q} = \frac{\Sigma\,(p\,d\,q)}{\Sigma\,(p\,q)} = \frac{d\,\Sigma\,(p\,q)}{\Sigma\,(p\,q)} \qquad (d\,p = 0)$$

On voit ainsi que les variation de l'index Q mettront en évidence, non seulement les accroissements ou les réductions de la consommation, mais encore les déplacements vers les produits les plus coûteux ou les moins chers ; en d'autres termes, l'utilisation de l'index Q permettra de tenir compte des différences de qualité, observables d'une époque à l'autre, et cette propriété présente un intérêt indiscutable car, ainsi que l'a montré Vilfredo Pareto tout accroissement du revenu individuel entraîne à la fois un accroissement de la quantité de biens consommés en vue de satisfaire un besoin déterminé, et une amélioration de la qualité de ces biens.

Il peut sembler étrange de recourir à la notion d'index pour étudier la loi de la Demande mais, si l'on y réfléchit bien, on remarque qu'en fait, tout prix n'est en somme qu'un index de prix, puisque dans la réalité, tout article présente des différences de qualité ou des modifications de prix telles qu'on soit obligé de recourir à l'usage d'un prix et d'un débit moyen. La notion d'index à laquelle nous faisons appel ici ne constitue donc une innovation que dans la mesure où elle contribue à donner plus de précision à ces valeurs moyennes.

Dans ses *Principes d'Economie politique*, Marshall envisage une étude qui porterait sur la Demande d'un ensemble de plusieurs articles, tels que le thé et le café ; il observe d'ailleurs qu'un même article comme la viande se présente en fait sous plusieurs formes : bœuf, mouton, porc, et qu'un produit comme l'eau peut servir à des usages très divers ; mais il ne donne aucune indication sur la façon dont pourrait être conduite l'étude d'une telle demande totale.

La recherche des lois de la Demande, portant non sur un seul article mais sur un groupe d'articles, a d'ailleurs été expressément recommandée par VILFREDO PARETO qui considérait cette méthode comme la seule propre à tenir compte de différences de qualité.

Il peut être intéressant de comparer notre classification des biens d'après l'intensité des besoins auxquels ils répondent, et l'étude de la Demande afférente à l'ensemble d'un groupe caractérisé par deux index économiques, à la notion de « paramètres économiques », établie à propos des phénomènes de production. Nous pensons qu'il y aurait dans cette étude corrélative de la consommation et de la production une manière d'ébauche d'une théorie des « ensembles économiques » ; cette notion d'ensemble a contribué à la création, peut-être inconsciente, d'une certaine terminologie dans le monde des affaires ou des transactions boursières, et, comme nous l'indiquons plus haut, la mise en œuvre d'index économiques serait de nature à préciser le sens de cette notion.

Ayant ainsi mis à notre disposition les instruments de mesure que constituent les index de quantité et de prix, nous admettrons, en nous conformant d'ailleurs aux enseignements de tous les auteurs précités, qu'il existe pour toute catégorie de biens une consommation individuelle maxima correspondant à la satiété ; dès que le revenu d'un individu lui permet d'acquérir cette quantité-limite, il affecte le surplus de son revenu à des biens placés dans la catégorie supérieure, pour passer ensuite à des catégories encore plus élevées, après obtention de la satiété.

Nous reconnaissons bien volontiers que cette hypothèse est loin d'être inattaquable et l'on pourrait fort bien concevoir l'existence d'une quantité-limite qui se serait plus constante, mais qui varierait avec l'importance du revenu ; certains pourraient même être tentés d'admettre que la fraction du revenu affectée à chaque catégorie de biens, ne varie pas lorsque le revenu s'élève, mais cette hypothèse paraît trop contraire aux faits, pour qu'il soit nécessaire de s'y arrêter. En étudiant la répartition des revenus, M. COLSON fait observer à ce propos que la fraction des ressources affectées au loyer varie beaucoup d'une classe à l'autre et diminue considérablement lorsque s'accroît le revenu.

La solution devrait donc être recherchée dans une hypothèse intermédiaire, mais nous pensons que, moyennant quelques précautions à prendre dans chaque cas particulier, l'hypothèse d'une quantité-limite, correspondant à la satiété des besoins, pour chaque catégorie d'articles, et indépendante de la valeur des revenus individuels, peut utilement servir de base à l'étude de la Demande ; d'ailleurs cette hypothèse est d'autant plus admissible que les revenus étudiés varient dans une proportion restreinte, ce qui paraît être le cas pour la plupart des articles, en raison de la prépondérance des petits et moyens revenus. Pour les revenus plus importants, la partie du revenu consacrée à la satiété des besoins les plus urgents, excède généralement la part des ressources affectées à ces mêmes besoins pour les revenus inférieurs, par suite des différences de qualité observables, même pour les produits de première nécessité ; mais nous ne pensons pas qu'une telle considération puisse avoir une influence sérieuse sur la loi de la Demande, en raison de la rapidité avec laquelle décroît le nombre des individus ayant un revenu déterminé, si l'on considère des revenus de plus en plus élevés.

D'ailleurs, nous ne prétendons pas édifier une théorie complète et définitive, et il suffirait de prendre pour point de départ une autre hypothèse conforme à l'observation des faits, pour aboutir à des conclusions plus voisines de la réalité.

En dernier lieu, nous ferons observer qu'à partir de la loi de la Demande, afférente à un groupe déterminé d'articles, il sera possible d'aboutir aux demandes correspondant à chaque article, moyennant certaines hypothèses consistant, par exemple, à admettre que l'ensemble des revenus affectés à chaque groupe, se répartit suivant des proportions invariables entre chaque article, ou telle autre hypothèse pouvant paraître mieux adaptée aux faits.

Nous reviendrons sur cette question, lorsque nous serons parvenus à l'établissement d'une loi générale pour la Demande.

Bien entendu, nous admettrons dans tout ce qui suit qu'aucune modification ne se produit, au cours de l'intervalle étudié, dans le chiffre de la population, dans celui des revenus, et dans leur loi de repartition.

Parmi les objets de consommation, une place toute spéciale doit être réservée à ceux qui sont acquis en vue de l'épargne ; pour ces biens, il ne serait certainement pas impossible d'établir une classification d'après l'importance des revenus qui en permettent l'acquisition, mais la classification paraît encore plus délicate à établir que pour les autres espèces de biens.

Les Objets de première nécessité.

Avant d'en arriver à l'étude des groupes, dont nous avons décrit le mode de constitution, nous ferons observer que si l'on embrasse la totalité des consommations, l'équation de la Demande est définie par la relation :

$$Q \,.\, P = R_0$$

Q et P désignant respectivement les index de quantité et de prix pour l'ensemble des consommations et R_0 le revenu total de la collectivité étudiée.

La loi de la Demande serait ainsi de nature hyperbolique, mais la courbe qui la représente ne constitue qu'une courbe moyenne, autour de laquelle se groupent les courbes relatives à la Demande correspondant à chaque groupe de produits particuliers.

L'équation dont il s'agit ne traduit donc qu'une tendance générale, et, pour nous faire une idée plus exacte des faits, nous devons en venir à la notion de groupe ; en premier lieu, nous rechercherons la loi de la Demande pour les objets de première nécessité, c'est-à-dire pour les produits dont l'acquisition est envisagée avant tout autre.

D'une façon générale, nous désignerons par q_l la valeur maxima de l'index de quantité pour chaque individu ; nous avons antérieurement admis que cette limite assurant la satiété pour le groupe de produits étudiés, était la même pour chaque individu, qu'elle ne dépendait pas en particulier du montant du revenu individuel. Cette limite q_l est telle que si l'index de prix correspondant au groupe étudié est égal à P, le revenu individuel assurant la satiété se trouve représenté par le produit $q_l\,P$.

Pour aboutir à la loi de la Demande des objets de première nécessité, nous observerons que l'ensemble des payements $Q.P$ relatif à cette catégorie d'objets, peut se répartir en deux groupes :

a) Groupe des personnes dont la consommation individuelle est imparfaite et reste par conséquent inférieure à la limite q_l.

b) Groupe des consommateurs dont le revenu individuel est suffisant pour permettre l'acquisition de la quantité-limite.

a) Pour le premier groupe, c'est-à-dire pour les individus dont le revenu est compris entre la limite inférieure r_0 et le revenu $q_l\,P$ permettant d'acquérir la quantité-limite, le total des revenus est affecté à l'acquisition des objets appartenant au groupe étudié, puisque ce sont, par hypothèse, des objets de première nécessité.

L'ensemble des payements ainsi envisagés peut donc être représenté par le symbole :

$$\int_{r_0}^{q_l\,P} r \,.\, n\,(r) \,.\, d\,r$$

b) Pour la seconde catégorie d'individus, ceux dont le revenu permet la satiété, l'ensemble des payements correspondants aux produits de première nécessité est égal au produit du revenu individuel $q_l P$ par le nombre des personnes dont le revenu est supérieur à cette quantité ; l'ensemble des payements dont il s'agit est donc représenté par l'expression :

$$q_l\ P \int_{q_l\ P}^{\infty} n\ (r)\ dr$$

Nous pouvons donc écrire :

$$Q\ .\ P = \int_{r_0}^{q_l\ P} r\ .\ n\ (r)\ .\ dr + q_l\ P \int_{q_l\ P}^{\infty} n\ (r)\ .\ d\ r\ .$$

La connaissance de la fonction $n(r)$ traduisant la loi de répartition des revenus permet d'exprimer le second membre de l'équation ci-dessus en fonction de l'index des prix P ; elle conduit par suite à la loi de la Demande.

Nous nous bornerons à utiliser la forme simplifiée correspondant à l'équation :

$$\int_{x}^{\infty} n\ (r)\ .\ d\ r = A\ .\ x^{-\ \alpha}$$

de laquelle on déduit :

$$n\ (r) = \alpha\ A\ .\ r^{-\ (\alpha\ +\ 1)}$$

On obtient ainsi :

$$Q\ .\ P = \alpha\ A\ .\int_{r_0}^{q_1 .\ P} r^{-\ \alpha}\ .\ d\ r + q_l\ P\ .\ A\ .\ (q_l\ P)^{-\ \alpha}$$

En intégrant, on aboutit à la relation :

$$Q\ .\ P = A\ .\ \frac{\alpha}{\alpha - 1} \cdot \left[r_0^{\ -\ \alpha\ +\ 1} - (q_l .\ P)^{-\ \alpha\ +\ 1}\right] + A\ q_l\ P\ .\ (q_l\ P)^{-\ \alpha}$$

$$Q\ .\ P = A\ .\ \frac{\alpha}{\alpha - 1}\ .\ r_0^{\ -\ \alpha\ +\ 1} - \frac{A}{\alpha - 1}\ .\ (q_l\ P)^{-\ \alpha\ +\ 1}$$

On obtient ainsi, en définitive, l'expression de la loi de la Demande :

$$Q = \frac{A}{\alpha - 1} \cdot \left[\frac{\alpha}{r_0^{\ \alpha\ -\ 1}} \cdot \frac{1}{P} - \frac{1}{q_l^{\ \alpha\ -\ 1}} \cdot \frac{1}{P^{\alpha}}\right]$$

Nous signalerons immédiatement que cette expression n'est valable que pour les valeurs de P supérieures à :

$$P_l = \frac{r_0}{q_l}$$

Dès que P atteint cette limite inférieure, le revenu minimum r_0 assure la satiété q_l pour tous les individus, et la consommation ne s'accroît plus pour les valeurs de P inférieures à P_l, le surplus des revenus étant affecté à la satisfaction de besoins rangés dans les catégories supérieures ; la valeur-limite de Q est ainsi fournie par l'expression :

$$Q_l = q_l \, . \, A \, . \, r_0^{\,-\alpha}$$

Connaissant maintenant l'expression du débit Q, nous étudierons la fonction qui le représente, en nous proposant notamment de rechercher le sens des variations et celui de la concavité pour la courbe représentative.

Etude de la fonction Q.

Nous rappelons tout d'abord l'expression du débit en fonction du prix :

$$Q = \frac{A}{\alpha - 1} \cdot \left[\frac{\alpha}{r_0^{\,\alpha - 1}} \cdot \frac{1}{P} - \frac{1}{q_l^{\,\alpha - 1}} \cdot \frac{1}{P^\alpha} \right]$$

En posant, pour simplifier :

$$\begin{cases} B = \dfrac{A}{\alpha - 1} \cdot \dfrac{1}{r_0^{\,\alpha - 1}} \\[2ex] P_l = \dfrac{r_0}{q_l} \end{cases}$$

On obtient :

$$Q = B \, . \left[\frac{\alpha}{P} - \frac{P_l^{\,\alpha - 1}}{P^\alpha} \right] \quad \text{ou encore :} \quad Q = \frac{B}{P} \left[\alpha - \left(\frac{P_l}{P} \right)^{\alpha - 1} \right]$$

Nous observons immédiatement, sur la première forme qui vient d'être donnée en dernier lieu, que la courbe représentative de la consommation en fonction du prix, pourrait être obtenue par soustraction des ordonnées de deux courbes, dont l'une serait une hyperbole en $\frac{1}{P}$ et l'autre, une courbe en $\frac{1}{P\alpha}$ dont l'allure est également hyperbolique, puisque α est voisin de $\frac{3}{2}$.

Mais nous ne pouvons nous contenter de cette indication et devons entreprendre l'étude approfondie de la fonction.

Nous remarquerons, avant d'étudier la dérivée de Q par rapport à P, que pour les valeurs élevées de P, le terme prépondérant est celui en $\frac{1}{P}$, puisque l'exposant α est toujours supérieur à l'unité ; il est d'autre part inutile de considérer les valeurs infiniment petites de P, puisqu'à partir de la limite P_l, la formule donnée pour Q n'est plus applicable, la consommation restant égale à la consommation-limite Q_l.

En prenant la dérivée de Q par rapport à P, on trouve :

$$\frac{dQ}{dP} = \alpha \, . \, B \, . \left[\frac{P_l^{\alpha-1}}{P^{\alpha+1}} - \frac{1}{P^2}\right]$$

$$\frac{dQ}{dP} = -\frac{\alpha B}{P^2} \cdot \left[1 - \left(\frac{P_l}{P}\right)^{\alpha-1}\right]$$

P restant toujours supérieur à P_l, $\frac{P_l}{P}$ reste constamment inférieur à l'unité, de telle sorte que la dérivée de Q par rapport à P reste elle-même négative, sauf pour P_l, valeur pour laquelle elle s'annule.

On voit ainsi que la consommation Q est une fonction décroissante du prix P, tant que celui-ci reste supérieur au prix-limite P_l ; pour les valeurs de P inférieures à cette limite, la consommation reste constamment égale à sa valeur maxima Q_l.

On peut également observer, en tenant compte du fait que la dérivée $\frac{dQ}{dP}$ s'annule pour le prix limite P_l, alors que la fonction Q a pour valeur principale un terme en $\frac{1}{P}$ pour les valeurs élevées de P, que la courbe représentative de cette fonction doit présenter un point d'inflexion pour une valeur de P supérieure à P_l ; mais nous ne pouvons rien affirmer à cet égard, tant que nous n'avons pas abordé l'étude de la dérivée seconde.

Il s'agit d'étudier la dérivée de la fonction :

$$y' = -\frac{1}{P^2} + \frac{P_l^{\alpha-1}}{P^{\alpha+1}}$$

On trouve pour la dérivée de cette fonction :

$$y'' = \frac{2}{P^3} - (\alpha + 1) \cdot \frac{P_l^{\alpha - 1}}{P^{\alpha + 2}}$$

$$y'' = \frac{1}{P^3} \cdot \left[2 - (\alpha + 1) \cdot \left(\frac{P_l}{P}\right)^{\alpha - 1} \right]$$

On constate sur cette expression que la dérivée seconde $\frac{d^2 Q}{d P^2}$ est bien positive pour les valeurs élevées de P, mais qu'il existe une valeur-critique P_c pour laquelle cette dérivée seconde s'annule ; cette valeur-critique est telle que :

$$2 - (\alpha + 1) \cdot \left(\frac{P_l}{P_c}\right)^{\alpha - 1} = 0$$

$$\left(\frac{P_l}{P_c}\right)^{\alpha - 1} = \frac{2}{\alpha + 1}$$

soit encore :

$$P_c = P_l \cdot \left(\frac{\alpha + 1}{2}\right)^{\frac{1}{\alpha - 1}} > P_l$$

Pour les valeurs de P supérieures à cette valeur critique, la concavité de la courbe représentative de Q en fonction de P, est orientée vers la direction positive de l'axe des quantités ; pour la valeur critique P_c, la courbe présente un point d'inflexion, et pour les valeurs comprises entre P_c et P_l, la concavité se trouve dirigée vers la partie négative de l'axe des quantités ; lorsque P atteint la valeur P_l la courbe présente un maximum et devient par suite tangente au palier horizontal qui représente la Demande pour les valeurs de P inférieures à P_l.

La figure II, à page 153 donne une idée de la courbe à laquelle nous aboutissons ; elle diffère assez sensiblement des courbes habituellement présentées, lesquelles admettent en général une concavité orientée vers la direction positive de l'axe des quantités ; mais il ne s'agit que d'une courbe théorique, et dans la pratique, il y aurait certainement des divergences à relever, aussi bien pour les valeurs de P inférieures au prix-limite P_l que pour les valeurs élevées de P, conformément aux remarques générales antérieurement faites à propos de la loi de substitution.

Quoi qu'il en soit, l'étude précédente paraît bien mettre en évidence le fait que, pour les objets de première nécessité, l'allure de la courbe représentative de la Demande, n'est pas forcément conforme aux tracés les plus usuels, en raison de l'existence du point d'inflexion I sur la courbe théorique ; dans les pays assez développés pour que les ressources de la très grande majorité des personnes permettent d'acquérir la totalité des objets procurant la satiété des besoins les plus urgents, il y a même de fortes chances pour que la loi du débit soit représentée par une courbe se rapprochant plus du palier MN ou du segment de courbe MI, que de la branche située à droite du point d'inflexion.

Nous observerons en dernier lieu, qu'en vertu de la formule :

$$P_c = P_l . \left(\frac{\alpha + 1}{2}\right)^{\frac{1}{\alpha - 1}}$$

l'allure même de la courbe représentative de la Demande se trouve intimement liée à la loi de répartition des revenus, laquelle ne dépend que de l'exposant α.

Il serait intéressant de soumettre nos conclusions à l'épreuve des faits, mais, comme nous l'avons déjà indiqué au début de cette étude, les données expérimentales en cette matière sont encore assez réduites et il faudrait entreprendre des recherches statistiques spéciales pour aboutir à des résultats précis.

Toutefois, à titre d'approximation, il serait intéressant de rapprocher les résultats obtenus par nos formules de ceux qu'indiquent Jevons et Marshall à propos de la Demande du blé, et que nous avons reproduits antérieurement.

Pour effectuer cette comparaison, il suffirait d'utiliser la relation ci-après, que nous avons établie avant d'aboutir à l'expression définitive de la loi de la Demande :

$$Q\,P = C + m\,.\,P^{-\alpha+1}$$

C désignant une constante et m un coefficient également constant.

On donnerait successivement, dans cette formule à Q les valeurs $1,\ 1 - \frac{1}{10}\ \ldots\ 1 - \frac{5}{10}$ et, à P les valeurs correspondantes indiquées par Marshall ; on éliminerait la constante C en comparant les différences des produits $Q.\ P$ aux différences des expressions

$P^{-\alpha+1}$; admettant, par exemple, pour α, la valeur $\frac{3}{2}$, le coefficient m résulterait d'une équation de la forme :

$$m = \frac{Q_2 P_2 - Q_1 P_1}{\frac{1}{\sqrt{P_1}} - \frac{1}{\sqrt{P_2}}}$$

On apprécierait ainsi les écarts obtenus dans le calcul de m, en adoptant pour P_1, Q_1, P_2, Q_2, les diverses combinaisons résultant des données susindiquées. Les divergences observées seraient partiellement imputables aux phénomènes de substitution.

Expression différentielle de la Demande. Elasticité.

Avant d'aborder l'étude de la consommation des produits autres que les objets de première nécessité, nous devons nous arrêter à une propriété fort importante de la loi de la Demande, propriété différentielle que nous retrouverons à propos des produits en général, et qui permet d'aboutir à une expression facilement interprétable pour les besoins de la pratique.

Nous partirons, à cet effet de l'équation qui nous a permis d'aboutir à la loi de la Demande pour les objets de première nécessité, soit :

$$Q\,P = \int_{r_0}^{q_l P} r \,.\, n\,(r) \,.\, d\,r + q_l\,P \int_{q_l\,P}^{\infty} n\,(r) \,.\, d\,r$$

On observe aisément, en dérivant sous le signe $\int$, que la dérivée du premier membre $Q.P$ par rapport à P se trouve précisément égale, dans le second membre, à :

$$q_l \int_{q_l\,P}^{\infty} n\,(r) \,.\, d\,r\,.$$

Nous n'insistons pas sur le calcul, d'ailleurs évident, car nous le reprendrons complètement, en étudiant la loi générale de la Demande, pour les produits autres que les objets de première nécessité.

On aboutit ainsi à la relation :

$$\frac{d\,(Q\,P)}{d\,P} = q_l \int_{q_l\,P}^{\infty} n\,(r) \,.\, d\,r = Q_s.$$

Q_s désignant la consommation totale des personnes dont le revenu est supérieur à $q_l P$ c'est-à-dire la consommation totale des personnes en mesure de s'assurer la satiété pour les produits envisagés.

Il en résulte qu'on peut écrire :

$$d\,(Q\,P) = Q_s\,d\,P$$

ou encore, en faisant apparaître les différentielles logarithmiques $\frac{d\,Q}{Q}$ et $\frac{d\,P}{P}$:

$$\frac{d\,Q}{Q - Q_s} = -\frac{d\,P}{P}$$

$$\frac{d\,Q}{Q} = -\frac{d\,P}{P} \cdot \left(1 - \frac{Q_s}{Q}\right) = -\frac{\lambda\,d\,P}{P}$$

Dans cette dernière équation, le coefficient

$$\lambda = 1 - \frac{Q_s}{Q} = \frac{Q - Q_s}{Q}$$

représente le rapport des consommations imparfaites à la consommation totale Q.

Cette dernière équation exprimant la différentielle logarithmique de Q en fonction de celle de P, par l'intermédiaire du coefficient λ, pourrait être aisément traduite en diagramme logarithmique, tel que pour les valeurs initialement considérées de Q et de P, la courbe représentative de la loi de la Demande soit, en première approximation, représentée par une droite dont le coefficient angulaire serait égal à $-\lambda$, en prenant pour déterminer λ les valeurs initiales de P et de Q.

L'expression simple ainsi trouvée pourrait être facilement utilisée en pratique, où l'on se propose le plus souvent d'exprimer des accroissements de prix ou de quantités en pourcentages de leurs valeurs initiales ; la considération du coefficient λ permet d'attribuer un sens précis à la notion d'élasticité envisagée par MARSHALL, et susceptible d'une interprêtation remarquable, en raison de l'expression de ce coefficient par rapport aux quantités Q et Q_s.

Etant donnée l'importance pratique de ce coefficient caractérisant l'élasticité de la Demande, nous estimons qu'il est indispensable de rechercher son expression, afin d'en déduire ses propriétés.

Or, en partant de l'équation qui le définit, soit :

$$\lambda = \frac{Q - Q_s}{Q}$$

on parvient à l'expression simple :

$$\lambda = \frac{\alpha . x^{\alpha - 1} - \alpha}{\alpha . x^{\alpha - 1} - 1}$$

expression déduite des valeurs sus-indiquées pour Q et Q_s, et dans laquelle nous représentons par x le rapport de l'index de prix P à l'index minimum de prix :

$$P_l = \frac{r_0}{q_l}$$

On observe immédiatement sur cette expression que λ s'annule pour $x = 1$ c'est-à-dire lorsque le prix est égal au prix-limite ; c'est un résultat déjà connu, puisque la consommation atteint son maximum pour ce prix-limite ; pour les valeurs de P supérieures à cette limite, le coefficient croît continuement en fonction de la variable x, pour tendre vers l'unité lorsque le prix augmente indéfiniment

Pour les valeurs usuelles de P, l'élasticité conserve une valeur qui, en première approximation, peut être regardée comme constante, tout au moins dans un champ de variation limité, susceptible de contenir les cas les plus fréquents d'application pratique.

La considération de ce coefficient, et sa recherche au moyen d'observations statistiques, permettrait donc à notre sens de préciser d'une façon déjà très satisfaisante la nature de la Demande pour le groupe de produits envisagés.

Nous ferons d'ailleurs observer qu'en vertu de l'équation :

$$\frac{dQ}{Q} = -\lambda \frac{dP}{P}$$

la dérivée de Q par rapport à P a elle-même pour expression :

$$\frac{dQ}{dP} = -\lambda \frac{Q}{P} < 0$$

λ étant positif en vertu de sa définition même.

On n'aurait cependant pas pu se dispenser de trouver l'expression complète de la dérivée de Q, puisqu'il était nécessaire de connaître également l'expression de la dérivée seconde.

D'autre part, nous pouvons mettre l'équation ;

$$\frac{dQ}{Q} = -\lambda \frac{dP}{P}$$

sous la forme :

$$\Sigma\,(p\,dq) = -\lambda\,\Sigma\,(q\,d\,p)$$

qui traduit d'une façon concrète la variation d'ensemble des quantités q en fonctions de la variation des prix p.

Enfin, l'avantage de l'équation différentielle de la demande est de faire intervenir des différentielles logarithmiques, c'est-à-dire en pratique des variations relatives, indépendantes des unités choisies ; le coefficient caractérisant l'élasticité est lui-même dénué de dimensions, comme un assez grand nombre de constantes physiques (indice de réfraction, coefficient de dilatation et coefficient de frottement.....).

Si l'on constate par exemple que, pour une certaine catégorie de produits, une variation relative v du prix moyen se traduit par un fléchissement relatif f de la consommation moyenne correspondante, nous admettrons en première approximation que l'elasticité de la Demande, relative à l'ensemble considéré, peut être exprimée par la formule

$$\lambda = \frac{f}{v}$$

ce qui s'interprête de la manière suivante : pour les produits envisagés, la proportion des consommations inférieures à celle qui procure la satiété, par rapport à la consommation totale, est exprimée par le rapport λ, et la loi de la Demande est approximativement représentée par l'équation :

$$\frac{d\,Q}{Q} = -\lambda\,\frac{d\,P}{P}$$

soit encore :

$$Q = Q_0 \cdot \left(\frac{P}{P_0}\right)^{-\lambda}$$

Cas de produits autres que les objets de première nécessité.

Pour les produits autres que les objets de première nécessité, que nous avons classés d'après l'intensité des besoins auxquels ils répondent, nous supposerons qu'il existe toujours une valeur-limite q_l de la consommation individuelle, susceptible d'assurer la satiété des besoins pour l'ensemble étudié, ladite limite étant indépendante du montant du revenu individuel.

Nous représenterons par ρ la partie du revenu individuel consacrée à l'acquisition de l'ensemble des produits classés dans les catégories inférieures à celle que nous étudions ; nous admettrons que ce revenu est supérieur à la limite inférieure r_0 du revenu individuel, de telle manière que l'acquisition du groupe de produits envisagé, ne puisse être réalisé que par une partie seulement de la population.

Comme pour les objets de première nécessité, le total des payements effectués pour l'acquisition des biens étudiés, se décompose en deux groupes, à savoir :

a) Payement effectués par les personnes dont le revenu ne permet pas l'acquisition de la quantité-limite q_l.

Pour cette première catégorie, le revenu individuel est compris entre les limites ρ et $(\rho + q_l P)$, P désignant le prix moyen de l'ensemble étudié ; le montant des payements effectués pour l'achat du groupe se trouve ainsi représenté par l'expression :

$$\int_{\rho}^{\rho + q_l P} (r - \rho) \, . \, n(r) \, . \, dr$$

b) Payements effectués par les personnes dont le revenu est supérieur à celui qui permet d'acquérir la quantité limite q_l.

Les payements relatifs à ce groupe sont représentés par l'expression :

$$q_l P \, . \int_{\rho + q_l P}^{\infty} n(r) \, . \, dr$$

En groupant ces deux ensembles de payements, nous aboutissons à l'équation :

$$Q\,P = \int_{\rho}^{\rho + q_l P} (r - \rho) \, . \, n(r) \, . \, dr + q_l \, P \int_{\rho + q_l P}^{\infty} n(r) \, . \, dr.$$

La connaissance de la fonction $n(r)$ traduisant la loi de répartition des revenus permet d'exprimer le second membre de l'équation précédente en fonction de l'index P, ce qui donne en définitive l'expression de la loi de la Demande.

Nous pourrions, comme pour les objets de première nécessité, utiliser dès maintenant la forme simplifiée :

$$\int_x^\infty n(r) \,.\, d\,r = A \,.\, x^{-\alpha}$$

mais nous préférons nous servir de la propriété sus-indiquée, de la dérivée du produit $Q.P.$, par rapport à P ; ce procédé nous permettra d'exposer plus simplement la suite des calculs conduisant à l'expression définitive de la loi de la Demande.

Pour calculer la dérivée de $Q.P.$ par rapport à P nous poserons :

$$z = \rho + q_l P$$

ρ et q_l ayant des valeurs constantes, nous en déduisons :

$$d\,z = q_l \,.\, d\,P$$

et par suite :

$$\frac{d\,(Q\,P)}{d\,P} = \frac{d\,(Q\,P)}{d\,z} \cdot \frac{d\,z}{d\,P} = q_l \,.\, \frac{d\,(Q\,.\,P)}{d\,z}$$

Or, l'expression du produit $Q.P$ peut encore s'écrire :

$$Q\,P = \int_\rho^{\rho + q_l P} r \,.\, n(r) \,.\, d\,r - \rho\, N(\rho) + (\rho + q_l P) \,.\, \int_{\rho + q_l P}^\infty n(r) \,.\, d\,r$$

$N(\rho)$ désignant le nombre total des personnes dont le revenu est supérieur à P et tel, par conséquent que :

$$N(\rho) = \int_\rho^\infty n(r)\, d\,r\,.$$

En introduisant la variable z ci-dessus définie, nous pouvons écrire :

$$Q\,P = \int_\rho^z r \,.\, n(r) \,.\, d\,r - \rho\, N(\rho) + z \,.\, \int_z^\infty n(r) \,.\, d\,r$$

En dérivant sous le signe $\int$, on obtient pour la dérivée de $Q.P$ par rapport à z :

$$\frac{d\,(Q\,P)}{d\,z} = z \,.\, n(z) + \int_z^\infty n(r) \,.\, d\,r - z \,.\, n(z) = \int_z^\infty n(r) \,.\, d\,r\,.$$

Nous en déduisons :

$$\frac{d(QP)}{dP} = q_l \cdot \int_{\rho + q_l P}^{\infty} n(r) \, . \, dr = Q_s \, .$$

Nous retrouvons ainsi, sous sa forme la plus générale la propriété différentielle de la Demande et nous allons nous en servir pour en tirer l'expression de Q en fonction de P.

Nous utiliserons à cet effet la forme simplifiée :

$$\int_{z}^{\infty} n(r) \, . \, dr = A \, . \, z^{-\alpha} \, .$$

On en déduit :

$$\frac{d(QP)}{dP} = q_l A (\rho + q_l P)^{-\alpha} \, .$$

En intégrant cette expression par rapport à P, on trouve :

$$QP = C + A \, q_l \int_{\rho + q_l P}^{\infty} (\rho + q_l P)^{-\alpha} \, dP$$

soit encore :

$$QP = C - \frac{A}{\alpha - 1} \cdot (\rho + q_l P)^{-\alpha + 1}$$

ce qui donne :

$$Q = \frac{C}{P} - \frac{A}{\alpha - 1} \cdot \frac{1}{P \, . \, (\rho + q_l P)^{\alpha - 1}}$$

Pour déterminer la constante d'intégration C, nous tiendrons compte du fait que, pour un prix P tendant vers 0, la consommation totale Q tend vers une limite finie, qui doit être précisément égale au produit par q_l de l'ensemble des personnes dont le revenu est supérieur à ρ.

Pour la valeur 0 de P, les coefficients de $\frac{1}{P}$ doivent donc être égaux, ce qui donne :

$$C = \frac{A}{\alpha - 1} \cdot \frac{1}{\rho^{\alpha - 1}} \, .$$

On en tire l'expression définitive de la consommation, en fonction du prix, soit :

$$Q = \frac{A}{\alpha - 1} \cdot \frac{1}{\rho^{\alpha - 1}} \cdot \frac{1}{P} \cdot \left[1 - \left(\frac{\rho}{\rho + q_l P}\right)^{\alpha - 1}\right].$$

En posant :

$$\begin{cases} B = \dfrac{A}{\alpha - 1} \cdot \dfrac{1}{\rho^{\alpha - 1}} \\ \gamma = \dfrac{q_l}{\rho} \end{cases}$$

on trouve comme expression de Q :

$$Q = B \cdot \frac{1}{P} \cdot \left[1 - \frac{1}{(1 + \gamma P)^{\alpha - 1}}\right].$$

On observera sur cette expression que, lorsque P tend vers o, Q tend vers la valeur $Q_l = A q_l \rho^{-\alpha}$. Il suffit à cet effet de remarquer que la valeur principale de :

$$1 - \frac{1}{(1 + \gamma P)^{\alpha - 1}}$$

est égale à

$$\gamma \cdot (\alpha - 1) P$$

et c'est bien là le résultat conforme à ce que nous avons déjà dit, au sujet de la consommation limite Q_l

Etude de la fonction Q.

Pour étudier la consommation en fonction du prix, nous conserverons l'expression trouvée en dernier lieu :

$$Q = B \cdot \frac{1}{P} \cdot \left[1 - \frac{1}{(1 + \gamma P)^{\alpha - 1}}\right]$$

Nous pouvons dès maintenant observer que, pour les valeurs infiniment grandes de P, la partie principale de Q est constituée par le terme $\dfrac{B}{P}$ puisque nous supposons toujours l'exposant α supérieur à l'unité ; la courbe admet donc l'axe des prix comme asymptote.

En second lieu, nous observons que la courbe représentative de Q en fonction de P peut être obtenue par soustraction des ordonnées de l'hyperbole en $\frac{1}{P}$ et de la courbe en $\frac{1}{P(1+\gamma P)^{\alpha-1}}$; ces deux courbes ayant une allure décroissante, nous ne pouvons rien affirmer à priori, quant au sens de variation de la fonction Q et il est nécessaire de passer à l'étude de sa dérivée par rapport à P.

En faisant abstraction du coefficient B, nous avons à étudier la fonction

$$y = \frac{1}{P} \cdot \left[1 - \frac{1}{(1+\gamma P)^{\alpha - 1}} \right].$$

Nous trouvons pour la dérivée par rapport à P :

$$y' = -\frac{1}{P^2} \cdot \left[1 - \frac{1}{(1+\gamma P)^{\alpha-1}} \right] + \gamma \,.\, (\alpha - 1) \,.\, \frac{1}{P \,.\, (1+\gamma P)^{\alpha}}$$

$$y' = -\frac{1}{P^2} + \frac{1}{P^2 (1+\gamma P)^{\alpha-1}} + \gamma \,.\, (\alpha - 1) \cdot \frac{P}{P^2 (1+\gamma P)^{\alpha}}$$

$$y' = -\frac{1}{P^2}\left[1 - \frac{1+\alpha\gamma P}{(1+\gamma P)^{\alpha}} \right]$$

La dérivé de y par rapport à P est donc le produit par la fonction négative $-\frac{1}{P^2}$ de l'expression :

$$1 - \frac{1+\alpha\gamma P}{(1+\gamma P)^{\alpha}}$$

expression dans laquelle α désigne, comme nous l'avons vu, un exposant supérieur à 1.

La fonction $(1 + \gamma P)^{\alpha}$ prend la valeur 1 pour $P = 0$, et croît constamment lorsque P augmente ; sa dérivée, pour $P = 0$ est égale au produit $\alpha \gamma$; elle a donc même valeur que la dérivée de la fonction linéaire $(1 + \alpha \gamma P)$ figurant au numérateur de la fraction entre crochets.

La courbe représentative de la fonction $(1 + \gamma P)^{\alpha}$ est donc tangente, pour $P = 0$ à la droite représentative de la fonction $(1 + \alpha \gamma P)$, et, comme α est supérieur à 1, la droite est constamment au-dessous de ladite courbe ; le dénominateur de la fraction entre crochets reste constamment supérieur au numérateur, de telle sorte que l'expression entre crochets est elle-même toujours positive.

La fonction Q est donc en définitive une fonction décroissante du prix, mais pour tracer la courbe, il est nécessaire d'étudier la dérivée seconde, à l'effet d'étudier le sens de la concavité.

Faisons observer auparavant que, pour $P = 0$, nous trouvons pour limite de la dérivée la valeur :

$$y' = -\alpha \cdot \frac{\alpha - 1}{2} \cdot \gamma^2.$$

Cette valeur étant négative, on est amené à penser que la concavité de la courbe est toujours orientée vers la partie positive de l'axe des quantités, mais nous ne pouvons rien affirmer, tant que nous n'avons pas calculé la dérivée seconde ; celle-ci a pour expression, en posant :

$$\left\{ \begin{aligned} u &= \frac{1 + \alpha \gamma P}{(1 + \gamma P)^{\alpha}} \\ y' &= -\frac{1}{P^2}(1 - u) \end{aligned} \right.$$

soit encore :

$$y'' = \frac{2}{P^3}(1 - u) + \frac{u'}{P^2}.$$

$$y'' = \frac{2}{P^3}\left[(1 - u) + \frac{P u'}{2}\right]$$

Il suffit maintenant de calculer u'. On trouve :

$$u' = \frac{\alpha \gamma}{(1 + \gamma P)^{\alpha}} - \frac{\alpha \gamma (1 + \alpha \gamma P)}{(1 + \gamma P)^{\alpha + 1}}$$

$$u' = \frac{\alpha \gamma}{(1 + \gamma P)^{\alpha + 1}} \cdot [(1 + \gamma P) - (1 + \alpha \gamma P)]$$

$$-u' = \frac{\alpha(\alpha - 1)\gamma^2 P}{(1 + \gamma P)^{\alpha + 1}}$$

On en déduit :

$$y'' = \frac{2}{P^3}\left[1 - \frac{1 + \alpha \gamma P}{(1 + \gamma P)^{\alpha}} - \frac{\alpha(\alpha - 1)}{2} \cdot \frac{\gamma^2 P^2}{(1 + \gamma P)^{\alpha + 1}}\right].$$

En mettant y'' sous la forme d'une fraction

$$\frac{2\,M}{P^3\,(1+\gamma\,P)^{\alpha+1}}$$

on observe que le numérateur M est représenté par l'expression :

$$M = \left\{(1+\gamma\,P)\cdot\left[(1+\gamma\,P)^{\alpha} - (1+\alpha\,\gamma\,P)\right] - \frac{\alpha\,(\alpha-1)}{2}\cdot\gamma^2\,P^2\right\}$$

Sous cette forme, on observe que le numérateur M est la différence de deux fonctions de degrés $\alpha + 1$ pour la première, et de degré 2 pour la seconde :

$$\alpha\,(\alpha-1)\cdot\gamma^2\,P^2.$$

L'exposant $(\alpha + 1)$ étant d'un degré supérieur à 2, la première fonction croît plus rapidement que la seconde, pour les valeurs positives de P.

On observe d'autre part que les deux fonctions s'annulent pour $P = 0$ et que la première est un infiniment petit du second ordre, équivalent à la fonction : $\frac{\alpha\,(\alpha-1)}{2}\,\gamma^2\,P^2$; le numérateur M est donc lui-même un infiniment petit du troisième ordre, dont le coefficient $\frac{\alpha\,(\alpha-1)\,(\alpha+1)}{6}\,\gamma^3$ est positif.

Ces deux fonctions constituant le numérateur M sont donc représentées par deux courbes passant par l'origine, tangentes à l'axe des prix en ce point et possédant également en ce point un contact du second ordre, de telle manière qu'au-delà de l'origine, pour les valeurs positives de P, la courbe représentative de la première fonction soit constamment au-dessus de la parabole représentative de la seconde.

y'' est donc positif pour toutes les valeurs positives de P, et tend vers une limite positive quand le prix s'annule. L'allure de la courbe représentative de la Demande est donc bien conforme à celle des courbes envisagées d'ordinaire, sauf cette particularité d'admettre l'axe des prix comme asymptote, pour les valeurs infiniment grandes de P ; nous reviendrons d'ailleurs sur cette particularité (voir Figure III, page 153).

En ce qui concerne la représentation graphique, une importante distinction doit donc être faite entre les objets de première nécessité et ceux dont l'acquisition suppose l'existence d'un revenu supérieur à celui qui est absorbé par les objets de première nécessité.

Nos conclusions sont, somme toute, conformes aux affirmations habituelles, mais elles se présentent ici comme l'aboutissement d'une théorie rationnelle et non plus comme de simples affirmations à priori.

Nous allons maintenant essayer de déduire des résultats acquis pour un ensemble d'articles, les lois propres à la Demande d'un article particulier.

DEMANDE RELATIVE À UN PRODUIT PARTICULIER. FORME PROBABLE DES COURBES RÉELLES.

Pour passer de la Demande d'un groupe de produits à celle d'un produit unique appartenant à ce groupe, il suffit de faire certaines hypothèses sur l'affectation des revenus à chaque produit du groupe, et d'utiliser les formules générales établies par ailleurs (*) pour le calcul des index économiques, dans chaque hypothèse.

Nous examinerons en particulier les deux cas traités à propos des index économiques, à savoir :

1) Pour l'ensemble des produits appartenant au groupe, la proportion consacrée à l'acquisition de chaque article est constante.

2) Les quantités relatives à chaque article varient toutes dans la même proportion.

Première hypothèse.

$\Sigma (q\,p)$ désignant le total des sommes consacrées à l'acquisition des articles appartenant au groupe, on admet par hypothèse que les rapports

$$\left\{\begin{aligned} \frac{q\,p}{\Sigma\,(q\,p)} &= a \\ \frac{q'\,p'}{\Sigma\,(q\,p)} &= b \end{aligned}\right.$$

sont constants.

Nous savons que, dans ce cas, l'index des prix P est lié aux prix individuels p, p',..... par l'équation logarithmique générale :

$$\log\left(\frac{P}{P_0}\right) = a \;\; \log\left(\frac{p}{p_0}\right) + b \;\; \log\left(\frac{p'}{p'_0}\right) + \text{ etc.}$$

(*) *Les index economiques* par RENÉ ROY, chez Sirey, 22 rue Soufflot, Paris V^e^.

Les prix individuels autres que p étant supposés constants, puisque nous nous proposons d'étudier la demande d'un article déterminé, l'équation précédente serait, en diagramme logarithmique, représentée par une droite permettant de faire correspondre à toute valeur de P une valeur bien déterminée de p et réciproquement.

D'autre part, les produits $Q.P$ des index de quantités et de prix restant constamment proportionnels à la somme $\Sigma\ (q\ p)$ en vertu de la définition même des index et le produit $q.\ p$ restant lui-même, par hypothèse, proportionnel à cette même somme $\Sigma\ (q\ p)$ les produits $Q.\ P$ et $q.\ p$ restent constamment proportionnels ; on obtient de cette façon une seconde équation qui, jointe à l'équation logarithmique, permet d'exprimer Q et P en fonction de q et p et, par suite, de déduire l'équation de Demande spéciale à l'article envisagé, de l'équation relative à l'ensemble du groupe.

Il serait facile de donner ici l'expression exacte de cette loi, mais il est absolument inutile d'en aborder la discussion, car l'hypothèse invoquée à l'appui des calculs serait certainement en défaut si le prix p s'écartait exagérément de sa valeur initiale.

Nous observerons que malgré la constance des prix autres que p, les quantités q', q'', . . . varient, puisque les produits $q'\ p'$, $q''\ p''$,..... sont eux-mêmes proportionnels à $\Sigma\ (q\ p)$; il serait même possible d'apprécier l'influence des variations du prix p d'un article du groupe sur la consommation des autres articles du groupe, en utilisant les relations auxquelles il vient d'être fait appel. Nous aurons l'occasion de traiter plus complètement par la suite le problème de l'influence des variations du prix d'un article ou d'un groupe d'articles sur la consommation d'autres articles ou d'autres groupes d'articles.

Deuxième hypothèse.

Si l'on admet que les quantités q, q', q'',..... restent proportionnelles à leurs valeurs de base, l'index de quantités Q reste lui-même proportionnel à chacune des quantités q, q', q'',..... ce qui donne une première relation fort simple entre Q et q.

D'autre part, nous savons que dans ce cas l'index de prix P est une fonction linéaire des prix p, p'', p''', . . . représentée par l'« équation générale linéaire » :

$$\frac{P}{P_0} = a\left(\frac{p}{p_0}\right) + b\left(\frac{p'}{p'_0}\right) + l\left(\frac{p''}{p''_0}\right) + \cdots\cdots$$

La courbe représentative de q en fonction de p pourra donc être déduite de celle qui traduit les variations de Q en fonction de P, par un simple changement d'axes, ce qui n'en modifiera pas l'allure générale.

On peut également dans ce cas obtenir immédiatement l'expression des quantités q', q'', . . en fonction de p, puisque ces quantités varient proportionnellement à l'index Q.

Il paraît bien difficile de choisir entre ces deux hypothèses, en vue de se rapprocher le plus possible des faits, et, seule une étude statistique pourrait nous renseigner à cet égard ; la question ne semble d'ailleurs pas présenter un grand intérêt si l'on se borne à n'envisager que des variations assez restreintes de prix, car dans ce cas, nous savons que la formule générale logarithmique et la formule générale linéaire se ramènent aisément l'une à l'autre, en première approximation ; les deux hypothèses sont pour ainsi dire équivalentes dans ce cas, et l'on peut se borner aux formules linéaires déduites de la seconde hypothèse.

Si l'on veut en venir à des variations de prix beaucoup plus étendues, il est bien difficile de prévoir ce qui se passera car, non seulement il se produira des irrégularités à l'intérieur de chaque groupe d'articles, par suite du jeu de la loi de substitution, mais l'établissement des groupes sera lui-même altéré par de telles variations. Il est de toute évidence en effet que la hiérarchie des besoins, dès que l'on n'envisage plus les objets de stricte nécessité, est lié à la grandeur relative des prix, conformément aux équations résultant des lignes d'indifférence, considérées par VILFREDO PARETO.

Tout ce que nous pouvons dire, c'est que la loi de substitution interviendra dès que la dispersion des prix deviendra grande et des irrégularités apparaîtront dans les courbes de Demande. En particulier, il semble bien qu'en pratique, on doive nécessairement envisager l'existence de prix limites, à partir desquels toute consommation cesse, si l'on considère, non plus un groupe, mais un seul produit.

Il y aurait lieu, pour tracer les courbes réelles de la Demande particulière à un produit bien déterminé, de remplacer la branche asymptotique à l'axe des prix, par une courbe limitée à un point de cet axe.

Il semble donc que la forme probable des courbes de Demande puisse être considérée comme suffisamment caractérisée par les deux schémas ci-après, qui concernent respectivement les objets de première nécessité et les produits autres que les objets de première nécessité (voir Figures IV et V, page 153).

Les formes que nous sommes conduits à admettre ne diffèrent en somme des courbes usuelles que pour la catégorie des objets de première nécessité, mais ce qu'il faut surtout retenir à notre avis, c'est la manière dont ces courbes ont été établies ; elles reposent sur des hypothèses sans doute discutables, mais voisines de la réalité, et sont elles-mêmes intimement liées à la loi de répartition des revenus, qui constitue le point de départ de toute la théorie.

Nous ne nous trouvons donc plus ici en présence d'affirmations à priori, simplement étayées sur de vagues justifications, et c'est surtout à ce titre que nous croyons devoir nous arrêter à ces conclusions.

Quant à la forme précise des équations de Demande, elle ne vaut que par celle de la répartition des revenus et par les hypothèses utilisées, mais ce qu'il nous faut surtout retenir, c'est la forme particulièrement simple de l'équation différentielle de la Demande, sur laquelle nous avons déjà insisté au cours de cette étude.

C'est sur la vérification de cette loi simple que devrait, à notre avis, porter l'effort des recherches statistiques en matière de distribution des revenus.

Avant d'en terminer avec cette étude, nous croyons bon d'indiquer la façon dont pourrait être précisée l'influence d'une baisse de prix des objets compris dans un certain groupe, sur la Demande des produits appartenant aux groupes supérieurs.

Influence de la variation des prix relatifs aux catégories inférieures de besoin. Modification des revenus individuels.

La variation des consommations relatives à un groupe de produits peut être due, non seulement à la variation des prix de ces articles, mais encore à celle des produits entrant dans la composition des groupes nécessaires à la satisfaction de besoins plus urgents. Nous avons d'ailleurs déjà donné une idée de la manière dont l'influence de telles variations de prix pourrait être étudiée à l'intérieur d'un groupe, en faisant appel aux propriétés des index de prix et de quantités.

Nous nous proposons actuellement de procéder à une étude plus générale de la question, non plus à l'intérieur d'un groupe, mais d'un groupe à l'ensemble des groupes inférieurs.

Reprenons à cet effet l'équation générale de la Demande, soit :

$$Q = \frac{A}{P\,(\alpha - 1)} \cdot \left[\frac{1}{\rho^{\alpha - 1}} - \frac{1}{(\rho + q_l P)^{\alpha - 1}} \right].$$

Dans l'expression du second membre, nous supposons ici que P ne varie pas, tandis que le revenu p correspondant à la satiété des besoins plus urgents que ceux du groupe étudié, est au contraire susceptible de variations.

Cette expression montre immédiatement que la consommation Q peut être représentée par la différence des ordonnées de deux courbes établies en fonction du revenu ρ : La première de ces courbes représente une expression de la forme $\frac{1}{\rho^{\alpha - 1}}$ et affecte l'allure d'une courbe régulière à forme hyperbolique, admettant les deux axes comme asymptotes ; la seconde se déduit facilement de la première par voie de translation parallèle à l'axe des revenus et d'amplitude négative $-q_l P$, cette dernière quantité représentant, comme nous le savons, la partie du revenu permettant d'assurer la satiété des besoins correspondant au groupe d'articles étudiés.

On peut donc aisément tracer la courbe représentative de la consommation Q en fonction du revenu ρ, cette courbe se rapprochant beaucoup de celle d'une expression en $\frac{1}{\rho^{\alpha - 1}}$. (voir Figure VI, page 153).

Si l'on veut étudier d'une manière plus précise l'allure du phénomène au voisinage d'un point correspondant à un revenu ρ et à une consommation Q, il nous faut recourir à l'équation générale sus-indiquée, en la différentiant.

On obtient ainsi :

$$Q = \frac{A}{P\,(\alpha - 1)} \cdot \frac{1}{\rho^{\alpha - 1}} \left[1 - \frac{1}{\left(1 + \frac{q_l P}{\rho}\right)^{\alpha - 1}} \right]$$

$$\frac{dQ}{Q} = -(\alpha - 1)\frac{d\rho}{\rho} - (\alpha - 1)\, q_l P \frac{d\rho}{\rho^2} \frac{1}{1 - \frac{1}{\left(1 + \frac{q_l P}{\rho}\right)^{\alpha - 1}}}$$

$$\frac{dQ}{Q} = -(\alpha - 1)\frac{d\rho}{\rho}\left[1 + \frac{q_1 P}{\rho} \cdot \frac{1}{1 - \frac{1}{\left(1 + \frac{q_1 P}{\rho}\right)^{\alpha - 1}}}\right].$$

Si nous représentons par μ le quotient $\frac{q_1 P}{\rho}$ de la fraction du revenu correspondant à la satiété du groupe étudié, par le revenu correspondant à la satiété des catégories inférieures de besoin, nous trouvons :

$$\frac{dQ}{Q} = -(\alpha - 1)\frac{d\rho}{\rho}\left[1 + \mu \cdot \frac{1}{1 - \frac{1}{(1 + \mu)^{\alpha - 1}}}\right].$$

Cette expression met nettement en évidence le lien existant entre une variation relative $\frac{d\rho}{\rho}$ du revenu ρ et celle de la consommation Q.

Si l'on suppose notamment que le rapport μ est assez petit vis-à-vis de l'unité, la partie principale de la fraction $\frac{\mu}{1 - \frac{1}{(1 + \mu)^{\alpha - 1}}}$ se trouve être égale à $+\frac{1}{\alpha - 1}$ de telle sorte qu'on trouve en première approximation :

$$\frac{dQ}{Q} = -(\alpha - 1)\frac{d\rho}{\rho}\left[1 + \frac{1}{\alpha - 1}\right] = -\alpha\frac{d\rho}{\rho}.$$

Cette expression simplifiée donne toute la mesure de l'importance prise par l'exposant α qui caractérise la loi de répartition des revenus ; mais en réalité, le rapport μ ne se trouve pas en général répondre à cette hypothèse, et la relation différentielle générale doit être conservée pour apprécier l'importance de l'influence prise par les variations de ρ sur celles de Q.

Une importante observation reste à faire au sujet des relations qui viennent d'être établies : nous pouvons observer qu'il revient absolument au même de supposer que la quantité Q varie dans une certaine proportion, ou de considérer que le revenu de chaque individu a été modifié dans le sens inverse, d'une égale variation. Cela revient à dire, par exemple, qu'une réduction sur le coût des objets de première nécessité équivaut, du point de vue de la consommation des produits autres que les objets de première nécessité, à une égale augmentation du revenu individuel.

Le problème ci-dessus traité pourrait donc être utilisé directement pour l'étude de l'influence des variations subies par le revenu individuel ; toutefois, l'étude générale de ce second phénomène devrait être conduite en admettant, non que les revenus individuels subissent une égale variation, mais plutôt qu'ils soient tous dilatés ou comprimés dans une égale proportion.

Il y aurait sans doute le plus grand intérêt à effectuer une telle étude, car à notre avis, les variations de débit attribuées aux variations du prix des articles en cause, sont souvent dues, en grande partie, soit à une réduction du prix des objets de consommation plus urgente, soit à un accroissement des revenus individuels.

Nous avons l'impression que de telles influence ont été bien suvent laissées dans l'ombre, dans les études relatives aux variations de la consommation des divers articles.

Conclusions.

Au terme de cette brève analyse, nous exprimons le regret de ne pouvoir être en mesure d'appuyer notre étude sur des observations statistiques précises et variées. De telles observations sont loin d'être impossibles à réaliser, mais elles exigeraient la mise en œuvre de moyens matériels qui nous font défaut, et le regret que nous exprimons ici ne nous est pas spécial, car nous l'avons maintes fois retrouvé dans d'autres études, issues des mêmes préoccupations.

Ces préoccupations tendraient en somme à faire sortir les études économiques du cadre exclusivement déductif où les ont enfermées jusqu'à présent les économistes de l'école mathématique ; seules, de minutieuses observations statistiques permettront de conférer à ce genre d'études le caractère concret qu'elles doivent avoir, si elles ne veulent pas se réduire à un exposé purement théorique et qualitatif des phénomènes ; ici comme ailleurs, nous devons faire appel au principe général posé par Henri Poincaré : « L'expérience seule est rénovatrice ».

Malgré son apparence par trop conventionnelle, nous espérons que notre étude pourra, dans une certaine mesure, contribuer à l'analyse des lois de la Demande, qui restent à la base de toute étude économique précise, et nous aurons atteint notre but si nous avons pu orienter les recherches vers l'étude du lien étroit qui existe entre la loi de répartition des revenus et celle de la Demande.

D'ailleurs, même si les résultats auxquels nous sommes parvenus ne se trouvaient pas vérifiés par l'observation des faits, il suffirait, en appliquant les données de l'observation à l'édification d'hypothèses différentes, de faire appel à la méthode que nous avons suivie pour parvenir à la relation qui doit nécessairement exister entre la Demande et la loi de répartition des revenus.

ANNEXE I.

Observations relatives au monopole des tabacs en France.

Dans une étude récente, M. Bérend, Ingénieur des Manufactures de l'Etat, a eu l'occasion de faire d'intéressantes observations sur le rendement du monopole des tabacs en France, et la répercussion des relèvements de tarifs sur la consommation.

En vue d'effectuer cette étude, M. Bérend, a tout d'abord, porté son attention sur la période 1904-1914, au cours de laquelle les prix de vente n'ont subi aucune modification ; l'indice adopté pour mesurer la consommation est la recette brute, dont les variations traduisent, non seulement celles de la consommation en poids, mais également les déplacements de cette consommation vers les produits de qualité supérieure ; nous savons, en effet, que lorsque les prix de vente sont constants, l'indice de la consommation peut être représenté par la recette brute, conformément aux équations :

$$\frac{dQ}{Q} = \frac{\Sigma\,(p\,dq)}{\Sigma\,(pq)} = \frac{d\,\Sigma\,(pq)}{\Sigma\,(pq)} = \frac{dR}{R} \qquad (dp = 0).$$

Il a été ainsi possible de dégager l'allure saisonnière de la consommation, ainsi que la loi de son accroissement annuel ; cette dernière loi est d'ailleurs conforme au rythme général des phénomènes de cette nature, dont l'expression mathématique est une fonction exponentielle du temps, correspondant à un accroissement relatif proportionnel à la période de temps à laquelle il s'applique.

Pour la période 1921-1926, au cours de laquelle ont été pratiqués de nombreux relèvement de tarifs, il a été tout d'abord nécessaire de tenir compte des stockages antérieurs à chaque relèvement et des sous-consommations postérieures à ces relèvements ; après élimination de ces phénomènes perturbateurs, il a été possible d'obtenir une recette virtuelle correspondant à la consommation effective du tabac et présentant des caractéristiques saisonnières identiques à celles de la période 1904-1914.

Pour effectuer la comparaison de la consommation correspondant à cette courbe régulière d'après-guerre, avec celle de la période 1904-1914, M. Bérend a divisé la recette effective par l'indice des prix de vente, rapportés aux prix moyens d'avant-guerre.

P_0 désignant le prix moyen d'avant-guerre, $pq, p'q', p''q''$ les prix et consommations en poids des diverses catégories de tabac, l'indice des prix défini comme il vient d'être dit, répond à la formule :

$$P = \frac{q}{\Sigma q} \cdot \frac{p}{P_0} + \frac{q'}{\Sigma q} \cdot \frac{p'}{P_0} + \dots$$

$$P = \frac{1}{P_0 \Sigma q} \cdot (q p + q' p' + \dots)$$

$$P = \frac{\Sigma q p}{P_0 \Sigma q}.$$

L'indice de consommation par rapport à la période antérieure à la guerre est donc :

$$Q = \frac{\Sigma q p}{P} = P_0 \Sigma q$$

L'indice choisi pour la période 1921-1926 est donc proportionnel à la consommation totale du tabac, ce qui peut se justifier en un certain sens si l'on considère la faible amplitude des variations subies par la répartition de la consommation suivant les diverses qualités de tabac.

Dans ces conditions, les observations faites à l'occasion des trois derniers relèvements de tarifs résulteraient des données ci-après :

a) *Relèvement d'avril* 1923 :

Valeur de l'indice de prix avant le relèvement	205
Valeur de l'indice de prix après le relèvement.	222
Importance relative du relèvement	7,96 %
Indice de consommation avant le relèvement .	69.933
Indice de consommation après le relèvement .	66.355
Perte de consommation	3.578
Importance relative de la perte de consommation	5,25 %
Valeur de l'élasticité	$\frac{5,25}{7,96} = 0,66.$

b) *Relèvement de mai* 1925 :

Valeur de l'indice de prix avant le relèvement	222
Valeur de l'indice de prix après le relèvement	302
Importance relative du relèvement	30,53 %
Indice de consommation avant le relèvement .	77.400
Indice de consommation après le relèvement .	66.337
Perte de consommation	15.063
Importance relative de la perte de consommation.	21,56 %
Valeur de l'élasticité	$\dfrac{21{,}56}{30{,}53} = 0{,}70.$

c) *Relèvement d'août* 1926 :

Valeur de l'indice de prix avant le relèvement	380
Valeur de l'indice de prix après le relèvement	490
Importance relative du relèvement	25,29 %
Indice de consommation avant le relèvement .	65.214
Indice de consommation après le relèvement .	63.244
Perte de consommation	11.970
Importance relative de la perte de consommation.	20,21 %
Valeur de l'élasticité	$\dfrac{20{,}21}{25{,}29} = 0{,}79.$

Nota : En raison de l'importance des augmentations de tarifs, le taux d'accroissement a été calculé sur la moyenne arithmétique, entre le tarif initial et le tarif majoré ; il en est de même pour le calcul de l'importance relative des réductions de consommation.

On constate donc, comme le laissait prévoir la théorie, que l'élasticité est fonction croissante du prix ; la valeur de cette élasticité ne se trouve d'ailleurs pas considérablement modifiée d'une époque à l'autre, malgré l'amplitude des relèvements de tarifs.

Il convient d'observer au surplus que le coefficient trouvé pour l'élasticité prend des valeurs assez élevées, se rapprochant même de l'unité pour la dernière augmentation ; si l'on considère, par exemple, la valeur médiane 0,70, indiquée ci-dessus, on peut l'interpréter en disant qu'à ce moment, 70 pour cent des consommations étaient constitués par des consommations incomplètes, ne procurant pas la satiété des besoins ; de tels coefficients se rencontrent pour les objets qui ne sont pas de stricte nécessité, et nous aurons lieu d'observer par la suite, que, pour l'usage des moyens de transport, on trouve des coefficients beaucoup plus faibles ; ce fait est l'indice d'une plus grande urgence dans les besoins, en matière de transport, qu'en matière de tabac.

ANNEXE II.

Observations relatives à l'industrie des transports.

En matière de transports, les lois de la Demande ne peuvent guère être envisagées qu'à propos du service des voyageurs, car le transport des marchandises relève, non de la consommation, mais de la production.

Cette observation a d'ailleurs été constamment utilisée pour les ajustements de tarifs auxquels il a été procédé au cours de ces dernières années, dans les pays à monnaie dépréciée ; les majorations appliquées aux tarifs de marchandises ont toujours été beaucoup plus élevées que celles des tarifs de voyageurs, en raison des fortes évasions de trafic auxquelles ont donné lieu les relèvements de tarifs pour voyageurs.

Pour une étude de la Demande, les transports ont l'avantage de ne pas donner lieu à stockage, ce qui en facilite l'analyse ; par contre, on observe de plus grandes irrégularités, en raison de l'influence de certains facteurs comme la température, les intempéries, etc.

Sur les réseaux d'intérêt général, en France, on observe, par exemple, que les deux majorations des 1er janvier et 1er mai 1926, qui s'élèvent ensemble à 21 pour 100 des prix de 1925, firent reculer de 3 pour 100 les recettes de base, c'est-à-dire le mouvement des voyageurs, tandis que dans les conditions normales, on eût dû trouver la progression habituelle de 5 pour 100.

La majoration massive du 16 août 1926, qui porta à 56 pour 100, l'augmentation sur les prix de 1925, fit baisser instantanément le trafic de 10,8 pour 100.

Pour les marchandises, au contraire, on trouve des résultats tout-à-fait différents, et le trafic paraît être surtout lié à la situation économique générale.

Les données ci-dessus nous permettent de calculer le coefficient d'élasticité sur les grands réseaux, pour les voyageurs, à savoir :

Pour l'ensemble des relèvements de janvier et mai 1926 :

Variation des tarifs . . . . $v = 19{,}91\ \%$
Fléchissement du trafic . . . . $f = 7{,}7\ \%$
Coefficient d'élasticité . . . . $\dfrac{f}{v} = 0{,}387$

Pour le relèvement du mois d'août 1926, on obtient :

Variation des tarifs . . . . $v = 25{,}27\ \%$
Fléchissement du trafic . . . . $f = 10{,}25\ \%$
Coefficient d'élasticité . . . . $\dfrac{f}{v} = 0{,}406$

On constate donc une permanence remarquable, qui confirme nos observations antérieures, sur les transports en commun des grandes villes.

Dans l'étude que nous avons consacrée aux voies ferrées d'intérêt local, nous avons en effet noté que le coefficient d'élasticité, relatif aux transports urbains, atteignait, aux environs de 1924, une valeur voisine de 0,25 pour la région parisienne, et qu'il dépassait 0,30 et même 0,40 pour des agglomérations moins importantes. On observe même des coefficients dépassant 0,50 dans les villes de moindre activité.

D'ailleurs, il est naturel que ce coefficient augmente, c'est-à-dire que la Demande soit plus élastique, quand il s'agit de services moins indispensables, comme c'est le cas des transports dans les petites villes.

Pour compléter ces indications, nous avons noté les répercussions des augmentations de tarifs effectuées en 1925 et 1926 sur les tramways de Marseille ; ce cas est fort intéressant, car les relèvements s'y sont succédés très rapidement ; il y en eut quatre entre septembre 1925 et octobre 1926 ; d'autre part, il n'existe pas de réseau souterrain à Marseille, ce qui élimine une possibilité de substitution ; enfin, contrairement à ce qui se passe dans la plupart des autres villes, les tarifs sont à peu près indépendants de la distance parcourue, de telle sorte que le nombre des voyageurs est un indice de trafic suffisant pour les besoins de la pratique.

Avant la guerre, le trafic affectait une allure saisonnière, grossièrement caractérisée par les chiffres de 22, 26, 27, 25, représentant pour chaque trimestre les pourcentages du trafic par rapport au trafic annuel ; le nombre des voyageurs était en progression régulière et s'accroissait environ de 4 pour 100 par an.

Depuis la guerre, il est difficile de dégager la loi du trafic, en raison des nombreux remaniements de tarifs, auxquels il a été procédé ; mais pour les années 1924 et 1927, au cours desquelles on n'a pas modifié les tarifs, le mouvement des voyageurs affecte une allure comparable à celle d'avant-guerre, en ce qui concerne l'importance du trafic de chaque trimestre ; pendant le premier semestre de 1925 et le début du second, c'est-à-dire immédiatement avant la série des relèvement de tarifs, l'accroissement du nombre des voyageurs par rapport à 1924, pouvait être évalué approximativement à 4 pour 100 comme avant la guerre.

Pour éviter de tenir compte des variations saisonnières, au sujet desquelles nous nous sommes limités à l'indication d'ordres de grandeur, nous étudierons la répercussion des augmentations de tarifs sur l'importance du trafic, en comparant le trafic des périodes comprises entre deux relèvements successifs, au trafic de la période correspondante pour l'année précédente.

Cette méthode offre évidemment l'inconvénient de ne pas faire apparaître directement le coefficient d'élasticité, qui devrait être calculé par comparaison du trafic postérieur à chaque relèvement avec le trafic antérieur au même relèvement ; nous serons en outre amenés, pour un but de simplification, à faire la somme de pourcentages d'augmentation, pour calculer un pourcentage global, mais les écarts imputables à la mise en œuvre de tels procédés ne nous semblent pas devoir être supérieurs à ceux qui peuvent résulter de l'utilisation de données portant sur des périodes très réduites, ainsi que de la considération de variations finies, pour calculer un coeffcient d'élasticité conçu comme rapport de variations infiniment petites.

Les diverses valeurs du coefficient d'élasticité correspondant à chacun des relèvements de tarifs, s'établissent comme il suit :

a) *Relèvement du* 13 *septembre* 1925.

Pourcentage d'augmentation	34
Rapport du trafic du troisième trimestre de 1925 à celui du troisième trimestre de 1924	0,93
Baisse	7 %
Manque à gagner	4 %
Total	11 %

Coefficient d'élasticité . . . . $\dfrac{11}{34} = 0{,}32.$

b) Relèvement du 21 *janvier* 1926.

Pourcentage d'augmentation	16
Rapport du trafic de février-mai 1926 à celui des mêmes mois de 1925	0,87
Baisse	13 %
Manque à gagner	4 %
Total	17 %
Relèvement survenu dans l'intervalle . . .	$34 + 16 = 50$
Coefficient d'élasticité	$\frac{17}{50} = 0{,}34.$

c) *Relèvement du* 4 *juin* 1926.

Pourcentage d'augmentation	11
Rapport du trafic de juin-août 1926 à celui des mêmes mois de 1925	0,85
Baisse	15 %
Manque à gagner	4 %
Total	19 %
Relèvements survenus dans l'intervalle . .	$34 + 16 + 11 = 61$
Coefficient d'élasticité	$\frac{19}{61} = 0{,}31.$

d) *Relèvement du* 18 *octobre* 1926.

Pourcentage d'augmentation	15
Rapport du trafic de novembre-décembre 1926 à celui des mêmes mois de 1925 .	0,91
Baisse	9 %
Manque à gagner	4 %
Total	13 %
Relèvements survenus dans l'intervalle . .	$16 + 11 + 15 = 42$
Coefficient d'élasticité	$\frac{13}{42} = 0{,}31.$

La stabilité du coefficient d'élasticité résultant de ces indications autorise, croyons-nous, à conclure à l'existence effective d'une loi de la Demande, pour le genre de consommation étudié, et ce fait constitue une présomption de plus en faveur de notre théorie.

On peut observer que la méthode utilisée est très rudimentaire, que d'autre part, la période étudiée fut marquée par des troubles monétaires d'une amplitude considérable, ce qui tendait évidemment à introduire des perturbations dans la marche normale des phénomènes.

Quoi qu'il en soit, la période à laquelle nous venons de faire allusion constituerait un champ d'observation fort intéressant pour la question qui nous occupe, mais en dehors des entreprises contrôlées par les pouvoirs publics, il est extrêmement difficile de se procurer des renseignements de cette nature.

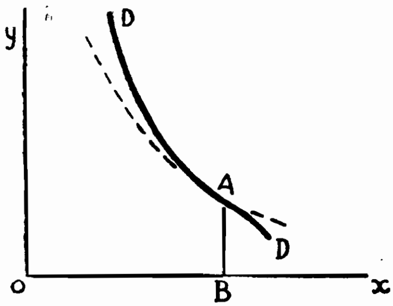


FIG. I.

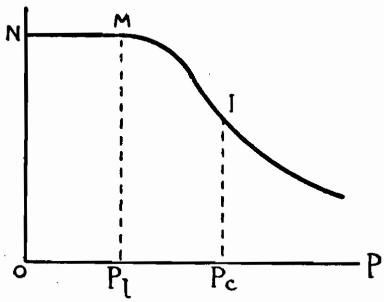


FIG. II.

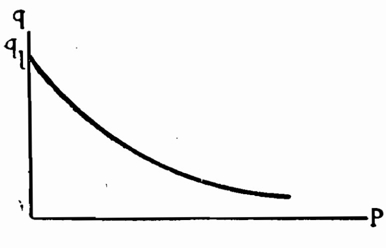


FIG. III.

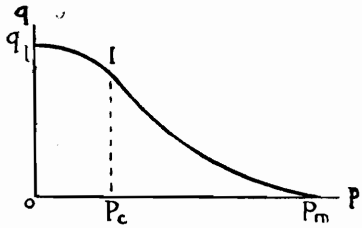


FIG. IV.

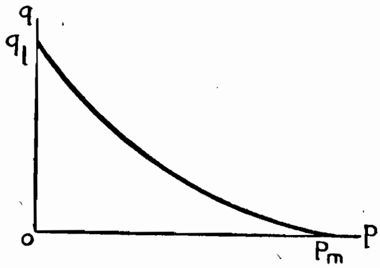


FIG. V.

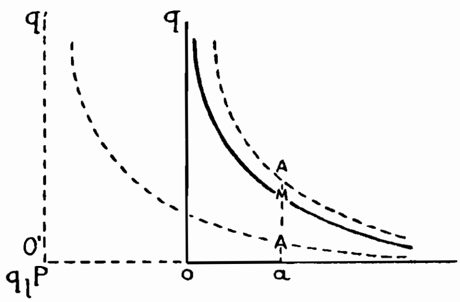


FIG. VI.

Prof. CORRADO GINI, *Direttore responsabile*.

Roma — « L'Universale » Tipografia Poliglotta.

The international Review of Statistics METRON is published in numbers. Four successive numbers make a volume of 700 to 800 pages in all.

It accepts original articles on statistical methods and on the applications of statistics to the different spheres of activity, and review or discussions of results obtained by statistical method in various fields of science, or such material as may be to interest to the statistician. A bibliography is annexed of all works or reviews presented or received in exchange.

Articles and reviews may be written in English, Italian, French or German. Manuscripts in English, French or German should be typewritten. Contributors will receive free of charge 25 copies of their publications issued.

Manuscripts submitted for publication should be addressed to *Prof. Corrado Gini, Istituto di Statistica, R. Università di Roma (Italy)*, or to the member of the Editorial Committee who represents the writers' country. Contributors are requested to retain one copy of each manuscript sent, as, in case of non acceptance, the Editors will not be responsable for the safe return of the original.

Proposals for exchange made by reviews or other periodicals and all publications sent in exchange, or as complimentary copies, should be addressed to Prof. Corrado Gini.

All applications of subscribers, as well as the sums for the subscriptions, are to be made payable to *Amministrazione del « Metron », Istituto di Statistica della R. Università di Roma, Italy.*

The subscription rate for each volume is **100 It. lire** and for single copies **30 It. lire,** each post paid.

Die Internationale Statistische Zeitschrift METRON erscheint in Heften. Vier aufeinanderfolgende Hefte bilden ein Band im Gesamtumfang von 700-800 Seiten.

Die Zeitschrift veröffentlicht Originalaufsätze über die Methode der Statistik und die Anwendung der Statistik auf die verschiedenen Zweige der Wissenschaften, sowie Uebersichten und Erörterungen über die Ergebnisse der statistischen Methode auf den verschiedenen Wissenschaftsgebieten, soweit die für den Statistiker von Interesse sind. Sie enthält ferner ein Verzeichnis aller unentgeltlich oder im Austauschverkehr eingehenden Bücher und Zeitschriften.

Die zur Veröffentlichung eingesandten Aufsätze und Mitteilungen können in deutscher, italienischer, französischer, und englischer Sprache verfasst sein. Deutsche, französische und englische Manuskripte müssen mit der Maschine geschrieben sein. Beiträge werden nicht honoriert. Jeder Verfasser erhält unentgeltlich 25 Sonderabdrücke seiner Arbeit.

Die Maniskripte, deren Veröffentlichung gewünscht wird, sind an Herrn *Prof. Corrado Gini, Istituto di Statistica, R. Università di Roma (Italien)* oder an das Mitglied des Direktion-Komitees, das den Staat des Mitarbeiters vertritt, zu richten.

Die Verfasser werden gebeten, eine Abschrift des eingesandten Manuskriptes zurückzubehalten, da die Schrifteilung für den Fall, dass die eingesandte Arbeit nicht veröffentlicht wird, keine Gewähr für deren Rücksendung übernimmt.

Austauschanträge für andere Zeitschriften und alle Veröffentlichungen, die unentgeltlich oder im Austausch zur Verfügung gestellt werden, sind an Herrn Prof. Corrado Gini zu richten.

Die neuen Abonnements-Anfragen, sowie die Zahlungen für die Abonnements, sind an *Amministrazione del « Metron », Istituto di Statistica della R. Università di Roma (Italien)* zu richten.

Der postfreie Bezugspreis für jeden Band ist **100 It. lire,** und **30 It. lire** für das einzelne Heft.

BIBLIOTECA DEL "METRON" - "METRON" LIBRARY
BIBLIOTHÈQUE DU "METRON" - "METRON'S" BIBLIOTHEK

SERIE ***A.*** — Problemi di attualità — Problèmes d'actualité — Gegenwärtige Fragen.

SERIES ***A.*** — Problems of the moment.

1. — A. ANDRÉADÈS — *La population anglaise avant, pendant et après la grande guerre.*
   10 lires pour l'Italie et les pays ayant un change plus défavorable **5 Frs. suisses** pour les autres pays.

2. — L. HERSCH. — *La population de la Palestine et les perspectives du Sionisme.*
   Lit. 3.

SERIE ***B.*** — Memorie scientifiche — Mémoires scientifiques — Wissenschaftliche Arbeiten.

SERIES ***B.*** — Scientific Memoirs.

1. — F. SCHINDLER — *Das Volksvermögen Voralbergs.*
   25 lires pour l'Italie **8 sh. autrich.** pour l'Autriche.
   8 Fr. suisses pour la Suisse et les autres pays.

2. — F. SAVORGNAN — *La scelta matrimoniale — Studi statistici.*
   12 lires pour l'Italie et les pays ayant un change plus défavorable.
   6 Frs. suisses pour les autres pays.

3. — F. V. FELLNER — *Die Verteilung des Volksvermögens und Volkseinkommens der Länder der Ungarischen Heiligen Krone zwischen dem heutigen Ungarn und den Successions-Staaten.*
   10 lires pour l'Italie et les pays ayant un change plus défavorable.
   5 Frs. suisses pour les autres pays.

4. — MARIO BALESTRIERI — *I consumi alimentari della popolazione italiana dal* 1910 *al* 1921 con prefazione del Prof. CORRADO GINI.
   15 lires.

Gli abbonati del *Metron* che domandano *direttamente* all'Amministrazione le opere pubblicate nella *Biblioteca del* « *Metron* » ricevono uno sconto, sul prezzo di copertina, del 30 %. Le spese di porto restano a carico dell'acquirente.

Les abonnés du *Metron*, qui commandent directement à l'Administration les ouvrages publiés par la *Bibliothèque du* « *Metron* » reçoivent un rabais de 30 % sur les prix indiqués. Les frais de port restent à la charge de l'acheteur.

Those subscribers to the *Metron* who obtain directly from the Administration works published in the « *Metron* » *Library* receive a discount, on the marked price, of 30 %. The cost of carriage must be borne by the buyer.

Den Abonnenten der Zeitschrift *Metron* welche die von der « *Metron* »'s *Bibliothek* veröffentlichen Werke daselbst beziehen, kommt ein Bonus von 30 % des angeschlagenen Preises zugute.