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 intèressant 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 écrits 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 andressera 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 — Via delle Terme di Diocleziano*, 10, *Roma, 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

Vol. XIII - N. 2 **28-II-1938.**

SOMMARIO — SOMMAIRE — CONTENTS — INHALT

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 reception*).

ARTICLES RECEIVED BY THE REVIEW WHICH WILL BE PUBLISHED IN FUTURE ISSUES.

(*According to date of receipt*)

G. A. Baker — *The probability that the standard deviation of a second sample will differ from the standard deviation of a first sample by a certain multiple of the standard deviation of the first sample.*

R. Cisbani, G. Zappa — *Contributi alla teoria delle medie.*

CORRADO GINI

Di una formula comprensiva delle medie

Distinguiamo anzitutto le infinite possibili medie di una serie in

A) medie *ferme*,

B) medie *lasche*,

secondo che il loro valore dipenda o meno da tutti i termini della serie.

Le medie lasche più spesso considerate nella statistica sono la mediana o valore centrale della serie, la moda o valore dominante, il valore divisorio, il valore poziore, il valore centrale del campo di variazione (1).

Le medie ferme possono suddividersi, a loro volta, in

a) medie *basali;*

b) medie *esponenziali;*

c) medie *baso-esponenziali*

a seconda che, nelle formule analitiche che le esprimono, i termini della serie figurano come basi o come esponenti o, ad un tempo, come basi e come esponenti.

Delle medie baso-esponenziali (di cui, se non erro, si parla qui per la prima volta) avremo occasione di dare esempi nel corso di questa stessa nota (cfr. pagg. 5-6).

---

(1) La differenza tra *medie ferme* e *medie lasche* non corrisponde esattamente alla differenza tra *media analitiche* e *medie di posizione*, in cui con medie analitiche si intendono le medie suscettibili di venire espresse mediante una formula analitica e con medie di posizione le medie che non sono suscettibili di tale espressione. Il valore centrale del campo di variazione, che è uguale alla semisomma dei valori estremi a_1 e a_n e solo da questi dipende, è una *media lasca ;* ma è una *media analitica*, la cui semplice formula è $\frac{a_1 + a_n}{2}$.

Le medie esponenziali sono utili per alcuni problemi della matematica finanziaria, ma le medie ferme di gran lunga più importanti per la statistica sono le medie basali. Rientrano in esse la media aritmetica, la geometrica, l'armonica, l'anti-armonica, la quadratica, la cubica, ecc. e per esse è quindi particolarmente desiderabile di possedere una formula generale quanto più possibile comprensiva.

In questa nota mi propongo appunto di arrivare, mediante successive generalizzazioni, a una formula comprensiva delle medie ferme basali.

Profitterò dell'occasione per proporre una classificazione razionale di tali medie.

* * *

Sia n il numero dei termini della serie; a_i $(i = 1, 2, \ldots n)$ un termine generico della serie; p e q esponenti reali qualunque a cui si elevano i termini della serie; $\binom{n}{c}$ [o rispettivamente $\binom{n}{d}$] il numero delle combinazioni degli n termini a c a c (o rispettivamente a d a d).

Fatte tutte le possibili combinazioni degli n elementi a_1, a_2, $\ldots$, a_n a c a c, e numeratele, chiamiamo $P^c{}_1\,(a_i)$ il prodotto dei termini della 1ª combinazione; $P^c{}_2\,(a_i)$ il prodotto dei termini della seconda combinazione; ... e in generale $P^c{}_l\,(a_i)$ il prodotto dei termini della l-esima combinazione.

L'espressione $\sum\limits_{l=1}^{\binom{n}{c}} P^c{}_l\,(a_i)$ ci indicherà allora la somma degli $\binom{n}{c}$ prodotti che si ottengono facendo tutte le possibili combinazioni degli n elementi a c a c e moltiplicando tra loro i termini di ciascuna combinazione. Analogamente nel caso che al posto di c vi sia d.

* * *

Prendiamo le mosse dalla nota formula generale

$$M^p = \sqrt[p]{\frac{\sum\limits_{i=1}^{n} a^p{}_i}{n}} \qquad \text{(I)}$$

che viene detta *media di potenze*, o, come noi preferiamo, *media potenziata*.

Per $p = -1, 1, 2, 3 \ldots, M^p$ si riduce rispettivamente alla media armonica, aritmetica, quadratica, cubica, mentre il suo valore tende alla media geometrica quando p tende a 0 (1).

Un'altra formula generale conosciuta è la seguente

$$M^c = \sqrt[c]{\frac{1}{\binom{n}{c}} \sum_{l=1}^{\binom{n}{c}} P^c{}_l\,(a_i)} \qquad \text{(II)}$$

che può chiamarsi *media delle combinazioni dei termini* o più brevemente *media combinatoria*. Per $c = n$, essa si riduce alla media geometrica; per $c = 1$, alla media aritmetica.

Se nella media combinatoria si innalza ciascun termine alla potenza p, si ottiene la *media combinatoria potenziata*:

$$M^{cp} = \sqrt[cp]{\frac{1}{\binom{n}{c}} \sum_{l=1}^{\binom{n}{c}} P^c{}_l\,(a^p{}_i)} \qquad \text{(III)}$$

che si riduce alla (I) per $c = 1$, e alla (II) per $p = 1$.

La (III), per p tendente a 0, tende alla media geometrica (1). Ciò ci porta a definire

$$M^{c0} = \sqrt[n]{\prod_{i=1}^{n} a_i} \qquad \text{(III')}$$

(1) Il Dunkel (*Generalized geometric means and algebraic equations.* « Annals of Mathematics », Vol. 11, 1909-1910) ha dimostrato che

$$\lim_{x \to 0} \sqrt[x]{\frac{\sum_{i=1}^{n} a_i^x}{n}} = \sqrt[n]{\prod_{i=1}^{n} a_i}\,.$$

Usando il medesimo procedimento, cioè passando ai logaritmi e applicando la regola dell'Hospital, si ottiene

$$\lim_{p \to 0} \sqrt[cp]{\frac{1}{\binom{n}{c}} \sum_{l=1}^{\binom{n}{c}} P^c{}_l\,(a_i{}^p)} = \sqrt[n]{\prod_{i=1}^{n} a_i}$$

Nella (III) (e naturalmente anche nella (I) e nella (II), che della (III) non sono che casi particolari) i termini figurano solo al numeratore; sono, potremo dire, tutti allo stesso piano. Queste si diranno quindi *medie monoplane*, in contrapposto a quelle, più comprensive, di cui passiamo a parlare, in cui i termini figurano ad un tempo al numeratore e al denominatore, e che si diranno pertanto *medie biplane*.

Nella (I), (II), (III) il simbolo M sta appunto a significare che si tratta di medie monoplane.

Generalizzando la (I) otteniamo la

$$\overset{p}{B^{q}} = \sqrt[p-q]{\frac{\sum\limits_{i=1}^{n} a^{p}_{i}}{\sum\limits_{i=1}^{n} a^{q}_{i}}} \qquad \text{(IV)}$$

che si dirà quindi *media biplana potenziata*.

---

come anche

$$\lim_{p \to q} \sqrt[p-q]{\frac{\sum\limits_{i=1}^{n} a_i^{p}}{\sum\limits_{i=1}^{n} a_i^{q}}} = \sqrt[\sum\limits_{i=1}^{n} a_i^{q}]{\prod_{i=1}^{n} a_i^{a_i^{q}}}$$

e ancora

$$\lim_{p \to 0} \sqrt[(c-d)\,p]{\frac{\binom{n}{d} \sum\limits_{l=1}^{\binom{n}{c}} P^{c}{}_{l}\,(a_i^{p})}{\binom{n}{c} \sum\limits_{l=1}^{\binom{n}{d}} P^{d}{}_{l}\,(a_i^{p})}} = \sqrt[n]{\prod_{i=1}^{n} a_i}$$

e infine

$$\lim_{p \to q} \sqrt[c\,(p-q)]{\frac{\sum\limits_{l=1}^{\binom{n}{c}} P^{c}{}_{l}\,(a_i^{p})}{\sum\limits_{l=1}^{\binom{n}{c}} P^{c}{}_{l}\,(a_i^{q})}} = \sqrt[c\left[\sum\limits_{l=1}^{\binom{n}{c}} P^{c}{}_{l}\,(a_i^{q})\right]]{\prod_{l=1}^{\binom{n}{c}} \left[P^{c}{}_{l}\,(a_i)\right]^{[P^{c}{}_{l}\,(a_i^{q})]}} =$$

$$= \sqrt[c\left[\sum\limits_{l=1}^{\binom{n}{c}} P^{c}{}_{l}\,(a_i^{q})\right]]{\prod_{i=1}^{n} a_i^{a_i^{q} \sum\limits_{l=1}^{\binom{n-1}{c-1}} P^{c-1}{}_{l}\,(a_1^{q}, \ldots, a^{q}_{i-1}, a^{q}_{i+1} \ldots a_n^{q})}}$$

Nel caso particolare $q = 0$, la (IV) si riduce alla media monoplana potenziata (formula I).

Nel caso particolare $q = p - 1$, la (IV) si riduce alla nota formula generale

$$B^{p}_{p-1} = \frac{\sum_{i=1}^{n} a^{p}_{i}}{\sum_{i=1}^{n} a_i^{p-1}} \qquad (IV')$$

detta *media di somme di potenze*, che, per $p = 0, 1, 2$, dà rispettivamente la media armonica, aritmetica, antiarmonica.

La (IV), per p tendente a q e a_i sempre positiva, tende al limite (vedi nota a pag. 5 e 6)

$$\sqrt[\sum_{i=1}^{n} a_i^q]{\prod_{i=1}^{k} a_i^{a_i^q}}\,.$$

Ciò ci porta a definire

$$B^{p}_{p} = \sqrt[\sum_{i=1}^{n} a_i^p]{\prod_{i=1}^{n} a_i^{a_i^p}} \qquad (V)$$

che ci darà l'espressione della *media biplana equipotenziata*.

La (V) per $p = 0$ si riduce alla media geometrica.

Generalizzando la (II), si ottiene la

$$B^{c}_{d} = \sqrt[c-d]{\frac{\binom{n}{d} \sum_{l=1}^{\binom{n}{c}} P^{c}_{l}(a_i)}{\binom{n}{c} \sum_{l=1}^{\binom{n}{d}} P^{d}_{l}(a_i)}} \qquad (VI)$$

che rappresenterà la *media biplana combinatoria*.

Per $c = n$, $d = n - 1$, essa si riduce alla media armonica.

Per $d = 0$, essa si riduce alla media monoplana combinatoria (formula II) in quanto si possono porre (1) le convenziòni

$$\binom{n}{0} = 1 \qquad \frac{1}{\binom{n}{0}} \sum_{l=1}^{\binom{n}{0}} P^o{}_l\,(a_i) = 1\,.$$

Se nella (VI) si eleva ciascun termine alla potenza p, si ottiene la *media biplana combinatoria equipotenziata*

$$B^{c\,p}_{d\,p} = \sqrt[(c-d)\,p]{\frac{\binom{n}{d} \sum_{l=1}^{\binom{n}{c}} P^c{}_l\,(a_i{}^p)}{\binom{n}{c} \sum_{l=1}^{\binom{n}{d}} P^d{}_l\,(a_i{}^p)}} \qquad \text{(VII)}$$

Per $p = 1$, la (VII) si riduce alla (VI).

Per $d = 0$, si ricade nella media monoplana combinatoria potenziata [formula (III)] in base alla convenzione (2) di cui in nota (1).

---

(1) La convenzione $\binom{n}{0} = 1$ è già introdotta, come è ben noto. Per giustificare la convenzione

$$\frac{1}{\binom{n}{0}} \sum_{l=1}^{\binom{n}{0}} P^o{}_l\,(a_i) = 1 \qquad (1)$$

si può osservare che, se $P^d{}_1\,,\ P^d{}_2\,,\ \ldots\ P^d{}_{\binom{n}{d}}$ (per $0 < d \leq n$) hanno il significato dato a pag. 2, si ha che gli $\binom{n}{d}$ numeri $P^d{}_1\,,\ P^d{}_2\,,\ \ldots\ P^d{}_{\binom{n}{d}}$ si ottengono, a meno dell'ordine, dividendo il prodotto di tutti gli a_i cioè $P^n{}_1$, per gli $\binom{n}{n-d} = \binom{n}{d}$ numeri $P^n{}_1{}^{-d}\,,\ P^n{}_2{}^{-d}\,, \ldots P^{n-d}_{\binom{n}{n-d}}$. Se vogliamo far sì che tale proprietà valga anche per $d = 0$, dobbiamo definire come unico $P^o{}_l$ il rapporto tra $P^n{}_1$ e $P^n{}_1$ stesso, cioè $P^o{}_1 = 1$. D'onde la convenzione (1).

La convenzione (1) si generalizza mediante la (2)

$$\frac{1}{\binom{n}{0}} \sum_{l=1}^{\binom{n}{0}} P^o{}_l\,(a^p{}_i) = 1 \qquad (2)$$

considerando come termine generico $a^p{}_i$, anzichè a_i.

Per p tendente a **0**, la (VII) tende alla media geometrica (vedi nota a pagg. 5-6). Ciò ci porta a definire

$$B^{c\,o}_{d\,o} = \sqrt[n]{\prod_{i=1}^{n} a_i} \; . \tag{VII'}$$

D'altra parte, se nella (IV) i termini si combinano a c a c sia al numeratore che al denominatore, si ottiene la *media biplana equicombinatoria potenziata*, data dalla formula

$$B^{c\,p}_{c\,q} = \sqrt[c(p-q)]{\frac{\sum_{l=1}^{\binom{n}{c}} P^c_l\,(a_i^p)}{\sum_{l=1}^{\binom{n}{c}} P^c_l\,(a_i^q)}} \; . \tag{VIII}$$

Per $c = 1$, essa si riduce alla (IV).

La (VIII), per p tendente a q e a_i sempre positivo, tende (vedi nota a pagg. 5-6) a

$$\sqrt[c\left[\sum_{l=1}^{\binom{n}{c}} P^c_l\,(a_i^q)\right]]{\prod_{l=1}^{\binom{n}{c}} [P^c_l\,(a_i)]^{[P^c_l\,(a_i^q)]}} =$$

$$= \sqrt[c\left[\sum_{l=1}^{\binom{n}{c}} P^c_l\,(a_i^q)\right]]{\prod_{i=1}^{n} a_i^{\,a_i^q \sum_{l=1}^{\binom{n-1}{c-1}} P^{c-1}_l\,(a_1^q,\, a_2^q \, \ldots \, a^q_{i-1},\, a^q_{i+1} \, \ldots \, a_n^q)}}$$

Ciò ci porta a definire

$$B^{c\,p}_{c\,p} = \sqrt[c\left[\sum_{l=1}^{\binom{n}{c}} P_l^c\,(a_i^p)\right]]{\prod_{i=1}^{n} a_i^{\,a_i^p \sum_{l=1}^{\binom{n-1}{c-1}} P^{c-1}_l\,(a_1^p,\, a_2^p \, \ldots \, a^p_{i-1},\, a^p_{i+1} \, \ldots \, a_n^p)}} \tag{IX}$$

che ci darà l'espressione della *media biplana equicombinatoria equipotenziata*.

Infine, la media geometrica ponderata delle $\overset{c\,p}{B^{d\,p}}$ e $\overset{d\,p}{B^{d\,q}}$, in cui si dia a $\overset{c\,p}{B^{d\,p}}$ un peso uguale a $(c-d)\,p$, e a $\overset{d\,p}{B^{d\,q}}$ un peso eguale a $d\,(p-q)$ dà la *media biplana combinatoria potenziata*, la cui formula è

$$\overset{c\,p}{B^{d\,q}} = \sqrt[pc-dq]{\frac{\binom{n}{d}\sum\limits_{l=1}^{\binom{n}{c}} P^{c}_{l}\,(a_i^{\,p})}{\binom{n}{c}\sum\limits_{l=1}^{\binom{n}{d}} P^{d}_{l}\,(a_i^{\,q})}} \qquad \text{(X)}$$

Alla stessa formula si perviene facendo la media geometrica ponderata di $\overset{c\,p}{B^{c\,q}}$ e di $\overset{c\,q}{B^{d\,q}}$, a cui si diano rispettivamente i pesi $c\,(p-q)$ e $(c-d)\,q$.

La (X) rappresenta la formula comprensiva per le medie ferme basali a cui siano pervenuti (1).

Per $c = d$, la (X) si riduce alla (VIII); per $p = q$, si riduce alla (VII); per $q = 0$ o $d = 0$ si riduce alla (III).

Per farci un'idea della generalità della (X), si noti che la media di potenze, la media di somme di potenze e la media combinatoria monoplana costituiscono le formule generali più comprensive finora proposte per le medie ferme basali (2). Ora le due

---

(1) Sarebbe facile generalizzare ulteriormente la (X), per esempio considerando, in luogo dei termini, certe loro funzioni che in casi particolari si riducono ai termini stessi. Ritengo però inutile procedere a tali ulteriori generalizzazioni, poichè la (X) già comprende, se non erro, tutte le medie ferme basali che sono note e che si possono usare utilmente nei casi pratici.

(2) Prescindiamo dalle formule generali delle medie, proposte da alcuni matematici, che lasciano indeterminata la funzione ed hanno pertanto un valore puramente teorico. Tale è, per esempio la formula

$$f\,(\xi) = \frac{\Sigma\, p_i\, f\,(a_i)}{\Sigma\, p_i}$$

che trovo in Bonferroni (*Elementi di statistica generale*, Anno accademico 1932-33, Bari. R. Istituto Superiore di Scienze Economiche. Torino, Litografia Gili, 1933, pag. 30), e Darmois (*Statistique mathématique*, Paris, Doin, 1928, pag. 31), o quella che ricorre nella seguente definizione del Chisini (*Sul concetto di media*. « Periodico di Matematica » 1929, n. 2, pag. 108):

« Data una funzione

$$y = f\,(x_1\,,\,x_2\,,\,.\,.\,.\,,\,x_n)$$

$x_1, x_2, \ldots, x_n$ rappresentando grandezze omogenee, dicesi « media » delle $x_1, x_2, \ldots, x_n$ rispetto alla funzione f quel numero M che soddisfa alla relazione

$$f\,(M\,,\,M\,\ldots\,M) = f\,(x_1\,,\,x_2\,,\,.\,.\,.\,,\,x_n) \text{ ».}$$

prime non sono che due casi particolari della (IV) e la terza non è che un caso particolare della (VI), mentre la (IV) e la (VI) rappresentano, a loro volta, due casi particolari della (VIII), la quale infine non è che un caso particolare della (X).

* * *

Mentre tutte le formule precedenti, quando i termini sono positivi (1), sono vere medie nel senso che non sono mai inferiori al termine minimo nè superiori al termine massimo, ciò non è sempre vero per la (X). Condizione sufficiente per che lo sia è che i due pesi $(c-d)\,p$ e $d\,(p-q)$, oppure $(c-d)\,q$ e $c\,(p-q)$ non abbiano segno contrario (2).

Questa condizione naturalmente non è sempre soddisfatta; per esempio, non lo è, per $n > 2$, quando sia $c = n$, $d = 1$, $p = 2/n$, $q = 1$.

In tal caso la (X) diviene

$$H^{II} = \frac{n \prod_{i=1}^{n} a_i^{\frac{2}{n}}}{\sum_{i=1}^{n} a_i} = \frac{G^2}{A}$$

dove G indica la media geometrica ed A la media aritmetica.

È questa una nota espressione che costituisce una generalizzazione (come vedremo, non corretta) della media armonica fra due termini a_1 e a_2 data dalla formula

$$H = \frac{2\,a_1\,a_2}{a_1 + a_2}$$

Quest'ultima formula può, invero, generalizzarsi per un numero qualunque di termini, in più modi.

Uno di questi dà luogo alla espressione

$$H^{I} = \frac{n}{\sum_{i=1}^{n} \frac{1}{a_i}}$$

(1) Quando qualcuno degli a_i è negativo, le formule da (I) a (X) talvolta perdono di significato; talvolta, pur avendo significato, non sono vere medie. Inoltre, come fu notato, la dimostrazione della (V) e della (IX) vale solo nel caso di a_i sempre positivo.

(2) La (X) non ha più significato quando $p\,c - q\,d = 0$. Essa, al tendere di p a $\frac{q\,d}{c}$, tende a $+\infty$, o a 0, a seconda che l'espressione sotto il segno di radice tende a un valore > 1 o < 1.

a cui si è convenuto di conservare la denominazione di media armonica.

Un'altra generalizzazione, che conduce appunto alla precedente espressione H^{II}, si è ottenuta in base alla proprietà della media armonica di due termini di essere terza proporzionale tra la media aritmetica e la media geometrica.

Il Messedaglia riteneva che questa seconda generalizzazione « corrisponderebbe pur sempre ad un vero e proprio valor medio » (1). Ma in questo caso il nostro autore, di consueto precisissimo, era effettivamente inesatto. È facile, invero, fornire esempi in cui l'espressione considerata porta a valori inferiori al termine minimo. Così, nella serie di tre termini: 1, 1, 64, in cui risulta $H^{II} = \frac{16}{22} < 1$, o nell'altra: 1, 2, 108, per cui è $H^{II} = \frac{36}{37} < 1$. Effettivamente la condizione che $c\,(p - q)$ e $(c - d)\,q$ non siano di segno contrario non si verifica per H^{II}, in quanto è $(p - q)\,c = 2 - n$ e $(c - d)\,q = n - 1$; e, mentre è sempre $n - 1 > 0$, è, per $n > 2$, $2 - n < 0$. E similmente sono di segno contrario $d\,(p - q)$ e $(c - d)\,q$, in quanto è $d\,(p - q) = \frac{2}{n} - 1$ e $(c - d)\,p = (n - 1)\,\frac{2}{n}$, e, mentre è sempre $(n - 1)\,\frac{2}{n} > 0$, è, per $n > 2$, $\frac{2}{n} - 1 < 0$.

In realtà la generalizzazione corretta, per n termini, della H non è rappresentata dalla

$$H^{II} = \frac{G^2}{A},$$

ma dalla

$$H^{III} = \frac{G^{\frac{n}{n-1}}}{A^{\frac{1}{n-1}}}$$

che è un caso particolare della (VI) per $c = n$, $d = 1$.

Altra generalizzazione corretta, per n termini, della H è data dalla

$$H^{IV} = \frac{a_1\,a_2 + a_1\,a_3 + \ldots + a_1\,a_n + a_2\,a_3 + \ldots + a_{n-1}\,a_n}{n\,(n - 1)\,A}$$

che è un altro caso particolare della (VI) per $c = 2$, $d = 1$.

Tanto la H^{III} che la H^{IV} si riducono alla H per $n = 2$.

(1) *Il calcolo dei valori medi e le sue applicazioni statistiche.* « Biblioteca dell'Economista » Serie V, Vol. XIX, pag. 227.

* * *

Un esempio numerico varrà a mettere in luce le proprietà che presentano alcune delle medie sopra indicate e ad indirizzare le ricerche atte a scoprire quelle delle altre medie.

Sia la serie di cinque termini 1,2, 10, 20, 100.

La Tavola I dà i valori delle medie monoplane combinatorie potenziate M^{cp} (formula III) per tutti i valori di c da $c = 1$ a $c = 5$ e per i valori di p da $p = -4$ a $p = 4$ (1).

Per $p = 0$, si applica la (III′) ed M^{cp} risulta allora uguale, qualunque sia c, alla media geometrica. M^{cp} risulta pure uguale alla media geometrica, in base alla (III) qualunque sia p, quando sia $c = n$ (nel caso nostro $c = 5$).

TAVOLA I.

Medie monoplane combinatorie potenziate $M^{cp} = \sqrt[cp]{\frac{1}{\binom{n}{c}} \sum_{l=1}^{\binom{n}{c}} P^c_l (a^p_i)}$

$c =$ \ $p =$	− 4	− 3	− 2	− 1	0	1	2	3	4
1	1,17	1,64	1,99	3,01	8,33	26,60	45,83	58,65	66,90
2	1,88	2,07	2,48	3,66	8,33	18,95	26,71	31,08	33,79
3	3,27	3,46	3,83	4,81	8,33	14,40	18,68	21,04	22,41
4	4,95	5,11	5,43	6,23	8,33	10,73	11,91	12,49	12,84
5	8,33	8,33	8,33	8,33	8,33	8,33	8,33	8,33	8,33

Dalla Tavola I si rileva anche che, per $c < n$, M^{cp}, a parità di c, cresce col crescere di p e che, per $p \neq 0$, M^{cp}, a parità di p, cresce col crescere di c per $p < 0$ e invece diminuisce col crescere di c per $p > 0$.

Queste proprietà hanno carattere generale e si possono dimostrare teoricamente (2).

(1) I calcoli per questa tavola e le successive furono eseguiti dal dott. Alberto Fenici.

(2) Il DUNKEL (art. cit.) aveva già dimostrato che M^{cp} cresce per $c = 1$ col crescere di p e diminuisce per $p = 1$ col crescere di c.

La Tavola I si può dividere in due parti, secondo che è $p \gtrless 0$. Nella parte destra in cui è $p > 0$, i valori situati sulle diagonali ascendenti, crescono ; ciò significa che M^{cp} per $p + c =$ costante diminuisce col crescere di p e col diminuire di c. Nella parte sinistra della tavola, invece, in cui $p < 0$, crescono i valori situati sulle diagonali discendenti : ciò significa che M^{cp} cresce per $p - c =$ costante col crescere di p e di c. Queste proprietà discendono immediatamente dalle precedenti, ed hanno pertanto esse pure carattere generale.

Non sempre, invece, diminuiscono per $p > 0$ i valori calcolati sulle diagonali discendenti che corrispondono a valori costanti di $p - c$, nè sempre diminuiscono per $p < 0$ i valori calcolati sulle diagonali ascendenti che corrispondono, a valori costanti di $p + c$.

* * *

Nella Tavola II sono indicati i valori delle medie biplane potenziate $\overset{p}{B^q}$, secondo la formula (IV) per i valori di p da $p = -4$ a $p = 4$ e per i valori di q pure da $q = -4$ a $q = 4$.

Per $p = q$, si applica la (V) che dà la media biplana equipotenziata.

La tavola mostra che $\overset{p}{B^q}$ cresce a parità di p col crescere di q e, analogamente, a parità di q col crescere di p.

Restando costante $p - q$, $\overset{p}{B^q}$ cresce col crescere di p ; crescono pertanto, nella tavola, i valori situati sulle diagonali discendenti. Questa proprietà discende immediatamente dalle precedenti.

I valori di $\overset{p}{B^q}$ sono simmetrici rispetto p e q. Essendo $p + q$ costante e non positivo, $\overset{p}{B^q}$, al crescere di p, diminuisce per p non positivo e, al contrario, cresce per p non negativo. Essendo, invece, $p + q$ costante e positivo, $\overset{p}{B^q}$, al crescere di p, cresce per p non positivo e, al contrario, diminuisce per p non negativo. Tale comportamento si rileva esaminando i valori situati sulle diagonali ascendenti.

La Tavola II può considerarsi come divisa in due settori simili dalla diagonale discendente massima che corrisponde alle coppie di valori $p = q$. Nel settore superiore è $p > q$ e in quello inferiore, invece, $p < q$. In ogni settore, diremo *diagonale discendente base* la diagonale discendente prossima alla massima, per cui è, nel

TAVOLA II.

Media biplane potenziate $\overset{p}{B^q} = \sqrt[p-q]{\dfrac{\sum_{i=1}^{p} a_i^p}{\sum_{i=1}^{q} a_i^q}}$

$q =$ \ $p =$	— 4	— 3	— 2	— 1	0	1	2	3	4
— 4	1,04	1,06	1,09	1,16	1,47	2,62	4,63	7,14	9,93
— 3	1,06	1,08	1,12	1,21	1,64	3,30	6,22	9,82	13,64
— 2	1,09	1,12	1,17	1,31	1,99	4,72	9,45	15,15	20,73
— 1	1,16	1,21	1,31	1,59	3,01	8,95	18,50	27,92	35,99
0	1,47	1,64	1,99	3,01	8,33	26,60	45,84	58,65	66,90
1	2,62	3,30	4,72	8,95	26,60	60,13	78,99	87,10	90,98
2	4,63	6,22	9,45	18,50	45,84	78,99	91,83	96,05	97,64
3	7,14	9,82	15,15	27,92	58,65	87,10	96,05	98,50	99,28
4	9,93	13,64	20,73	35,99	66,90	90,98	97,64	98,28	99,72

settore superiore, $p - q = 1$ e, nel settore inferiore, $q - p = 1$. Ora tutti i valori di $\overset{p}{B^q}$ situati sulle altre diagonali del settore si possono ottenere mediante medie geometriche dai valori consecutivi situati sulla rispettiva diagonale base. E precisamente, se si prende una serie di m valori consecutivi situati sulla diagonale base, la loro media geometrica si trova, nella tavola, al punto di incrocio tra la colonna che finisce ad un estremo della serie considerata di m valori e la linea che finisce all'altro estremo di detta serie. Per esempio, la media geometrica dei due primi valori della diagonale base, 1,06 e 1,12 è 1,09; la media geometrica dei tre successivi valori situati su detta diagonale 1,31, 3,01 e 26,60 è 4,72; la media geometrica di tutti i cinque valori anzi detti è 2,62 e, infine, la media geometrica di tutti gli otto valori situati sulla diagonale base è 9,93.

Le anzidette proprietà hanno carattere generale e si possono dimostrare teoricamente.

* * *

La Tavola III dà i valori delle medie biplane combinatorie secondo la formula (VI).

Per $c = d$ la formula non ha valore determinato.

TAVOLA III.

Medie biplane combinatorie $\overset{c}{B}{}^{d} = \sqrt[c-d]{\dfrac{\binom{n}{d} \sum\limits_{l=1}^{\binom{n}{c}} P^{c}{}_{l}\,(a_i)}{\binom{n}{c} \sum\limits_{l=1}^{\binom{n}{d}} P^{d}{}_{l}\,(a_i)}}$

c = \\ d =	0	1	2	3	4	5
0		26,60	18,95	14,40	10,73	8,33
1	26,60		13,50	10,60	7,93	6,23
2	18,95	13,50		8,31	6,08	4,81
3	14,40	10,60	8,31		4,45	3,66
4	10,73	7,93	6,08	4,45		3,01
5	8,33	6,23	4,81	3,66	3,01	

Dalla tavola si rileva che $\overset{c}{B}{}^{d}$ diminuisce a parità di c col crescere di d e, similmente, diminuisce a parità di d col crescere di c.

I valori situati sulle diagonali discendenti, corrispondenti a coppie di valori c e d per cui resta costante la differenza $c - d$, diminuiscono progressivamente, proprietà che discende dalle precedenti.

I valori di $\overset{c}{B}{}^{d}$ sono simmetrici rispetto a c e d. La tavola si può pertanto dividere in due settori simili, secondo che è $d > c$ (settore superiore) e $d < c$ (settore inferiore). Nel settore superiore, i valori situati sulle diagonali ascendenti, che corrispondono a valori di c e d per cui è $c + d =$ costante, crescono per $c + d = \lesssim 5$ e diminuiscono, invece, per i valori di $c + d > 5$. Il contrario avviene nel settore inferiore.

Anche in questa tavola, come nella precedente, chiameremo *diagonale discendente base* nel settore superiore quella per cui è $d - c = 1$ o nel settore inferiore quella per cui è $c - d = 1$. E anche in questa tavola si rileva che i valori situati sulle altre diagonali si possono ottenere mediante medie geometriche dai valori consecutivi situati sulla diagonale base del settore. Ad esempio, la media geometrica di tutti i cinque valori situati sulla diagonale base è 8,33, valore che si trova all'angolo inferiore e all'angolo superiore della tavola.

Le proprietà anzidette hanno valore generale e si possono dimostrare teoricamente.

Si avverta come i valori di $\overset{c}{B^{d}}$ per $d = 5$, contenuti nell'ultima colonna e quelli per $c = 5$ contenuti nell'ultima linea della tavola, riproducano, in ordine inverso, i valori delle medie monoplane combinatorie potenziate che si ottengono facendo $p = -1$ e che sono indicati nella linea 4ª della Tavola I. Ciò dipende dal fatto che è

$$\frac{\sum\limits_{l=1}^{\binom{n}{n-c}} P_l^{n-c}(a_i)}{\binom{n}{n-c} \prod\limits_{i=1}^{n} (a_i)} = \frac{1}{\binom{n}{c}} \sum_{l=1}^{\binom{n}{c}} P^c{}_l(a_i^{-1})$$

come si desume dalle osservazioni contenute nella nota (1) a pag. 6.

* * *

Passiamo a considerare i valori delle medie biplane combinatorie equipotenziate, secondo la formula (VII).

Per $p = 0$, si applica la formula (VII') che è in ogni caso uguale alla media geometrica.

Se si tiene fermo p e si fanno variare c e d, si rientra nel caso precedente, potendosi considerare $(a_i)^p$ come termine generico di una nuova serie.

Teniamo invece fermo d e vediamo come varii $\overset{c\,p}{B^{d\,p}}$ al variare di p e di c. I risultati sono esposti nella Tavola IV per $d = 1$ e

Medie biplane combinatorie equipotenziate $B^{cp}_{dp} = \sqrt[(c-d)p]{\dfrac{\binom{n}{d} \sum\limits_{l=1}^{\binom{n}{c}} P^c_l (a_i^p)}{\binom{n}{c} \sum\limits_{l=1}^{\binom{n}{d}} P^d_l (a_i^p)}}$

TAVOLA IV.

$d = 1$

$c =$ \ $p =$	— 4	— 3	— 2	— 1	0	1	2	3	4
0	1,47	1,64	1,99	3,01	8,33	26,60	45,83	58,65	66,90
2	2,41	2,62	3,08	4,44	8,33	13,50	15,56	16,47	17,06
3	4,88	5,01	5,31	6,08	8,33	10,61	11,92	12,60	12,96
4	7,40	7,46	7,59	7,94	8,33	7,94	7,59	7,46	7,40
5	12,84	12,49	11,94	10,73	8,33	6,23	5,44	5,11	4,94

TAVOLA V.

$d = 3$

$c =$ \ $p =$	— 4	— 3	— 2	— 1	0	1	2	3	4
0	3,27	3,46	3,83	4,81	8,33	14,40	18,65	21,04	22,41
1	4,87	5,02	5,31	6,08	8,33	10,61	11,92	12,60	12,92
2	9,85	9,64	9,14	8,31	8,33(1)	8,31	9,14	9,64	9,85
4	17,07	16,47	15,56	13,50	8,33	4,45	3,08	2,61	2,43
5	33,80	31,09	26,71	18,95	8,33	3,66	2,47	2,07	1,88

(1) Per $d = 3$, $c = 2$ e $p = \begin{cases} +0,5 \\ -0,5 \end{cases}$ è $B^{cp}_{dp} = 8,18$; per $d = 3$, $c = 2$ e $p \begin{cases} +0.333 \\ -0.333 \end{cases}$ è $B^{cp}_{dp} = 8,24$.

nella Tavola V per $d = 3$ (Analoghi sarebbero i risultati tenendo fermo c e facendo variare p e d, poichè i valori di $\overset{cp}{B}{}^{dp}$ sono simmetrici rispetto a c e d).

Seguendo i valori situati sulla stessa colonna, si vede che, nelle due tavole, a parità di p, i valori di $\overset{cp}{B}{}^{dp}$ crescono con c per $p < 0$, restano costanti (uguali alla media geometrica) per $p = 0$, diminuiscono per $p > 0$.

Seguendo i valori delle diagonali, si rileva che, nella parte sinistra delle tavole, per cui è $p < 0$, i valori crescono lungo le diagonali discendenti e diminuiscono lungo le diagonali ascendenti; mentre il contrario avviene nella parte destra per cui è $p > 0$.

È probabile che tali proprietà abbiano carattere generale.

Più complicate sono le leggi che regolano il variare di $\overset{cp}{B}{}^{dp}$, a parità di d e c, col variare di p.

Per $d = 1$, i valori di $\overset{cp}{B}{}^{dp}$ crescono con p per $c = 0$, $c = 2$, $c = 3$; diminuiscono per $c = 5$; prima crescono e poi diminuiscono per $c = 4$, restando simmetrici rispetto al valore centrale corrispondente a $p = 0$.

Per $d = 3$, invece, i valori di $\overset{cp}{B}{}^{dp}$ crescono con p per $c = 0$ e $c = 1$; diminuiscono per $c = 4$ e $c = 5$; prima diminuiscono e poi crescono per diminuire di nuovo e poi crescere per $c = 2$, restando anche in questo caso simmetrici rispetto al valore centrale corrispondente a $p = 0$.

In entrambi i casi, l'andamento simmetrico si verifica nella colonna per cui è $d + c = 5$.

* * *

Passiamo a considerare le medie biplane equicombinatorie potenziate secondo la formula (VIII).

Per $p = q$ si applica la formula (IX) della media biplana equicombinatoria equipotenziata.

Se teniamo fermo c e facciamo variare p e q, si rientra nel caso della media biplana potenziata, perchè ogni prodotto $P^{c}{}_{l}\,(a_i{}^{p}) = [P^{c}{}_{l}\,(a_i)]^{p}$ può considerarsi come la potenza p^{ma} del termine generico $P^{c}{}_{l}\,(a_i)$ di una nuova serie.

Teniamo quindi fermo q facendo variare p e c (Agli stessi ri-

Medie biplane equicombinatorie potenziate $\overset{cp}{B^{cq}} = \sqrt[(p-q)c]{\dfrac{\sum_{l=1}^{\binom{n}{c}} P^c_l\,(a_i^p)}{\sum_{l=1}^{\binom{n}{c}} P^c_l\,(a_i^q)}}$

TAVOLA VI.

$q = -1$

$c =$ \ $p =$	— 4	— 3	— 2	— 1	0	1	2	3	4
1	1,16	1,22	1,31	1,59	3,01	8,95	18,50	27,92	36,01
2	1,51	1,56	1,68	2,05	3,66	8,33	13,77	18,21	21,66
3	2,87	2,93	3,04	3,37	4,81	8,32	11,89	14,54	16,47
4	4,58	4,63	4,74	4,99	6,23	8,17	9,59	10,50	11,11
5	8,33	8,33	8,33	8,33	8,33	8,33	8,33	8,33	8,33

TAVOLA VII.

$q = 1$

$c =$ \ $p =$	— 4	— 3	— 2	— 1	0	1	2	3	4
1	2,62	3,30	4,72	8,95	26,60	60,13	78,98	87,10	90,98
2	2,99	3,60	4,88	8,33	18,95	32,37	37,65	39,80	40,97
3	4,40	4,94	5,95	8,32	14,40	21,79	24,23	25,43	25,96
4	5,77	6,15	6,81	8,17	10,73	12,62	13,20	13,27	13,62
5	8,33	8,33	8,33	8,33	8,33	8,33	8,33	8,33	8,33

sultati si perviene tenendo fermo p e facendo variare q e c, perchè il valore di $\overset{cp}{B^{cq}}$ è simmetrico rispetto a p e q).

Le due Tavole VI e VII contengono i valori di $\overset{cp}{B^{cq}}$ per $q = 1$ e $q = -1$ quando c assume i valori da 1 a 5 e p i valori da -4 a 4.

L'esame delle linee delle due tavole mostra che, a parità di p e q, $\overset{cp}{B^{cq}}$ cresce col crescere di p per i valori di c da 1 a 4 e resta costante (uguale alla media geometrica) per $c = 5$.

L'esame delle colonne della Tavola VI mostra che, per $q = 1$, $\overset{cp}{B^{cq}}$ cresce, al crescere di c, per $p = -4$, $p = -3$, $p = -2$, $p = -1$ e $p = 0$; diminuisce per $p = 2$, $p = 3$, $p = 4$, e presenta un andamento irregolare per $p = 1$.

I risultati contenuti nella Tavola VII sono analoghi, con la differenza che qui la colonna dei valori irregolari corrisponde a $p = -1$. Nelle colonne a destra di questa, i valori di $\overset{cp}{B^{cq}}$ diminuiscono al crescere di c; nelle colonne a sinistra, crescono.

L'esame delle diagonali ascendenti e discendenti mostra un andamento dei valori diverso nella Tavola VI e nella VII e, in ciascuna di esse, non uniforme per tutte le diagonali.

Le leggi di variazione dei valori di $\overset{cp}{B^{cq}}$ al variare di c, di p e di q sono evidentemente complicate: meriterebbero uno studio teorico a sè.

* * *

La Tavola VIII, infine, dà i valori per le medie biplane combinatorie potenziate, secondo la formula (X). per $p = 2$, $q = 1$, quando c e d assumono ciascuno i valori da 1 a 5.

La condizione di sufficienza perchè $\overset{cp}{B^{dq}}$ sia una vera media (cfr. pag. 9) si verifica quando sia $d \lesssim c$, vale a dire, nel settore della tavola posto al di sotto della linea in grassetto. In tale settore, i valori di $\overset{cp}{B^{dq}}$ diminuiscono, lungo la stessa colonna, a parità di d, col crescere di c; crescono, col crescere di c, lungo le diagonali ascendenti per cui è $c + d =$ costante e diminuiscono lungo le diagonali discendenti per cui è $c - d =$ costante; sono irregolari lungo talune linee, e cioè, quando d varia, a parità di c, Il

TAVOLA VIII.

Medie biplane combinatorie potenziate $\overset{cp}{B^{dq}} = \sqrt{\dfrac{\binom{n}{d}\sum_{l=1}^{\binom{n}{c}} P^{c}{}_{l}\,(a^{p}{}_{i})}{\binom{n}{c}\sum_{l=1}^{\binom{n}{d}} P^{d}{}_{l}\,(a^{q}{}_{i})}}$

$p = 2 \quad q = 1$

c = \ d =	1	2	3	4	5
1	78,98	+ ∞ (1)	1,42	2,51	2,67
2	26,75	37,65	170,54	+ ∞ (1)	0,079
3	13,92	18,54	24,24	56,59	1063,11
4	10,61	10,20	10,62	13,21	13,60
5	7,33	6,78	6,58	7,89	8,33

(1) Cfr. nota (2) a pag. 11.

fatto che i valori decrescono regolarmente lungo le colonne e risultano irregolari lungo le linee, fa pensare che anche le regolarità riscontrate in questo settore possano non verificarsi per altri valori di p e di q.

Nell'altro settore della tavola, per cui è $d \geq c$, i valori presentano la maggiore irregolarità. Per quanto per essi non si verifichi la condizione sufficiente di internalità, 5 di essi sono tuttavia interni ai limiti estremi della serie (1 e 100) e 5, invece, sono esterni, 4 superando il massimo ed uno restando inferiore al minimo.

La complicatezza delle leggi che regolano il variare di $\overset{cp}{B^{dq}}$ al variare di p, q, c, e d, ha consigliato a non dilungarci in questa nota sulle applicazioni numeriche della formula (X) che merita uno studio a parte.

RENZO CISBANI

Contributi alla teoria delle medie

Sotto questo titolo, ho riunito alcune note relative ad argomenti disparati appartenenti al vasto campo delle medie. Esse sono state stese per risolvere alcune questioni particolari su cui ha richiamato la mia attenzione il Prof. Gini in occasione di di una trattazione sistematica di tale materia che è in corso.

In questo numero della rivista compaiono le prime due note:

I. — *Di alcune medie continue;*

II. — *Determinazione univoca della mediana nelle serie cicliche.*

1. — Di alcune medie continue.

SOMMARIO. — L. Galvani ha esteso al continuo una formula del Dunkel, esprimente una media in senso generale di n quantità discrete a_i. Qui si determina una formula più generale per quella media continua, nella distribuzione dei valori della quale è immersa la distribuzione dei valori di quella data dal Galvani. Infine si studiano alcune interessanti ed elementari proprietà di quella distribuzione.

POSIZIONE DEL PROBLEMA.

§ 1. Dunkel (1) ha proposto quale « media generalizzata » delle quantità (discrete):

$$a_1 , a_2 , \dots a_n$$

la funzione continua per qualunque valore dell'argomento x

$$\begin{cases} y_n(x) = \left[\dfrac{\Sigma a_i^x}{n}\right]^{\frac{1}{x}} & \text{per } x \neq 0 \\ y_n(x) = \sqrt[n]{a_1 a_2 \dots a_n} & \text{per } x = 0 \end{cases}$$

(1) DUNKEL, *Generalized geometric means and algebraic equations.* « Annals of Mathematics », 1909-1910.

La funzione $y_n(x)$ gode infatti delle seguenti proprietà, che giustificano la proposta ora fatta:

a) per ogni sistema di numeri a_i, è funzione simmetrica di essi;

b) ha valori sempre compresi fra i due limiti:

$$a = min\,(a_i) \qquad b = max\,(a_i)$$

c) rappresenta le medie analitiche usualmente impiegate, per determinati valori dell'argomento; p. e. per $x = -1$, $x = 0$, $x = 1$, $x = 2$, rappresenta le ordinarie medie armonica, geometrica, aritmetica, quadratica;

d) è sempre:

$$\lim_{n=-\infty} y_n(x) = a\,; \qquad \lim_{n=+\infty} y_n(x) = b$$

L. Galvani (2) estende tale concetto di « media generalizzata » al caso di una totalità di numeri reali compresi fra i limiti, a, b; ed ottiene come espressione di essa la formula:

$$\begin{cases} \bar{y}(x) = \left[\dfrac{b^{x+1} - a^{x+1}}{(x+1)\,(b-a)}\right]^{\frac{1}{x}} & \text{per } x \neq 0\,, -1 \\[2ex] \bar{y}(x) = \dfrac{b-a}{log\,b - log\,a} & \text{per } x = -1 \\[2ex] \bar{y}(x) = \dfrac{1}{e}\left[\dfrac{b^b}{a^a}\right]^{\frac{1}{b-a}} & \text{per } x = 0 \end{cases}$$

La media ora definita gode delle proprietà analoghe alle *a*), *b*), *c*), *d*), e per ... $x = -1$, $x = 0$, $x = 1$, ..., dà le corrispondenti delle medie..., armonica, geometrica, aritmetica..., che risultano espresse rispettivamente dalle formule:

$$\ldots, \frac{b-a}{log\,b - log\,a}, \frac{1}{e}\left[\frac{b^b}{a^a}\right]^{\frac{1}{b-a}}, \frac{b+a}{2}, \ldots \quad (*)$$

L'estensione di L. Galvani non conduce però ad un risultato del tutto generale: vedremo infatti come una più completa estensione al caso continuo del concetto di « media generalizzata » del Dunkel porta a definire una media in senso generale della variabile continua t in (a, b), che contiene, come caso particolare la

(2) L. Galvani, *Dei limiti a cui tendono alcune medie*. « Boll. Un. Mat. Italiana », 1927.

media di Galvani, e nella distribuzione dei valori della quale, la distribuzione dei valori (*) della $\bar{y}$ risulta immersa.

Il motivo di ciò è da ricercarsi nel fatto che L. Galvani ha determinato la forma della sua media continua passando al limite a partire da una successione discreta di punti le cui ascisse sono in progressione aritmetica. Mentre questo non è che uno degli infiniti modi mediante i quali si può eseguire il passaggio al limite ora detto.

Ricerca della media continua in senso generale.

§ 2. Ciò posto, supponiamo di dividere l'intervallo (a, b), in n parti, le ascisse degli estremi delle quali, sieno in progressione. . ., armonica, aritmetica, quadratica. . ., in corrispondenza ai successivi valori interi relativi di j. Le ascisse degli estremi saranno della forma

$$[a^j + i\,h]^{\frac{1}{j}} \qquad j = \ldots -1, 0, 1, 2, \ldots$$

Studiamo il comportamento della funzione:

$$y_j(x) = \left[\frac{1}{n} \Sigma\, [a^j + i\,h]^{\frac{x}{j}}\right]^{\frac{1}{x}} \qquad \text{per } j \neq 0 \; x \neq 0 \quad (1)$$

al variare di x, j.

Vediamo se possiamo assumere come « media continua » della variabile t in (a, b), nel senso del Dunkel, l'espressione:

$$\bar{y}_j(x) = \lim_{n=\infty} y_j(x)$$

Si ha:

$$\sqrt[j]{a^j + nh} = b\,,\; h = \frac{b^j - a^j}{n}$$

e quindi:

$$\bar{y}_j(x) = \left[\frac{1}{b^j - a^j} \int_{a^j}^{b^j} t^{\frac{x}{j}}\, d\,t\right]^{\frac{1}{x}}$$

(1) Ora è nel seguito si intenda che le considerazioni svolte si limitano al campo reale.

$$\begin{cases} \bar{y}_j(x) = \left[\dfrac{b^{x+j} - a^{x+j}}{\left(\dfrac{x}{j} + 1\right)(b^j - a^j)}\right]^{\frac{1}{x}} & \text{per } x \neq -j \\ \bar{y}_j(x) = \left[\dfrac{b^j - a^j}{j\,(log\,b - log\,a)}\right]^{\frac{1}{j}} & \text{per } x = -j \end{cases}$$

Posto

$$\bar{y}_j(x) = \frac{1}{e^{\frac{1}{j}}}\left[\frac{b^{b^j}}{a^{a^j}}\right]^{\frac{1}{b^j - a^j}} \qquad \text{per } x = 0,$$

è facile mostrare che la $\bar{y}_j(x)$ è funzione continua per qualunque valore x dell'argomento.

È intanto $\bar{y}_j(x)$ continua per $x \neq 0$, $x \neq -j$.

Si ha poi, applicando la regola dell'Hôpital

$$log \lim_{x=0} \bar{y}_j(x) = \lim_{x=0} \frac{1}{x}\left[log\,(b^{x+j} - a^{x+j}) - log\,(b^j - a^j) - \right.$$

$$\left. - log\left(\frac{x}{j} + 1\right)\right] = \lim_{x=0}\left[\frac{b^{x+j}\,log\,b - a^{x+j}\,log\,a}{b^{x+j} - a^{x+j}} - \frac{\frac{1}{j}}{\frac{x}{j} + 1}\right] =$$

$$= \frac{b^j\,log\,b - a^j\,log\,a}{b^j - a^j} - \frac{1}{j}$$

e quindi

$$\lim_{x=0} \bar{y}_j(x) = \frac{1}{e^{\frac{1}{j}}}\left[\frac{b^{b^j}}{a^{a^j}}\right]^{\frac{1}{b^j - a^j}}$$

Si ha infine:

$$\lim_{x=-j} \bar{y}_j(x) = \lim_{x=-j}\left[\frac{b^{x+j} - a^{x+j}}{(b^j - a^j)\left(\dfrac{x}{j} + 1\right)}\right]^{\frac{1}{x}} =$$

$$= \lim_{x=-j}\left[\frac{b^{x+j}\,log\,b - a^{x+j}\,log\,a}{\dfrac{1}{j}\,(b^j - a^j)}\right]^{-\frac{1}{j}} = \left[\frac{j\,(log\,b - log\,a)}{b^j - a^j}\right]^{-\frac{1}{j}}$$

e quindi

$$\lim_{x=-j} \bar{y}_j(x) = \left[\frac{b^j - a^j}{j\,(log\,b - log\,a)}\right]^{\frac{1}{j}}$$

La $\bar{y}_j(x)$ gode inoltre di tutte le proprietà enunciate nel § 1. Qui basterà di mostrare che gode della proprietà *d*). Si ottiene infatti successivamente:

$$\log \lim_{x=-\infty} \bar{y}_j(x) = \lim_{x=-\infty} \frac{1}{x}\left[\log(b^{x+j} - a^{x+j}) - \log\left(\frac{x}{j}+1\right) - \right.$$

$$\left. -\log(b^j - a^j)\right] = \lim_{x=-\infty} \frac{1}{x}\left[(x+j)\log a + \log\left[\left(\frac{b}{a}\right)^{x+j} - 1\right] - \right.$$

$$\left. -\log\left(\frac{x}{j}+1\right) - \log(b^j - a^j)\right] = \lim_{x=-\infty}\left[\log a + \frac{\left(\frac{b}{a}\right)^{x+j}(\log b - \log a)}{\left(\frac{b}{a}\right)^{x+j} - 1} - \frac{\frac{1}{j}}{\frac{x}{j}+1}\right] = \log a$$

e analogamente

$$\log \lim_{x=+\infty} \bar{y}_j(x) = \log b$$

Dunque la funzione $\bar{y}_j(x)$, per $j \neq 0$, è atta a rappresentare una « media continua » nel senso di Dunkel: e si noti subito che, per $j = 1$, essa coincide con la « media generalizzata » di Galvani.

§ 3. Per fare assumere significato alla $\bar{y}_j(x)$, anche per $j = 0$, occorre studiare il comportamento della funzione

$$y_0(x) = \left[\frac{1}{n} \Sigma (a k^i)^x\right]^{\frac{1}{x}} \qquad \text{per } x \neq 0$$

per $n = \infty$; posto $\bar{y}_0(x) = \lim_{n=\infty} y_0(x)$ si ha

$$a k^n = b \quad , \quad k = \left(\frac{b}{a}\right)^{\frac{1}{n}}$$

e quindi

$$\bar{y}_0(x) = \lim_{n=\infty}\left[\frac{1}{n} \Sigma\, a^{\frac{x}{n}(n-i)}\, b^{\frac{x}{n}i}\right]^{\frac{1}{x}}$$

Posto

$$\frac{\log b - \log a}{n} = h\,, \quad \text{è} \quad a^{\frac{x}{n}(n-i)}\, b^{\frac{x}{n}i} = e^{x(\lg a + ih)};$$

quindi

$$y_0(x) = \left[\lim_{n=\infty} \frac{1}{n} \sum_0^{n-1} e^{x(\log a + ih)}\right]^{\frac{1}{x}} = \left[\frac{1}{\log b - \log a}\int_{\log a}^{\log b} e^{xt}\, dt\right]^{\frac{1}{x}}$$

che conduce immediatamente alla formula

$$\bar{y}_0(x) = \left[\frac{b^x - a^x}{x(\log b - \log a)}\right]^{\frac{1}{x}} \qquad \text{per } x \neq 0$$

Posto ora $\bar{y}_0(x) = \sqrt{ab}$ per $x = 0$, poichè si ottiene facilmente

$$\lim_{x=0} \bar{y}_0(x) = \sqrt{ab}$$

la $\bar{y}_0(x)$ risulta continua per qualunque valore x dell'argomento. Con calcoli analoghi a quelli del paragrafo precedente, si ottiene

$$\lim_{x=-\infty} \bar{y}_0(x) = a\,; \qquad \lim_{x=\infty} \bar{y}(x) = b$$

e si vede che le altre proprietà richieste per la media in senso generale del Dunkel son tutte verificate anche per la $\bar{y}_0(x)$.

Poniamo

$$\bar{y}_j(x) = \bar{y}_0(x) \qquad \text{per } j = 0.$$

Se dimostriamo che $\lim\limits_{j=0} \bar{y}_j(x) = \bar{y}_0(x)$ la funzione $\bar{y}(x)$ risulta continua non solo per qualunque valore x dell'argomento, ma anche per qualunque valore j dell'indice.

A tal uopo, si noti che, per $x \neq 0$, $x \neq -j$, si ha

$$\lim_{j=0} \bar{y}_j(x) = \left[\frac{b^x - a^x}{x \lim\limits_{j=0} \dfrac{b^j - a^j}{j}}\right]^{\frac{1}{x}} = \left[\frac{b^x - a^x}{x(\log b - \log a)}\right]^{\frac{1}{x}};$$

per $x = 0$, si ha

$$\log \lim_{j=0} \bar{y}_j(x) = \log \lim_{j=0} \frac{1}{e^{\frac{1}{j}}} \left[\frac{b^{b^j}}{a^{a^j}}\right]^{\frac{1}{b^j - a^j}} = \lim_{j=0} \frac{b^j \log b - a^j \log a}{b^j - a^j} - \frac{1}{j} =$$

$$= \lim_{j=0} \frac{j(b^j \log b - a^j \log a) - (b^j - a^j)}{j\,(b^j - a^j)} = \lim_{j=0} \frac{b^j (\log b)^2 - a^j (\log a)^2}{\dfrac{b^j - a^j}{j} + b^j \log b - a^j \log a} = \frac{\log a + \log b}{2}$$

e che infine, per $j = 0$, si ricade nel caso precedente.

Dunque la funzione

$$z(x, j) = \left[\frac{b^{x+j} - a^{x+j}}{\left(\dfrac{x}{j} + 1\right)(b^j - a^j)}\right]^{\frac{1}{x}} \qquad \text{per } x \neq 0,\ x \neq -j\,;\ j \neq 0$$

$$z(x, j) = \left[\frac{b^x - a^x}{x(\log b - \log a)}\right]^{\frac{1}{x}} \qquad \text{per } x \neq 0,\ j = 0$$

$$z(x, j) = \left[\frac{b^j - a^j}{j(\log b - \log a)}\right]^{\frac{1}{j}} \qquad \text{per } x = -j,\ j \neq 0$$

$$z=\left[\frac{b^{x+j}-a^{x+j}}{\left(\frac{x}{j}\right)(b^j-a^j)}\right]^{\frac{1}{x}}$$

Fig. 1

$x \backslash j$	∞	...	3	2	1	0	−1	−2	−3	...	$-\infty$
$-\infty$		a	a	a	a	a	a	a	a	a	
...	b	...	...	... K	...	...	...	...	...	...	a
−3	b	...	$\sqrt[3]{\frac{b^3-a^3}{3(\lg b-\lg a)}}$	$\sqrt[3]{\frac{(a+b)ab}{2}}$	$\frac{ab}{\sqrt[3]{\frac{(a+b)ab}{2}}}$	$\frac{ab}{\sqrt[3]{\frac{b^3-a^3}{3(\lg b-\lg a)}}}$	$\frac{ab}{\sqrt[3]{\frac{(a+b)(a^2+b^2)}{4}}}$	$\frac{ab}{\sqrt[3]{\frac{2(b^5-a^5)}{5(b^2-a^2)}}}$	$\frac{ab}{\sqrt[3]{\frac{b^3+a^3}{2}}}$	...	a
−2	b	...	$\sqrt{\frac{b^2+ab+a^2}{3}}$	$\sqrt{\frac{b^2-a^2}{2(\lg b-\lg a)}}$	$\sqrt{ab}$	$\frac{b}{\sqrt{\frac{b^2-a^2}{2(\lg b-\lg a)}}}$	$\frac{ab}{\sqrt{\frac{a^2+ab+b^2}{3}}}$	$\frac{ab}{\sqrt{\frac{a^2+b^2}{2}}}$	$\frac{ab}{\sqrt{\frac{3(b^5-a^5)}{5(b^3-a^3)}}}$	...	a
−1	b	...	$\frac{2}{3}\frac{b^2+ab+a^2}{a+b}$	$\frac{a+b}{2}$	$\frac{b-a}{\lg b-\lg a}$	$\frac{b}{\frac{b-a}{\lg b-\lg a}}$	$\frac{ab}{\frac{a+b}{2}}$	$\frac{ab}{\frac{2}{3}\cdot\frac{b^2+ab+b^2}{a+b}}$	$\frac{ab}{\frac{3(b+a)(a^2+b^2)}{4(a^2+ab+b^2)}}$	...	a
0	b	... H	$\frac{1}{\sqrt[3]{e}}\left(\frac{b^{b^3}}{a^{a^3}}\right)^{\frac{1}{b^3-a^3}}$	$\frac{1}{\sqrt{e}}\left(\frac{b^{b^2}}{a^{a^2}}\right)^{\frac{1}{b^2-a^2}}$	$\frac{1}{e}\left(\frac{b^b}{a^a}\right)^{\frac{1}{b-a}}$	$\sqrt{ab}$	$\frac{ab}{\frac{1}{e}\left(\frac{b^b}{a^a}\right)^{\frac{1}{b-a}}}$	$\frac{ab}{\frac{1}{\sqrt{e}}\left(\frac{b^{b^2}}{a^{a^2}}\right)^{\frac{1}{b^2-a^2}}}$	$\frac{ab}{\frac{1}{\sqrt[3]{e}}\left(\frac{b^{b^3}}{a^{a^3}}\right)^{\frac{1}{b^3-a^3}}}$	H ...	a
1	b	...	$\frac{3(b+a)(b^2+a^2)}{4(a^2+ab+b^2)}$	$\frac{2}{3}\frac{b^2+ab+a^2}{a+b}$	$\frac{a+b}{2}$	$\frac{b-a}{\lg b-\lg a}$	$\frac{ab}{\frac{b-a}{\lg b-\lg a}}$	$\frac{ab}{\frac{a+b}{2}}$	$\frac{ab}{\frac{2}{3}\frac{b^2+ab+b^2}{a+b}}$	...	a
2	b	...	$\sqrt{\frac{3(b^5-a^5)}{5(b^3-a^3)}}$	$\sqrt{\frac{a^2+b^2}{2}}$	$\sqrt{\frac{a^2+ab+b^2}{3}}$	$\sqrt{\frac{b^2-a^2}{2(\lg b-\lg a)}}$	$\sqrt{ab}$	$\frac{ab}{\sqrt{\frac{b^2-a^2}{2(\lg b-\lg a)}}}$	$\frac{ab}{\sqrt{\frac{b^2+ab+b^2}{3}}}$	...	a
3	b	...	$\sqrt[3]{\frac{b^3+a^3}{2}}$	$\sqrt[3]{\frac{2(b^5-a^5)}{5(b^2-a^2)}}$	$\sqrt[3]{\frac{(a+b)(a^2+b^2)}{4}}$	$\sqrt[3]{\frac{b^3-a^3}{3(\lg b-\lg a)}}$	$\sqrt[3]{\frac{(a+b)ab}{2}}$	$\frac{ab}{\sqrt[3]{\frac{(a+b)ab}{2}}}$	$\frac{ab}{\sqrt[3]{\frac{b^3-a^3}{3(\lg b-\lg a)}}}$	...	a
...	...	...	...	...	...	...	...	K ...	...	...	a
∞		b	b	b	b	b	b	b	b	b	

$$z\,(x\,,j) = \frac{1}{e^j}\left[\frac{b^{b^j}}{a^{a^j}}\right]^{\frac{1}{b^j + a^j}} \qquad \text{per } x = 0\,,\ j \neq 0$$

$$z\,(x\,,j) = \sqrt{a\,b} \qquad \text{per } x = 0\,,\ j = 0$$

è funzione continua per tutti i valori x, j, (1), e soddisfa a tutte le condizioni richieste per la media in senso generale del Dunkel; essa è quindi atta a rappresentare la più generale « media continua » della variabile t in (a, b), media che risulta appunto la cercata estensione di quella proposta dal Dunkel nel caso di quantità a_i discrete.

Proprietà della media proposta e applicazioni.

§ 4. Può ora essere interessante studiare la distribuzione di z, in corrispondenza dei valori interi delle due variabili x, j da cui essa dipende.

Riuniremo in un quadro a duplice entrata tali valori di z, riportando in ordinata la variabile x e in ascissa la variabile j.

Otterremo la distribuzione della Figura 1ª; in corrispondenza di un qualunque valore di j, e di un qualunque valore di x si hanno, all'incontro delle rispettive colonne e righe, i valori corrispondenti della « media continua » prima definita.

Nei paragrafi precedenti già è stato dimostrato che i limiti superiore ed inferiore delle medie di ciascuna colonna son fissi e rispettivamente eguali ad a, b, cioè agli estremi dell'intervallo; e, inoltre, le medie di ogni colonna crescono al crescere di x, in accordo con la nota proprietà, che si verifica nel caso di quantità discrete, che le medie aumentano il loro valore al crescere dell'esponente x della formula del Dunkel.

Lo stesso fatto ora notato per le colonne vale anche per le righe del quadro: si ha cioè con calcoli analoghi a quelli dei paragrafi precedenti

$$\lim_{j=-\infty} z\,(x\,,j) = a\,;\ \lim_{j=+\infty} z\,(x,j) = b$$

per qualunque valore di x.

---

(1) S'intende che nel testo si è dimostrato solo che la $z\,(x\,,j)$ è continua rispetto a ciascuna delle due variabili presa separatamente.

Dunque le varie righe e la varie colonne del quadro costituiscono altrettanti sistemi di valori della nostra media in senso generale della variabile t in (a, b), aventi per valori estremi, gli estremi di quell'intervallo.

I valori $z\,(x, j)$ per i quali è $x = j$ danno le medie. . ., armonica, geometrica, aritmetica . . . degli estremi. È noto infatti che inserendo fra due quantità a, b, un qualunque numero n di medi. . ., armonici, geometrici, aritmetici. . ., la media. . ., armonica, geometrica, aritmetica,. . . risulta sempre, rispettivamente, la media. . ., armonica, geometrica, aritmetica. . ., degli estremi a, b.

L'esame del quadro porta ad enunciare alcune proprietà che si possono esprimere geometricamente nel seguente modo.

Si tracci una linea che passi per i centri dei rettangoli la cui media corrispondente è $\sqrt{a\,b}$, linea che risulta essere una retta $(K\,K)$, e che diremo retta fondamentale. Centro della distribuzione diremo il centro del quadro, la cui media corrispondente è $z\,(0\,,0) = \sqrt{a\,b}$, asse la linea dei centri dei rettangoli delle medie corrispondenti ai valori $z\,(0\,,j)$, cioè la retta $(H\,H)$.

Le proprietà di cui è il discorso si possono allora così esprimere:

a) le medie corrispondenti ai rettangoli per i centri dei quali passa una parallela alla retta fondamentale, e simmetrici rispetto l'intersezione di essa con l'asse, coincidono;

b) le medie z', z'' corrispondenti ai rettangoli i centri dei quali giacciono su una retta qualunque per il centro del quadro, e simmetrici rispetto ad esso, sono legate dalla semplice relazione

$$z' \cdot z'' = a\,b$$

Sinteticamente, si hanno le formule:

$$a) \qquad z\,(x\,, -2\,x) = \sqrt{a\,b}\,,\ z\,(x, j) = z\,(-x\,, x + j)$$

$$b) \qquad z\,(x\,,j)\,z\,(-x\,,-j) = a\,b$$

che esprimono le proprietà prima dette, e delle quali daremo ora la semplice giustificazione analitica.

Consideriamo la superficie z

$$z\,(x\,,j) = \left[\frac{b^{x+j} - a^{x+j}}{\left(\frac{x}{j} + 1\right)(b^j - a^j)}\right]^{\frac{1}{x}} \ (*)$$

resa continua per tutti i valori di x, j, (1), mediante le posizioni fatte nel § 3, e riferiamola ad un sistema di assi cartesiani (z, x, j).

La retta rappresentata, parametricamente, dalle equazioni

$$x = t \, , j = - \frac{1}{2} t \, , z = \sqrt{a\,b} \quad (**)$$

appartiene alla superficie, come si può vedere immediatamente, sostituendo i valori ora detti di x, j, z nella (*), che risulta identicamente soddisfatta.

Una parallela alla (**), le cui equazioni parametriche saranno

$$x = t + h \, , j = - \frac{1}{2} t + h \, , z = l \, ,$$

taglia la superficie z in punti, le cui coordinate son date dai valori di t soluzioni dell'equazione

$$\frac{b^{\frac{1}{2}t + 2h} - a^{\frac{1}{2}t + 2h}}{\dfrac{\frac{1}{2}t + 2h}{-\frac{1}{2}t + 2h}\left(b^{-\frac{1}{2}t + h} - a^{-\frac{1}{2}t + h}\right)} = l^{t+h} \quad (***)$$

Senza risolvere completamente questa equazione, possiamo studiare il comportamento delle soluzioni di essa, per ogni fissato valore di l, cioè del valore della media generalizzata, e di h.

Sia $t = \bar{t}$ una soluzione corrispondente ad un fissato valore di l, $l = \bar{l}$, compreso tra a e b, e di h, $h = \bar{h}$; allora anche il valore $t = -\bar{t} - 2\bar{h}$ è soluzione corrispondente agli stessi valori fissati $\bar{l}$, $\bar{h}$.

Infatti, sostituendo nella (***), al posto di t, il valore detto, si ottiene

$$\frac{b^{-\frac{1}{2}\bar{t} + \bar{h}} - a^{-\frac{1}{2}\bar{t} + \bar{h}}}{\dfrac{-\frac{1}{2}\bar{t} + \bar{h}}{\frac{1}{2}\bar{t} + 2\bar{h}}\left(b^{\frac{1}{2}\bar{t} + 2\bar{h}} - a^{\frac{1}{2}\bar{t} + 2\bar{h}}\right)} = \bar{l}^{-\bar{t} - \bar{h}}$$

---

(1) La $z\,(x\,,j)$ è una superficie rappresentativa dei valori interni all'intervallo (a, b). Non deve stupire il fatto che una superficie, cioè un ente geom. a due dimensioni, rappresenti i punti di un intervallo. Dalla teoria dei numeri transfiniti sappiamo che tanto l'intervallo $(a\,,b)$ quanto la superficie hanno la stessa potenza puntuale [c].

equazione identicamente soddisfatta, se $t = \bar{t}$, è una soluzione corrispondente ai predetti valori fissati $\bar{l}$, $\bar{h}$.

Le soluzioni si distribuiscono dunque a coppie sulle parallele (**) ; e cadono precisamente in punti simmetrici rispetto l'intersezione di quelle parallele con la parallela all'asse j condotta per il punto $(0, 0, l)$: il che si può facilmente ottenere ricorrendo ad una nota formula di geometria analitica.

Infine può facilmente dedursi dall'identità :

$$z(x, j) \cdot z(-x, -j) = a\, b$$

di facile dimostrazione, che le ordinate z di punti della superficie le cui proiezioni sul piano $x\,y$ sono simmetriche rispetto l'origine, sono inversamente proporzionali.

In tal modo le proprietà a), b) sono state completamente giustificate.

§ 5. — Una delle proprietà della distribuzione studiata nel paragrafo precedente potrebbe essere utilmente applicata ai problemi di indole pratica del tipo che ora esporremo.

Supponiamo di conoscere che due serie hanno intensità in progressione, ad es., rispettivamente, aritmetica e geometrica ; le formule del paragrafo precedente ci dicono che la media $z(0, -1)$ coincide con la media $z(1, -1)$, il che significa che, al limite per il numero dei termini tendente all'infinito, devono coincidere la media armonica della prima serie e la media aritmetica della seconda.

Di modo che, se il numero dei termini delle due serie è abbastanza alto, e si sia venuti a conoscere, ad es., la media armonica della serie i cui termini sono in progressione aritmetica, il valore di essa può essere assunto, con discreta approssimazione, come valore della media aritmetica della seconda serie, i cui termini crescono in progressione geometrica.

La condizione che il numero dei termini sia abbastanza alto ha effetto notevole sul grado dell'approssimazione numerica che si vuole ottenere.

Un esempio numerico illustrerà meglio dei calcoli algebrici, che dovrebbero esser fatti sui limiti

$$\lim_{n=\infty} \left[\frac{1}{n} \left\{ a + \sqrt[n-1]{a^{n-2}\, b} + \ldots + \sqrt[n-1]{a\, b^{n-2}} + b \right\} \right] =$$

$$= \lim_{n=\infty} \left[\frac{n}{\frac{1}{a} + \frac{n-1}{(n-2)\,a + b} + \ldots \frac{n-1}{a + (n-2)\,b} + \frac{1}{b}} \right] = \frac{b-a}{\log b - \log a},$$

il grado di approssimazione che si può ottenere.

Posto $a = 1$, $b = 2$, si ottiene il grafico a tratto unito della fig. 2, ove la curva inferiore rappresenta la media armonica della prima serie, la curva superiore la media aritmetica della seconda serie; la retta, il valore limite cui le due medie convergono per $n = \infty$.

Come si vede, la convergenza non è molto rapida, ma si può presumere che per ottenere una coincidenza alla seconda cifra decimale (approssimazione a meno di $\frac{1}{100}$ nella determinazione della media della seconda serie mediante il valore della prima) occorre al più disporre di una cinquantina di termini.

Un altro esempio è riportato nel grafico tratteggiato della fig. 2.

Si è posto ancora $a = 1$, $b = 2$; e si è supposto di essere in presenza di due serie i cui termini sono in progressione aritmetica e quadratica.

La curva rappresentativa della media aritmetica della prima serie è una retta coincidente con la retta limite, $y = 1{,}5$; come doveva accadere ricordando che, qualunque sia il numero dei termini, la media aritmetica di una progressione aritmetica è la media aritmetica degli estremi.

La curva rappresentativa della media armonica della seconda serie converge lentamente, mantenendosi sempre inferiore, al valor limite 1,5, secondo la formula:

$$\lim_{n=\infty}\left[\frac{n}{\frac{1}{a}+\frac{1}{\sqrt{\frac{(n-2)\,a^2+b^2}{n-1}}}+\dots+\frac{1}{\sqrt{\frac{a^2+(n-2)\,b^2}{n-1}}}+\frac{1}{b}}\right]=\frac{a+b}{2}=1{,}5.$$

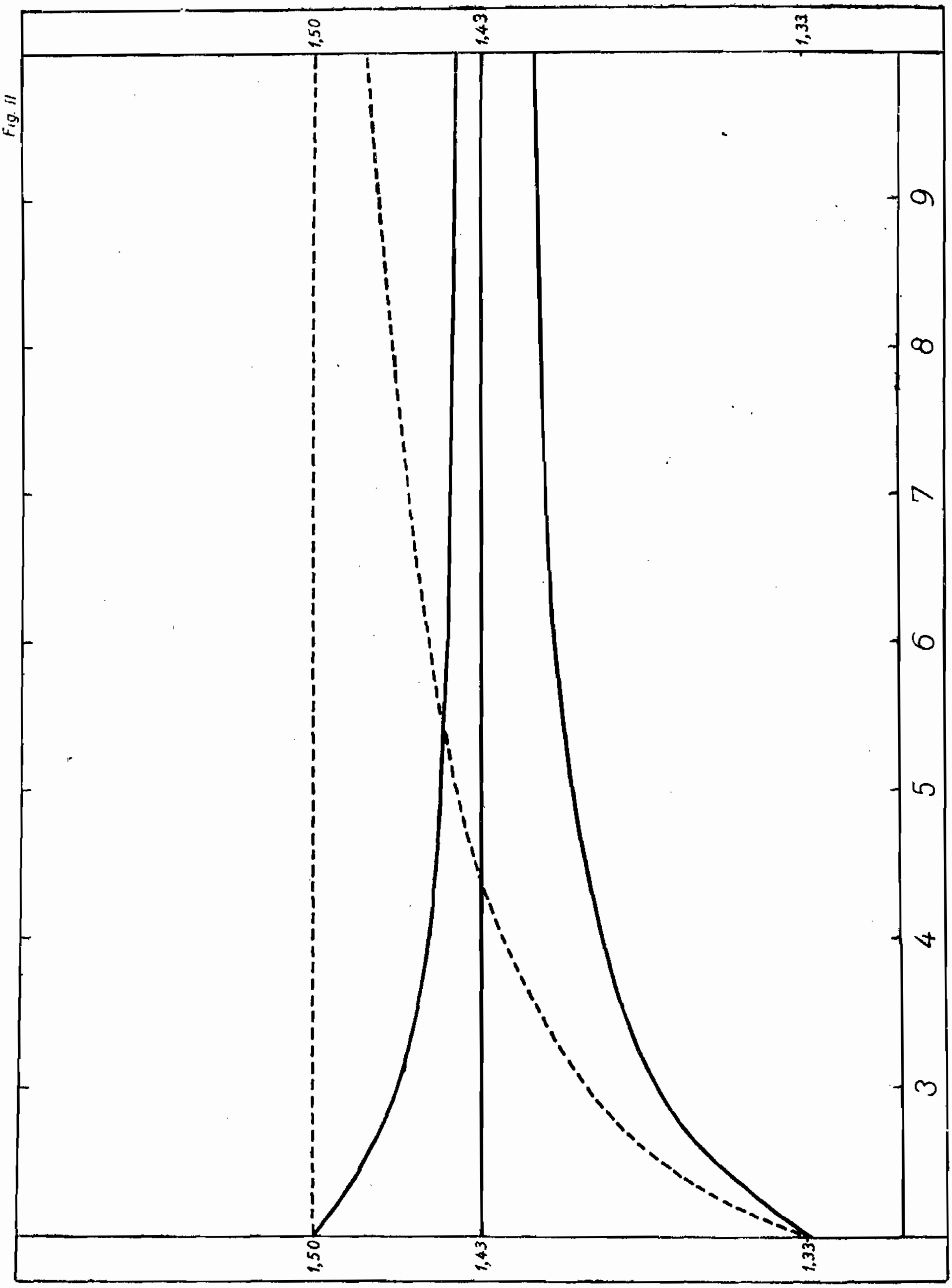


Fig. II

II. — Determinazione univoca della mediana nelle serie cicliche.

SOMMARIO. — Jackson ha ideato un artificio teorico atto a determinare in modo univoco la mediana di una successione di numeri. Qui si studia l'estensione del detto artificio ad una successione di modalità qualitative che danno luogo a serie cicliche.

Si dà, in pari tempo, un processo di approssimazione mediante il quale, sia per una successione di numeri, sia per una successione di modalità qualitative, si può approssimare rapidamente il valore cercato della mediana in caso di indeterminazione, cosicchè quella ricerca, che sembrava aver solo utilità teorica, diviene applicabile, sia per le seriazioni, che per le serie cicliche, a casi che assumono importanza pratica nella statistica.

POSIZIONE DEL PROBLEMA

1. — I valori medi delle serie cicliche sono stati studiati da C. Gini ed L. Galvani nella monografia « *Di talune estensioni del concetto di media ai caratteri qualitativi* » (1) indagine completa e rigorosa dell'argomento, fondata sul principio matematico della conservazione delle leggi formali.

In particolare, i paragrafi 27-30 della monografia sono dedicati alla definizione ed alla ricerca della mediana di una serie ciclica.

Si definisce usualmente mediana delle quantità $a_1\ a_2 \ldots a_n$ il valore x^* che bipartisce la graduatoria degli a_i posti in ordine crescente (o decrescente) di grandezza.

Ora, in base a tale definizione, la determinazione della mediana può condurre ad una indeterminazione, nel senso che tutti i valori di un intervallo tra due quantità possono essere mediane delle n quantità considerate. È questo il caso, quando si ha una successione di quantità diverse

$$a_1, a_2, \ldots, a_n$$

con n pari, $n = 2k$; oppure quando le varie quantità si presentano ripetutamente come nello schema seguente

Quantità

$$x_1, x_2, \ldots, x_n$$

Frequenze

$$y_1, y_2, \ldots, y_n$$

(1) « Metron », 1927, 1°, pagg. 1-209.

con

$$\sum_{1}^{m} y_i = \sum_{m+1}^{n} y_i \quad \sum_{1}^{m-1} y_i < \sum_{m}^{n} y_i .$$

La stessa possibilità di indeterminazione si presenta quando la mediana viene definita come il valore che rende minima la somma

$$\sum_{1}^{n} | x - x_i | y_i .$$

Tale possibilità sussiste pertanto anche per le serie cicliche, poichè è precisamente in base alla proprietà di rendere minima la somma $\sum_{1}^{n} | x - x_i | y_i$ che il concetto di mediana è stato esteso da Gini e Galvani al caso delle serie cicliche.

Il Jackson ha proposto per una successione di quantità un procedimento per eliminare la indeterminazione più sopra descritta. Primo scopo di questa nota è di estendere tale procedimento al caso delle serie cicliche, eliminando anche per queste serie l'indeterminazione in parola.

2. — Secondo il procedimento del Jackson, il valore X mediana univoca del sistema di numeri $[a_i]$ è l'unica soluzione compresa fra a_k e a_{k+1} (supposti limiti dell'intervallo di indeterminazione) dell'equazione

$$(x - a_1)(x - a_2) \ldots (x - a_k) = (a_{k+1} - x) \ldots (a_n - x) \quad [1]$$ (1).

Il Jackson non ha dato alcuna indicazione sulla via da seguire per la risoluzione pratica dell'equazione del testo. Daremo, nei successivi paragrafi, un procedimento semplice e rapido atto a risolvere numericamente la [1], e sopra tutto le sue analoghe, che determineremo nei seguenti paragrafi, e che si presentano quando si applichi il procedimento a effettive serie statistiche (2).

---

(1) Si noti che l'equazione sopra scritta solo per il caso $n = 2$ conduce alla formula $X = \frac{a_1 + a_2}{2}$; per $n > 2$ appare così pienamente arbitrario l'assumere, come ordinariamente si fa, per mediana, nel caso di indeterminazione, la media aritmetica dei valori limiti dell'intervallo di indeterminazione.

(2) Notiamo anche, fin da ora, che gli ordinari metodi di approssimazione delle radici non conducono che in casi molto semplici, o molto particolari, a risultati soddisfacenti, se applicati nella attuale ricerca.

Determinazione univoca della mediana nelle seriazioni statistiche.

3. — Applicheremo prima il procedimento del Jackson al caso di una seriazione del tipo
Quantità

$$x_1, x_2, \ldots, x_n$$

Frequenze

$$y_1, y_2, \ldots, y_n$$

che rappresentano il caso che comunemente si riscontra in statistica.

Diciamo $x^* = x_k$ la mediana della seriazione secondo la ordinaria definizione, cioè quella (o quelle) quantità per la quale valgono insieme le due disuguaglianze :

$$\sum_1^{k-1} y_i < \sum_k^n y_i ; \sum_1^k y_i \geq \sum_{k+1}^n y_i .$$

Posto

$$S_p(x) = \sum_1^n | x - x_i |^p y_i$$

$S_p(x)$ è funzione continua con derivata $S_p'(x)$ continua e sempre crescente al crescere di x espressa, per $x_i \leq x \leq x_{i+1}$ della formula

$$\frac{1}{p} S'_p(x) = \sum_{1=j'}^{i} (x - x_{j'})^{p-1} y_{j'} - \sum_{i+1=j''}^{n} (x_{j''} - x)^{p-1} y_{j''}$$

Il minimo valore di $S_p(x)$ si avrà dunque per $x = x^*_p$, se $S'_p(x^*_p) = 0$.

Assumiamo ora come « mediana » della seriazione il valore $X = \lim_{p=1} x^*_p$.

È

$$\lim_{p=1} \frac{1}{p} S'_p(x_k) = \sum_1^{k-1} y_i - \sum_{k+1}^n y_i$$

$$\lim_{p=1} \frac{1}{p} S'_p(x_k + \varepsilon) = \sum_1^k y_i - \sum_{k+1}^n y_i$$

$$\lim_{p=1} \frac{1}{p} S'_p(x_k - \varepsilon) = \sum_1^{k-1} y_i - \sum_k^n y_i$$

ove ε indica un numero positivo convenientemente piccolo, del resto arbitrario.

Se supponiamo esistere una quantità x_k, per la quale

$$\sum_{1}^{k-1} y_i = \sum_{k+1}^{n} y_i$$

è $\lim_{p=1} S'_p(x_k) = 0$; quindi x_k è la mediana cercata, e si ha facilmente

$$X = \lim_{p=1} x^*_p = x_k = x^*$$

Altrimenti esisterà un primo valore di x, sia ancora, per comodità di scrittura, x_k, per il quale sono verificate insieme le disuguaglianze:

$$\sum_{1}^{k-1} y_i < \sum_{k}^{n} y_i \qquad \sum_{1}^{k} y_i > \sum_{k+1}^{n} y_i$$

oppure

$$\sum_{1}^{k-1} y_i < \sum_{k}^{n} y_i \qquad \sum_{1}^{k} y_i = \sum_{k+1}^{n} y_i \qquad (1)$$

Nel primo caso si ha $\lim_{p=1} S'_p(x_k + \varepsilon) > 0$, $\lim_{p=1} S'_p(x_k - \varepsilon) < 0$; quindi, ε essendo una quantità piccola ad arbitrio, $\lim_{p=1} S'_p(x_k) = 0$ e si vede ancora facilmente che $X = \lim_{p=1} x^*_p = x_k = x^*$.

Il valore di X coincide dunque con il valore x^* in tutti i casi in cui x^* è unicamente determinato.

Nel caso successivo si ha infatti: $\lim_{p=1} S_p'(x_k + \varepsilon) = 0$; vi è indeterminazione, e precisamente ogni valore x, dell'intervallo $x_k\, x_{k+1}$, (2) annulla il limite della derivata $S_p'(x)$ per $p \rightarrow 1$.

La stessa indeterminazione accade, ovviamente, per x^*.

---

(1) Si noti che le disuguaglianze:

$$\sum_{1}^{k-1} y_i > \sum_{k}^{n} y_i, \ \sum_{1}^{k} y_i < \sum_{k+1}^{n} y_i$$

sono incompatibili; e che, se per un valore di k, k', sussistono insieme le disuguaglianze

$$\sum_{1}^{k'-1} y_i = \sum_{k'}^{n} y_i, \ \sum_{1}^{k'} y_i > \sum_{k'+1}^{u} y_i$$

esiste anche un valore di k, k'', per il quale sussistono le

$$\sum_{1}^{k''-1} y_i < \sum_{k''}^{n} y_i, \ \sum_{1}^{k''} y_i \leq \sum_{k''}^{n} y_i.$$

(2) È da escludersi x_{k+1} perchè $\lim_{p=1} S'_p(x_{k+1}) \neq 0$.

Si ricorre allora all'artificio, già usato dal Jackson, fondato sullo sviluppo in serie esponenziale:

$$[f(x)]^p = e^{p \log [f(x)]} = 1 + p \, log [f(x)] + \frac{p^2}{2} \, log^2 [f(x)] + \dots$$

valevole per $p > 0$, $f(x) > 0$.

Si ottiene facilmente, per $x_k \leq x < x_{k+1}$,

$$\frac{1}{(p-1)\,p} S'_p(x) = \sum_1^k y_{j'} \, log \, (x - x_{j'}) - \sum_{k+1}^n y_{j''} \, log \, (x_{j''} - x) +$$

$$+ (p-1)\,\varphi(x, y, p) = log \frac{(x - x_1)^{y_1} \dots (x - x_k)^{y_k}}{(x_{k+1} - x)^{y_{k+1}} \dots (x_n - x)^{y_n}} + (p-1)\,\varphi(x, y, p)$$

ove φ è una funzione che rimane finita per x, y fissati in modo qualsiasi.

Se $\bar{x}_p$ è il valore che annulla il secondo membro della equazione ultima scritta, per un certo valore di p, e X il valore unicamente determinato che annulla il logaritmo che in quel secondo membro compare, si vede facilmente che $\lim_{p=1} \bar{x}_p = X$. X può dunque giustificatamente assumersi come mediana della seriazione nel caso che x^* sia indeterminato; il suo valore può desumersi dalla equazione:

$$log \frac{(x - x_1)^{y_1} \dots (x - x_k)^{y_k}}{(x_{k+1} - x)^{y_{k+1}} \dots (x_n - x)^{y_n}} = 0$$

cioè

$$f(x) = (x - x_1)^{y_1} \dots (x - x_k)^{y_k} - (x_{k+1} - x)^{y_{k+1}} \dots (x_n - x)^{y_n} = 0 \quad [1].$$

L'equazione che determina X si riconduce evidentemente a quella data dal Jackson quando tutti i pesi siano uguali; essa ha per grado il numero $\sum_1^k y_i = \sum_{k+1}^n y_i$ (1). Perciò la determinazione del vero valore di X non potrà farsi, in generale (2) che mediante la teoria dell'approssimazione. Nel paragrafo successivo si vedrà come si possono semplificare i calcoli spesso complicati che questa teoria richiede.

---

(1) Si noti che i pesi y_i possono in generale sempre ridursi a numeri interi.

(2) Per $n = 2$ non si può avere indeterminazione a meno che non sia $y_1 = y_2$ e allora si ricade nel caso esaminato dal Jackson e $X = \frac{x_1 + x_2}{2}$; per $n = 3$ si giunge ad un'equazione del grado 3° al meno; per $n \geqslant 4$ sempre ad equazioni di grado superiore al 4° (a meno che non si abbiano particolari gruppi di x_i, per i quali il grado si può abbassare di una o più unità).

Applicazioni numeriche per le seriazioni.

4. — Crediamo ora opportuno soffermarci sull'equazione che determina la mediana, per la risoluzione pratica della quale, come già è stato detto, occorre impiegare la teoria dell'approssimazione.

Il metodo di approssimazione da impiegarsi dovrà anzitutto permetterci di giungere alla determinazione approssimata della radice cercata senza sviluppare i prodotti dei binomi che l'equazione [1] contiene, sviluppo che riesce laboriosissimo quando i pesi non sono i primi numeri interi. Se si nota che in (x_k, x_{k+1}) la $f(x)$ ha in generale dei flessi e che comunque il calcolo della sua derivata riesce molto complicato — anche quando non si debbano sviluppare i binomi — dovremo ricorrere al metodo così detto delle secanti, metodo che qui brevemente riassumiamo: si congiunga il punto P_{x_k} della curva, di ascissa x_k, con il punto $P_{x_{k+1}}$ di ascissa x_{k+1}, e si determini il valore x^*_1 in cui la retta P_{x_k}, $P_{x_{k+1}}$, taglia l'asse x; x^*_1 dà un primo valore approssimato della radice cercata; detto $P_{x^*_1}$ il punto sulla curva la cui ascissa è x^*_1, si congiunga $P_{x^*_1}$ con P_{x_k} o con $P_{x_{k+1}}$ secondo che $f(x^*_1) > 0$ oppure $f(x^*_1) < 0$; si determini il valore x^*_2 dell'ascissa del punto di incontro della retta $P_{x^*_1} P_{x_k}$ $(P_{x^*_1} P_{x_{k+1}})$ con l'asse x, valore che dà una seconda e migliore approssimazione della radice cercata (Fig. 1).

L'iterazione del procedimento conduce ad un'approssimazione quanto si voglia spinta del valore vero della radice.

Il metodo ora descritto è molto lento, e del resto potrà essere utilmente impiegato solo quando i pesi y sono ancora relativamente piccoli. Un'applicazione si troverà nel paragrafo dedicato alle serie cicliche.

5. — Descriveremo ora un rapido procedimento di approssimazione, valevole per pesi y qualunque; si vedrà come l'approssimazione delle radici della [1] possa spingersi a qualsivoglia cifra decimale mediante successive equazioni di 2° grado.

Invece di risolvere la [1] possiamo risolvere l'equazione seguente

$$log\, K(x) = log \frac{(x - x_1) \ldots (x - x_k)}{(x_{k+1} - x) \ldots (x_n - x)} = 0$$

ove il logaritmo è assunto in base e.

Sviluppando in serie di potenze i fattori dell'equazione precedente, secondo le formule

$$x_0 = \frac{x_k + x_{k+1}}{2}$$

$$log\,(x - x_i) = log\,[(x_0 - x_i) + (x - x_0)] = log\left[1 + \frac{x_0 - x_i}{x - x_0}\right] +$$

$$+ log\,(x_0 - x_i) = log\,(x_0 - x_i) + \frac{x - x_0}{x_0 - x_i} - \frac{(x - x_0)^2}{2\,(x_0 - x_i)^2} + \ldots$$

$$i = 1, \ldots, k$$

$$log\,(x_j - x) = log\,[(x_j - x_0) + (x - x_0)] = log\left[1 - \frac{x - x_0}{x_j - x_0}\right] +$$

$$+ log\,(x_j - x_0) = log\,(x_j - x_0) - \frac{x - x_0}{x_j - x_0} - \frac{(x - x_0)^2}{2\,(x_j - x_0)^2} - \ldots$$

$$j = k + 1, \ldots, n$$

e fermando gli sviluppi al 2° termine, si ottiene la seguente equazione di 2° grado in $(x - x_0)$:

$$\frac{1}{2}\,(x - x_0)^2\left[\sum_{k+1}^{n}\frac{y_j}{(x_j - x_0)^2} - \sum_{1}^{k}\frac{y_i}{(x_0 - x_i)^2}\right] + (x - x_0)\left[\sum_{k+1}^{k}\frac{y_j}{x_j - x_0} +\right.$$

$$\left. + \sum_{1}^{k}\frac{y_i}{x_0 - x_i}\right] + \left[\sum_{1}^{k} y_i\, log\,(x_0 - x_i) - \sum_{k+1}^{n} y_j\, log\,(x_j - x_0)\right] = 0 \quad [2]\ (1)$$

Se con

$$\varepsilon_l = x_l - x_0;\ l = 1 \ldots n$$

indichiamo gli scostamenti algebrici da x_0 delle successive modalità $x_1 \ldots x_n$, e con z lo scostamento approssimato da x_0 della radice cercata x^*, l'equazione può anche scriversi

$$\frac{1}{2}\,z^2 \sum_{1}^{n}\frac{y_l\,[s\,g\,n\,.\,\varepsilon_l]}{\varepsilon_l^{\,2}} + z\sum_{1}^{n}\frac{y_l}{|\,\varepsilon_l\,|} - \sum_{1}^{n} y_l\,[s\,g\,n\,.\,\varepsilon_l]\; log\,|\,\varepsilon_l\,| = 0 \qquad [a]$$

la radice x_0 di questa equazione fornisce un primo valore x' approssimato di x^*

$$x' = x_0 + (x' - x_0) = x_0 + z_0$$

(1) Se $x^* = x_0$, deve essere $\Sigma\, y_i\, l\,g\,(x_0 - x_i) - \Sigma\, y_j\, l\,g\,(x_j - x_0) = 0$; cioè esser nullo il 3° coefficiente della [2] e viceversa. Quindi, se il 3° coefficiente risulta 0 è $x^* = x_0$ e il procedimento à qui termine; altrimenti è $x^* \neq x_0$, e gli sviluppi impiegati hanno sempre significato.

Questa prima approssimazione non è molto esatta a motivo del fatto, ben noto, che gli sviluppi in serie di *log* $(1 \pm x)$ sono lentamente convergenti per valori di x non molto vicini a 0. Si può ovviare a questa difficoltà iterando il procedimento nel modo seguente: si ponga

$$log\,(x - x_i) = log\left[1 + \frac{x - x'}{x' - x_i}\right] + log\,(x' - x_i)$$

$$log\,(x_j - x) = log\left[1 - \frac{x - x'}{x_j - x'}\right] + log\,(x_j - x')$$

$$\frac{x - x'}{x' - x_i}, \frac{x - x'}{x_j - x'}$$

sono allora in modulo minori di 1 (1) e abbastanza vicini a 0, perchè gli sviluppi in serie dedotti da essi secondo le formule già date precedentemente, siano rapidamente convergenti.

---

(1) Ciò implica che siano contemporaneamente verificate le condizioni

$$-1 < \frac{x^* - x'}{x' - x_i} < 1\,;\; -1 < \frac{x^* - x'}{x_j - x'} < 1 \qquad [*]$$

ove con x^* si è indicato la vera radice della equazione $log\,K(x) = 0$.

Valutiamo l'errore che si commette trascurando i resti degli sviluppi in serie dei termini

$$y_i\,log\left(1 + \frac{x - x_0}{x_0 - x_i}\right); - y_i\,log\left(1 - \frac{x - x_0}{x_j - x_0}\right)$$

a partire dal terzo termine incluso in poi.

Tali resti sono rispettivamente

$$\frac{y_i}{3}\left(\frac{1}{\frac{x_0 - x_i}{x - x_0} + \Theta_i}\right)^3 ; \frac{y_j}{3}\left(\frac{1}{\frac{x_j - x_0}{x - x_0} + \Theta_j}\right)^3$$

con $|\Theta_i| < 1$, $|\Theta_j| < 1$.

Posto

$$log\;K(x) = f(x) + \Delta(x)$$

ove $f(x)$ è l'equazione [2], $\Delta(x)$ la somma dei resti ora determinati, sarà

$$log\;K(x') = f(x') + \Delta(x') = \Delta(x')$$

e quindi secondochè $\Delta(x') \gtrless 0$ è anche $x' \gtrless x$; inoltre è

$$\Delta(x') = \Sigma\,\frac{y_i}{3}\left(\frac{1}{\frac{x_0 - x_i}{x' - x_0} + \Theta_i}\right)^3 + \Sigma\,\frac{y_j}{3}\left(\frac{1}{\frac{x_j - x_0}{x' - x_0} + \Theta_j}\right)^3$$

Se ora supponiamo $x' > x_0$, i denominatori dei cubi che compaiono nelle sommatorie sono > 0 e quindi è $\Delta(x') > 0$.

Sviluppando in serie, e fermando gli sviluppi al 1° termine, si ottiene un'equazione della stessa forma della [1], e precisamente ancora

$$\frac{1}{2} w^2 \sum_1^n \frac{y_l [sgn . \eta_l]}{\eta_l^2} + w \sum_1^n \frac{y_l}{|\eta_l|} - \sum_1^n y_l [sgn . \eta_l] \log |\eta_l| = 0 \quad [b]$$

ove w è lo scostamento approssimato da x' della radice cercata x^*, e η_l gli scostamenti esatti da x' delle successive modalità $x_1 \ldots x_n$.

La soluzione w_0 dell'equazione ultima scritta è la correzione da apportarsi a x' per ottenere un secondo miglior valore approssimato x'' di x^*, dato dalla formula

$$x'' = x' + (x'' - x') = x' + w_0;$$

x'' dà un'approssimazione molto soddisfacente di x, come si vedrà nel paragrafo seguente.

6. — Sia data la seriazione:

Modalità	1	2	2,25	4	8	12
Frequenze	120	15	18	30	100	83

Si vede facilmente che ogni valore compreso fra 4 e 8 è mediana secondo l'usuale definizione.

La mediana risulta poi determinata in modo univoco dalla unica soluzione compresa fra 4 e 8 dell'equazione

$$\frac{(x-1)^{120}(x-2)^{15}(x-2{,}25)^{18}(x-4)^{30}}{(x-8)^{100}(x-12)^{83}} = 1$$

La media aritmetica dei limiti dell'intervallo di indeterminazione è 6.

L'equazione in z ottenuta mediante la formula (a) risulta essere

$$6{,}394\, z^2, + 111{,}633\, z + 40{,}363 = 0$$

La sua unica radice compresa in $(-2, +2)$, $z_0 = -0{,}369$, dà la correzione da eseguire su x_0 per avere il primo valore approssimato cercato,

$$x' = x_0 + z_0 = 5{,}631$$

Se invece supponiamo $x' < x_0$, è $\Delta(x') < 0$.

Dunque è sempre o: $x' > x_0$ e contemporaneamente $x' > x$

oppure: $x' < x_0$ e contemporaneamente $x' < x$.

In ambedue i casi si può facilmente vedere che vale sempre il sistema di diseguaglianze [*] e quindi gli sviluppi del testo sono sempre convergenti.

Iterando il procedimento secondo la formula (*b*) si ottiene l'equazione in w

$$0{,}1397\, w^2 + 109{,}0047\, w - 0{,}0902 = 0$$

e per secondo valore approssimato

$$x'' = x' + x_0 = 5{,}631 + 0{,}0082 = 5{,}6392$$

Ora è

$$f\,(5{,}6392) < 0\,;\; f\,(5{,}6393) > 0$$

quindi x'' approssima x^* a meno di 1/10000 per difetto.

Estensione alle serie cicliche.

7. — In accordo al principio di conservazione delle leggi formali, si definisce mediana della serie ciclica a modalità equispaziate (al qual caso ci si può sempre ridurre):
modalità

$$X_{t-k+1}, \ldots, X_{t-1}, X_t, X_{t+1}, \ldots X_{t+k}$$

frequenze

$$y_{t-k+1}, \ldots, y_{t+1}, y_t, y_{t+1}, \ldots, y_{t+k}$$

la modalità X (o le modalità, effettive o di conto) ove la funzione

$$\Theta\,(X) = \Sigma \mid X X_i \mid y_i$$

ha un minimo assoluto.

Con il simbolo $\mid X X_i \mid$ si vuole indicare il valore assoluto della diversità base da X a X_i nel senso di Gini e Galvani (1).

La funzione $\Theta\,(X)$ è continua, ma non altrettanto accade per la sua derivata, che presenta $n - 1$ discontinuità di seconda specie.

Non si possono quindi caratterizzare i minimi della $\Theta\,(X)$ mediante lo studio delle sue successive derivate; occorre ridursi all'esame dell'effettivo comportamento della $\Theta\,(X)$ nei vari punti modalità del ciclo.

Si può dimostrare che la mediana cercata può non essere unica, ma comunque cade sempre in una modalità effettiva, salvo

---

(1) Loc. cit. pag. 29.

il caso in cui essa è una qualunque delle modalità, effettive o di conto, di un intervallo tra due modalità effettive consecutive.

È facile convincersi che ciò accade quando la rappresentazione grafica della $\Theta(x)$ ha l'andamento descritto nella figura 2, e quindi la sua derivata $\Theta(x)$ quello della figura 3 (1).

La derivata della $\Theta(x)$ ha quindi, nel caso di indeterminazione, il segmento corrispondente all'intervallo $x_t\, x_{t+1}$ di indeterminazione, coincidente con un segmento dell'asse X.

Assunta l'origine X_t arbitraria, studiamo ora la funzione

$$\Theta_p(X) = \Sigma \mid X X_i \mid^p y_i$$

La funzione è continua, e la sua forma è, nei successivi intervalli dell'asse X, la seguente:

....

per X variabile da X_{t-1} a X_t

$$(-x)^p y_t + (1-x)^p y_{t+1} + \dots +$$
$$+ (k - x - 1)^p y_{t+k-1} + (k+x)^p y_{t+k} + \dots + (x+1)^p y_{t-1}$$

per X variabile da X_t a X_{t+1}

$$(1-x)^p y_{t+1} + \dots + (k-x)^p y_{t+k} +$$
$$+ (k+x-1)^p y_{t-k+1} + \dots + x^p y_t$$

per X variabile da X_{t+1} a X_{t+2}

$$(2-x)^p y_{t+2} + \dots + (k-x+1)^p y_{t-k+1} +$$
$$+ (k+x-2)^p y_{t-k+2} + \dots + (x-1)^p y_{t+1}$$

....

ove x indica lo scostamento di X da X_t (2).

---

(1) La serie impiegata per il grafico è la seguente:

S	O	N	E
1	2	2	1

assunta come origine la modalità O, si sono calcolati i valori di $\Theta(x)$, $\Theta'(x)$, ottenendo i seguenti risultati, ove x indica lo scostamento di X da O:

per X *in*	$\Theta(X)$	$\Theta'(X)$
$S\ O$	$-2x+5$	-2
$O\ N$	5	0
$N\ E$	$2x+3$	2
$E\ S$	7	0

(2) Le funzioni $\Theta_p(X)$, $\Theta(X)$, ecc., sono valutabili numericamente solo quando si ricorra agli scostamenti x da un'origine fissa. Perciò non si

La forma della derivata di $\Theta_p(x)$, che indicheremo con $\Theta'_p(x)$ è allora, nei successivi intervalli del ciclo, la seguente:

per X variabile da X_{t-k} a X_{t-k+1}

$$p \{ -(-k-x+1)^{p-1} y_{t-k+1} - \ldots - (-x)^{p-1} y_t + \\ + (2k-1+x)^{p-1} y_{t-1} + \ldots + (k+x)^{p-1} y_{t+k} \}$$

. . . .

per X variabile da X_{t-1} a X_t

$$p \{ -(-x)^{p-1} y_t - \ldots - (k-1-x)^{p-1} y_{t+k-1} + \\ + (k+x)^{p-1} y_{t+k} + \ldots + (x+1)^{p-1} y_{t-1} \}$$

per X variabile da X_t a X_{t+1}

$$p \{ -(1-x)^{p-1} y_{t+1} - \ldots - (k-x)^{p-1} y_{t+k} + \\ + (k+x-1)^{p-1} y_{t-k+1} + \ldots + x^{p-1} y_t \}$$

per X variabile da X_{t+1} a X_{t+2}

$$p \{ -(2-x)^{p-1} y_{t+2} - \ldots - (k-x+1)^{p-1} y_{t-k+1} + \\ + (x+k-2)^{p-1} y_{t-k+2} + \ldots + (x-1)^{p-1} y_{t+1} \}$$

. . . .

per X variabile da X_{t+k-1} a X_{t+k}

$$p \{ -(k-x)^{p-1} y_{t+k} - \ldots - (2k-1-x)^{p-1} y_{t+1} + \\ + x^{p-1} y_t + \ldots + (x-k+1)^{p-1} y_{t+k-1} \}$$

La funzione $\Theta'_p(x)$ ha dunque $n-1$ discontinuità di seconda specie: anzi, se si immagina l'ultima modalità congiunta ciclicamente con la prima, n discontinuità di seconda specie.

Immaginando di percorrere il ciclo nel verso antiorario, si vede facilmente che la $\Theta'_p(x)$ diminuisce, all'atto di attraversare le modalità

$$X_{t-k+1}, \ldots, X_{t-1}, X_t$$

il suo valore rispettivamente di

$$2\,p\,k^{p-1} y_{t+1}, \ldots, 2\,p\,k^{p-1} y_{t+k-1}, 2\,p\,k^{p-1} y_{t+k}$$

e all'atto di attraversare le modalità

$$X_{t+1}, \ldots, X_{t+k-1}, X_{t+k},$$

---

dà luogo ad alcuna ambiguità sostituendo, nel testo, alle $\Theta_p(X)$, $\Theta(X)$, ecc. le corrispondenti $\Theta_p(x)$, $\Theta(x)$, ecc., che denotano, in sostanza, le stesse funzioni.

il suo valore di

$$2\,p\,k^{p-1}\,y_{t-k+1},\ 2\,p\,k^{p-1}\,y_{t-1},\ (2\,p\,k^{p-1}\,y_t)$$

cioè l'intensità algebrica dei successivi salti è il peso della modalità opposta a quella che si attraversa moltiplicato per il coefficiente $-2\,p\,k^{p-1}$.

In ciascuno degli intervalli $X_j\,X_{j+1}$ consécutivi la $\Theta'_p(x)$ è funzione sempre crescente; infatti la sua derivata ordinaria calcolata, ad esempio, in $X_t\,X_{t+1}$ è data dalla formula:

$$\frac{d}{d\,x}\Theta'_p(x) = p\,(p-1)\left\{(1-x)^{p-2}\,y_{t+1} + \ldots + (k-x)^{p-2}\,y_{t+k} + \right.$$

$$\left. + (k+x-1)^{p-2}\,y_{t-k+1} + \ldots + x^{p-2}\,y_t\right\}$$

e quindi è sempre positiva.

8. — Il grafico della $\Theta'_p(x)$ (Fig. 4) offre una chiara rappresentazione delle proprietà ora viste (1).

Nell'interno di ognuno degli intervalli considerati la funzione $\Theta_p(x)$ ha al più un minimo relativo, e tale minimo, se esiste, è dato dal valore x^*_p che annulla la $\Theta'_p(x)$.

Supponiamo ora che vi sia indeterminazione nel calcolo della mediana mediante la funzione $\Theta(x)$ e che l'intervallo di indeterminazione sia per esempio $X_t\,X_{t+1}$. Vogliamo dimostrare che in tal caso la $\Theta_p(x)$ ammette, in $(X_t\,X_{t+1})$, qualunque sia il valore di p, un minimo; basterà dimostrare che l'equazione che si ottiene uguagliando la $\Theta'_p(x)$ a 0 nel corrispondente intervallo di variabilità per la x, $(0,1)$, ha sempre una e una sola radice reale compresa fra i limiti di quell'intervallo.

---

(1) La serie impiegata per la costruzione del grafico è la seguente:

S O N E
1 2 3 4

Assunta come origine la modalità *O*, si son calcolati i valori di $\Theta'(X)$, $\Theta'_3(X)$ nei quattro intervalli di variabilità per X, ottenendo i risultati seguenti:

per X *in*	$\Theta'(X)$	$\Theta'_3(X)$
S O	0	$3\,(24\,x + 14)$
O N	-4	$3\,(-4\,x^2 + 24\,x - 18)$
N E	0	$3\,(16\,x - 22)$
E S	4	$3\,(4\,x^2 + 32\,x + 42)$

Il fatto ora enunciato si ottiene facilmente, per i valori interi del grado $p - 1$, mediante l'applicazione del teorema di Sturm.

Si formi la serie di Sturm relativa all'equazione $\Theta'_p(x) = 0$: Si otterranno $p - 2$ funzioni razionali intere

$$f(x) = \Theta'_p(x), f_1(x) = \frac{d}{dx}\Theta''_p(x) \ldots, f_{p-2}(x)$$

Poichè la seconda di esse, $f_1(x)$ ha segno costante nell'intervallo (0,1) la serie di Sturm può essere limitata a quella funzione (1).

È

$$f(0) < 0, f_1(0) > 0 ; f(1) > 0, f_1(1) > 0 \qquad (2),$$

quindi nell'intervallo (0,1) si perde una sola variazione, e perciò l'equazione proposta ha una e una sola radice reale x^*_p in (0,1).

L'esistenza e l'unicità della radice delle successive equazioni $\Theta'_p(x) = 0$ per $p = n \ldots 3.2$ ci suggerisce che essa deve, per $p = 1$, tendere ad un limite determinato.

A noi interessano dunque ora i valori di p prossimi e maggiori di 1. Per conseguire il risultato enunciato, si impiega lo stesso artificio del paragrafo 2. Si sviluppano in serie esponenziale i termini della $\Theta'_p(x)$; si ottiene successivamente:

$$(1 - x)^{p-1} = 1 + (p-1)\, log\,(1-x) + (p-1)^2 \varphi_1(x, p)$$

....

(1) CESARO, *Analisi algebrica*, pag. 195.

(2) Se tutti i valori di X dell'intervallo $X_t\, X_{t+1}$, sono mediane (in senso lato o in senso stretto) la $\Theta'(X)$ è ivi come si è visto, un segmento dell'asse X, quindi deve essere nullo identicamente in $X_t\, X_{t+1}$ il coefficiente della x in $\Theta(x)$, che è funzione lineare di x.

Ma tale coefficiente (cfr. GINI e GALVANI, loc. cit., pagg. 85, 86) vale in $X_t\, X_{t+1}$, $\sum_{t-k+1}^{t} y_i - \sum_{t+1}^{t+k} y_i$; quindi deve essere $\sum_{t-k+1}^{t} y_i = \sum_{t+1}^{t+k} y_i$. Ciò posto, si noti che il calcolo di $f(0)$ e $f(1)$ conduce alle formule:

$$f(0) = \lim_{x \to 0+0} \Theta'_p(x) = -y_{t+1} - 2^{p-1}; y_{t+2} - \ldots - k^{p-1} y_{t+k} +$$

$$+ (k-1) k^{p-1} y_{t-k+1} + \ldots + y_{t-1} ; (1) = \lim_{x \to 1-0} \Theta'_p(x) = -y_{t+2} -$$

$$- \ldots - (k-1)^{p-1} y_{k+t} + k^{p-1} y_{t-k+1} + \ldots + y_t .$$

Tenendo conto della uguaglianza precedentemente stabilita fra i valori y_i si vede con facilità che $f(0) < 0$ $f(1) > 0$.

$$(k - x)^{p-1} = 1 + (p - 1)\, log\, (k - x) + (p - 1)^2\, \varphi_p\, (x, p)$$

. . . .

$$x^{p-1} = 1 + (p - 1)\, log\, x + (p - 1)^2\, \varphi_n\, (x, p)$$

$$\frac{1}{p\,(p - 1)}\, \Theta'_p\, (x) = log\, \frac{(k + x - 1)\, y_{t+1} \ldots (k - x)\, y_{t+k}}{(1 - x)\, y_{t+1} \ldots (k - x)\, y_{t+k}} +$$

$$+ (p - 1)\, \varphi\, (x, p, y)$$

ove $\varphi\, (x, p, y)$ rimane finita per valori fissati di x, y, qualsiansi.

Per un qualunque valore di p fissato, la frazione del secondo membro cresce quando x cresce da 0 a 1 : esisterà un valore unicamente determinato x^* in (0,1) in cui quella frazione vale 1 e il suo logaritmo 0.

Se p è sufficientemente vicino a 1,

$$\Theta_p\, (x^* + \varepsilon) > 0,\; \Theta'_p\, (x^* - \varepsilon) < 0\,;$$

allora esiste un valore x^*_p che annulla $\Theta'_p\, (x)$, e questo è certo compreso, per p sufficientemente vicino a 1, tra $x^* + \varepsilon$ e $x^* - \varepsilon$, come volevamo mostrare.

9. — Il fatto che x^*_p sia compreso, per p sufficientemente vicino a 1, fra $x^* + \varepsilon$ e $x^* - \varepsilon$, mostra che è: $\lim_{p=1} x^*_p = x^*$.

x^* è dunque il valore unicamente determinato a cui tende per $p = 1$, la successione dei valori che danno il minimo della $\Theta'_p\, (x)$ nell'intervallo d'indeterminazione. Sia X^* la modalità avente lo scostamento x^* dall'origine.

E' giustificato assumere X^* come modalità mediana della serie assegnata ? Dobbiamo dimostrare:

a) *che la successione delle modalità X^*_p che minimizzano le funzioni $\Theta_p\, (x)$ tende per $p = 1$ alla modalità $\overline{X}$ che minimizza la $\Theta\, (x)$ nel caso in cui non vi sia l'indeterminazione del tipo studiato :* cioè occorre dimostrare che $\lim_{p=1} x^*_p = \overline{x}$ per ogni valore $\overline{x}$ (e sappiamo che possono essercene più di uno, ma sempre in numero finito) che minimizza $\Theta\, (x)$;

b) *che i valori x^*_p, nel caso di indeterminazione, realizzano, almeno per p sufficientemente vicino a 1 il minimo (o uno dei minimi) assoluto della $\Theta_p\, (x)$.*

Supponiamo, a tale scopo, dapprima, che una modalità, necessariamente effettiva, X_{t+1} per esempio, realizzi il minimo assoluto della $\Theta\, (x)$.

La differenza $\Theta\, (X_j) - \Theta\, (X_{t+1})$ è allora positiva per $j \neq t + 1$

ed il limite inferiore di essa al variare di j è un numero positivo ∂.

Ora si ha:

$$\Theta_p(X_j) - \Theta_p(X_{t+1}) \geq \Theta(X_j) - \Theta_p(X_{t+1})$$

qualunque sia j, come si può facilmente dedurre dalla forma delle due *funzioni* $\Theta_p(X)$, $\Theta(X)$.

Di qui segue

$$\Theta_p(X_j) - \Theta_p(X_{t+1}) \geq [\Theta(X_j) - \Theta(X_{t+1})] - [\Theta_p(X_{t+1}) - \Theta(X_{t+1})] \geq$$
$$\geq \partial - [\Theta_p(X_{t+1}) - \Theta(X_{t+1})]$$

Ma

$$\Theta_p(X_{t+1}) - \Theta(X_{t+1}) = 2(2^{p-1} - 1)(y_{t+3} + y_{t+1}) +$$
$$+ \ldots + k\, y_{t-k+1}(k^{p-1} - 1)$$

e quindi possiamo sempre trovare un valore $p_1 > 1$ tale che per $1 < p \leq p_1$ sia

$$\Theta_p(X_{t+1}) - \Theta(X_{t+1}) \leq \frac{\partial}{2}$$

Per tali valori di p è

$$\Theta_p(X_j) - \Theta_p(X_{t+1}) \geq \partial - \frac{\partial}{2} > 0$$

cioè

$$\Theta_p(X_j) > \Theta_p(X_{t+1}),\ j \neq t + 1$$

il che dimostra che X_{t+1} è un minimo anche per la Θ_p (X), relativamente alle modalità effettive del ciclo.

Ciò posto, si noti che, non essendovi nel caso attuale indeterminazione per il minimo della $\Theta(X)$, la rappresentazione grafica di questa funzione deve, a sinistra e a destra di X_{t+1} esser costituita da due segmenti di coefficiente angolare rispettivamente negativo e positivo, che indicheremo con $-\omega$, $+\omega'$.

Ora la $\Theta_p(X)$ converge in ogni punto del ciclo alla $\Theta(X)$, come si vede esaminando le differenze $\Theta_p(X) - \Theta(X)$. Lo stesso accade per la $\Theta'_p(X)$, considerando separatamente ciascun segmento $X_j X_{j+1}$, in ogni punto dei quali essa converge alla $\Theta'(X)$. In particolare, nei due segmenti adiacenti a X_{t+1}, $(X_t X_{t+1})$, $(X_{t+1} X_{t+2})$, la $\Theta'_p(X)$ converge ai due segmenti paralleli all'asse x, di ordinata costante $-\omega$, $+\omega'$.

Si può allora determinare un valore $p = p_2 < p$ sufficientemente vicino a 1, tale che per $p_2 \geq p > 1$ le $\Theta'_p(X)$ siano tutte comprese in striscie rettangolari parallele all'asse x, di ordinate estreme $(-\omega \pm \varepsilon < 0)$, $(+\omega' \pm \varepsilon > 0)$ e di ascisse estreme $(-1, 0)$, $(0,1)$ rispettivamente. Allora la $\Theta_p(X)$, se p soddisfa alle limitazioni $p_2 \geq p > 1$, è sempre, in ogni punto dei seguenti $(X_t X_{t+1})$, $(X_{t+1} X_{t+2})$, rispettivamente decrescente e crescente, e quindi non può assumere, in nessuna modalità di conto dei due segmenti $(X_t X_{t+1})$, $(X_{t+1} X_{t+2})$, valori minori che nell'estremo comune ai due segmenti.

Se ora esistesse una modalità X^{**} (necessariamente di conto per quello che prima si è dimostrato) nel ciclo, appartenente ad un segmento $X_j X_{j+1}$, in cui la $\Theta_p(X)$ avesse un valore minore che in X_{t+1}, in essa dovrebbe annullarsi la $\Theta'_p(X)$. Ma si può sempre determinare un valore $p_3 < p_2$ di p così vicino ad 1, che per $p_3 \geq p > 1$, sia sempre, in $(X_j X_{j+1})$, $\Theta'_p(X) \neq 0$, purchè nello stesso segmento, sia $\Theta'(X) = k \neq 0$; il che esclude che la $\Theta_p(X)$ possa avere in X^{**} un minimo relativo e quindi anche assoluto.

Se invece, nello stesso segmento, è $\Theta'(X) = 0$, e se è Δ la differenza fra il valore costante T di $\Theta(X)$ in $(X_j X_{j+1})$, e il valore $\Theta(X_{t+1})$ si può sempre trovare ancora un valore $p_3 < p_2$ tale che per $p_3 \leq p < 1$ sia $\Theta_p(X) > T - \Delta$ in tutto $(X_j X_{j+1})$; il che porta ancora ad escludere che $\Theta_p(X^{**})$ possa essere minore di $\Theta_p(X_{t+1})$.

Se ne conclude che la $\Theta_p(X)$, per p sufficientemente vicino ad 1, realizza il suo minimo assoluto nella stessa modalità effettiva in cui realizza il suo minimo la $\Theta(X)$, il che è a dire, ritornando alle notazioni precedentemente introdotte,

$$\lim_{p=1} x_p^* = \bar{x}$$

Piccole modifiche al ragionamento ora fatto sono necessarie per giungere all'asserto nel caso che vi sia più di una modalità (ma sempre in numero finito) che realizzi il minimo assoluto di $\Theta(X)$.

In tal caso vi saranno due o più successioni di valori minimizzanti $[x^*_p]$, $[x^{*\prime}_p]$, . . . , per ciascuna delle quali avverrà che

$$\lim_{p=1} x^*_p = \bar{x}, \lim_{p=1} x^{*\prime}_p = \bar{x}', \ldots$$

ove $\bar{x}$ $\bar{x}'$, . . . sono i valori che minimizzano $\Theta(X)$. Così la *a*) è completamente dimostrata.

Analoghe considerazioni di continuità portano pure a concludere che, nel caso di indeterminazione, la funzione $\Theta_p(X)$ realizza effettivamente il suo minimo assoluto (o uno dei suoi minimi assoluti) nel valore x^*_p in cui $\Theta'_p(x) = 0$ per sufficientemente piccoli valori di p.

Basta ricordare che nel caso attuale la $\Theta(x)$ è costituita nell'interno di $X_t\, X_{t+1}$ dal segmento parallelo all'asse dell'ascisse, in cui cade la mediana, mentre i due segmenti adiacenti, devono necessariamente avere coefficienti angolari costanti di segno diverso.

È dunque completamente giustificato l'assumere come mediana, nel caso di indeterminazione, il valore x^* soluzione della equazione:

$$(k + x - 1)^{y_{t-k+1}} \dots x^{y_t} = (1 - x)^{y_t} \dots (k - x)^{y_{t+k}}$$

valore che, si noti, è indipendente dall'origine X_t fissata (1).

10. — Se n è dispari, $n = 2k + 1$, si vede con facilità che l'indeterminazione studiata non può mai presentarsi (2).

Si è dunque ricondotta l'indeterminazione nella ricerca della mediana di una serie ciclica a quella proveniente dalla definizione stessa della funzione $\Theta(X)$, che può ammettere per la sua stessa natura più minimi assoluti, in corrispondenza di modalità effettive in numero finito, modalità che possono sempre tutte essere considerate mediane in senso stretto della serie data.

Casi particolari e applicazioni numeriche.

11. — L'equazione che determina la mediana in modo univoco ha per le serie cicliche la forma

$$f(x) = (k + x - 1)^{y_{t-k+1}} \dots x^{y_t} - (1 - x)^{y_{t+1}} \dots (k - x)^{y_t} = 0. \quad [1']$$

o può sempre a questa forma ridursi cambiando convenientemente, quando occorra, l'origine, l'assunzione della quale era arbitraria.

---

(1) Infatti supponiamo di spostare l'origine, per es., da X_t a X_{t+1}; diciamo x' lo scostamento $(X\, X_{t+1})$. Il ragionamento ripetuto a partire dalla nuova origine, conduce alla equazione

$$(-x')^{y_{t+1}} (1 - x')^{y_{t+2}} \dots (k - x' - 1)^{y_{t+k}} = (k + x')^{y_{t-k-1}} \dots (x' + 1)^{y}$$

Si noti ora che x' è legata a x dalla relazione

$$x = (X\, X_t) = (X\, X_{t+1}) + 1 = x' + 1$$

sostituendo, nella equazione del testo, a x, il valore $x' + 1$, si ottiene l'equazione in x' ora stabilita. Il che giustifica l'affermazione fatta della indipendenza del valore x^* trovato, dalla origine fissata.

(2) Cfr. Gini e Galvani, loc. cit. pag. 90. Quanto è detto nel testo si deduce dalla forma dei coefficienti di x in $\Theta(X)$ negli intervalli consecutivi

$$X_{t+1}\, X'_{t+k+1},\ X'_{t+k+1}\, X_{t+1},\ X_{t+1}\, X'_{t+k+2}:$$

supponendo che nel 2° intervallo quel coefficiente sia nullo identicamente, gli altri due hanno sempre lo stesso segno, e quindi nessun punto dell'intervallo considerato può esser punto di minimo. L'estensione agli altri intervalli è ovvia.

Nei paragrafi precedenti già è stato dimostrato che quella equazione ha sempre una radice reale x^* nell'intervallo (0,1).

Descriveremo prima alcuni casi così semplici da non richiedere l'impiego della teoria dell'approssimazione; daremo loro più il significato di conferma alla parte teorica, svolta nei paragrafi precedenti, che di vera applicazione numerica.

Sia data la serie

N	*NO*	*O*	*SO*	*S*	*SW*	*W*	*NW*
1	1	0	0	0	0	0	0

e assumiamo come origine la modalità *N*.

È evidente che in questo caso la modalità mediana ha scostamento $\frac{1}{2}$ da *N* e pertanto essa è la modalità di conto *N NO*.

Vediamo se allo stesso risultato si giunge per mezzo della equazione (1): essa si riduce, in questo caso, alla seguente:

$$x = 1 - x$$

e quindi $x = \frac{1}{2}$, come si era previsto.

Si può anche notare che nel caso attuale è costantemente $x = \frac{1}{2}$ il valore che minimizza le funzioni $\Theta_p(X)$. E pertanto coincidono in *N NO* mediana e media aritmetica nel senso $\mu_2 = min (1)$ della serie assegnata.

Passiamo ad un secondo caso, ugualmente semplice:

N	*NO*	*O*	*SO*	*S*	*SW*	*W*	*NW*
2	2	1	1	0	0	1	1

Si vede facilmente che vi è indeterminazione nell'intervallo (*N*, *NO*) i cui punti modalità sono tutti mediane della serie data; accade anche qui lo stesso fatto descritto nel caso precedente; e precisamente che il valore che minimizza la $\Theta(X)$ minimizza anche le $\Theta_p(X)$. L'equazione da risolvere è adesso:

$$(2 + x)(1 + x)x^2 - (1 - x)^2(2 - x)(3 - x) = 0$$

che, posta sotto forma normale, vale:

$$10x^3 - 5x^2 + 17x - 6 = 0$$

Di questa è unica radice reale in (0,1), il valore $X = \frac{1}{2}$ valore che minimizza per esempio anche la $\Theta_2(X)$.

Si ha infatti :

$$\Theta_2(X) = (1 - x)^2 . 2 + (2 - x)^2 . 1 + (3 - x)^2 . 1 + $$
$$+ (x + 2)^2 . 1 + (x + 1)^2 . 1 + x^2 . 2 = 8 x^2 - 8 x + 14$$
$$\Theta'_2(X) = 16 x - 8; \Theta''_2(X) = 16 > 0$$

e quindi $x = \frac{1}{2}$ minimizza anche $\Theta_2(X)$.

Gli esempi ora portati, autorizzano a generalizzare quanto ora si è visto ad una serie ciclica del tipo

$$\begin{matrix} (1) & (2) & \dots & (k) & (k+1) & \dots & (2k) \\ X_t & X_{t+1} & \dots & X_{t+k} & X_{t-k+1} & \dots & X_{t-1} \\ y_t & y_{t+1} & \dots & y_{t+k} & y_{t-k+1} & \dots & y_{t-1} \end{matrix}$$

con $n = 2k$, della quale si supponga che i pesi y_j possano disporsi in due graduatorie :

$$\begin{matrix} y_{t+j} \, , \, y_{t+j-1} \, , \, \dots \, , \, y_{t+j-k+1} \\ y_{t+j+1} \, , \, y_{t+j+2} \, , \, \dots \, , \, y_{t+j+k} \end{matrix} \qquad (1)$$

i cui termini siano nelle singole graduatorie, decrescenti o almeno non crescenti e in cui, inoltre, quelli di posto corrispondente siano eguali.

Per serie di tipo ora detto, vale la proprietà che la modalità di conto il cui scostamento è la media aritmetica fra gli scostamenti di X_{t+j}, X_{t+j+1}, $\left(\frac{1}{2}\right.$, se si assume X_{t+j} come origine) è la mediana della serie e insieme la media aritmetica nel senso $\mu_2 = min$.

12. — Supponiamo ora di dover risolvere casi in cui non si verifichi la simmetria prima descritta : dovremo in generale ricorrere alla teoria dell'approssimazione. Può talvolta essere impiegato con successo il metodo delle secanti già ricordato al paragrafo 4. Ad esempio, quando i pesi di cui sono affette le modalità X_t X_{t+1}, limiti dell'intervallo d'indeterminazione, sono alti rispetto ai pesi di cui sono affette le altre modalità e inoltre uguali o molto vicini

---

(1) Vedi GINI e GALVANI, loc. cit., μ_2 indica il momento secondo della serie assegnata.

(1) Si ricordi che $y_i = y_k$ quando $i \equiv k \pmod{n}$; e si confronti con l'alinea 5 pg. 95 di GINI e GALVANI, loc., cit.

fra loro; si noti infatti che i due termini dell'equazione [1'] contengono rispettivamente i fattori x^{y_t}, $(1 - x)^{y_{t+1}}$ potenze la cui base è minore di 1, mentre tutti gli altri fattori hanno basi maggiori e relativamente vicine ad 1.

Si intende allora che i fattori detti hanno la maggiore importanza nella determinazione dei valori dei due termini della [1']; e poichè si ha $x = 1 - x$ quando $x = 0{,}5$, il valore della mediana oscillerà, nelle varie serie soddisfacenti alle condizioni anzidette, attorno al valore 0,5; pertanto, sarà di grande utilità partire da questo valore nella ricerca della radice approssimata delle radici, procedendo poi con il metodo di approssimazione già indicato.

Sia data ad esempio la serie:

N	*NO*	*O*	*SO*	*S*	*SW*	*W*	*NW*
4	4	3	2	1	0	2	4

Assunta come origine la modalità *N*, la funzione $\Theta(x)$ definita nel paragrafo 3 vale:

Nell'intervallo (*NW*, *N*)	$-6x + 28$
nell'intervallo (*N*, *NO*)	28
nell'intervallo (*NO*,*O*)	$8x + 20$

e si vede facilmente, calcolandone il valore nelle successive modalità effettive *N* . . . *NW*, che essa raggiunge il suo valore minimo in tutti i punti dell'intervallo (*N*, *NO*).

Siamo dunque di fronte ad una indeterminazione del tipo studiato; inoltre i pesi di cui sono affette le modalità estreme dell'intervallo d'indeterminazione sono eguali, e alti rispetto a quelli delle altre modalità della serie: la radice dell'equazione che fornisce la mediana, dovrà dunque essere vicina a 0,5.

L'equazione da risolvere è:

$$f(x) = (2 + x)^2 (1 + x)^4 x^4 - (1 - x)^4 (2 - x)^3 (3 - x)^2 (4 - x) = 0.$$

Si ha

$$f(0{,}5) = -2{,}52,$$

mentre sappiamo che

$$f(0) < 0\,,\ f(1) > 0\,;$$

poichè il valore di $f(x)$ in 0,5 è negativo e molto piccolo, presumibilmente il valore di $f(x)$ in 0,6 sarà positivo. Infatti si ha $f(0{,}6) = 4{,}42$; segue che la radice cercata è certo compresa fra 0,5 e 0,6. Applicando il procedimento descritto al paragrafo 4 si ottengono successivamente i due valori approssimati

$$x' = 0{,}53\,;\ x'' \doteq 0{,}53.$$

È

$$f(0{,}53) = -1{,}55\,;\; f(0{,}54) = 0{,}06\,,$$

quindi il valore approssimato per difetto della radice a meno di un centesimo è 0,53, per eccesso 0,54.

In generale, come già è stato detto, il metodo delle secanti non conduce però a risultati soddisfacenti e per la lentezza dell'approssimazione e per la laboriosità dei calcoli per alti pesi.

13. — Applichiamo ora il procedimento di approssimazione, valevole anche per pesi aventi molte cifre intere o decimali, descritto al paragrafo 5.

Invece di risolvere la [1'], possiamo risolvere l'equazione seguente:

$$log\,\frac{(k+x-1)^{y_{t-k+1}}\ldots x^{y_t}}{(1-x)^{y_{t+1}}\ldots(k-x)^{y_{t+k}}} = 0 \qquad [1'']$$

ove il logaritmo è assunto in base e.

Si noti che si può sempre porre

$$log\,(a+x) = log\left[\left(a+\frac{1}{2}\right)+\left(x-\frac{1}{2}\right)\right] = log\left(a+\frac{1}{2}\right)+log\left[1+\frac{x+\frac{1}{2}}{a+\frac{1}{2}}\right]$$

$$log\,(a-x) = log\left[\left(a-\frac{1}{2}\right)-\left(x-\frac{1}{2}\right)\right] = log\left(a-\frac{1}{2}\right)+log\left[1-\frac{x-\frac{1}{2}}{a-\frac{1}{2}}\right]$$

$$log\,(x) = log\left[\frac{1}{2}+\left(x-\frac{1}{2}\right)\right] = log\left(\frac{1}{2}\right)+log\left[1+\frac{x-\frac{1}{2}}{\frac{1}{2}}\right]$$

eseguendo queste sostituzioni nella equazione [1'] sviluppando in serie i logaritmi dei vari termini, e fermando lo sviluppo al 2° termine si ottiene facilmente l'equazione di 2° grado in $\left(x-\frac{1}{2}\right)$, analoga alla [$a$] del paragrafo 5.

$$\frac{1}{2}\left(x-\frac{1}{2}\right)^2\sum_1^k\frac{1}{\left(i-\frac{1}{2}\right)^2}\,(y_{t+1}-y_{t-i+1})\;+$$

$$+\left(x-\frac{1}{2}\right)\sum_1^k \frac{1}{i-\frac{1}{2}}(y_{t-i+1}+y_{t+i})+$$

$$+\sum_1^k (y_{t-i+1}-y_{t+i})\, log\left(i-\frac{1}{2}\right)=0 \qquad [a']$$

i cui coefficienti sono di calcolo molto rapido.

La soluzione z_0 di questa equazione fornisce un primo valore approssimativo $x'=\frac{1}{2}+z_0$. Iterando il procedimento, si ottiene una equazione della stessa forma della [a'] ove però $\frac{1}{2}$ è sostituito dal valore di x' ora determinato e che può scriversi sotto la forma seguente:

$$\frac{1}{2}(x-x')^2\sum_1^k\left[\frac{y_{t+i}}{(i-x')^2}-\frac{y_{t-k+i}}{(k-i+x')^2}\right]+$$

$$+(x-x')\sum_1^k\left[\frac{y_{t-k+i}}{k-i+x'}+\frac{y_{t+i}}{i-x'}\right]+$$

$$+\sum_1^k\left[y_{t-k+i}\, log\,(k-i+x')-y_{t+i}\, log\,(i-x')\right]=0 \qquad [b']$$

Qualora si desideri un'approssimazione più spinta, basterà iterare ancora il procedimento.

14. — Il metodo dà anche per le serie cicliche ottimi risultati: riprendiamo l'esempio dato precedentemente, quale applicazione del metodo delle secanti. Sostituendo nella (a') alle costanti i valori numerici si ottiene rapidamente l'equazione.

$$0{,}181\left(x-\frac{1}{2}\right)^2-22{,}552\left(x-\frac{1}{2}\right)+0{,}847=0;\; x'-\frac{1}{2}=0{,}043$$

Non occorre nella risoluzione di questa prima equazione conservare nei calcoli molte cifre decimali: il valore x' va determinato soltanto per permettere, mediante l'artificio escogitato, di convergere rapidamente. La seconda equazione, nella quale si deve tener conto di molte cifre decimali per avere un risultato soddisfacente, fornisce per la correzione w_0, il valore $-0{,}006$ (1).

L'approssimazione a $\frac{1}{100}$ è assicurata, coincidendo la seconda

(1) Ottenuto conservando costantemente 4 cifre decimali nei calcoli.

cifra decimale del valore approssimato ora trovato 0,53 con quello (0,53 p. d, 0,54 p. e) già trovato per altra via.

Si vede che, anche per le serie cicliche, non vi sono difficoltà ad applicare il procedimento di approssimazione descritto quando i pesi sono molto alti, fatto che, come si è notato, rende in generale impossibile utilizzare gli ordinari metodi di approssimazione.

Anzi, il procedimento impiegato suggerisce, che il grado di approssimazione che si ottiene mediante esso migliora sensibilmente al crescere dei pesi dei quali sono affette le modalità della serie assegnata: osservazione, questa, che, come s'intende facilmente, vale anche per la ricerca della mediana in una seriazione.

Si abbia infatti la serie ciclica:

N	*O*	*S*	*E*
6001	3520	5751	3770

La determinazione della mediana conduce ad una indeterminazione, cadente nell'intervallo *OS*. Assunta la modalità *O* come origine, si ottiene, sostituendo nella (*a'*) alle costanti i valori numerici, l'equazione

$$3966{,}22\, z^2 + 25056 + 2450{,}33 = 0$$

$$z_0 = -0{,}0994;\ x' = \frac{1}{2} + z_0 = 0{,}4006$$

Iterando il procedimento, mediante la (*b'*) si ottiene l'equazione in w

$$3754{,}9064\, w^2 - 25023{,}1406\, w + 31{,}0537 = 0$$

$$w_0 = 0{,}0011;\ x'' = x' + w = 0{,}4017. \qquad (1)$$

È

$$f\,(0{,}4017) > 0\ ,\ f\,(0{,}4016) < 0$$

quindi x'' approssima il valor vero x^* a meno di $\frac{1}{10000}$ per eccesso; mentre nell'equazione data nel paragrafo precedente il valore x'' approssimava il valor vero x soltanto a meno di $\frac{1}{500}$.

(1) Ottenuto conservando costantemente 4 cifre decimali nei calcoli.

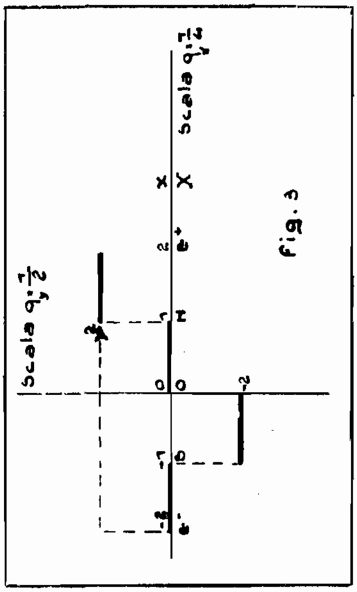


fig. 3

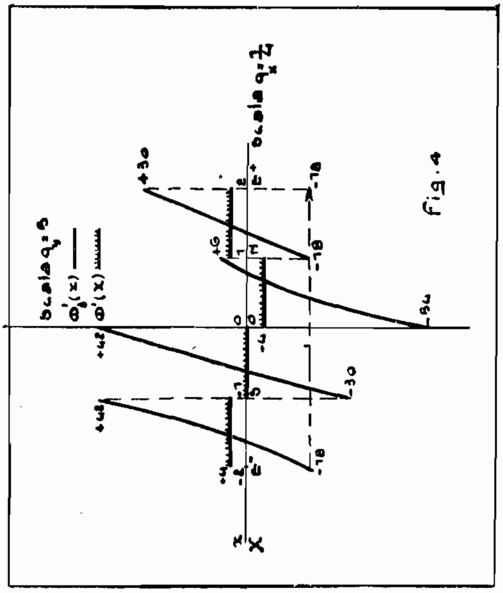


fig. 4

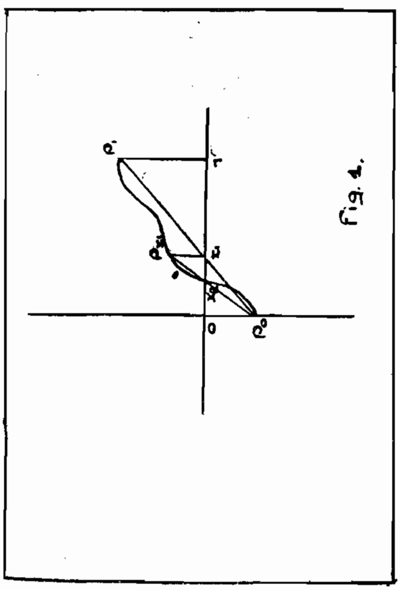


fig. 1

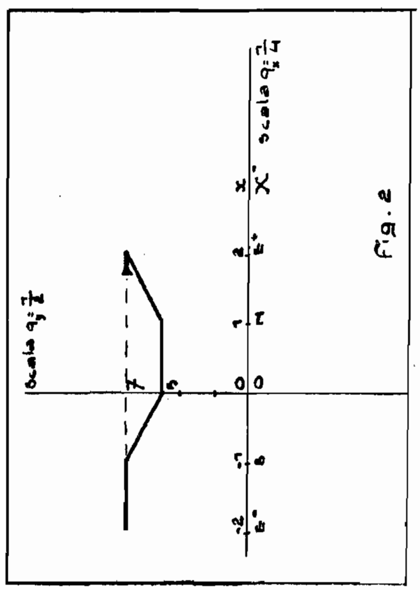


fig. 2

CURTIS BRUEN

Methods for the Combination of Observations: Modal Point or Most Lesser-Deviations, Median Loci or Least Deviations, Mean Loci or Least Squares, and Mid-Point of Least Range or Least Greatest-Deviation

A power-mean, as defined by Fechner (**1**) (1), is a value, relative to which the sum of the absolute deviations of the individual values, raised to a given power, is the least possible. The p^{th} -order power-mean of a set of observations, x_i ($i = 1, 2, 3, \ldots, n$), is that value, x, which makes the sum, $\Sigma \mid x_i - x \mid^p$, a minimum. The mode is the null-order power-mean, for, as shown by Foster (**2**), it is the limiting value of the power-mean as p approaches 0, or $\lim_{p \to 0} x =$ the mode. The median is the first-order power-mean (**1**). When defined simply as the origin from which the sum of the absolute deviations is a minimum, the median remains indeterminate, if, in the sequence, $x_1 \leqq x_2 \leqq \ldots \leqq x_n$, $n = 2k$ and $x_k \neq x_{k+1}$, being any value, x, belonging to the interval, $x_k \leqq x \leqq x_{k+1}$. But, as shown by Jackson (**3**), the median is always determinate, if taken as that power-mean which approaches a definite limit as p approaches 1, and its value, $\lim_{p \to 1} x$, is characterized in the otherwise indeterminate case by the general equation,

$$(x - x_1) \ldots (x - x_k) = (x_{k+1} - x) \ldots (x_n - x),$$

from which special algebraic expressions for sequences of any given even number of terms may be derived. The arithmetic mean is the second-order power-mean (**20, 1**). The succession of further power-means of higher orders (**1**) ultimately reaches the infinite-order power-mean, originally due to Laplace (**4**), which is the mid-point of the range (**12, 3**) : $\lim_{p \to \infty} x = (x_1 + x_n)/2$.

(1) The numbers printed in heavy type refer to the bibliography at the end of the article.

The concept of the power-mean may be generalized from its origin in sequences of direct observations so as to become applicable to sets of indirect observations or of implicit functional observations as the method of least power-sums of the absolute values of the deviations. With indirect observations, the observed values of a dependent variable, z_i, are at once some unknown function of the observed values of the independent variables, $x_i, y_i, \ldots, u_i, v_i$, and known functions of the unknown parameters, $a, b, \ldots, l, m$, so that the observation equations are

$$z_i = f(x_i, y_i, \ldots, u_i, v_i) \equiv f_i(a, b, \ldots, l, m),$$

and the sum of the absolute differences between the observed values of the dependent variable and the required function of the observed values of the independent variables, or the known functions of the unknown parameters, raised to the p^{th} power, as

$$\sum_{i=1}^{n} | z_i - [f(x_i, y_i, \ldots, u_i, v_i) \equiv f_i(a, b, \ldots, l, m)] |^p,$$

is to be rendered a minimum. With implicit functional observations, some unknown implicit function of the observed values of the variables, $x_i, y_i, \ldots, u_i, v_i$, which are known functions of the unknown parameters, $\alpha, \beta, \ldots, \pi, \rho$, is equal to zero, so that the observation equations are

$$\Phi(x_i, y_i, \ldots, u_i, v_i) \equiv \Phi_i(\alpha, \beta, \ldots, \pi, \rho) = 0,$$

and the p^{th} power-sum of the absolute values of the differences of the function of the observed values of the variables, or of the functions of the parameters, from zero, as

$$\sum_{i=1}^{n} | \Phi(x_i, y_i, \ldots, u_i, v_i) |^p, \quad \text{or} \quad \sum_{i=1}^{n} | \Phi_i(\alpha, \beta, \ldots, \pi, \rho) |^p,$$

is to be rendered a minimum.

In the case of indirect observations, when the functions with respect to the parameters in the observation equations are of the first degree, the problem of the determination of the parameters for a minimum sum of absolute deviations raised to a given power reduces to the method of least p^{th} powers for a set of n simultaneous linear equations in m unknowns, the existence and uniqueness of the solutions for which, when $n > m$, for a general exponent, $0 < p \leqq \infty$, has been treated by Jackson (**5**). The linear equation in slope-intercept form which is to be obtained through the combination of indirect observations, when set up as observation equations assumes the form, $y_i = a + b\,x_i + \ldots$, and the problem becomes to determine the parameters, a, b, . . . , so that the power-sums of the absolute values of the deviations of the observed values of the dependent variable relative to the functions of the parameters, as

$$\sum_{i=1}^{n} | y_i - (a + b\,x_i + \ldots) |^p,$$

will be rendered a minimum for a given exponent p. In the case of implicit functional observations, the linear equation, which is to be obtained through their combination, when in normal form, and set up as observation equations assumes the form, $x_i \cos \alpha + y_i \sin \alpha - \rho = 0$, and the problem becomes to determine the parameters, α, ρ, so that the power-sums of the deviations of these functions of the parameters relative to zero, as

$$\sum_{i=1}^{n} | x_i \cos \alpha + y_i \sin \alpha - \rho |^p,$$

will be rendered a minimum for a given exponent.

The methods of solution used for the determination of the values of the specific power-means, the mode, the median, the arithmetic mean, and the mid-point of the range, of a sequence of direct observations, can be extended directly to apply to the determination of the parameters in the equations obtained from sets of indirect observations which give observation equations linear with respect to the parameters, and of implicit functional observations which yield equations linear with respect to their unknowns, and the solution by each method will result in minimum sums of the absolute values of the deviations raised to the power of the corresponding power-mean.

Propaedeutic.

When the linear equation in x, y, . . in slope-intercept form, $y = a + b\,x + \ldots$, is set up as an observation equation, $y_i = a + b\,x_i + \ldots$, by the substitution of sets of concomitantly observed values, x_i, y_i, . . , for its unknowns, x, y, . . , it remains an equation of the first degree, though in a, b, . . , and in general form with a unit coefficient for its first unknown and a transposed constant term. But, when the linear equation in normal form, such as $x \cos \alpha + y \sin \alpha - \rho = 0$, is set up as an observation equation, $x_i \cos \alpha + y_i \sin \alpha - \rho = 0$, by the substitution of observed values, x_i, y_i, for its unknowns, x and y, it does not remain an algebraic equation, but is converted into a polar equation, not reducible to an equation of the first degree (though expressible as a quadratic function of the form, $x^2 + y^2 + \ldots - x_i x - y_i y - \ldots = 0$, which is the equation of a circle or sphere with center, $C\left(\frac{x_i}{2}, \frac{y_i}{2}, \ldots\right)$, and radius, $r = \sqrt{x_i^2 + y_i^2 + \ldots /2)}$.

The locus of a linear observation equation is a straight line or a plane. The locus of the observation equation, $y_i = a + b\, x_i$, is a straight line, the a-intercept of which in rectangular coordinates is located on the y-axis since it is equal to the observed value of the dependent variable, y_i, while the b-intercept is located on the x-axis and is equal to the observed value of the dependent variable divided by that of the independent variable, or y_i/x_i. The locus of the observation equation, $z_i = a + b\, x_i + c\, y_i + \ldots$, is a plane, the intercepts of which are $a = z_i$, $b = z_i/x_i$, $c = z_i/y_i, \ldots$

The locus of a polar observation equation may be regarded as a polar vectorial epicycloid or epispheroid generated respectively through the linear translation and addition of radial vectors from a point where the diameters and tangents of two circles or the diametral and tangent planes of three spheres are mutually perpendicular.

The locus of the polar observation equation, $x_i \cos \alpha + y_i + \sin \alpha = \rho$, the variables of which are the vectorial angle, α, and the radius vector, ρ, will be a polar vectorial epicycloid. Let XX' and YY' be a plane system of rectangular axes; let their intersection, O, be the origin, and the positive x-axis, OX, the prime direction, of a system of polar coordinates (ρ, α) (Fig. 1). Then $x_i \cos \alpha = r_{x_i}$ is the equation of a circle with its center on the x-axis at the point, $x_i/2$, and with radius, $x_i/2$; $y_i \sin \alpha = r_{y_i}$, the equation of a circle with its center on the y-axis at $y_i/2$, and radius, $y_i/2$; and $r_{x_i} + r_{y_i} = \rho$, the equation of the sum of the two component radius vectors at each vectorial angle. The resulting polar vectorial epicycloid is equivalent to a circle with the intercepts, x_i and y_i, on the respective axes, and passing through the origin.

The locus of the polar observation equation, $x_i \cos \alpha + y_i$

$\cos\beta + z_i \cos\gamma = \rho$, will be a polar vectorial epispheroid. Let the positive *x*-, *y*-, and *z*-axes, *OX*, *OY*, *OZ*, of the system of rectangular axes, *XX'*, *YY'*, *ZZ'*, be the prime directions from which the respective vectorial angles, α, β, γ, are measured; the position of a point $(\rho, \alpha, \beta, \gamma)$, is then located by its radius vector, ρ, and the angles its radius vector makes with the three rectangular axes, subject to the condition, $\cos^2\alpha + \cos^2\beta + \cos^2\gamma = 1$,

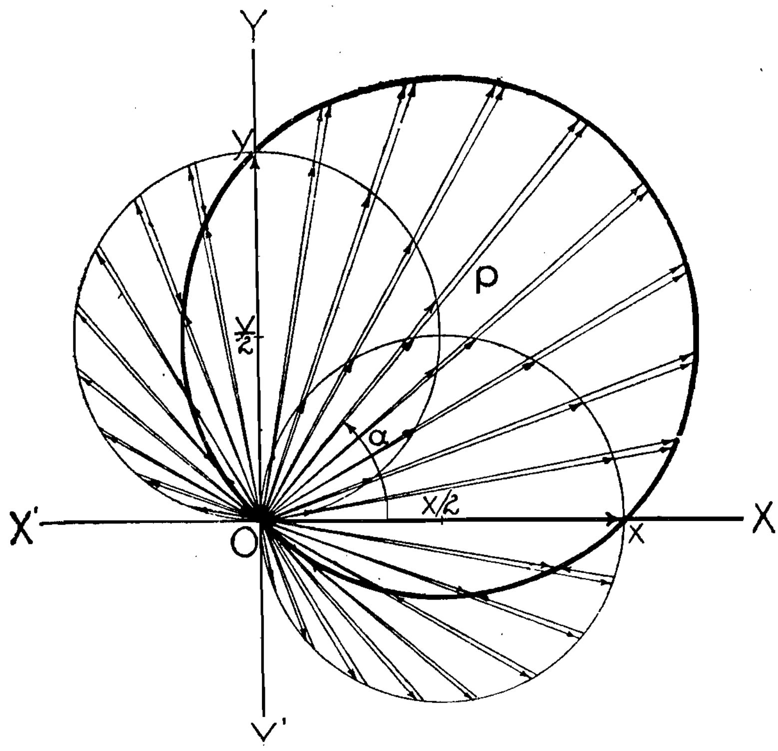


Fig. 1

whereby any two given vectorial angles determine the third. When their loci are not restricted to a plane, the equations, $x_i \cos\alpha = r_{x_i}$, $y_i \cos\beta = r_{y_i}$, and $z_i \cos\gamma = r_{z_i}$, represent spheres with their centers on the respective axes at the points and with the radii, $x_i/2$, $y_i/2$, $z_i/2$ (Fig. 2). The sum of the three component radius vectors, r_{x_i}, r_{y_i}, r_{z_i}, at each combination of the vectorial angles,

$$0^0 \leqq \alpha, \beta, \gamma \; (= \cos^{-1}\sqrt{1 - \cos^2\alpha - \cos^2\beta}) \leqq 180^0,$$

defines the polar vectorial epispheroid which is equivalent to a sphere with the intercepts, x_i, y_i, z_i, on the respective axes, and passing through the origin (1).

Regarding the loci of polar observation equations from the aspect of their particular mode of generation as polar vectorial epicycloids or epispheroids rather than from the aspect of their

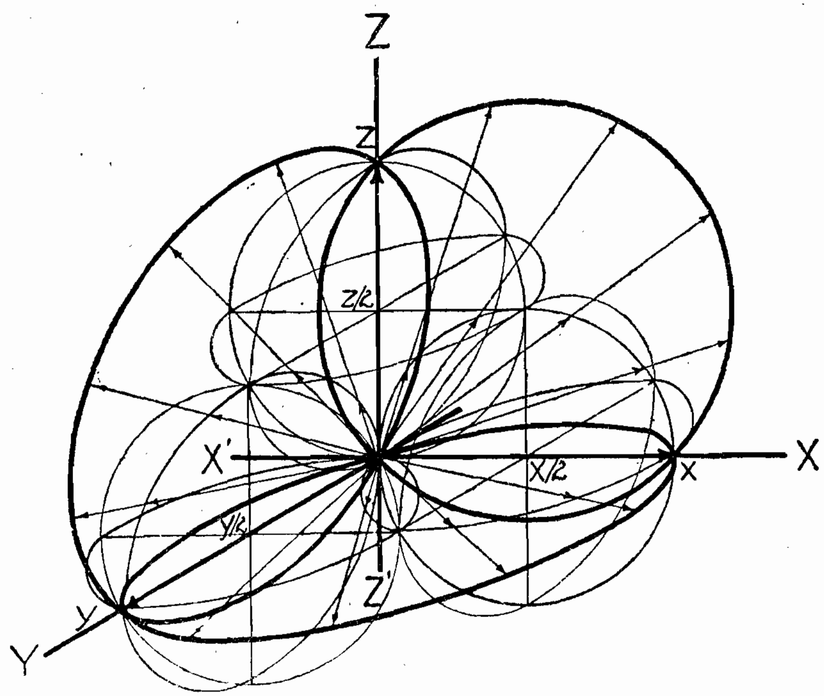


Fig. 2

general geometrical form as circles or spheres serves to emphasize the mode of variation of their radius vectors in its relation to the variation of the several component radius vectors with the vectorial angles.

The distance from the locus of a polar observation equation to an arbitrary polar parameter point, P (ρ, α, . .), is the difference

(1) The extension to polar vectorial epihyperspheroids would involve considerations of an n-dimensional coordinate system and goniometry comparable to J. McMahon's, *Hyperspherical Goniometry*, . « Biometrika » 15: 173-208 (1923).

between the radius vector of the polar observation equation, $\rho_{(x_i, y_i, \ldots)}$, and the assumed radius vector, ρ, at the assumed vectorial angle, $\alpha, \ldots$, or $(\rho_{(x_i, y_i, \ldots)} - \rho)_{\alpha, \ldots}$. The distance from the locus of a linear equation in normal form defined by an arbitrary set of polar parameters, $\rho, \alpha, \ldots$, such as the line, $x \cos \alpha + y \sin \alpha - \rho = 0$, to a given observation point, $P_i (x_i, y_i, \ldots)$, is found by substituting observed values for the unknowns in the equation, as $d = x_i \cos \alpha + y_i \sin \alpha - \rho$. Since the left member of a polar observation equation is equal to its radius vector, as $x_i \cos \alpha + y_i \sin \alpha = \rho_{(x_i, y_i)}$, the distance from the locus of a polar observation equation to an arbitrary polar parameter point is equal to the distance from the line or plane defined by the arbitrary set of parameters represented by the parameter point to the observation point represented by the equation. The deviation attributable to an implicit linear functional observation is accordingly the perpendicular distance of its observation point from the line or plane to be determined.

The distance from the locus of a linear equation in general form, $Ax + By + \ldots + M = 0$, to a given point, $P_i (x_i, y_i, \ldots)$, becomes, by changing the equation to normal form, $d = \frac{A x_i + B y_i + \ldots \pm M}{\mp \sqrt{A^2 + B^2 + \ldots}}$. The distance from the locus of a linear equation in slope-intercept form, $y = a + b x + \ldots$, to an observation point, $P (x_i, y_i, \ldots)$, is then $\frac{- b x_i + y_i - \ldots - a}{\sqrt{b^2 + 1^2 + \ldots}}$. Multiplication of this perpendicular distance by the secant of the angle between the normals to the locus and the vertical, that is, by $\sqrt{b^2 + 1^2 + \ldots}/1$, transforms the perpendicular deviation into deviation with respect to the dependent variable, or $\delta_y = y_i - (a + b x_i + \ldots)$. This deviation in the dependent variable is the distance from the locus of a linear observation equation, $y_i = a + b x_i + \ldots$, to an arbitrary rectangular coordinates parameter point, $P(a, b, \ldots)$, and is equal to the difference between the first parameter of the observation equation, $a_{(y_i, x_i, \ldots)}$, and the assumed value of the first parameter, a, at the assumed values of the succeeding parameters, $b, \ldots$, as $(a_{(y_i, x_i, \ldots)} - a)_{b, \ldots}$. Thus the deviation attributable to an indirect observation giving an observation equation linear with respect to its parameters is the distance in the direction of the axis of the dependent variable of an observation point, P_i $(y_i. x_i, ..)$, from the curve to be determined.

1. Modal Point or Most Lesser-Deviations.

The mode is that value about which the individual values collect most densely (1). In the absence of analytic criteria, that parameter point about which the greatest number of the loci of a set of linear observation equations traversed by a line parallel to the axis of the dependent variable, or of polar observation equations traversed by a polar ray with the origin preferentially excluded, sensibly tend to gather, cluster most closely, and concentrate in greatest density, so that they most definitely establish and determine a mode, is to be selected as constituting the empirical modal point of the set of observation equations. The combination of observations by the method of the location of the mode or the modal point may be designated the method of most lesser-deviations, for it results in the greatest possible numerical preponderance of deviations of magnitudes approaching zero.

2. Median Loci or Least Deviations.

Given a set of observation equations, $w_i y = x_i$, in which the coefficients of y are all positive, if of the several values of $y = x_i/w_i$, arranged in the order of magnitude, as $x_1/w_1 \leqq x_2/w_2 \leqq \ldots \leqq x_n/w_n$, that value, $y = x_m/w_m$, is selected, which is derived from the m^{th} equation which is indentified by means of the inequalities, $\sum_{i=1}^{m} w_i > \sum_{m+1}^{n} w_i$, $\sum_{i=1}^{m-1} w_i < \sum_{m}^{n} w_i$, this value renders the sum of the absolute deviations of the other values a minimum (1). This

(1) For if this value of y be increased by the increment Δy, the sum of the absolute deviations, $\sum_{i=1}^{n} | x_i - w_i y |$, becomes $\sum_{i=1}^{n} | x_i - w_i (y + \Delta y)|$, and, since the sum of the absolute values of the negative deviations and of the zero deviation is increased by the quantity, $\Delta y \sum_{i=1}^{m} w_i$, and that of the positive deviations diminished by the quantity, $\Delta y \sum_{m+1}^{n} w_i$, the sum of the absolute deviations is increased by a positive quantity, $\sum_{i=1}^{n} w_i \Delta y =$

procedure constitutes Laplace's method of situation (**6**), which, despite its priority and greater generality, may well be assimilated to Fechner's concept of the median (**1**), to which it reduces when the weights of the several observations are equal, by interpreting it as the method of the weighted median.

The weighted median renders the sum of the weighted absolute deviations a minimum. Given a series of values of $y = x_i/w_i$ arranged in the order of magnitude, let those in the interval beginning with the lowest and including $y = x_k/w_k$ give negative deviations, while those in the interval beginning with $y = x_{k+1}/w_{k+1}$ and including the greatest give positive deviations. The sum of the weighted absolute deviations is then

$$\sum_{i=1}^{n} w_i \,|\, x_i/w_i - y \,| = \sum_{i=1}^{k} (w_i y - x_i) + \sum_{k+1}^{n} (x_i - w_i y) =$$

$$= \left(\sum_{i=1}^{k} w_i - \sum_{k+1}^{n} w_i\right) y + \left(\sum_{k+1}^{n} x_i - \sum_{1}^{k} x_i\right).$$

Differentiating this expression for the sum of the weighted absolute deviations with respect to y, and equating the derivative to zero, gives

$$\frac{d}{dy} \sum_{i=1}^{n} |\, x_i - w_i y \,| = \sum_{i+1}^{k} w_i - \sum_{k+1}^{n} w_i = 0, \text{ or } \sum_{i=1}^{k} w_i = \sum_{k+1}^{n} w_i .$$

The weighted median is, therefore, that value among the several values of $y = x_i/w_i$, or a value in the interval between a pair of the values, arranged in the order of magnitude, which is characterized by its equal partition of the sum of the weights of all the values into the sums of the weights of the values below and above its own value. Since, for the simple, or equally weighted, median, $w_i \propto 1$, the summation of the weights in the derivative reduces to $k = n - k$, or $k = n/2$; the simple median is therefore that value or a value in the interval between a pair of the values arranged in the order of magnitude characterized by its equal partition of the number of the values (**7**). When the weighted me-

---

$= \Delta y \left(\sum_{i=1}^{m} w_i - \sum_{m+1}^{n} w_i\right)$, while if the values of y be diminished by Δy, the sum of the absolute deviations is again increased by a positive quantity, $- \Delta y \left(\sum_{i=1}^{m-1} w_i - \sum_{m}^{n} w_i\right)$ (**6, 16**).

dian assumes the value of $y = x_m/w_m$, the sum of the weighted absolute deviations becomes

$$\sum_{i=1}^{n} |x_i - w_i y| = \left(\sum_{i=1}^{m-1} w_i - \sum_{m+1}^{n} w_i\right) y + \left(\sum_{m+1}^{n} x_i - \sum_{1}^{m-1} x_i\right),$$

and its derivative with respect to y equated to zero gives

$$\sum_{i=1}^{m-1} w_i = \sum_{m+1}^{n} w_i,$$

from which follow the inequalities,

$$\sum_{i=1}^{m} w_i > \sum_{m+1}^{n} w_i, \ \sum_{i=1}^{m-1} w_i < \sum_{m}^{n} w_i, \text{ or } \sum_{i=1}^{m} w_i > \sum_{1}^{n} w_i/2 > \sum_{m+1}^{n} w_i,$$

which then define the weighted median. For the simple median the derivative reduces to $m - 1 = n - m$, and the definition to $m = \frac{n+1}{2}$ (1).

The application of the median can be extended from weighted observations to indirect observations (2). Edgeworth's method of median loci (**12, 13**) constitutes such a generalization of Laplace's method of situation from the determination of the value of weighted observations to the determination of the parameters of linear observation equations. The procedure originally proposed was to determine for each parameter a locus which would contain the medians of the sets of values of that parameter which would result from the substitution in the given set of observation equations of any set of assigned values of the other parameters, and to locate the point of intersection of the several median loci. Given a set of observation equations, $y_i = a + b\,x_i$: substitute

(1) The proof of the minimum deviation property of the median by the elimination of enveloping pairs (**8**), as adapted to values of varied frequencies (**9**), will apply to the weighted median if weights are read for frequencies.

(2) Thus Laplace applied it in the analytic solution he substituted for the original graphic solution of Boscovich's method for the combination of observations, the earliest method proposed to effect a definite disposition of the errors of indirect observation, namely, according to the conditions that the sum of the negative deviations should be equal to the sum of the positive deviations, and that the sum of all the deviations, negative as well as positive, should be the least possible (**10**). Given a set of observation equations, $y_i = a + b\,x_i$, of weights, w_i: the sum of

a series of values of b in the equations, determine the simple median of the set of values of a for each value of b, and plot this series of medians to form the a-locus; substitute a series of values of a in the equations and plot the resulting series of weighted medians of b as the b-locus; locate the intersection of the loci. For any number of parameters determine the medians of each m^{th} parameter for sets of assigned values of the other $m - 1$ parameters; the set of m values which occurs identically for all the parameters would be the required solution.

The combination of observations by the method of median loci renders the sum of the absolute deviations a minimum. Let the set of linear observation equations, $y_i = a + b\,x_i + \ldots$, give deviations, $y_i - (a + b\,x_i + \ldots)$, for some set of arbitrary values of the parameters, a, b, .., which, when arranged in the order of magnitude, are negative including the equation k, and positive beginning with the equation $k + 1$. The sum of the absolute deviations is then

$$\sum_{i=1}^{n} | y_i - (a + b\,x_i + \ldots) | = \sum_{i=1}^{k} (a + b\,x_i + \ldots - y_i) +$$

$$+ \sum_{k+1}^{n} (y_i - a - b\,x_i - \ldots) = \left(\sum_{k+1}^{n} y_i - \sum_{1}^{k} y_i \right) + (k - (n - k))\,a +$$

$$+ \left(\sum_{1}^{k} x_i - \sum_{k+1}^{n} x_i \right) b + \ldots$$

The differentiation of this expression with respect to the parameters, a, b,.., and equation of the resulting partial derivatives

---

the observation equations multiplied by their respective weights, divided by the sum of the weights, gives an equation,

$$\Sigma\, w_i\, y_i / \Sigma\, w_i = a + b\, \Sigma\, w_i\, x_i / \Sigma\, w_i \,,$$

which satisfies the condition that the sum of the algebraic deviations be equal to zero; the substraction of this equation of condition from each of the original observation equations gives a set of secondary equations,

$$(y_i - \Sigma\, w_i\, y_i / \Sigma\, w_i) = (x_i - \Sigma\, w_i\, x_i / \Sigma\, w_i)\, b \,,$$

which yield the weighted median of b, which, substituted in the equation of condition, gives the weighted median of a, satisfying the condition that the sum of the absolute deviations be a minimum (**11, 38**).

See also the observations of Gini referred to by AMOROSO (*Contributo al metodo delle minime differenze* in « Giornale degli Economosti e Rivista di Statistica », nota 1, pagg. 60-61) relative to the case of an interpolation curve depending from a single parameter.

to zero, shows that the determination of a parameter point where the simple or weighted medians of the several parameters coincide is the general analytic condition for the least sum of absolute deviations (1).

The median loci which satisfy the general analytic condition with respect to the separate parameters may be most readily determined by the graphic method devised by Turner (**14**). The loci of the observation equations, $y_i = a + b\,x_i$, are designated observation lines. The *a*-locus is a simple median locus, traced, if the number of observation lines is odd, by locating a median line segment by counting in a direction parallel to the *a*-axis from the 1^{st} to the $\left(\frac{n+1}{2}\right)^{th}$ line, and traversing a chain of such segments formed by the network of the intersecting observation lines, and, if the number of observating lines is even, by locating an area between the $\left(\frac{n}{2}\right)^{th}$ and $\left(\frac{n}{2}+1\right)^{th}$ lines and traversing a chain of such areas. The *b*- locus is a weighted median locus, traced by weighting the observation lines with the numerical values of the coefficients of *b* in their equations, locating a weighted median line segment by adding the weights of successive lines in a direction parallel to the *b*-axis until a line is reached such that the relationships between the sum of the weights of the antecedent lines, $\sum_{i=1}^{m-1} |x_i|$, the weight of the line itself, $|x_m|$, and the sum of the weights of the succeeding lines, $\sum_{m+1}^{n} |x_i|$, satisfy the inequalities, $\sum_{i=1}^{m} |x_i| > \sum_{m+1}^{n} |x_i|$, $\sum_{i=1}^{m-1} |x_i| < \sum_{m}^{n} |x_i|$, which characterize the weighted median, and traversing the chain of such segments formed by the intersecting lines of the network. Where the m^{th} line is intersected by the $m+1^{th}$ line, the locus turns to the new line, if, when the sum of the antecedent lines remains unchanged, that of the succeeding lines changes to $\sum_{m,\,m+2}^{n} |x_i|$, and the weight

---

(1) If a parameter have negative coefficients in certain of the observation equations, the signs of all the terms in such equations must be changed before summation of the observed values, substitution of the sums in the expression for the sum of the absolute deviations, and differentiation with respect to the particular parameter, in order that the weights which are to define its weighted median shall all be positive.

of the median line to $|x_{m+1}|$, the first inequality becomes $\sum_{i=1}^{m-1,\,m+1} |x_i| > \sum_{m,\,m+2}^{n} |x_i|$ and the second remains $\sum_{i=1}^{m-1} |x_i| < \sum_{m}^{n} |x_i|$. If, when its members are reconstituted in this way, the first inequality is not satisfied, and the sign of their inequality is reversed, the locus continues to coincide with the m^{th} line, and the inequalities become $\sum_{i=1}^{m+1} |x_i| > \sum_{m+2}^{n} |x_i|$ and $\sum_{i=1}^{m-1,m+1} |x_i| < \sum_{m,m+2}^{n} |x_i|$. But if the inequality becomes an equality, $\sum_{i=1}^{m-1,m+1} |x_i| = \sum_{m,m+2}^{n} |x_i|$, the locus follows neither line, but enters the sector between them, and occupies it as far as the intersection which closes its area. Thus median plane loci are broken lines, chains of linked areas, or chains of linked line segments and areas, formed or bounded by segments of the network of observation lines. The parameter point of the intersection of the loci, if it is singular, constitutes the final solution. But the loci may intersect in several points, coincide through one or more segments, or even be identical throughout. The loci of observation equations containing more than two parameters are observation planes. Median spatial loci are therefore broken planes, chains of contiguous polyhedrons, or chains of polygonal plane segments and polyhedrons, formed or bounded by segments of the network of observation planes, and may intersect in one or more points, or coincide through one or more line segments, plane segments, or polyhedral spaces. The unsupplemented graphic solution will to the extent that the several median loci are concomitant be indeterminate.

When the graphic method does not render a unique solution, the employment of a special analytic criterion (**15, 16**) is required to arrive at the final solution. The sum of the absolute deviations, D, from a given point, P $(a, b, ..)$, on the locus of a given observation equation, $y_m = a + b\,x_m + \ldots$, is

$$D = \sum_{i=1}^{n} |\, y_i - (a + b\,x_i + \ldots) \,| = [(m - 1) - (n - m)]\,a +$$
$$+ \left(\sum_{i=1}^{m-1} x_i - \sum_{m+1}^{n} x_i\right) b + \ldots + \left(\sum_{m+1}^{n} y_i - \sum_{1}^{m-1} y_i\right).$$

The ratio of the increment to the sum of deviations from a given point on the locus of a given equation,

$$\Delta D = [(m - 1) - (n - m)]\, \Delta a + \left(\sum_{i=1}^{m-1} x_i - \sum_{m+1}^{n} x_i\right) \Delta b + \ldots,$$

for any set of parameter increments, Δa, Δb, . . , subject to the condition, $\Delta a + x_m \Delta b + \ldots = 0$, or from a given point not on the locus of an equation,

$$\Delta D = [(k - (n - k)]\, \Delta a + \left(\sum_{i=1}^{k} x_i - \sum_{k+1}^{n} x_i\right) \Delta b + \ldots,$$

for any set of parameter increments, to the corresponding distance on the locus or in an unrestricted direction, $\Delta s = \sqrt{\Delta^2 a + \Delta^2 b + \ldots}$, serves as a means toward a determinate solution when the ratio, $\Delta D/\Delta s$, is evaluated for successive segments of the median loci. The coefficients of the parameter increments, Δa, Δb, . . , in the numerator of the expression for $\Delta D/\Delta s$ are the algebraic differences between the sums of the algebraic values of the coefficients of the respective parameters in the observation equations of the loci antecedent to and succeeding the particular median locus, and correspond to, or, when all the coefficients are positive, equal, the sums of the absolute values of the coefficients obtained in determining the paths of the median loci (1). To apply the ana-

---

(1) , When the median loci are observation lines, segments of the median lines themselves are to be tested by application of the increment ratio. With a median observation line represented by the observation equation, $y_m = a + x_m b$, as a result of the condition, $\Delta a + x_m \Delta b = 0$, if the parameter increment, $\Delta a = \pm x_m$, then the parameter increment, $\Delta b = \mp 1$. Division of each parameter increment by the square root of the sum of the squares of the parameter increments would give the sine and cosine of the slope-angle, θ, of the observation line, or $\Delta' a = x_m/\sqrt{x_m^2 + 1^2} = \sin\theta$ and $\Delta' b = -1/\sqrt{x_m^2 + 1^2} = \cos\theta$, as parameter increments, the square root of the sum of the squares of which gives 1 as the denominator of the increment ratio.

When the median loci are observation planes, the lines of intersection of the three intersecting median planes represented by the observation equations, $z_{m_\iota} = a + x_{m_\iota} b + y_{m_\iota} c$ $(\iota = i, j, k)$, taken two at a time are to be tested by application of the increment ratio. Each set of parameter increments will be a set of direction components of the line of intersection of two planes, as $\Delta a = \pm (x_{m_i} y_{m_j} - x_{m_j} y_{m_i})$, $\Delta b = \pm (y_{m_i} - y_{m_j})$ and $\Delta c = \pm (x_{m_j} - x_{m_i})$. Division of each direction component by the square root of the sum of the squares of the several direction components will give the direction cosines of the line of intersection of the two planes, $\Delta' a = \pm \Delta a/\sqrt{\Delta^2 a + \Delta^2 b + \Delta^2 c}$, $\Delta' b = \pm \Delta b/\sqrt{\Delta^2 a + \Delta^2 b + \Delta^2 c}$, and $\Delta' c = \pm \Delta c/\sqrt{\Delta^2 a + \Delta^2 b + \Delta^2 c}$, as the set of parameter increments, the lytic criterion, test the segments of the median loci radiating from the points of intersection of the median loci, and segments of concomitant median loci, so as to eliminate points and segments from which there is any path of departure with a negative $\Delta D/\Delta s$, and arrive at a point, or a segment of zero ratio, located or delimited solely by loci with positive $\Delta D/\Delta s$'s. Such a parameter point represents a determinate solution. The analytic criterion will by its sign indicate the direction to points on the median locus giving larger or smaller deviations sums except in the case of a segment for which all the coefficients of the parameter increments are zero. The final solution will be left indeterminate within the interval of such a segment, if the segment is a line segment intersected, or an area or space segment bordered, by segments giving positive ratios for all outward directions. Only if the observation equations represent a system of parallel lines or planes, which can occur only if all the sets of the observed values of the independent variables are identical for all the observed values of the dependent variable, will no solution exist.

The solution obtainable by the graphic method, supplemented, if need be, by the analytic criterion, can likewise be obtained by the exclusively analytic method developed by Rhodes (**17**). Select any member of the set of observation equations, solve it for the first parameter, substitute this value in each of the remaining

---

square root of the sum of the squares of which gives 1 for the denominator of the increment ratio. The coefficients of the parameter increments are the differences of the sums of the coefficients of the respective parameters in the observation equations antecedent to and succeeding either of the two intersecting observation planes.

When three observation planes or surfaces intersect, if elsewhere in the region of the intersection than at the intersection itself two planes or surfaces are always on the same side of the third and median plane or surface, the median plane or surface is always on one and the same extreme of the three intersecting planes or surfaces, and the median locus traverses the three faces which form a trihedron with its concave aspect on the side of the median planes opposite to the two remaining planes. But, if elsewhere in the region of the intersection than at the intersection itself the two remaining planes or surfaces are always on opposite sides of the median plane or surface, the median plane or surface is always the midmost of the three intersecting planes or surfaces, and the median locus traverses the six middle sectors formed when a third plane cuts across two intersecting planes.

equations, and determine the weighted median of the second parameter. Substitute the observation equation from which this weighted median was derived for the observation equation initially selected, and repeat the procedure to identify a third observation equation. Continue until two observation equations, either of which will lead to the identification of the other, are isolated. Their simultaneous solution gives the required values of the parameters. For the extension of the method to m parameters, select $m - 1$ observation equations in the course of the successive elimination of the first $m - 1$ parameters, and identify an m^{th} equation which gives the weighted median of the m^{th} parameter on the line the $m - 1$ equations determine in m-dimensional space. Substitute this equation for one of the equations originally selected, and repeat the procedure. Continue until m equations, any $m - 1$ of which will identify the m^{th}, are segregated. Determine their simultaneous solution (1). The combination of indirect observations by the method of median loci is the method of least deviations in the dependent variable.

The combination of implicit functional observations so as to render the sum of their absolute deviations a minimum constitutes an extension of the application of the median locus from linear observation equations to polar observation equations. Let the set of equations of polar vectorial epicycloids, $x_i \cos \alpha + y_i \sin \alpha = \rho$, give deviations, $x_i \cos \alpha + y_i \sin \alpha - \rho$, for a set of arbitrary values of the parameters, α, ρ, which, arranged in the order of magnitude, are negative including equation k, and positive beginning with equation $k + 1$. The sum of the absolute deviations is then

$$\sum_{i=1}^{n} | x_i \cos \alpha + y_i \sin \alpha - \rho | = \sum_{i=1}^{k} (\rho - x_i \cos \alpha - y_i \sin \alpha) +$$

(1) If the final solution be indeterminate within a line segment, two equations will severally identify a third, yet the third will not identify either of these two, but merely indicate them as defining the limits of an identified interval; if within an area, any one of three or more equations will indicate two, other equations; if within a plane segment, a number of different sets of $m - 1$ equations will all identify the same m^{th} equation, yet replacement of one of the equations in any such set by this equation will not identify the equation for which it was substituted, but merely indicate a pair of equations; and if within a volume, a number of different sets of $m - 1$ equations will each indicate two additional equations.

$$+\sum_{k+1}^{n}(x_i\cos\alpha+y_i\sin\alpha-\rho)=\left(\sum_{i=k+1}^{n}x_i-\sum_{1}^{k}x_i\right)\cos\alpha+$$
$$+\left(\sum_{k+1}^{n}y_i-\sum_{1}^{k}y_i\right)\sin\alpha+[k-(n-k)]\rho\,.$$

Differentiate this expression for the sum of the absolute deviations with respect to the radius vector, ρ, and equate the resulting partial derivative to zero :

$$\frac{\partial}{\partial\rho}\sum_{i=1}^{n}|x_i\cos\alpha+y_i\sin\alpha-\rho|=(k-(n-k))=0\,.$$

This derivative is the general expression for the simple median. Differentiate the expression for the sum of the absolute deviations with respect to the vectorial angle, α :

$$\frac{\partial}{\partial\alpha}\sum_{i=1}^{n}|x_i\cos\alpha+y_i\sin\alpha-\rho|=$$
$$=\left(\sum_{i=1}^{k}x_i-\sum_{k+1}^{n}x_i\right)\sin\alpha+\left(\sum_{k+1}^{n}y_i-\sum_{1}^{k}y_i\right)\cos\alpha\,.$$

As the vectorial angle varies, the radius vector, which satisfies the condition defined by equation of the partial derivative with respect to the radius vector to zero, describes a simple median locus. As this locus is traversed, the value of the partial derivative with respect to the vectorial angle varies. That polar parameter point on the ρ-locus where the sign of the α-derivative changes from minus to plus, while the sum of the absolute deviations preferentially remains greater than zero, represents the required solution.

The locus of the radius vector may be determined for discrete vectorial angles by the method of substitution, or in its continuity by the graphic method (1). The polar vectorial epicycloids which are the loci of the polar observation equations are designated observation curves. The ρ-locus is then a simple median locus, traced, if the number of observation curves is odd, by locating a median arc by counting from the 1^{st} to the $\left(\frac{n+1}{2}\right)^{th}$ curve as the curves intersect a polar ray at points other than the origin, which

(1) The construction is merely a matter of plotting the point $\left(\frac{x_i}{2},\frac{y_i}{2}\right)$ as center, and drawing a circle through the origin.

is excluded since it is indeterminate with respect to the vectorial angle, and traversing the circuit of such arcs formed by the network of intersecting observation curves, and, if the number of observations is even, by locating an area between the $\left(\frac{n}{2}\right)^{th}$ and $\left(\frac{n}{2}+1\right)^{th}$ curves and traversing the circuit of such areas. The locus is a closed chain of arcs, intersecting in salient points, and formed by segments of the network of observation curves, or a closed chain or polar cluster of areas bounded by such segments.

The derivative with respect to the vectorial angle may be evaluated for successive salient points or linkages on the ρ-locus. If the number of observations is odd, and the median locus follows the arc of the m^{th} observation curve, the summations of the coefficients in each term of the derivative extend from $m + 1$ instead of from $k + 1$, and to $m - 1$ instead of to k. The evaluation of the derivative at successive salient points or vertices requires the determination of the vectorial angles of the intersections of the arcs or the linkages of the areas composing the median locus. The vectorial angle of the intersection of two polar vectorial epicycloids which are the loci of observation equations i and j, derived from the expression, $\tan\alpha = \frac{\sin\alpha}{\cos\alpha} = -\frac{x_i - x_j}{y_i - y_j}$, obtained in their simultaneous solution, is $\alpha = \tan^{-1}\frac{x_j - x_i}{y_i - y_j}$ in the quadrant of their graphic intersection, or the quadrant for which, when the values of cos α and sin α are substituted in equations i and j, the resulting values of ρ are positive and equal. At a vectorial angle where the derivative is negative, the sum of the absolute deviations decreases as the angle increases; where the derivative is positive, the deviations sum increases as the angle increases. The derivative is therefore evaluated for both proximal and distal branches of the median locus at successive salient points or linkages until an intersection is identified, preferentially at a point other than the origin, where the sign of the derivative is minus for the proximal branch and plus for the distal branch. The values of cos α and sin α for the vectorial angle of this intersection are substituted in the equation of either of its loci which is solved for ρ. This in general constitutes the complete and final solution. But in the event that more than one such intersection at points other than the origin occur, the sum of the absolute deviations is evaluated for each intersection and the polar parameter point giving the least deviations sum greater than zero selected. The final solution will be completely determinate unless there be two or more such such points giving equal deviations sums in which case the solution will not be unique but will be any of two or more alternative lines.

For a set of equations of polar vectorial epispheroids, $x_i \cos\alpha + y_i \cos\beta + z_i \cos\gamma = \rho$, the expression for the sum of the absolute deviations becomes

$$\sum_{i=1}^{n} | x_i \cos\alpha + y_i \cos\beta + z_i \cos\gamma - \rho | = \left(\sum_{i=k+1}^{n} x_i - \sum_{1}^{k} x_i \right) \cos\alpha +$$

$$+ \left(\sum_{k+1}^{n} y_i - \sum_{1}^{k} y_i \right) \cos\beta + \left(\sum_{k+1}^{n} z_i - \sum_{1}^{k} z_i \right) \cos\gamma + (k - (n - k))\, \rho.$$

Differentiation with respect to ρ gives the same partial derivative as with polar vectorial epicycloids and its equation to zero defines a simple median locus of the observation surfaces. The number of vectorial angles involved in the equation of a polar vectorial epispheroid can be reduced by means of the condition, $\cos^2\alpha + \cos^2\beta + \cos^2\gamma = 1$, to two essential angles. Thus if $\sqrt{1 - \cos^2\alpha - \cos^2\beta}$ is substituted for $\cos\gamma$ in the equations prior to differentiation, the partial derivatives with respect to α and β become

$$\frac{\partial}{\partial\alpha} \sum_{i=1}^{n} | x_i \cos\alpha + y_i \cos\beta + z_i \sqrt{1 - \cos^2\alpha - \cos^2\beta} - \rho | =$$

$$= \left(\sum_{i=1}^{k} x_i - \sum_{k+1}^{n} x_i \right) \sin\alpha + \left(\sum_{k+1}^{n} z_i - \sum_{1}^{k} z_i \right) \frac{\sin 2\alpha}{2\cos\gamma},$$

$$\frac{\partial}{\partial\beta} \sum_{i=1}^{n} | x_i \cos\alpha + y_i \cos\beta + z_i \sqrt{1 - \cos^2\alpha - \cos^2\beta} - \rho | =$$

$$\left(\sum_{i=1}^{k} y_i - \sum_{k+1}^{n} y_i \right) \sin\beta + \left(\sum_{k+1}^{n} z_i - \sum_{1}^{k} z_i \right) \frac{\sin 2\beta}{2\cos\gamma}.$$

To this set of partial derivatives with respect to the two vectorial angles, α and β, remaining on elimination of γ, the respective members of which may be denoted as $\Phi_{\alpha(\gamma)}$ and $\Phi_{\beta(\gamma)}$, correspond two other sets of partial derivatives with respect to the angles remaining on elimination of β and of α, the members of which may be denoted as $\Phi_{\alpha(\beta)}$, $\Phi_{\gamma(\beta)}$ and as $\Phi_{\beta(\alpha)}$, $\Phi_{\gamma(\alpha)}$. The cosines of the vectorial angles of the intersection of three polar vectorial epispheroids which are the loci of observation equations i, j, and k, are given by the following general expressions: by subtraction ρ is eliminated and two homogeneous linear equations in three unknowns are obtained, and by cross multiplication the ratios cos α : cos β : cos y in terms of the coefficients of the homogeneous equations are determined; then, denoting these ratios by the auxilliary symbol, k,

$$\cos\alpha = ((y_i - y_j)(z_i - z_k) - (y_i - y_k)(z_i - z_j))\,\mathrm{k},$$
$$\cos\beta = ((z_i - z_j)(x_i - x_k) - (z_i - z_k)(x_i - x_j))\,\mathrm{k},$$
$$\cos\gamma = ((x_i - x_j)(y_i - y_k) - (x_i - x_k)(y_i - y_j))\,\mathrm{k},$$

and, substituting in the equation of condition,

$$\mathrm{k} = 1/\sqrt{((y_i - y_j)(z_i - z_k) - (y_i - y_k)(z_i - z_j))^2 + ((z_i - z_j)(x_i - x_k) - (z_i - z_k)(x_i - x_j))^2 + ((x_i - x_j)(y_i - y_k) - (x_i - x_k)(y_i - y_j))^2}.$$

The intersection where the signs of all the members of the complete set of six partial derivatives, $\Phi_{\alpha(\beta)}$, $\Phi_{\alpha(\gamma)}$, $\Phi_{\beta(\alpha)}$, $\Phi_{\beta(\gamma)}$, $\Phi_{\gamma(\alpha)}$, $\Phi_{\gamma(\beta)}$, change from minus to plus as the respective vectorial angle of each increases constitutes the solution. The combination of implicit functional observations by the method of the median locus is the method of least normal deviations.

The methods of median loci and of the median locus can with slight modifications be extended to weighted observations. If the members of a set of indirect observations are of unequal weights, w_i, to render the sum of their weighted absolute deviations, $\sum_{i=1}^{n} w_i \mid y_i - (a + b - x_i + \ldots) \mid$, a minimum, involves weighting the observation lines or planes, plotted indifferently for either the unweighted or the weighted linear observation equations, with the weighted dependent variable and weighted coefficients of the parameters, $w_i\, y_i$, w_i, $w_i\, x_i$, . . . , and using sums of their numerical values in tracing median loci and of their algebraic values in evaluating increment ratios and deviations sums (1). If the

(1) The sum of the weighted absolute deviations in the dependent variable is

$$D_{w\,(y\,.\,x,\,..)} = \sum_{i=1}^{n} w_i \mid y_i - (a + b\,x_i + \ldots) \mid = \left(\sum_{i=1}^{m-1} w_i - \sum_{m+1}^{n} w_i\right) a +$$

members of a set of implicit functional observations are of unequal weights, to render the sum of their weighted absolute deviations, such as $\sum_{i=1}^{n} w_i \mid x_i \cos \alpha + y_i \sin \alpha - \rho \mid$, a minimum, involves weighting the observation curves or surfaces, plotted for the unweighted polar observation equations, and tracing their weighted median locus, and using sums of the weighted coefficients in evaluating deviations sums and their partial derivatives with respect to the vectorial angles (1).

The methods of least normal deviations and of least deviations in the dependent variable are related as methods for the combination of simple observations and of weighted observations. The sum of the deviations in the dependent variable is equal to the sum of the perpendicular distances, from a rectangular coordinates parameter point to the several observation lines or planes, each weighted by the square root of the sum of the squares of the

---

$$+ \left(\sum_{1}^{m-1} w_i x_i - \sum_{m+1}^{n} w_i x_i \right) b + \ldots + \left(\sum_{m+1}^{n} w_i y_i - \sum_{1}^{m-1} w_i y_i \right),$$

and its increment ratio,

$$\Delta D_{w}/\Delta s = \left[\left(\sum_{i=1}^{m-1} w_i - \sum_{m+1}^{n} w_i \right) \Delta a + \left(\sum_{1}^{m-1} w_i x_i - \sum_{m+1}^{n} w_i x_i \right) \Delta b + \right.$$
$$\left. + \ldots \right] / \sqrt{\Delta^2 a + \Delta^2 b + \ldots} .$$

(1) The sum of the weighted absolute normal deviations

$$D_{w(x,y)} = \sum_{i=1}^{n} w_i \mid x_i \cos \alpha + y_i \sin \alpha - \rho \mid = \left(\sum_{i=m+1}^{n} w_i x_i - \right.$$
$$\left. - \sum_{1}^{m-1} w_i x_i \right) \cos \alpha + \left(\sum_{m+1}^{n} w_i y_i - \sum_{1}^{m-1} w_i y_i \right) \sin \alpha + \left(\sum_{1}^{m-1} w_i - \right.$$
$$\left. - \sum_{m+1}^{n} w_i \right) \rho = \left(\sum_{m+1}^{n} w_i x_i - \sum_{1}^{m-1} w_i x_i \right) \cos \alpha + \left(\sum_{m+1}^{n} w_i y_i - \right.$$
$$\left. - \sum_{1}^{m-1} w_i y_i \right) \sin \alpha + \left(\sum_{1}^{m-1} w_i - \sum_{m+1}^{n} w_i \right) (x_m \cos \alpha + y_m \sin \alpha) ,$$

and its partial derivative,

$$\frac{\partial D_w}{\partial \alpha} = \left(\sum_{i=1}^{m-1} w_i x_i - \sum_{m+1}^{n} w_i x_i \right) \sin \alpha + \left(\sum_{m+1}^{n} w_i y_i - \sum_{1}^{m-1} w_i y_i \right) \cos \alpha +$$
$$+ \left(\sum_{1}^{m-1} w_i - \sum_{m+1}^{n} w_i \right) (y_m \cos \alpha - x_m \sin \alpha).$$

coefficients of the parameters in its respective observation equation, as

$$D = \Sigma \sqrt{1^2 + x_i^2 + y_i^2 + \ldots} \left| \frac{-a - x_i b - y_i c - \ldots + z_i}{\sqrt{1^2 + x_i^2 + y_i^2 + \ldots}} \right| \quad \textbf{(15, 18, 19)}.$$

If eache perpendicular distance is regarded as the projection of the line segment, parallel to the axis of, and representing the deviation in, the dependent variable, the deviation in the dependent variable, which may be designated the inverse projection of the perpendicular distance, will equal the perpendicular distance multiplied by the secant of the angle between the perpendicular and the axis of the dependent variable. Deviations in the dependent variable are inverse projections of normal deviations. The sum of the uniformly weighted normal deviations,

$$\Delta_{\omega\,(x,y,z,..)} = \Sigma \mid \omega\,(x_i \cos\alpha + y_i \cos\beta + z_i \cos\gamma + \ldots - \rho) \mid ,$$

will therefore equal the sum of the deviations in the dependent variable, as $D_{z.x,y,..} = \Sigma \mid z_i - (a + b\,x_i + c\,y_i + \ldots) \mid$, if their weights are the secant of the angle between the normals and the axis of the dependent variable, as $\omega = \sec\gamma$, and will equal the least sum of deviations in the dependent variable if their sum when so weighted is rendered a minimum (1). Only one set of polar observation equations,

$$x_i \cos\alpha + y_i \cos\beta + z_i \cos\gamma + \ldots - \rho_{(x_i, y_i, z_i, ..)} = 0 ,$$

---

(1) The sum of the normal deviations of uniformly weighted observation equations,

$$\Delta_{\omega\,(x,y)} = \sum_{i=1}^{n} \mid \omega\,(x_i \cos\alpha + y_i \sin\alpha - \rho) \mid ,$$

equals the sum of the deviations in the dependent variable, x, when $\omega = \sec\alpha$. The sum of the uniformly weighted normal deviations and of the deviations in the dependent variable,

$$D_{x \cdot y} = \sum_{i=1}^{n} \mid \sec\alpha\,(x_i \cos\alpha + y_i \sin\alpha - \rho) \mid = \sum_{1}^{n} \mid x_i + y_i \tan\alpha -$$

$$- \rho \sec\alpha \mid = \left(\sum_{m+1}^{n} x_i - \sum_{1}^{m-1} x_i \right) + \left(\sum_{m+1}^{n} y_i - \sum_{1}^{m-1} y_i \right) \tan\alpha + [(m-1) -$$

$$- (n - m)]\,\rho \sec\alpha,$$

and the partial derivative of the sum of the uniformly weighted normal deviations from the median locus with respect to the vectorial angle,

$$\frac{\partial}{\partial\alpha} \sum_{i=1}^{n} \mid x_i + y_i \tan\alpha - \rho\sec\alpha \mid = \left(\sum_{m+1}^{n} y_i - \sum_{1}^{m-1} y_i \right) \sec^2\alpha.$$

exists for any set of observations involving a given number of variables, but, since any variable may be treated as the dependent variable, as many possible sets of linear observation equations,

$$x_i = a_{(x\cdot y,z,\ldots)} + y_i\, b_{(x,y\cdot z,\ldots)} + z_i\, c_{(x,z\cdot y,\ldots)} + \ldots,$$
$$y_i = a_{(y\cdot x,z,\ldots)} + x_i\, b_{(y,x\cdot z,\ldots)} + z_i\, c_{(y,z\cdot x,\ldots)} + \ldots,$$
$$z_i = a_{(z\cdot x,y,\ldots)} + x_i\, b_{(z,x\cdot y,\ldots)} + y_i\, c_{(z,y\cdot x,\ldots)} + \ldots,$$
$$\cdot \quad \cdot \qquad \cdot\ \cdot \qquad \cdot\ \cdot \qquad \ldots,$$
$$\cdot \quad \cdot \qquad \cdot\ \cdot \qquad \cdot\ \cdot \qquad \ldots,$$
$$\cdot \quad \cdot \qquad \cdot\ \cdot \qquad \cdot\ \cdot \qquad \ldots,$$

With individually weighted observation equations,

$$\Delta_{\omega w (x,y)} = \sum_{i=1}^{n} |\, \omega\, w_i\, (x_i \cos\alpha + y_i \sin\alpha - \rho)\, |,$$

$$D_{w(x.y)} = \left(\sum_{i=m+1}^{n} w_i\, x_i - \sum_{1}^{m-1} w_i\, x_i\right) + \left(\sum_{m+1}^{n} w_i\, y_i - \sum_{1}^{m-1} w_i\, y_i\right) \tan\alpha +$$
$$+ \left(\sum_{1}^{m-1} w_i - \sum_{m+1}^{n} w_i\right)(x_m + y_m \tan\alpha),$$

and

$$\frac{\partial}{\partial\alpha}\sum_{i=1}^{n} |\, w_i\, (x_i + y_i \tan\alpha - \rho \sec\alpha)\, | = \left[\left(\sum_{m+1}^{n} w_i\, y_i - \sum_{1}^{m-1} w_i\, y_i\right) -\right.$$
$$\left. - y_m \left(\sum_{1}^{m-1} w_i - \sum_{m+1}^{n} w_i\right)\right] \sec^2\alpha.$$

The sum of the uniformly weighted normal deviations equals the sum of deviations in the dependent variable, y, when $\omega = \csc a$:

$$D_{y\cdot x} = \sum_{i=1}^{n} |\, \csc\alpha\, (x_i \cos\alpha + y_i \sin\alpha - \rho)\, | = \sum_{1}^{n} |\, x_i \cot\alpha +$$
$$+ y_i - \rho \csc\alpha\, | = \left(\sum_{m+1}^{n} x_i - \sum_{1}^{m-1} x_i\right)\cot\alpha + \left(\sum_{m+1}^{n} y_i - \sum_{1}^{m-1} y_i\right) +$$
$$+ [(m - 1) - (n - m)]\, \rho \csc a,$$

and

$$\frac{\partial}{\partial\alpha}\sum_{i=1}^{n} |\, x_i \cot\alpha + y_i - \rho\csc\alpha\, | = \left(\sum_{1}^{m-1} x_i - \sum_{m+1}^{n} x_i\right)\csc^2\alpha.$$

Similarly the sum of the uniformly weighted normal deviations,

$$\Delta_{\omega(x,y,z)} = \sum_{i=1}^{n} |\, \omega\, (x_i \cos\alpha + y_i \cos\beta + z_i \cos\gamma - \rho)\, |,$$

equals the sum of deviations in a dependent variable, as z, when ω equals the secant of the corresponding vectorial angle, as γ, when

$$D_{z.x,y} = \sum_{i=1}^{n} |\, \sec\gamma\, (x_i \cos\alpha + y_i \cos\beta + z_i \cos\gamma - \rho)\, | =$$
$$= \sum_{1}^{n} \left|\, x_i \frac{\cos\alpha}{\cos\gamma} + y_i \frac{\cos\beta}{\cos\gamma} + z_i - \frac{\rho}{\cos\gamma}\, \right|.$$

exist as the number of variables. The solution by the method of least normal deviations will mediate between the respective solutions by the method of least deviations in the several alternative dependent variables by rendering the sum of the normal deviations themselves, in contradistinction to the sums of the inverse projections of the normal deviations on the axis of a dependent variable, or of the normal deviations weighted with the secant of the angle between the normals and the axis, as

$$D_{x \cdot y, z, \ldots} = \Sigma \mid \sec \alpha \, (x_i \cos \alpha + y_i \cos \beta + z_i \cos \gamma + \ldots - \rho) \mid ,$$
$$D_{y \cdot x, z, \ldots} = \Sigma \mid \sec \beta \, (x_i \cos \alpha + y_i \cos \beta + z_i \cos \gamma + \ldots - \rho) \mid ,$$
$$D_{z \cdot x, y, \ldots} = \Sigma \mid \sec \gamma \, (x_i \cos \alpha + y_i \cos \beta + z_i \cos \gamma + \ldots - \rho) \mid ,$$

a minimum. The method of least deviations in a dependent variable is a method of weighted least normal deviations.

3. Mean Loci or Least Squares.

The original method of least squares, of Legendre (**20**) and Gauss (**21**), for deviations in the dependent variable, depends upon partial derivatives with respect to the several parameters, which, when written in the form,

$$\frac{\partial}{\partial a} \Sigma \, (z_i - (a + b x_i + c y_i + \ldots))^2 = \Sigma \quad (z_i - (a + x_i b + y_i c + \ldots)) = 0 \, ,$$

$$\frac{\partial}{\partial b} \Sigma \, (z_i - (a + b x_i + c y_i + \ldots))^2 = \Sigma \, x_i (z_i - (a + x_i b + y_i c + \ldots)) = 0 \, ,$$

$$\frac{\partial}{\partial c} \Sigma \, (z_i - (a + b x_i + c y_i + \ldots))^2 = \Sigma \, y_i (z_i - (a + x_i b + y_i c + \ldots)) = 0 \, ,$$

clearly represent weighted sums, and are proporzional to weighted arithmetic means, obtained for each derivative by weighting the observations with the coefficients of its respective parameter (**12**). The simultaneous solution of the normal equations is equivalent to the intersection of the loci of the weighted sums or arithmetic means. Thus the method of least squares in the dependent variable is the method of mean loci.

The method of least squared normal deviations, introduced by Pearson (**22**) as the method of closest fit (1), is the method of

(1) The derivation advanced by Pearson (**22**) was based on analytic mechanics and related to frequency surfaces for correlated variables. A direct analytic deiivation was given by Reed (**23**) for the straight line. The analytic derivation for any number of variables was modified by

the mean locus. Differentiation of the sum of the squared normal deviations with respect to ρ, division of the partial derivative by twice the number of observations, and equation to zero, gives, in the special case of two variables,

$$\frac{\partial}{\partial \rho} \Sigma (x_i \cos \alpha + y_i \sin \alpha - \rho)^2 = \frac{\Sigma x_i}{n} \cos \alpha + \frac{\Sigma y_i}{n} \sin \alpha - \rho = 0.$$

Snow (**24**) to apply to restricted lines and planes passing through a fixed point or points where the values of all the variables are exactly known, and by Glauert (**25**) to apply to lines and planes connecting variables equally liable to error associated with variables the values of which are exactly known. Finally a generalized derivation was elaborated by Rhodes (**26**) bor variables not equally liable to error.

The slope and position of the straight line was determined by Pearson from the expression for the directions of the principal axes of inertia calculated for the centroid appropriated from analytic mechanics. The claims of least squared normal deviations had originally been argued much earlier from considerations of statics (*The Correct Method of Least Squares*, « Analyst », 1 : 64 (1874)) and of geometrical probability (*Note on the Method of Least Squares*, idem, 4 : 183-184 (1877), 5 : 21-22 (1878)) by R. J. Adcock who sought explicit expressions for the parameters in the equation of the line, $y = a + b\,x$, which would render the squared normal deviations sum, $\Sigma \frac{(y_i - a - b\,x_i)^2}{1 + b^2}$, a minimum (*A Problem in Least Squares*, ibid., 53-54), and indicated a procedure for determining the parameters in the equation of the plane, $z = a + b\,x + c\,y$, which would render the squared normal deviations sum, $\Sigma \frac{(z_i - a - b\,x_i - c\,y_i)^2}{1 + b^2 + c^2}$, a minimum (Ibid., 122, 149-150). A general solution of observation equations in which w_{x_i} is the weight of x_i, and w_{y_i} the weight of y_i, by the method of least squares had been given by C. H. Kummel (*Reduction of Observation Equations which Contain More than One Observed Quantity*. idem, 6 : 97-105 (1879)) and implemented by expressions for the parameters in the equation for the straight line which would render the generalized squared deviations sum, $\Sigma \frac{(y_i - a - b\,x_i)^2}{1/w_{y_i} + b^2/w_{x_i}}$, a minimum under the conditions that the weights are different in each observation equation and their ratio is constant in all observation equations; as $0 < w_{x_i} \leq \infty$, $w_{y_i} = k\,w_{x_i}$, so that if x is taken to be exactly known and y to be liable to error, $w_{x_i} = \infty$, and $k = w_{y_i}/w_{x_i} = 0$, and the parameters have the values obtained by the original method of least squares when y is treated as the dependent variable, if x and y are taken to be equally liable to error. and the observation equations of equal weight, $w_{x_i} = w_{y_i} = 1$ and $k = 1$, and they have the values obtained from the corrected forms of Adcock's expressions, and if y is taken to be

The ρ-locus is therefore the mean polar vectorial epicycloid of the set of observations. Differentiation of the squared deviations from the mean locus, obtained by substituting its mean value for ρ in each of the observation equations, with respect to α, gives the derivative,

$$\frac{\partial}{\partial \alpha} \Sigma \left(\left(x_i - \frac{\Sigma x_i}{n}\right) \cos \alpha + \left(y_i - \frac{\Sigma y_i}{n}\right) \sin \alpha\right)^2,$$

---

exactly known and x to beliable to error, $w_{y_i} = \infty$, and $k = \infty$, the values by the method of least squares when x is the dependent variable, and by M. Merriman (*The Determination, by the Method of Least Squares, of the Relation between Two Variables, Connected by the Equation $Y = A X + B$, Both Variables being Liable to Errors of Observation*, « Rep. U. S. Coast & Geod. Surv. », 1890 : 687-690 (1891)) undei the more restricted conditions that both the weights and their ratios in all observation equations are constant, or $w_y = 1$ and $w_x = \frac{1}{k} = \infty$, 1, and 0 respectively. The general solution was later restated independently of the predecessors and successors of Pearson and applied to specific problems by M. R. Stewart (*The Adjustment of Observations*, « Phil. Mag.», 6th, 40 : 219-222, 224-227 (1920)), and its development for the straight line comprehensively elaborated by H. S. Uhler, (*Method of Least Squares and Curve Fitting*, « J. Opt. Soc. Amer. & Rev. Sci. Instrum. », 7 : 1043-1066 (1923)) who later generalized and extended Pearson's derivation for planes (*Determination of the Minimum Plane in Four-Dimensional Space with Respect to a System of Non-Coplanar Points*, « Phil. Mag. », 6th, 49 : 1260-1271 (1925)), and formulated a basic theorem concerning the *Least Distance from a Point to a Linear $(n - k)$ Space, Both in a Linear n-Space*, (« Ann. Math. », 27 : 65-68 (1925)). And finally, following Stewart and Uhler, the derivations of Snow, Glauert, and Rhodes were in effect placed on a more general foundation, reproduced, and widely extended by W. E. Deming (*The Application of Least Squares*, « Phil. Mag. », 7th, 11 : 146-158 (1931), 17 : 804-829 (1934)).

Prompted by Reed's derivation, C. Gini (*Sull'interpolazione di una retta quando i valori della variabile indipendente sono affetti da errori accidentali*, « Metron », 1 (3) : 63-82 (1921)) by algebraic treatment of the errors of a large number of observations also derived expressions for the parameters in the linear equations, $y = a + b x$ and $x = \left(\frac{-a}{b}\right) + \frac{1}{b} y$, as in turn did G. Pietra (*Interpolating Plane Curves*, idem, 3 (3-4) : 311-328 (1924); *Dell'interpolazione parabolica nel caso in cui entrambi i valori delle variabili sono affetti da errori accidentali*, idem, 9 (3-4) : 77-86 (1932)) for the parameters in the parabolas, $y = a + b x + c x^2$ and $x = a + b y + c y^2$, for any ratio of the weights of the variables.

The mathematical derivations for two variables connected with the methods of least squares for deviations in the dependent variable and for which, on equation to zero and simplification, reduces to

$$\alpha = \frac{\tan^{-1}}{2} \frac{2\left(\Sigma x_i y_i - \frac{(\Sigma x_i)(\Sigma y_i)}{n}\right)}{\left(\Sigma x_i^2 - \frac{(\Sigma x_i)^2}{n}\right) - \left(\Sigma y_i^2 - \frac{(\Sigma y_i)^2}{n}\right)},$$

or, divided by n, to $\alpha = \frac{\tan^{-1}}{2} \frac{2 p_{xy}}{\sigma_x^2 - \sigma_y^2}$, if p_{xy} is the product-

---

normal deviations were comprehensively elaborated by W. WIRTH (*Spezielle psychophysische Massmethoden*, « Handbuch der biologischen Arbeitsmethoden », her. E. Abderhalden, Berlin, 1920, VI, A, 1 : 5-15, 71-81, 105-117, 126-142, 145-146) in the statistical setting of linear equations of regression and of the representation of relationship, and the theoretical concepts involved developed in a critical discussion by E. CZUBER (*Zur Theorie der linearen Korrelation*, « Arch. f. d. ges. Psychol. », 41 : 310-334 (1921)) and a rejoinder (*Bemerkungen zu der vorangehenden Abhandlung von Herrn Prof. E. Czuber, über die Theorie der linearen Korrelation*, ibid., 334-352). The statistical connotations were further elucidated by S. KOLLER (*Die Analyse der Abhängigkeitsverhältnisse in zwei Korrelationssystemen*, « Metron », 12 (4) : 73-105 (1936)) under the categories of correlation, regression, and the representation of the relationship that lies at the foundation of the correlation system. The line for the representation of relationship had the disadvantage that it is not independent of the selection of the units of measurement of the correlates, bu might be made independent of the selection of the units of measurement of the correlates by taking their standar deviations as the fixed common unit of measurement, as had previously been done as a special case under Rhodes' general solution (26). Thus, if the variates x_i and y_i are divided by their standard deviations, the denominators in the expressions in the text for tan 2α equal 0, $\tan 2\alpha = \pm \infty$, $2\alpha = 90°$ or $270°$, $\alpha = 45°$ or $135°$, and $\cos \alpha = \pm 0.70711$, $\sin \alpha = 0.70711$, $\tan \alpha = \pm 1$, so that the final explicit equation is $\frac{y - m_y}{\sigma_y} = \pm \frac{x - m_x}{\sigma_x}$ according to whether the numerator is positive or negative.

The fact that the method of least squared normal deviations is not independent of the choice of a particular coordinates system became the point of departure for the derivation of a method of least squared deviations independent of a coordinates system outlined by C. F. Roos and A. OPPENHEIMER (*A Symmetric Method of Fitting Lines and Planes*, « Bul. Amer. Math. Soc. », 34 : 140-141 (1928)) and developed by C. F. Roos (*A General Invariant Criterion of Fit for Lines and Planes Where All Variates are Subject to Error*, « Metron », 13 (1) : 3-20 (1937)). For the set of observations, (x_i, y_i), of weights, w_i, and the line, $a\,x + b\,y + c = 0$, the most general function, $\Sigma\, w_i f_i (x_i, y_i, a, b, c)$, which remains invariant under homogeneous strain, translation, and rotation is $\Sigma\, w_i \mid a\,x_i + b\,y_i + c \mid^p$, where p is any number, but specifically 2, which is made subject moment and σ_x and σ_y are the standard deviations, in the quadrant, letting $0^0 \leq \alpha \leq 360^0$, for which, when cos α and sin α are substituted in the equation of the mean polar vectorial epicycloid, ρ is positive. In the general case of any number of variables partial differentiation of the sum of the squared normal deviations with respect to ρ gives

$$\frac{\partial}{\partial \rho} \Sigma (x_i \cos \alpha + y_i \cos \beta + z_i \cos \lambda + \ldots - \rho)^2 =$$
$$= \frac{\Sigma x_i}{n} \cos \alpha + \frac{\Sigma y_i}{n} \cos \beta + \frac{\Sigma z_i}{n} \cos \gamma + \ldots - \rho = 0 .$$

Hence the ρ-locus is according the number of variables the mean polar vectorial epicycloid, epispheroid, or epihyperspheroid of the observations. Substitution of the mean value of ρ from the equation of its locus in each of the observation equations, differentiation of the sum of the resulting squared deviations from the ρ-locus,

$$\Sigma \left(\left(x_i - \frac{\Sigma x_i}{n}\right) \cos \alpha + \left(y_i - \frac{\Sigma y_i}{n}\right) \cos \beta + \left(z_i - \frac{\Sigma z_i}{n}\right) \cos \gamma + \ldots\right)^2 ,$$

with the condition equation, $\cos^2 \alpha + \cos^2 \beta + \cos^2 \gamma + \ldots = 1$, multiplied by an undetermined coefficient, $\varkappa$, added, or

$$\Sigma \left(\left(x_i - \frac{\Sigma x_i}{n}\right) \cos \alpha + \left(y_i - \frac{\Sigma y_i}{n}\right) \cos \beta + \left(z_i - \frac{\Sigma z_i}{n}\right) \cos \gamma + \ldots\right)^2$$
$$+ \varkappa (\cos^2 \alpha + \cos^2 \beta + \cos^2 \gamma + \ldots - 1) ,$$

---

to the condition that $\Sigma w_i \mid a x_i + b y_i + c \mid^p$ is the sum of the distances in the direction, $\tan \alpha = w_{xi}/w_{yi} = k$, where w_{yi} and w_{xi} are the respective weights of x_i and y_i, and is implemented with an explicit general solution and explicit solutions for its special cases. The method is extended to planes and hyperplanes. Geometrical interpretations of Roos' general solution and some of the special cases and demonstrations of its property of invariance under transformations of the coordinates were advanced by H. E. Jones (*Some Geometrical Considerations in the General Theory of Fitting Lines and Planes*, ibid., 21-30).

Though formulae were available for the determination of the linear relationship between two variables for any absolute and relative weights of the variables, sanction was lacking for the assumption of other than the values of unity and infinity as the weights and of the resulting values of zero, unity, and infinity as the ratios until an approach was made by W. R. Cook (*On Curve-Fitting by means of Least Squares*, idem, 7th, 12 : 1025-1039 (1931)) and by Beryl M. Dent (*On Observations of Points Connected by a Linear Relationship*, « Proc. Phys. Soc. », 47 : 92-106 (1935)) to methods for deriving from the data an indication of the relative weights to be attached to the variables and of the errors of the resulting parameters.

with respect to the cosine of each of the vectorial angles, and equation of the partial derivatives to zero, give

$$\left(\Sigma\left(\left(x_i - \frac{\Sigma x_i}{n}\right)\cos\alpha + \left(y_i - \frac{\Sigma y_i}{n}\right)\cos\beta + \left(z_i - \frac{\Sigma z_i}{n}\right)\cos\gamma + \ldots\right)\right.$$

$$\left.\left(x_i - \frac{\Sigma x_i}{n}\right)\right) + \varkappa\cos\alpha = 0\,,$$

$$\left(\Sigma\left(\left(x_i - \frac{\Sigma x_i}{n}\right)\cos\alpha + \left(y_i - \frac{\Sigma y_i}{n}\right)\cos\beta + \left(z_i - \frac{\Sigma z_i}{n}\right)\cos\gamma + \ldots\right)\right.$$

$$\left.\left(y_i - \frac{\Sigma y_i}{n}\right)\right) + \varkappa\cos\beta = 0\,,$$

$$\left(\Sigma\left(\left(x_i - \frac{\Sigma x_i}{n}\right)\cos\alpha + \left(y_i + \frac{\Sigma y_i}{n}\right)\cos\beta + \left(z_i - \frac{\Sigma z_i}{n}\right)\cos\gamma + \ldots\right)\right.$$

$$\left.\left(z_i - \frac{\Sigma z_i}{n}\right)\right) + \varkappa\cos\gamma = 0\,,$$

. . . .

Multiplication of each of these equations by the corresponding cosine, and addition of the several equations with regard for the condition equation, gives

$$\left(\Sigma\left(\left(x_i - \frac{\Sigma x_i}{n}\right)\cos\alpha + \left(y_i - \frac{\Sigma y_i}{n}\right)\cos\beta +\right.\right.$$

$$\left.\left. + \left(z_i - \frac{\Sigma z_i}{n}\right)\cos\gamma + \ldots\right)^2\right) + \varkappa = 0\,.$$

The undetermined multiplier is therefore the negative of the squared-deviations sum at the sets of critical values of the cosines of the vectorial angles. On letting the negative of the squared-deviations sum be denoted by the negative of the undetermined multiplier, the equations become

$$\left(\left(\Sigma\left(x_i - \frac{\Sigma x_i}{n}\right)^2\right) - \varkappa\right)\cos\alpha + \left(\Sigma\left(x_i - \frac{\Sigma x_i}{n}\right)\left(y_i - \frac{\Sigma y_i}{n}\right)\right)\cos\beta +$$

$$+ \left(\Sigma\left(x_i - \frac{\Sigma x_i}{n}\right)\left(z_i - \frac{\Sigma z_i}{n}\right)\right)\cos\gamma + \ldots = 0\,,$$

$$\left(\Sigma\left(x_i - \frac{\Sigma x_i}{n}\right)\left(y_i - \frac{\Sigma y_i}{n}\right)\right)\cos\alpha + \left(\left(\Sigma\left(y_i - \frac{\Sigma y_i}{n}\right)^2\right) - \varkappa\right)\cos\beta +$$

$$+ \left(\Sigma\left(y - \frac{\Sigma y_i}{n}\right)\left(z - \frac{\Sigma z_i}{n}\right)\right)\cos\gamma + \ldots = 0\,,$$

$$\left(\Sigma\left(x_i - \frac{\Sigma x_i}{n}\right)\left(z_i - \frac{\Sigma z_i}{n}\right)\right)\cos\alpha + \left(\Sigma\left(y_i - \frac{\Sigma y_i}{n}\right)\left(z_i - \frac{\Sigma z_i}{n}\right)\right)\cos\beta +$$
$$+ \left(\left(\Sigma\left(z_i - \frac{\Sigma z_i}{n}\right)^2\right) - \varkappa\right)\cos\gamma + \ldots = 0\,,$$

. . . .

Simplifying the coefficients, the eliminant of the cosines of the vectorial angles is the determinantal equation,

$$\begin{vmatrix} \left(\Sigma x_i^2 - \frac{(\Sigma x_i)^2}{n}\right) - \varkappa & \Sigma x_i y_i - \frac{(\Sigma x_i)(\Sigma y_i)}{n} & \Sigma x_i z_i - \frac{(\Sigma x_i)(\Sigma z_i)}{n} & \ldots \\ \Sigma x_i y_i - \frac{(\Sigma x_i)(\Sigma y_i)}{n} & \left(\Sigma y_i^2 - \frac{(\Sigma y_i)^2}{n}\right) - \varkappa & \Sigma y_i z_i - \frac{(\Sigma y_i)(\Sigma z_i)}{n} & \ldots \\ \Sigma x_i z_i - \frac{(\Sigma x_i)(\Sigma z_i)}{n} & \Sigma y_i z_i - \frac{(\Sigma y_i)(\Sigma z_i)}{n} & \left(\Sigma z_i^2 - \frac{(\Sigma z_i)^2}{n}\right) - \varkappa & \ldots \\ \cdot & \cdot & \cdot & \ldots \\ \cdot & \cdot & \cdot & \ldots \\ \cdot & \cdot & \cdot & \ldots \end{vmatrix} = 0,$$

the smallest root of which is the least squared-deviations sum, or, on transforming this equation into an equation the roots of which are the roots of this equation divided by n, the eliminant of the cosines is the determinantal equation,

$$\begin{vmatrix} \sigma_x^2 - \left(\frac{\varkappa}{n}\right) & p_{xy} & p_{xz} & \ldots \\ p_{xy} & \sigma_y^2 - \left(\frac{\varkappa}{n}\right) & p_{yz} & \ldots \\ p_{xz} & p_{yz} & \sigma_z^2 - \left(\frac{\varkappa}{n}\right) & \ldots \\ \cdot & \cdot & \cdot & \ldots \\ \cdot & \cdot & \cdot & \ldots \\ \cdot & \cdot & \cdot & \ldots \\ \cdot & \cdot & \cdot & \ldots \end{vmatrix} = 0,$$

the smallest root of which is the least mean squared deviation (1).

---

(1) Recently alternatives to Pearson's formula for the mean squared deviation from the straight line have been given by Y. K. Wong (*On Standard Error for the Line of Mutual Regression*, « Ann. Math. Stat. », 7 : 47-50 (1936)) together with the correction of the non sequitur in the printed statement of its derivation.

On insertion of the value of the least squared-deviations sum in the equations, or of the least mean squared deviation in the equations divided through by n, the ratios of the cosines of the vectorial angles may be determined from any combination of equations one less than the number of equations, and the values of the cosines from the condition equation.

4. MID-POINT OF LEAST RANGE OR LEAST GREATEST-DEVIATION.

The method of the least greatest-deviation, the earliest method proposed for the combination of indirect observations to effect a consistent disposition of the errors of observation, was originally discovered by Laplace (**27**), and recently rediscovered by Goedseels in his method of minimum approximation (**28**), and by Smith in his mid-course method (**29**). General analytic methods have been advanced by Laplace (**27, 38**), Goedseels (**28, 30**), and de la Vallee-Poussin (**31**) for the solution of sets of linear observation equations so as to obtain the least greatest-absolute-deviation in the dependent variable (1). But the method of graphic selection, supplemented

(1) In Laplace's general solution (**27, 38**) the n equations for the deviations corresponding to the given observation equations, $y_i - a - bx_i - \ldots - mv_i = \delta_i$, by elimination of the parameters yield $n - m$ simultaneous indeterminate linear equations in the deviations, $E_i = p_i \delta_1 + q_i \delta_2 + \ldots + o_i \delta_n$, which are to be solved for the condition that the $m + 1$ greatest deviations which are equal in absolute value shall be the least possible; in his special solution for the case of two variables (**11, 38**) the deviation equations are subjected to arithmetically more expeditious but deductively more involved manipulations.

In Goedseels' primitive solution (**28**) for his method of minimum approximation, the deviation equations, since the individual deviations are at once equal to or less than an undetermined extreme positive deviation, δ, and equal to or greater than an undetermined extreme negative deviation, $-\delta$, are converted into pairs of inequalities.

$$y_i - a - b\,x_i - \ldots \leqq \delta \quad \text{and} \quad -\delta \leqq y_i - a - b\,x_i - \ldots \ (i = 1, 2, \ldots, n),$$

which by transposition of δ and a are transformed into

$$y_i - b\,x_i - \ldots - \delta \leqq a \text{ and } a \leqq y_i - b\,x_i - \ldots + \delta,$$

which through juxtaposition of their outer members each to each by the elimination of a produce n^2 inequalities,

$$y_i - b\,x_i - \ldots - \delta \leqq y_j - b\,x_j - \ldots + \delta \ (i = 1, 2, \ldots, n;\ j = 1, 2, \ldots, n)$$

which on solution for b produce the pairs of inequalities,

$$\frac{y_i - y_j - \ldots + \ldots - 2\,\delta}{x_i - x_j} \leqq b \text{ and } b \leqq \frac{y_j - y_i + \ldots - \ldots + 2\,\delta}{x_j - x_i},$$

by analytic criteria, and implemented by algebraic solutions or special formulae, developed in a form adaptable to indefinite generalization in a certain specific direction by Smith (**29**), is more amenable than any of the analytic methods to development in a form capable of extension to implicit functional observations.

To render the greatest deviation in the dependent variable

---

which through juxtaposition of their outer members each to each by elimination of b produce $4\,n^4$ inequalities,

$$\frac{y_i - y_j - \ldots + \ldots - 2\,\delta}{x_i - x_j} \leqq \frac{y_1 - y_k + \ldots - \ldots + 2\,\delta}{x_1 - x_k},$$

which, if no further parameters remain to be eliminated, on solution for δ produce the inequalities,

$$\delta \geqq \frac{(y_i - y_j)(x_1 - x_k) + (y_k - y_1)(x_i - x_j)}{2\,[(x_i - x_j) + (x_1 - x_k)]},$$

selection of the largest right member of which defines the extreme absolute deviation, substitution of which in the inequalities for b determines its value, substitution of which together with that of the extreme deviation in the inequalities for a determines its value ; in his simplified solution (**30**) the procedure is abbreviated through the omission of certain members of the sets of intermediary inequalities from further consideration.

In de la Vallée-Pouissin's solution (**31**) for $m + 1$ deviation equations in m parameters,

$a + bx_\iota + c\,y_\iota + \ldots + m\,v_\iota - z_\iota = -\upsilon_\iota\,\delta\ (\iota = 1, 2, \ldots, m, m+1\,;\ \upsilon = \pm 1)$,
the least greatest-deviation is determined from the determinant, $\Delta = |\,1_1\,x_2 \ldots v_m\,z_{m+1}\,|$, and its minors, Z_ι, when that sign is attributed to each unit, υ_ι, which will render the product, $\upsilon_\iota Z_\iota$, the same in sign as the determinant, Δ, as

$$\delta = \frac{\Delta}{\upsilon_1 Z_1 + \upsilon_2 Z_2 + \ldots + \upsilon_m Z_m + \upsilon_{m+1} Z_{m+1}}$$

or

$$\frac{|\,\Delta\,|}{|\,Z_1\,| + |\,Z_2\,| + \ldots + |\,Z_m\,| + |\,Z_{m+1}\,|}$$

and the parameters from the minors, as

$$a = \frac{\upsilon_1\,1_1 + \upsilon_2\,1_2 + \ldots + \upsilon_m\,1_m + \upsilon_{m+1}\,1_{m+1}}{\upsilon_1 Z_1 + \upsilon_2 Z_2 + \ldots + \upsilon_m Z_m + \upsilon_{m+1} Z_{m+1}}$$

and

$$b = \frac{\upsilon_1 X_1 + \upsilon_2 X_2 + \ldots + \upsilon_m X_m + \upsilon_{m+1} X_{m+1}}{\upsilon_1 Z_1 + \upsilon_2 Z_2 + \ldots + \upsilon_m Z_m + \upsilon_{m+1} Z_{m+1}}$$

or the simultaneous solution of any m of the $m + 1$ deviation equations with the proper values of $\upsilon_\iota\,\delta$ inserted, and in his solution for n equations when $n > m + 1$ as the solution for that combination among all the possible combinations of the n equations taken $m + 1$ at a time which gives the greatest least-greatest-deviation.

least, the straight line, $y = a + b\,x$, for a given set of observation points, $P_\iota\,(y_\iota\,.\,x_\iota)$, must be equidistant in the direction parallel to the axis of the dependent variable from those three observation points, $P_{i\,\pm\,\delta}\,(y_i\,.\,x_i)$, $P_{j\,\pm\,\delta}\,(y_j\,.\,x_j)$, $P_{k\,\mp\,\delta}\,(y_k\,.\,x_k)$, where $x_i \leqq x_k \leqq x_j$, which give deviations, $\delta = y_\iota - (a + b\,x_\iota)$, all of which are equal in absolute value to the least greatest-deviation for the set of observation points, and of which two are the same in sign and the third is of the opposite sign. The least greatest-absolute-deviation will then be half the distance in the direction parallel to the axis of the dependent variable from the point with the deviation of the opposite sign to the line containing the two points with deviations of the same sign, or

$$\delta = \left| \frac{(y_i - y_j)\,x_k + (x_j - x_i)\,y_k + (x_i\,y_j - x_j\,y_i)}{2\,(x_j - x_i)} \right| ,$$

and the required equation will be

$$y = \left(\frac{x_i\,y_j - x_j\,y_i}{x_i - x_j} \mp \delta\right) + \frac{y_i - y_j}{x_i - x_j}\,x$$

in which the sign before δ is the sign of the deviation of the point not contained in the line (1). To give the least greatest-

(1) The method of the least-greatest deviation, which was originally applied by Laplace by means of analytic solutions (**27, 11, 38**) and subsequently by J. B. J. Fourier by means of a proposed geometric solution (*Analyse des equations déterminées*, Paris, 1831, 81-84) to the determination of the parameters of linear equations, and by J. V. Poncelet to the determination of linear approximations to be substituted for the original expressions in the evaluation of radicals (*Sur la valeur approchée lineaire et rationelle des radicaux de la forme $\sqrt{a^2 + b^2}$, qui expriment la resultante de deux ou de plusieurs forces*, « Cours de mécanique appliquée aux machines », Metz, 1826, Sec. III, Note 1 ; *Sur la valeur approchée lineaire et rationelle des radicaux de la forme $\sqrt{a^2 + b^2}$, $\sqrt{a^2 - b^2}$, . . . ,* ibid., 3me ed., 1832; « J. f. d. Math. reine u. angew. », 13 : 277-291 (1835)) and of series (*Recherches sur le Calcul des Séries, ou application de la méthode de moyennes a la transformation ou calcul numérique et a la détermination des limites du reste des séries*, « Mém. pres. a l'Acad. roy. d. sc. », 6 : 785-872 (1835)), was finally applied by P. L. Tchebychev to the power polynomial (*Théorie des mécanismes connus sous le nom de parallélogrammes*, « Mém. pres. a l'Acad. imp. d. sc. de St. Pétersbourg », 7 : 539-568 (1854) ; *Sur les questions de minima qui se rattachent à la representation approximative des fonctions*, « Bul. de la Clas. phys.-math. de l'Acad. imp. d. sci. de St. Petersbourg », 16 : 145-150 (1858), « Mém. de l'Acad. imp. d. sci. de St. Petersbourg », 6oo Ser., Sci. math., phys. et nat., 9, Sci. math. et phys., 7 : 199-291 (1859)).

deviation in the dependent variable, the plane, $z = a + b x + c y$, for a set of observation points, $P_\iota (z_\iota . x_\iota , y_\iota)$, must be equidistant in the direction parallel to the axis of the dependent variable form the four observation points,

$$P_i (z_i . x_i , y_i), \; P_j (z_j . x_j , y_j), \; P_k (z_k . x_k , y_k), \; P_1 (z_1 . x_1 , y_1),$$

which give deviations, $\delta = z_\iota - (a + b x_\iota + c y_\iota)$, equal in absolute value to the least greatest-deviation for the set of observation points. If of the group of four observation points, $P_{i \pm \delta}$, $P_{j \pm \delta}$, $P_{k \pm \delta}$, $P_{1 \mp \delta}$, where (x_1 , y_1) lies within, or nearer one side of, the plane triangle the vertices of which are (x_i, y_i), (x_j, y_j), and (x_k, y_k), than any other of the points to a side of any other triangle, three give deviations which are the same in sign and the fourth gives a deviation which is of the opposite sign, the least greatest-deviation will be half the distance in the direction parallel to the axis of the dependent variable from the point with the deviation of the opposite sign to the plane containing the three points with deviations of the same sign, or

---

The mathematical theory of the solution of the least greatest-deviation polynomial was successively developed by P. KIRCHBERGER (*Ueber Tchebychefsche Annäherungsmethoden*, Inaug-Diss., Göttingen, 1902, 5-96, « Math. Ann. », 57 : 509-540 (1903)), E. BOREL (*Méthode d'approximation de Tchebicheff*, « Leçons sur les fonctions de variables réelles et les developpements en séries de polynomes », Paris, 1905, 82-92), J. W. YOUNG (*General Theory of Approximation by Functions Involving a Given Number of Arbitrary Parameters*, « Trans. Amer. Math. Soc. », 8 : 331-344 (1907)), L. TONELLI (*I polinomi d'approssimazione di Tchebychev*, « Ann. mat. », Ser. 3, 15 : 47-119 (1908)), F. SIBIRAMI (*Sulla rappresentazione approsimata delle funsioni, idem*, 16 : 208-231 (1909)), G. POLYA (*Sur an algorithme toujours convergent pour obtenir les polynomes de meilleure approximation de Tchebychef pour une fonction continue quelconque*, « Compt. rend. l'Acad. d. sci. », 157 : 840-843 (1913)), C. de la VALLÉE-POUISSIN (*Polinome d'approximation minimum*, « Leçons sur l'approximation des fonctions d'une variable réelle », Paris, 1919, 75-92), D. JACKSON (*On Functions of Closest Approximation*, « Trans. Amer. Math. Soc. », 22 : 117-128 (1921) ; *On Approximation by Functions of Given Continuity*, idem, 25 : 449-458 (1924) ; (5) ; *The Theory of Approximation*, « Amer. Math. Soc. Colloquium Pub. », 11 : 77-108 (1930)), and S. BERNSTEIN (*Lecons sur les propriétés extrémales et la meilleure approximation des fonctions analytiques d'une variable reélle*, Paris, 1926). In his mid-course method Smith (**29**) rediscovered the method of the least greatest-deviation and such of its theorems as were required as a basis for the first practical procedure for the numerical computation of the least greatest-deviation polynomial.

$$\delta = \left| \frac{\begin{array}{c}(y_i z_j + y_j z_k + y_k z_i - y_k z_j - y_j z_i - y_i z_k)\, x_1 + (x_k z_j + \\ + x_j z_i + x_i z_k - x_i z_j - x_j z_k - x_k z_i)\, y_1 + (x_i y_j + x_j y_k + \\ + x_k y_i - x_k y_j - x_j y_i - x_i y_k)\, z_1 + (x_k y_j z_i + x_j y_i z_k + \\ + x_i y_k z_j - x_i y_j z_k - x_j y_k z_i - x_k y_i z_j)\end{array}}{2\,(x_i y_j + x_j y_k + x_k y_i - x_k y_j - x_j y_i - x_i y_k)} \right| ,$$

and the equation,

$$z = \left(\frac{(x_i y_j z_k + x_j y_k z_i + x_k y_i z_j - x_k y_j z_i - x_j y_i z_k - x_i y_k z_j)}{(x_i y_j + x_j y_k + x_k y_i - x_k y_j - x_j y_i - x_i y_k)} \mp \delta\right) +$$
$$+ \frac{y_k z_j + y_j z_i + y_i z_k - y_i z_j - y_j z_k - y_k z_i}{x_i y_j + x_j y_k + x_k y_i - x_k y_j - x_j y_i - x_i y_k}\, x +$$
$$+ \frac{x_i z_j + x_j z_k + x_k z_i - x_k z_j - x_j z_i - x_i z_k}{x_i y_j + x_j y_k + x_k y_i - x_k y_j - x_j y_i - x_i y_k}\, y ,$$

in which the sign before δ is that of the deviation of the point not contained in the plane. If of the four observation points, $P_{i_{\pm\delta}}$, $P_{j_{\pm\delta}}$, $P_{k_{\mp\delta}}$, $P_{1_{\mp\delta}}$, where (x_i , y_i), (x_j , y_j), (x_k , y_k), and (x_1 , y_1) are the vertices of a trapezium, the two pairs give deviations of opposite sign, and determine two non-parallel lines, $P_{i_{\pm\delta}}$ $P_{j_{\pm\delta}}$ and $P_{k_{\mp\delta}}$ $P_{1_{\mp\delta}}$, the least greatest-deviation will be half the distance in the direction parallel to the axis of the dependent

---

The literal application of this recents mid-course method to the straight line would in effect require the determination of a parabola $y' = a + bx + c\,x^2$, for the selected points, $P_{i_{\pm\delta}}\,(y_i \cdot x_i)$, $P_{k_{\mp\delta}}\,(y_k \cdot x_k)$, and $P_{j_{\pm\delta}}\,(y_j \cdot x_j)$, giving these values of the dependent variable at these values of the independent variable, and of a second parabola, $\upsilon = \alpha + \beta\,x + \gamma\,x^2$, for the assumed points, $P(\mp 1, x_i)$, $P(\pm 1, x_k)$, and $P(\mp 1, x_j)$, giving the value of unity with the sign opposite to that of the deviation at these values of the independent variable, the linear combination of which obtained by putting $c = \delta\,\gamma$ so as to eliminate their quadratic terms,

$$y = [y' - \delta\,\upsilon = a + b\,x + c\,x^2 - \delta\,(\alpha + \beta\,x + \gamma\,x^2)] = (a - \delta\,\alpha) + (b - \delta\,\beta)\,x,$$

would be the equation of the line and $\delta = c/\gamma$ the least greatest-deviation. On the same basis the least greatest-deviation and the parameters of an n^{th} degree parabola may be determined for $n + 1$ points, selected to lie alternately on opposite extremes of the central trend of the observation points as the independent variable increases, from explicit expressions utilizing the Lagrangean polynomial for two variables, and might be determined for an $n, m, \ldots^{th}$ degree paraboloid from the polynomial for several variables (J. F. Steffensen, *Interpolation*, Baltimore, 1927, 21-22, 205, 216) contingent upon the elaboration of a rule and graphic artifice for the location and selection of the required points.

variable from one of the points, as $P_{k \mp \delta}$, to the plane containing the line, $P_{i \pm \delta}\, P_{j \pm \delta}$, and hence either of its points, as $P_{i \pm \delta}$, and parallel to the line, $P_{k \pm \delta}\, P_{1 \pm \delta}$ or

$$\delta = \left| \frac{\begin{array}{c}((y_j - y_i)(z_1 - z_k) - (y_1 - y_k)(z_j - z_i))(x_k - x_i) + \\ + ((z_j - z_i)(x_1 - x_k) - (z_1 - z_k)(x_j - x_i))(y_k - y_i) + \\ + ((x_j - x_i)(y_1 - y_k) - (x_1 - x_k)(y_j - y_i))(z_k - z_i)\end{array}}{2\,((x_j - x_i)(y_1 - y_k) - (x_1 - x_k)(y_j - y_i))} \right|,$$

and the equation,

$$z = \left(\frac{\begin{array}{l}((y_j - y_i)(z_1 - z_k) - (y_1 - y_k)(z_j - z_i))\,x_i + \\ + ((z_j - z_i)(x_1 - x_k) - (z_1 - z_k)(x_j - x_i))\,y_i\end{array}}{(x_j - x_i)(y_1 - y_k) - (x_1 - x_k)(y_j - y_i)} + z_i \mp \delta \right) -$$
$$- \frac{(y_j - y_i)(z_1 - z_k) - (y_1 - y_k)(z_j - z_i)}{(x_j - x_i)(y_1 - y_k) - (x_1 - x_k)(y_j - y_i)}\,x -$$
$$- \frac{(z_j - z_i)(x_1 - x_k) - (z_1 - z_k)(x_j - x_i)}{(x_j - x_i)(y_1 - y_k) - (x_1 - x_k)(y_j - y_i)}\,y,$$

in which the sign before δ is that of the deviation of the point not contained in the plane. If the pairs of observation points determine two parallel lines, and the set contain only four observation points, the deviations would be zero, and the plane represented by the required equation would contain both lines, but if the set contain more than four observation points, (x_i, y_i), (x_j, y_j), (x_k, y_k), and (x_1, y_1) would lie in a straight line, the least greatest-deviation would be half the distance in the direction parallel to the axis of the dependent variable between the two lines, and the plane represented by the required equation would pass through two points equidistant from, and be perpendicular to the plane containing, both lines. The procedure for the determination of the least greatest-deviation involves inspection of the set of plotted observation points, selection of trial combinations of the proper number of outlying observation points, and substitution of their values in the indicated deviation formula. As a criterion in the determination of the least greatest-deviation of a set of observation points such formulae have comparative value but lack directive efficacy. A trial group of observation points will be shown to give a deviation greater or less than the deviations of other previously tested groups with which it is compared. But it remains necessary to test all likely groups separately and to ascertain by elimination which is the required group.

The least greatest-deviation in the dependent variable can be determined more systematically through the selection of extreme observation lines or planes because the analytic criteria applicable to the sensible results then have the requisite directive efficacy to insure a definitive solution (1). The sum of the infinite powers of the absolute deviations in the dependent variable of a set of indirect observations, $\sum_{i=1}^{n} | y_i - (a + b\, x_i \ldots) |^{\infty}$, is rendered a minimum, when the greatest absolute deviation, $\delta = | y_\iota - - (a + b\, x_\iota + \ldots) |$, is rendered least. For a set of observation equations in any number of variables whatsoever, at least two deviations must occur equal in absolute value to the least greatest-deviation, which are of opposite signs. Let the equations, $y_1 = a + bx_1$ and $y_n = a + b\, x_n$, from a given set of observation equations, $y_i = a + b\, x_i$, arranged in the order of the algebraic magnitude of their deviations, represent a pair of extreme observation lines, for any set of values of the parameters, a, b, between these extreme observation lines and within the interval for which they continue to be the extreme observation lines. The sum of the extreme absolute deviations, D, is the sum of the greatest positive deviation and the greatest negative deviation with its sign changed, or

$$D = (y_n - a - b\, x_n) + (a + b\, x_1 - y_1) = (y_n - y_1) + (x_1 - x_n)\, b\,.$$

The increment to the extreme deviations sum, ΔD, for a

---

(1) In his introduction to an unfinished posthumous treatise, FOURIER (op. cit.), outlined a geometric construction for the solution of the least greatest-deviation linear equation. The deviation equations, $\delta = y_i - a - x_i b$, corresponding to the given observation equations, $y_i = a + x_i b$, express the algebraic deviation as a function of the values of the parameters. Let a and b be represented by a parameter point in a horizontal coordinate plane and δ by a deviation point on a vertical coordinate axis. The deviations equations may then be represented by deviations planes. But only the absolute values of the deviations come into consideration. The absolute deviations equations, $\delta = | y_i - a - x_i\, b |$, are represented by dihedrons consisting of the portions of the original deviations planes above the (a, b) plane and the reflection of these planes in the lines of their intersection with the (a, b) plane. The extreme planes of the system of absolute deviations dihedrons form a polyhedron with its convexity directed toward the horizontal plane. The vertex of this polyhearon is both the parameter point of the solution for the least greatest — deviation linear equation and the deviation point of the least greatest-deviation.

parameter increment, Δb, is $\Delta D = (x_1 - x_n) \Delta b$. The ratio of the extreme deviations sum increment to the parameter increment is simply $\Delta D / \Delta b = (x_1 - x_n)$, and the interval between a pair of extreme observation lines is increasing or decreasing as b increases according to whether the sign of $(x_1 - x_n)$ is plus or minus. In the region where the observation lines tend to converge that intersection of two extreme observation lines is to be selected which appears to lie nearest in a direction parallel to the axis of the dependent variable to an opposite extreme observation line. The values of the independent variable in the equations of the three observation lines, taken two at a time, each of the intersecting lines with the opposite line, are substituted in the expression for the increment ratio. If the sign of the ratio changes from minus to plus at the intersection, the extreme deviations sum becomes a minimum. With a set of observation planes, $z_i = a + b x_i + c y_i$, the extreme deviations sum is

$$D = (z_n - z_1) + (x_1 - x_n) b + (y_1 - y_n) c ,$$

and the ratio of the extreme deviations sum increment to the distance, Δs, in the (b, c)-plane, corresponding to the parameter increments, Δb, Δc,

$$\Delta D / \Delta s = \frac{(x_1 - x_n) \Delta b + (y_1 - y_n) \Delta c}{\sqrt{\Delta^2 b + \Delta^2 c}} .$$

In the region where the observation planes converge, the four extreme planes in nearest approach may consist of three extreme planes which intersect to form a trihedron with its vertex directed toward the opposite extreme plane, or of two pairs of extreme planes which intersect to form two opposite dihedrons with their apices directed toward each other. If for each member of the triad of intersecting observation planes taken with the opposite plane, the ratio is positive for all directions of departure from the intersection included in the sector of the (b, c)-plane in which any individual member of the triad of intersecting planes is an extreme plane, or if for each member of one pair of intersecting observation planes taken with each member of the opposite pair of intersecting planes, the ratio is positive for all directions of departure from the intersection of the projecting planes of the lines of intersection of the pairs of intersecting observation planes on the (b, c)-plane, represented by linear combinations of each pair of intersecting observation planes which eliminate a between their equations, included in the sector in which any two observation planes are both extreme planes, the extreme deviations sum becomes a minimum. Hence, taking three extreme observation lines, or four extreme observation planes, two at a time, each intersecting line or plane successively with the opposite line or plane, or each of one pair of intersecting planes with each of the opposite pair of intersecting planes, if the ratio is positive for each pair of opposite lines or planes for all directions of departure along the b-axis or in the (b, c)-plane, the distance between the intersection and the opposite line or plane, or the distance between the lines of intersection of each pair of intersecting planes at the intersection of their projecting planes on the (b, c)-plane, is the least extreme deviations sum or the least range in the dependent variable for the set of observation equations. Half of the least extreme deviations sum or the least half-range is the least greatest-deviation. The parameter point which is the mid-point of the least range is the required solution. If two opposite extreme lines or planes were parallel, the solution would be indeterminate within the interval for which both were extreme observation lines or planes. Only if all the observation lines or planes were parallel would no solution exist.

Both of these methods for indirect observations can readily be extended to apply to implicit functional observations. By the method employing observation points, the straight line, $x \cos \alpha + y \sin \alpha - \rho = 0$, for a given set of observation points, $P_\iota (x_\iota, y_\iota)$, must be equidistant in a perpendicular direction from the three outlying observation points, $P_{i \pm d} (x_i, y_i)$, $P_{j \pm d} (x_j, y_j)$, $P_{k \mp d} (x_k, y_k)$; hence the least greatest-absolute-normal-deviation will be

$$d = \left| \frac{(y_i - y_i)\, x_k + (x_j - x_i)\, y_k + (x_i\, y_j - x_j\, y_i)}{2\sqrt{(y_i - y_j)^2 + (x_j - y_i)^2}} \right| ,$$

and the equation,

$$\frac{(y_i - y_j)\, x + (x_j - x_i)\, y - (x_i\, y_j + x_j\, y_i)}{\sqrt{(y_i - y_j)^2 + (x_j - x_i)^2}} + d = 0\, .$$

Similar substitutions of the square root of the sum of the squared coefficients for the coefficient of the dependent variable in their denominators will convert the formulae for the least greatest-deviation in the dependent variable and the equations in slope-intercept form for the plane into formulae for the least greatest-normal-deviation and equations in normal form.

The least greatest-normal-deviation can be determined more systematically by a method employing observation curves or surfaces. The sum of the infinite powers of the absolute normal deviations, as $\sum_{i=1}^{n} | x_i \cos \alpha + y_i \sin \alpha - \rho |^{\infty}$, is rendered a minimum, when the greatest absolute deviation, as $d = | x_i \cos \alpha + y_i \sin \alpha - \rho |$, is rendered least. At least two deviations occur equal in absolute value to the least greatest-deviation and of opposite signs. If the polar vectorial epicycloids, $x_1 \cos \alpha + y_1 \sin \alpha = \rho_{(x_1, y_1)}$ and $x_n \cos \alpha + y_n \sin \alpha = \rho_{(x_n, y_n)}$, from a given set of observation equations, $x_i \cos \alpha + y_i \sin \alpha - \rho = 0$, are the extreme observation curves intersecting a ray at points other than the origin, the sum of the extreme absolute normal deviations for any set of values of the parameters, α within the sector for which both are extreme observation curves, and $\rho_{(x_1, y_1)} \leqq \rho \leqq \rho_{(x_n, y_n)}$, is

$$(x_n \cos \alpha + y_n \sin \alpha - \rho) + (\rho - x_1 \cos \alpha - y_1 \sin \alpha) =$$
$$= (x_n - x_1) \cos \alpha + (y_n - y_1) \sin \alpha .$$

The derivative of the extreme deviations sum with respect to α is

$$\frac{d}{d\alpha} (x_n - x_1) \cos \alpha + (y_n - y_1) \sin \alpha = (x_1 - x_n) \sin \alpha + (y_n - y_1) \cos \alpha .$$

The radial distance from the intersection of two extreme observation curves to the opposite extreme observation curve where the sign of the derivative changes from minus to plus is the least range; the least half-range, the least greatest-deviation; and the parameter point, which is the mid-point of the least range, the solution. With polar vectorial epispheroids, the extreme normal deviations sum is

$$(x_n \cos \alpha + y_n \cos \beta + z_n \cos \gamma - \rho) + (\rho - x_1 \cos \alpha - y_1 \cos \beta -$$
$$- z_1 \cos \gamma) = (x_n - x_1) \cos \alpha + (y_n - y_1) \cos \beta + (z_n - z_1) \cos \gamma .$$

The partial derivatives with respect to two essential vectorial angles are

$$\frac{\partial}{\partial \alpha}(x_n - x_1)\cos\alpha + (y_n - y_1)\cos\beta + (z_n - z_1)\sqrt{1 - \cos^2\alpha - \cos^2\beta} =$$

$$= (x_1 - x_n)\sin\alpha + (z_n - z_1)\frac{\sin 2\alpha}{2\cos\gamma}.$$

and

$$\frac{\partial}{\partial \beta}(x_n - x_1)\cos\alpha + (y_n - y_1)\cos\beta + (z_n - z_1)\sqrt{1 - \cos^2\alpha - \cos^2\beta} =$$

$$= (y_1 - y_n)\sin\beta + (z_n - z_1)\frac{\sin 2\beta}{2\cos\gamma}.$$

The method of least greatest-deviation in the dependent variable and the method of least greatest-normal-deviation are the methods of the mid-point of the least range (1).

Numerical Examples.

1. When an uncharged droplet falling in an electrostatic field becomes charged with an electron, the resulting change in its rate of fall corresponds to an alteration in its apparent weight. The electrostatic charge of an electron, e, may be determined from the electrostatic field strength, F, and the alteration in apparent weight, W, by means of the equation, $F e = W$. Treated as weighted observations, the values of $e = W_i/F_i$ calculated from a set of determinations of F and W (**32**) and arranged in the order of their magnitude, together with their respective weights, F_i, locate a weighted median (Table I) (2).

---

(1) A variant of the method of the least greatest-deviation or the mid-point of the least range, which, though apart from the succession of methods of least power-sums of the absolute deviations, might be of some significance. is the method of the least quartile or median deviation (so-called probable error) or the mid-point of the least interquartile range. The nearest approach of the first and the third quartile loci of a given set of observation lines, curves, planes, or surfaces defines their least interquartile range. Half of the least interquartile range or the least semi-interquartile range is the least quartile or median deviation, and the mid-point of the least interquartile range the solution.

(2) The weight decrement in grams is derived by calculation from the experimental data by means of the formula,

$$W = \frac{4}{3}\pi\left(\frac{9 s \eta}{2 t_g}\right)^{\frac{3}{2}}\left(\frac{1}{g(\sigma - \rho)}\right)^{\frac{1}{2}}\frac{t_g - t_f}{t_f},$$

in which s is the distance of fall in centimeters, t_g the duration in seconds of fall under gravity, t_f the duration of fall under the influence of the field,

TABLE I.

$W_i = 1.47$	1.14	3.70	2.04	3.27×10^{-9}
$F_i = 3.65$	2.11	6.67	2.85	4.10
$e = 4.02$	5.38	5.55	7.18	7.97×10^{-10}
$\sum_{i=1}^{m} F_i = 12.43 > 6.95 = \sum_{m+1}^{n} F_i$				
$\sum_{1}^{m-1} F_i = 5.76 < 13.62 = \sum_{m}^{n} F_i$				

2. In physiological experiments an entire animal is often subjected to some physical or chemical influence when only a single action on a specific tissue or organ or a single effect on a specific function is to be determined. When quantitative results are required, in this situation where it is in the nature of the case impossible to secure that selective constancy of all relevant conditions generally obtainable in the conduct of physical measurements, the employment of the weighted median may be singularly appropriate. In biological assay, where a representative value is essential, the animals will vary not only in size but also in tolerance or susceptibility to the action of the given drug. Preliminary to an assay of digitalis bodies by the cat method it may be advisable to determine the minimal lethal dose per unit weight of intravenous crystalline ouabain for a sample of the population of cats to be used in the standardization **(33)** (Table II).

3. The constants for a second approximation to the generalized law of error with its origin at the median may be so determined by the method of percentiles that the observed and calculated fractions of a frequency distribution which lie between the median and certain arbitrary limits shall be as nearly

---

η the coefficient of viscosity of air, g the acceleration of gravity. σ the density of the oil, and ρ the density of air, which involves a systematic error in as much as Stokes' law for the radius of a sphere falling through a resisting medium is not quite valid for ultra-microscopic dimensions, and the electrostatic field strength from $F = V/d$ in which V is the electromotive force in statvolts, and d the distance in centimeters between the plates.

TABLE II.

MLD (*mg./kg.*) =	0.084 . . .	0.085	0.09	*0.09*	0.09	0.098	0.099	0.110
W_i (*kg.*) =	2.2 . . .	4.32	2.0	2.04	2.38	1.47	1.61	1.44
			$\sum_{i=1}^{m-1} w_i = 8.52$	2.04	$6.84 = \sum_{m+1}^{n} w_i$			

median, $x_M = 0.5$, and the several class-limits of the tabulated data, x_i, represent values of the probability integral, $F\ (\xi_i)$; the corresponding arguments, $\mp \xi_i\ (= \mp x/c)$, enter into a sequence of equations, $|\, x_M - x_i \,| = c\, \xi_i + 2/3\ j\ c\ \xi_i^2$, where c is the modulus, equal as possible (**34,16**). The decimal fractions between the and j, a measure of skewness. Subtraction of each equation from its successor yields a set of equations, $x_{i+1} - x_i = c\,(\xi_{i+1} - \xi_i) + 2/3\, j\, c\,(\xi_{i+1}^2 - \xi_i^2)$, or taking a class-limit as origin and the class-interval as unity, $1 = c\,(\xi_{i+1} - \xi_i) + 2/3\ j\ c\ (\xi_{i+1}^2 - \xi_i^2)$, and dividing through by $(\xi_{i+1} - \xi_i)$, $\frac{1}{\xi_{i+1} - \xi_i} = c + 2/3\, j\, c\,(\xi_{i+1} + \xi_i)$, to be solved for the values of the constants.

The final set of equations, $\frac{1}{\xi_{i+1} - \xi_i} = a + (\xi_{i+1} + \xi_i)\ b$, obtained by putting $a = c$ and $b = 2/3\ j\ c$, is numerically of the simplest form expressing the relationships between the arguments derived from the data and the constants required for the curve. But to treat its left members as values of the dependent variable is to attribute error to the reciprocals of the differences of the successive arguments, $\frac{1}{\xi_{i+1} - \xi_i}$, and to deny it to their sums, $\xi_{i+1} + \xi_i$. Hence it would seem antecedently preferable to enter the values of the sums and the reciprocals of the differences in an equation of the form, $(\xi_{i+1} + \xi_i) \cos \alpha + \frac{1}{\xi_{i+1} - \xi_i} \sin \alpha - - \rho = 0$, in order the distribute the error somewhat more equitably between them.

The percentile distribution of the number of men paid at different wage rates in the principal industries in a selected year (**35**) after these transformations and substitutions yields the respective sets of linear (34,16) and polar observation equations,

$$1.23 = a - 3.59\,b \qquad\qquad -3.59 \cos\alpha + 1.23 \sin\alpha = \rho$$
$$\cdots\cdots \quad \text{and} \quad \cdots\cdots$$
$$2.99 = a + 2.47\,b \qquad\qquad 2.47 \cos\alpha + 2.99 \sin\alpha = \rho$$

represented by the network of observation lines (Fig. 3) and observation curves (Fig. 4), and soluble for the rectangular coordinates and polar parameters by any of the four graphic-analytic methods.

For most lesser-deviations in the dependent variable and for most lesser-normal-deviations their solution requires the method of the modal point and depends on the proper selection of the respective modal points, M ($a = 1.735$, $b = 0.345$) (Fig. 3), and in the region of the mutual intersections of three curves M ($\alpha = 110^0$, $\rho = 1.6$) (Fig. 4) which is equivalent to ($a = 1.6 \sec\alpha = 1.702$, $b - \cot\alpha = 0.364$).

For least deviations in the dependent variable the solution requires the method of median loci (**16**). The a-locus is a chain of linked areas, the b-locus, a broken line; the point of coincidence of the a-and b-loci, the intersection, (1.774, 0.3426), of equations, $2.36 = a + 1.71\,b$ and $1.13 = a - 1.88\,b$, is the solution (Fig. 3). The increment ratio for the direction of departure from the intersection for the segment of the b-locus on the line representing the first equation is

$$\frac{\Delta D}{\Delta s} = \frac{1.71\ (0 + 1) - 1\ (4.38 - 4.71)}{\sqrt{1.71^2 + (-1)^2}} = 1.03\,;$$

for the segment on the line representing the second equation,

$$\frac{\Delta D}{\Delta s} = \frac{-1.88\ (0 - 1) - 1\ (4.38 - 4.54)}{\sqrt{(-1.88)^2 + (-1)^2}} = 0.956\,;$$

for the segments of the a-locus,

$$\frac{\Delta D}{\Delta s} = (-1.45 - 0.59)/-1 = 2.04 \text{ and } \frac{\Delta D}{\Delta s} = (2.14 + 3.00)/1 = 5.14:$$

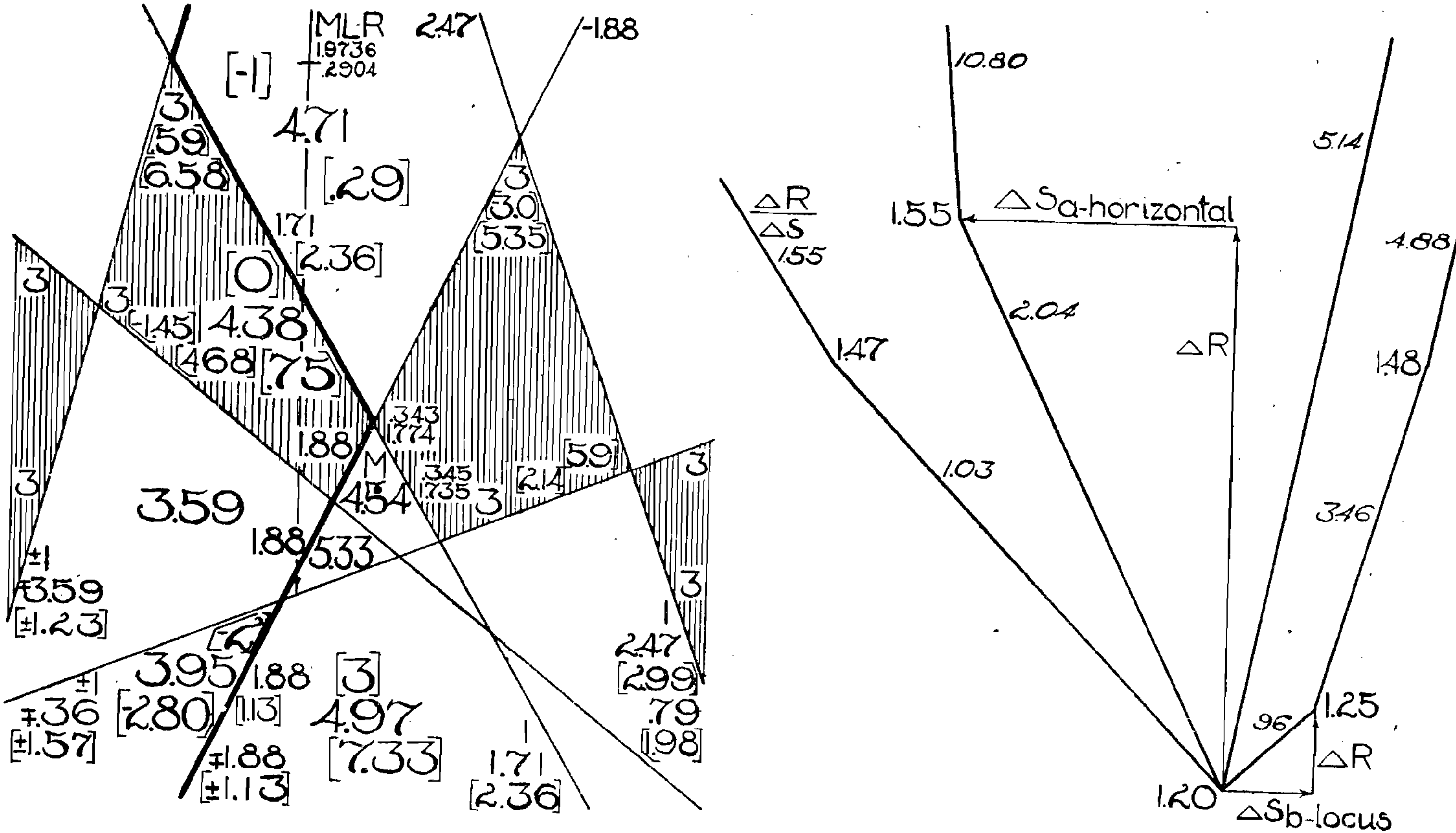
hence the intersection, with the deviations sum,

$$D_{y\,.\,x} = (0.29 - 0.75) + 1.774\,(0 + 1) + 0.3426\,(4.38 - 4.71) = 1.201,$$

is the minimum point (Fig. 3) (1).

---

(1) The linear equation, $Ax + By = 1$, in what may be designated reciprocal-intercept form, since its intercepts are $x = 1/A$ and $y = 1/B$, when set up as an observation equation by the substitution of observed values for its unknowns becomes $Ax_i + By_i = 1$ with the intercepts $A = 1/x_i$ and $B = 1/y_i$. Solution of a set of observation equations based on



ML-R
1.9736
.2904
2.47
-1.88
[-1]
4.71
[.29]
3
[.59]
[6.58]
1.71
[2.36]
[0]
4.38
3
[-1.45]
[468]
[.75]
3.43
1.774
1.88
M
4.54
3.45
1.735
3
[2.14]
[5.9]
3.59
1.88
5.33
±1
3.59
[±1.23]
3.95
[-2.80]
1.88
[3]
[1.13]
4.97
[7.33]
±.36
[±1.57]
∓1.88
[±1.13]
1.71
[2.36]
2.47
[2.99]
.79
[1.98]
3
[3.0]
[5.35]
△R
△S
1.55
1.47
1.03
10.80
1.55
△Sa-horizontal
2.04
△R
5.14
4.88
1.48
3.46
.96
1.25
△R
1.20
△Sb-locus

Fig. 3

For least normal deviations the solution requires the method of the median locus. The ρ-locus is a circuit of linked areas (Fig. 4). The linkage of the two smaller curvilinear quadrilaterals is the intersection of equations, 1.71 cos α + 2.36 sin α = ρ and — 1.88 cos α + 1.13 sin α = ρ, for which tan α = $\frac{-1.88 - 1.71}{2.36 - 1.13} = -2.9187$ or α = 108° 54′ 45″. The derivative at this angle is

$$\frac{\partial D}{\partial \alpha} = 0.94602\,(-1.45 - 0.59) - 0.32412\,(6.58 - 4.68) = -2.55$$

on arrival, and

$$\frac{\partial D}{\partial \alpha} = 0.94602\,(2.14 + 3.00) - 0.32412\,(5.35 - 5.91) = 5.05$$

on departure, changing from negative to positive, thus establishing a minimum, with a deviations sum,

$$D_{x,y} = -0.32412\,(0.59 + 1.45) + 0.94602\,(6.58 - 4.68) = 1.13\,,$$

which is the only minimum occurring at any point other than the origin (Fig. 4). The polar parameter point, (108° 54′ 45″, 1.6983), of the solution for least normal deviations is equivalent

---

the linear equation in slope-intercept form by the method of median loci renders the sum of the deviations of the values of $(y_i - bx_i)$ from their median, which is equal to the sum of the deviations in the dependent variable, a minimum, and also renders the sum of the deviations of the values of $\frac{y_i - a}{x_i}$ from their median, which is without intrinsic significance, a minimum; solution of a set of observation equations based on the linear equation in reciprocal-intercept form by the method of median loci renders the sums of the deviations of the values of $\frac{1 - Ax_i}{y_i}$ and of $\frac{1 - By_i}{x_i}$ from their medians a minimum. The solution by the method of median loci of the set of observation equations in reciprocal-intercept form involved in the present problem,

$$1 = c\,(\xi_{i+1} - \xi_i) + 2/3\, j\, c\,(\xi_{i+1}^2 - \xi_i^2)\,,$$

numerically

$$0.886\,a - 1.669\,b = 1$$

$$\cdots\cdots\cdots\cdots$$

$$0.335\,a + 0.826\,b = 1\,,$$

for equations of unitary and of different weights was begun but left indeterminate by Edgeworth (**18**) and completed by Bowley (**16**) for equations of equal weight.

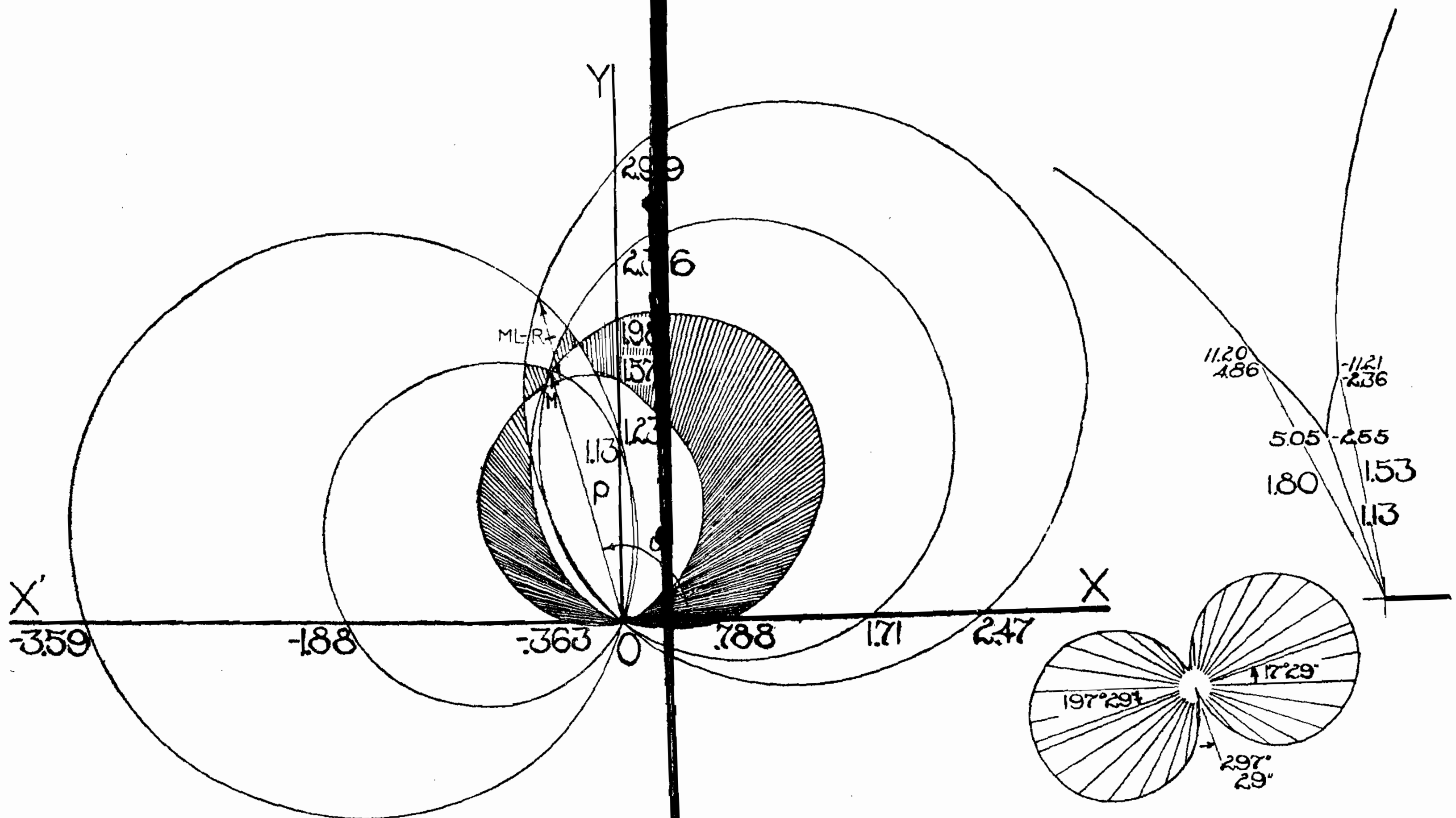

Y
X'
X
2.99
2.36
1.98
1.57
1.23
1.13
P
ML-R+
M
-3.59
-1.88
-.363
O
.788
1.71
2.47
11.20
4.86
-11.21
-2.36
5.05
-2.55
1.80
1.53
1.13
17°29'
197°29'
297°
29'

Fig. 4

to the rectangular coordinates parameter point, (1.774, 0.3426), of the solution for least deviations in the dependent variable.

For least greatest-deviation in the dependent variable or least greatest-normal-deviation the solution requires the method of the mid-point of the least range. The least range in the depend variable occurs in the region between the intersections of the two pairs of lines, $1.13 = a - 1.88\ b$ and $1.57 = a - 0.36\ b$, and $2.99 = a + 2.47\ b$ and $1.23 = a - 3.59\ b$ (Fig. 3). Taking the first equation of each pair, the increment ratio, $\Delta D/\Delta b = - 1.88 - 2.47 = - 4.35$; the second equation of the first pair with the first of the second, $\Delta D/\Delta b = - 0.36 - 2.47 = - 2.83$; and the second equation of each pair, $\Delta D/\Delta b = - 0.36 + 3.59 = 3.23$. The least range, which occurs between the intersection of the second pair, (2.2727, 0.2904), and the opposite point on the second equation of the first pair, (1.6745, 0.2904), is 0.5982; the least half-range, 0.2991; and the mid-point of the least range, (1.9736, 0.2904) (Fig. 3).

The least normal range occurs in the region between the intersections of $- 1.88 \cos \alpha + 1.13 \sin \alpha = \rho$ and $- 0.36 \cos \alpha + 1.57 \sin \alpha = \rho$ at 106° 8′ 42″ and of $2.47 \cos \alpha + 2.99 \sin \alpha = \rho$ and $- 3.59 \cos \alpha + 1.23 \sin \alpha + \rho$ at 106° 11′ 41″ (Fig. 4). Taking the first equation of each pair at the angle of the intersection of the first pair, the derivative,

$$\frac{dD}{d\alpha} = 0.96056\,(- 1.88 - 2.47) - 0.27807\,(2.99 - 1.13) = - 4.70;$$

the second equation of the first pair with the first of the second at the same angle,

$$\frac{dD}{d\alpha} = 0.96056\,(- 0.36 - 2.47) - 0.27807\,(2.99 - 1.57) = - 3.11;$$

and at the angle of the intersection of the second pair,

$$\frac{dD}{d\alpha} = 0.96032\,(- 0.36 - 2.47) - 0.27890\,(2.99 - 1.57) = - 3.12;$$

and the second equation of each pair at this angle,

$$\frac{dD}{d\alpha} = 0.96302\,(- 0.36 + 3.59) - 0.27890\,(1.23 - 1.57) = 3.20.$$

The least range, which occurs between the intersection of the second pair, (106° 11′ 41″, 2.187), and the opposite point on the second equation of the first pair, (106° 11′ 41″, 1.609), is 0.578; the least half-range, 0.289; and the mid-point of the least range, the polar parameter point, (106^0 11′ 41″, 1.898), equivalent to the rectangular coordinates parameter point, (1.9736, 0.2904).

4. The personal equation of an astronomical observer measuring the vertical diameter of Venus with a specific instrument, y, consists of a constant correction, a, and a relative correction, b, which varies linearly with the actual diameter, x (**36**). Each measurement gives an observation equation for the difference, y_i, between an observed diameter and the corresponding diameter according to the Nautical Almanac, x_i, of the form, $y_i = a + b\, x_i$. In the solution of a set of 97 such equations for an individual member of the staff of the Greenwich Observatory (**37**) by the method of median loci (**14**) the otherwise divergent a- and b-loci intersect at (1.372, 0.0256), coincide throughout the segment from (1.243, 0.0297) to (1.218, 0.0302), and come into close juxtaposition in the region of (1.109, 0.0350) and (1.106, 0.0351). But the increment ratios for the relevant segments of the two loci locate the parameter point, (1.243, 0.0297), as the solution for least deviations (Fig. 5). Intersection or coincidence of the median loci merely represent points or tracts of equal deviation sums from the several loci. The increment ratio designates the singular point of the final solution.

5. A geometric criterion has been proposed for locating the singular point of the final solution when the median loci coincide through two or more segments :

« Let [a] segment of [a] line ... be a locus of double medians : meaning that [it] will be an [a]-median ; and will also be a [b]-median. Let [a segment of a second] line [which intersects that of the first] be likewise a locus of double medians. And let there be no other double median points or loci. The required point then must be on either [the one line segment] or [the other] ; and it will be on both, at their intersection... For if it were at any other point on either of the segments, the sum of the weighted residuals would be greater by the weighted perpendicular let fall from that point on the other line. Now let [a segment of a third line intersecting one of the segments of the first two lines] also be a locus of double medians. By parity the required point must be either [one] or [the other of the two intersections of the three line segments]. It will be the intersection of those two lines for which the weighted perpendicular let fall from the intersection on the third line is the less ». (**18**).

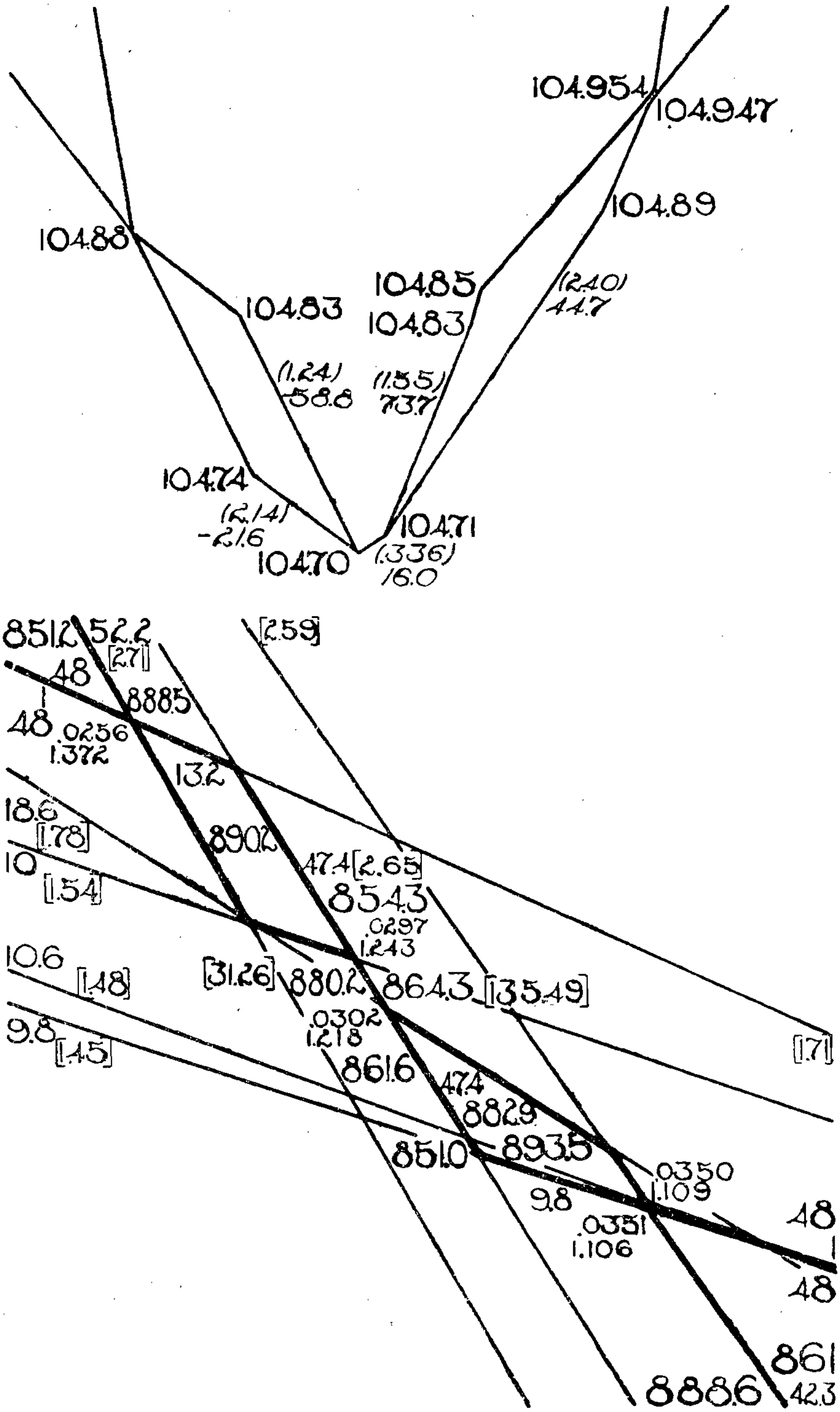

104.954
104.947
104.89
104.88
104.83
104.85
104.83
(2.40)
44.7
(1.24)
58.8
(1.55)
73.7
104.74
(2.14)
-21.6
104.71
104.70
(.336)
16.0
851.2
52.2
[2.59]
[2.7]
48
888.5
48
.0256
1.372
13.2
18.6
[1.78]
890.2
47.4
[2.65]
10
[1.54]
854.3
.0297
1.243
10.6
[1.48]
[31.26]
880.2
864.3
[35.49]
.0302
1.218
9.8
[1.45]
[1.7]
861.6
47.4
882.9
851.0
893.5
.0350
1.109
9.8
.0351
1.106
48
48
861
888.6
42.3

Fig. 5

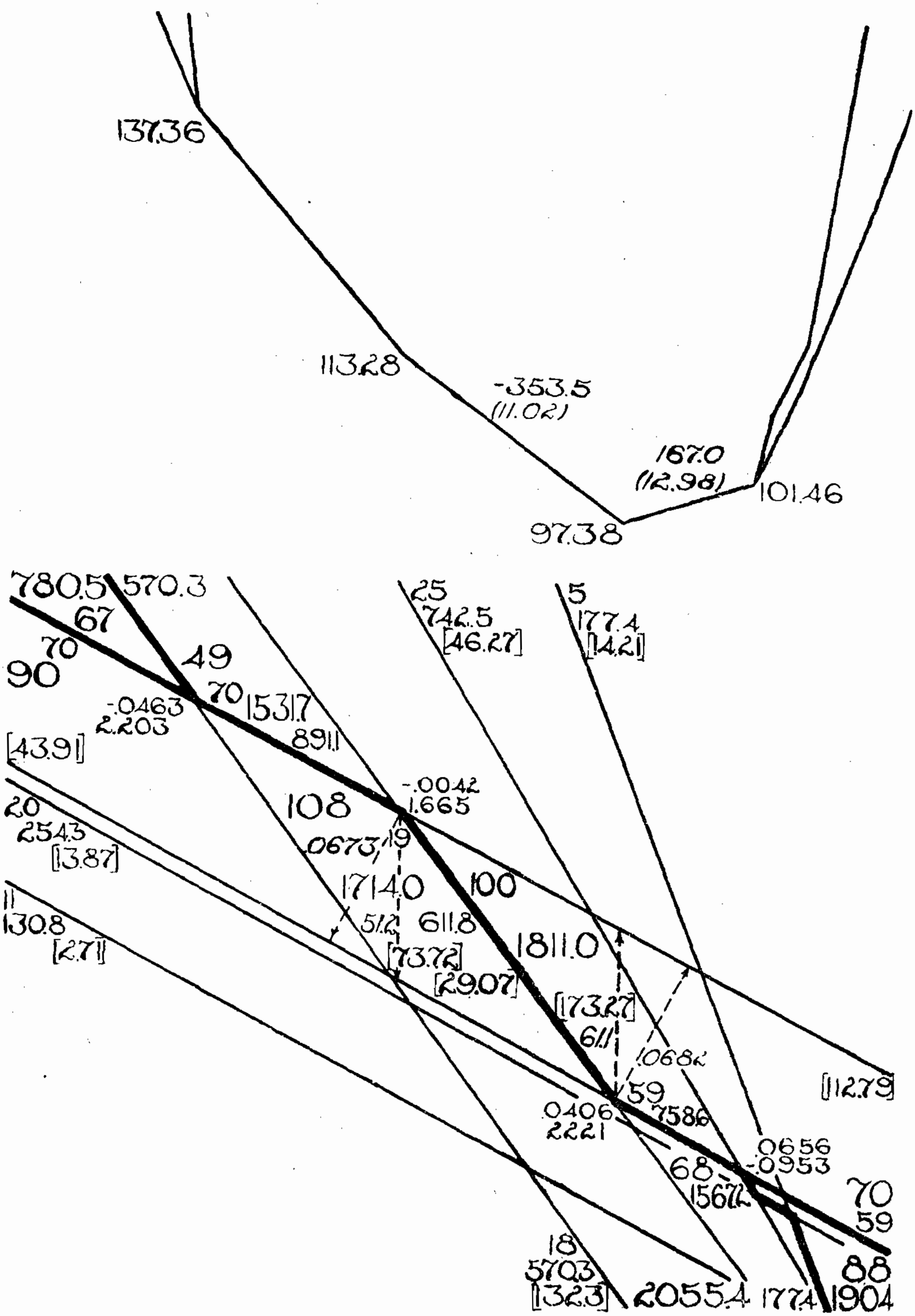

137.36
113.28
-353.5
(11.02)
167.0
(12.98)
101.46
97.38
780.5
570.3
67
70
90
49
70
1531.7
-.0463
2.203
891.1
[43.91]
108
-.0042
1.665
20
254.3
[13.87]
.0673
19
1714.0
100
51.2
611.8
11
130.8
[27.1]
[73.72]
[29.07]
1811.0
[173.27]
61.1
25
742.5
[46.27]
5
177.4
[14.21]
1068.2
[112.79]
59
.0406
2.221
758.6
68
1567.2
.0656
-.0953
70
59
88
18
570.3
[132.3]
2055.4
177.4
1904

Fig. 6

This criterion may be put to its crucial test by applying it to the eight equations, formed by the summation of two groups of the observation equations of the preceding example for each of four different observers (**36**), employed in its original demonstration. The *a*-and *b*-loci coincide throughout the broken line, (2.203, — 0.0463) (1.665, — 0.0042) (0.2221, 0.0406) (— 0.0953, 0.0656) ; the analytic criterion designates (0.2221, 0.0406) the solution (Fig. 6). The weighted perpendicular, or deviation in the dependent variable, from the point, (1.665, — 0.0042), to the line, 43.91 = 59 *a* + 758.6 *b*, or

$$\delta_{y_i \cdot x_i} = \left| \sqrt{59^2 + 578.6^2}\, \frac{59\,(1.665) + 758.6\,(-0.0042) - 43.91}{\sqrt{59^2 + 578.6^2}} \right| = 51.2,$$

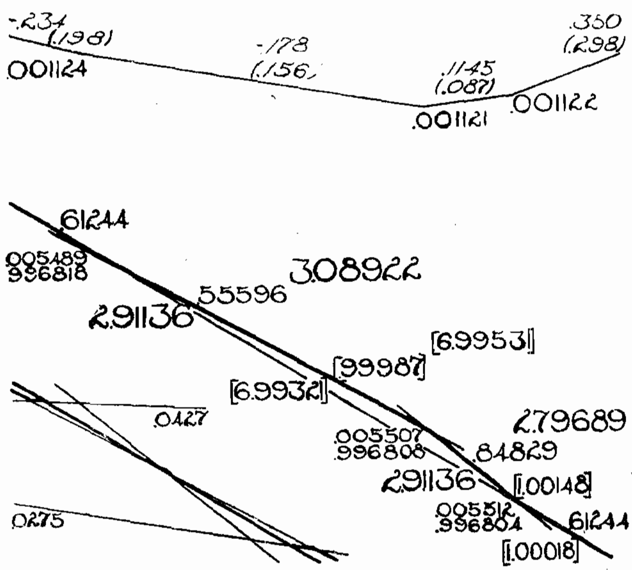


Fig. 7

nevertheless is smaller than that from (0.2221, 0.0406) to 112.79 = 70 *a* + 891.1 *b*, or

$$\delta_{y_i \cdot x_i} = |\, 70\,(0.2221) + 891.1\,(0.0406) - 112.79\,| = 61.1.$$

The geometric criterion, deduced with reference exclusively to the lines immediately involved in the coincident median loci, will function in any particular case only by sufferance of the remaining lines which must be grossly symmetrical in distribution relative to the loci. Obviously in the present example the outlying lines, $46.27 = 25\,a + 742.5\,b$ and $14.21 = 5\,a + 177.4\,b$, since they are both on the same side of the median loci, are able by virtue of the disproportionate steepness of their slopes to counteract and reverse the local influence of the relationships of the segments of the median loci. The proposed geometric criterion is therefore a generalization of limited and uncertain applicability.

6. Fifteen observations of the ratio of the length of a seconds pendulum to that of the standard seconds pendulum, r, and the

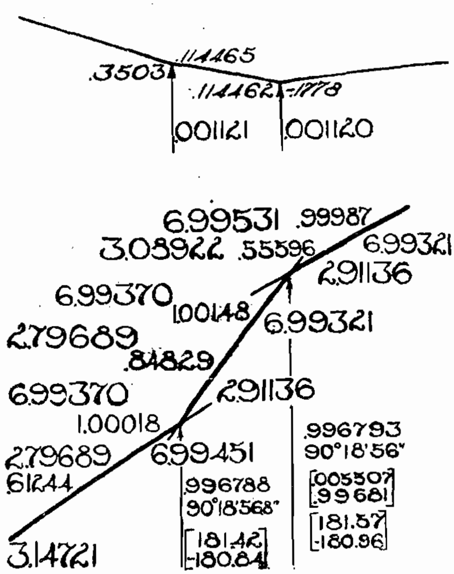


Fig. 8

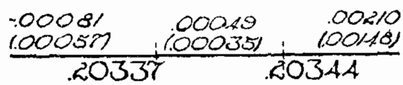
square of the sine of the latitude, Φ, used in a classical calculation of the ellipticity of the earth (**38**), substituted in observation equations treating the ratio as the dependent variable (**18**), the sine squared latitude and the ratio as conjoint variables, and the sine squared latitude as the dependent variable, solved by the method of median loci (Figs. 7,9), or the method of the median locus (Fig. 8), when, since geometric construction is impracticable because the range of values of r is too narrow relative to that of $\sin^2\,\Phi$ to permit the requisite representation on the same scale, the ρ-locus is plotted by substitution of successive trial values of α, yield corresponding equations, as

$$r = a_{r \cdot \sin^2 \Phi} + \sin^2 \Phi \, b_{r \cdot \sin^2 \Phi}$$
$$\sin^2 \Phi \cos \alpha + r \sin \alpha - \rho_{\sin^2 \Phi, r} = 0 ,$$
$$\sin^2 \Phi = a_{\sin^2 \Phi \cdot r} + r \, b_{\sin^2 \Phi \cdot r} ,$$

and

$$r = 0.99681 + 0.0055074 \sin^2 \Phi ,$$
$$- 0.0055073 \sin^2 \Phi + 0.99998 \, r - 0.99673 = 0 ,$$
$$\sin^2 \Phi = - 180.843 + 181.423 \, r ,$$

the first two of which are equivalent (1).



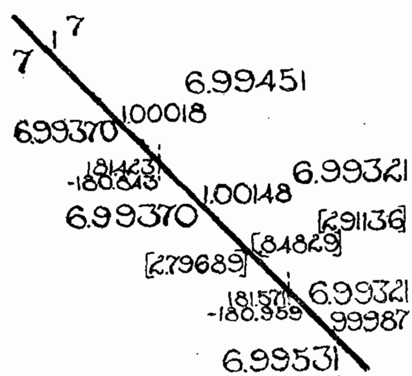


Fig. 9

7. The interrelationship of the supply and the price of a farm crop was investigated through the employment as data of

---

(1) The solution for the Boscovich-Laplace method as corrected by Bowditch (**38**), $r = 0.996838 + 0.005472 \sin^2 \Phi$, is the intersection of the mean observation line, $0.99923 = a + 0.43710\, b$ with the particular observation line, $1.00148 = a + 0.84829\, b$; the inverse solution, $\sin^2 \Phi = - 182.13329 + 182.75111\, r$, the intersection of the mean line, $0.43710 = a + 0.99923\, b$, with the particular line, $0.84829 = a + 1.00148\, b$: the inverse solutions are in this case equivalent equations with reciprocal

the percentile deviation of its yearly production from the preceding-five-year average and that of its purchasing power represented by its farm price multiplied by the ratio of its farm price index to a corresponding general wholesale commodity price index (1). The power equation was selected to conciliate economic generalities with an analytic function for the concomitant variation of the resulting indices of the production of the crop and the purchasing power of its farm price (**39**).

The sensible solution for most lesser-normal-logarithmic-deviations by the method of the modal point is (32° 15', 2.7970) (Fig. 10, Table III).

The solution for least normal deviations by the method of the median locus is distinct from and intermediate to the solutions for least deviations in either dependent variable by the method of median loci (Fig. 10, Table III) (2).

---

slopes and identical intercepts of which that for the axis of the respective dependent variable is explicit.

The solution of Laplace (**38**), for the least greatest-deviation, 0.00018, $r = 0.99687 + 0.0055399 \sin^2 \Phi$, is the midpoint of the interval between the intersection of the lines, $0.99669 = a$ and $1.00137 = a + 0.84478\ b$, and the line, $0.99877 = a + 0.31142\ b$.

(1) That the mathematical validity and economic significance of such indices have been vigorously impugned (B. Jones, *Horses and Apples, A Study of Index Numbers*, New York, 1934) does not impair the abstract interest of the present problem as a numerical example.

(2) The solution for the least y Σ deviations sum given by the intersection of the observation curves,
$2.0828 \cos \alpha + 1.9294 \sin \alpha = \rho$ and $1.9542 \cos \alpha + 2.0899 \sin \alpha = \rho$ (Fig. 10), where

$$\frac{\partial D}{\partial \alpha} = 1.59930^2\,(38.5171 - 38.5947) = -0.1988\,,$$

$$\frac{\partial D}{\partial \alpha} = 1.59930^2\,(38.6457 - 38.5947) = 0.1306\,,$$

and

$$D = 1.2481\,(38.5947 - 38.5171) + (38.8822 - 37.2771) = 1.7019\,,$$

is equivalent to the intersection of the observation lines, $1.9294 = a + 2.0828\ b$ and $2.0899 = a + 1.9542\ b$, where

$$\frac{\Delta D}{\Delta s} = \frac{-(38.5171 - 38.5947)}{\sqrt{1^2 + 2.0828^2}} = 0.0336\,,$$

$$\frac{\Delta D}{\Delta s} = \frac{(38.6457 - 38.5947)}{\sqrt{1^2 + 1.9542^2}} = 0.0233\,,$$

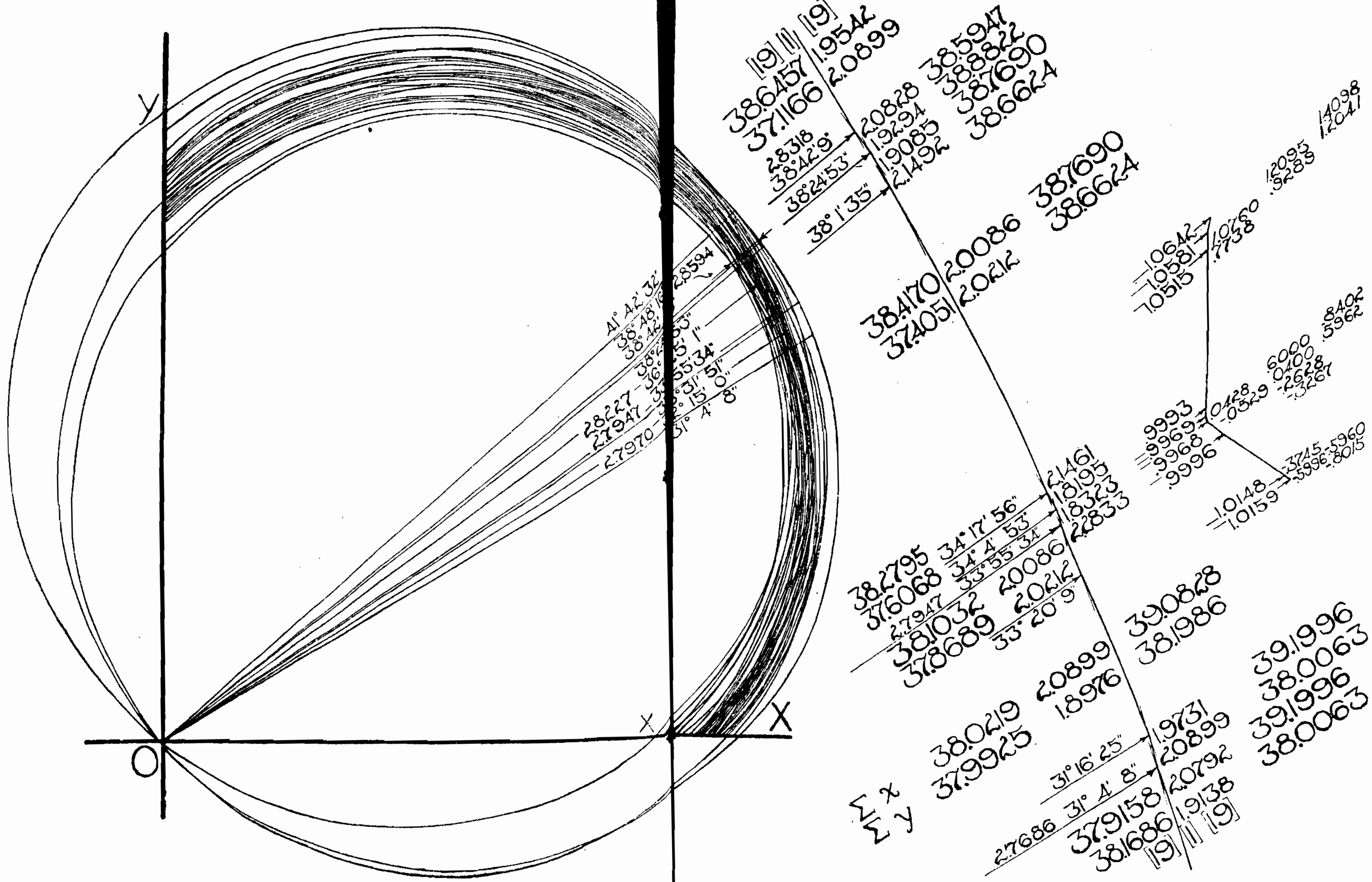


Fig. 10

Table III.

Most Lesser-Deviations	$0.84573 \log x + 0.53361 \log y - 2.7970 = 0$	$32^\circ\ 15'$	$y = 174404\ x^{-1.58493}$ $x^{0.84573}\ y^{0.53361} = 626.6$ $x = 2028.3\ y^{-0.63094}$	
Least Deviations		$38^\circ\ 24'\ 53''$	$y = 35967\ x^{-1.2610}$ (Rhodes)	$D_y = 395$
	$\log y = 4.52886 - 1.24806 \log x$	$38^\circ\ 42'\ 9''$	$y = 33795\ x^{-1.2481}$	$D_{\log y} = 1.7019$ (0.0436)
	$0.82076 \log x + 0.55813 \log y - 2.7947 = 0$	$33^\circ\ 55'\ 34''$	$y = 101777\ x^{-1.48645}$ $x^{0.82976}\ y^{0.55813} = 623.3$ $x = 2334.3\ y^{-0.67263}$	$\Delta_{\log \varrho} = 0.9968$ (0.0256)
	$\log x = 3.23226 - 0.60250 \log y$	$31^\circ\ 4'\ 8''$	$x = 1707.1\ y^{-0.60250}$	$D_{\log x} = 1.1860$ (0.0305)
Least Squares	$\log y = 4.28066 - 1.12205 \log x$ (Warren-Pearson)	$41^\circ\ 42'\ 32''$	$y = 19083\ x^{-1.1220}$ (Warren-Pearson)	$S_{\log y} = 0.04879$
	$\log y' = 4.24409 - 1.12205 \log x$ (Warren-Pearson adjusted)		$y' = 17539\ x^{-1.1220}$ (Warren-Pearson adjusted)	
	$0.80472 \log x + 0.59365 \log y - 2.8227 = 0$	$36^\circ\ 25'\ 1''$	$y = 56859\ x^{-1.3555}$ $x^{0.80472}\ y^{0.59365} = 664.8$ $x = 3219.2\ y^{-0.73772}$	$S_{\log \varrho} = 0.03073$
	$\log x = 3.35746 - 0.66266 \log y$	$33^\circ\ 31'\ 51''$	$x = 2277.5\ y^{-0.66266}$	$S_{\log x} = 0.03750$
Least Greatest Deviation	$0.77929 \log x + 0.62666 \log y - 2.8594 = 0$	$38^\circ\ 48'\ 16''$	$y = 36550\ x^{-1.24357}$ $x^{0.77929}\ y^{0.62666} = 723.4$ $x = 4668.4\ y^{-0.80415}$	$LGD_{\log \eth} = 0.0632$

Notwithstanding that a representation of an interrelationship was the desideratum in the original inquiry, purchasing power was treated as the dependent variable in the equation computed by the method of least squares (Table III) (**39**). In order to render it indicative of the effect of supply on price this equation was subjected to an adjustment consisting of the application of the transformation, $y' = y/r$, where r is the ratio of the value of the base of the independent variable index to the corresponding value of the dependent variable index, which, however, leaves the exponent, the essential measure of the relationship, unchanged (Table III) (**39**). The more directly significant solution would have been that for least squared normal deviations (Table III). The equation by least squares with supply treated as the dependent variable completes the set (Table III).

The solution for the least greatest-normal-deviation by the method of the midpoint of the least range is located between $2.1673 \cos\alpha + 1.9685 \sin\alpha = \rho$ and the intersection of $2.0492 \cos\alpha + 1.9138 \sin\alpha = \rho$ and $1.8751 \cos\alpha + 2.1303 \sin\alpha = \rho$ (Fig. 10, Tables III).

The angles of the normals to the equations for x-deviations, normal deviations, and y-deviations derived by any of the type methods, and to corresponding forms of equation solved for most lesser-deviations, least deviations (1), least squares, and least grea-

---

and

$$D = 4.52886\,(19 - 19) - 1.24806\,(38.6457 - 38.5947) + \\ + (38.8822 - 37.1166) = 1.7019\,;$$

the solution for the least x-deviations sum given by the observation curves, $2.0792 \cos\alpha + 1.9138 \sin\alpha = \rho$ and $1.9731 \cos\alpha + 2.0899 \sin\alpha = \rho$ (Fig. 10), where

$$\frac{\partial D}{\partial \alpha} = 1.16749^2\,(38.0063 - 38.1686) = -0.221\,,$$

$$\frac{\partial D}{\partial \alpha} = 1.16749^2\,(38.0063 - 37.9925) = 0.0188\,,$$

and

$$D = (39.1996 - 37.9158) + 0.60250\,(39.0063 - 38.1686) = 1.1860,$$

is equivalent to that given by the intersection of the observation lines,

$$2.0792 = a + 1.9138\,b \text{ and } 1.9731 = a + 2.0899\,b.$$

(1) The theorem that the geometric median, or the antilogarithm of the median of the sequence of the logatithms of the values, is identical with the median of the values themselves, applies only to the simple median (**1**), and not to the weighted median. Since the logarithmic weights increase geometrically, and the numerical weights arithmetically, the weighted median and the weighted geometric median will not necessarily oc-

test-deviation, are successively greater, because increasingly influenced by the disparity of the dispersions of the observed values of the two variables (Table III, Fig. 10).

8. Seven observations of the length of a degree of the meridian and the square of the sine of the latitude, substituted in

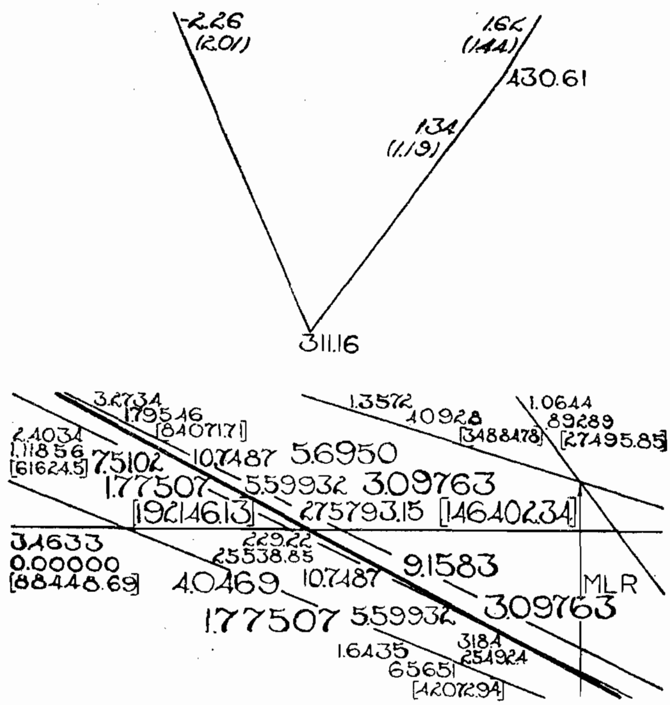


Fig. 11

observation equations treating the length of a degree as the dependent variable, weighted by the lengths of the measured arcs, used in a classical calculation of the ellipticity of the earth (**38**), may be solved by the method of median loci (Fig. 11). The *a*-

cupy corresponding positions in the numerical and logarithmic sequences and have equivalent values. The solution for least numerical deviations, determined by Rhodes by his analytic method of median loci, generalized to render the deviations sum, $\Sigma \mid z_i - f(x_i, y_i, \ldots, u_i, v_i) \mid$, a minimum for any function of the independent variables, and applied concretely to the power function, $y = Bx^a$, by the use of a specially derived system of approximate weighting in the location of the medians of the parameters (**17**), is therefore slightly different from the solution of least logarithmic deviations (Table III; Fig. 10).

and b-loci coincide with the locus of a single equation through several segments where this equation so outweights the other equations that its locus constitutes a double median locus. The analytic criterion nevertheless designates the intersection of this double median locus with the locus of one of the other equations as the solution.

Following the pattern of Laplace's analytic solution for Boscovich's method (**11**), Edgeworth solved this problem by eliminating a between the equation of the double median locus and each of the other equations and determining the weighted median of b (Table IV) (**18**) (1). This procedure, generalized to apply to any selected observation equation, rather than restricted to an observation equation known to represent a median locus, and to any number of parameters, rather than restricted to two parameters, constitutes Rhodes' analytic method of median loci (**17**).

TABLE IV.

$b = \frac{w_i y_i - w_i y_m}{w_i x_i - w_i x_m} = -382.896$	229.224	318.422	482.928	530.599	906.666
$w_i x_i - w_i x_m = 0.29773$.	1.80413	0.13344	0.19664	0.33841	0.09025

These observations, substituted in weighted observation equations treating the length of a degree and the sine squared latitude as conjoint variables, may be solved by the method of the median locus (Fig. 12). The weighted median locus of the observation curves is plotted by the method of substitution. The final solution is determined by means of the analytic criter-

(1) Through an inadvertant sinistroluxation of the decimal point in the second member of the sequence of weights Edgeworth however obtained an aberrant result. The solution for Boscovich's conditions, $\alpha = 25538.85 + 246.93 \sin^2 \Phi$, is in Laplace's method the intersection of the weighted mean line, $25646.80 = a + 0.43717\ b$, with the weighted particular line, $88448.69 = 3.4633\ a$, or the equivalent unweighted line, $25538.85 = a$ (**38**). The solution for the least greatest-deviation, $\alpha = 25525.10 + 308.202 \sin^2 \Phi$, is the mid-point between the intersection of the unweighted lines, $25666.65 = a + 0.30156\ b$ and $25832.25 = a + 0.83887\ b$, and the unweighted line. $25599.60 = a + 0.39946\ b$ (Fig. 11) (**38**).

ion for weighted polar observation equations. The values of the derivative at the crucial angle are

$$\frac{\partial D_w}{\partial \alpha} = 0.00313\,(1.77507 - 3.09763) - 0.99999\,(234851.03 -$$
$$- 103697.44) + (25658.28. - 0.99999 - 0.52093.0.00313)\,(4.0469 -$$
$$- 9.1583) = - 3.9$$

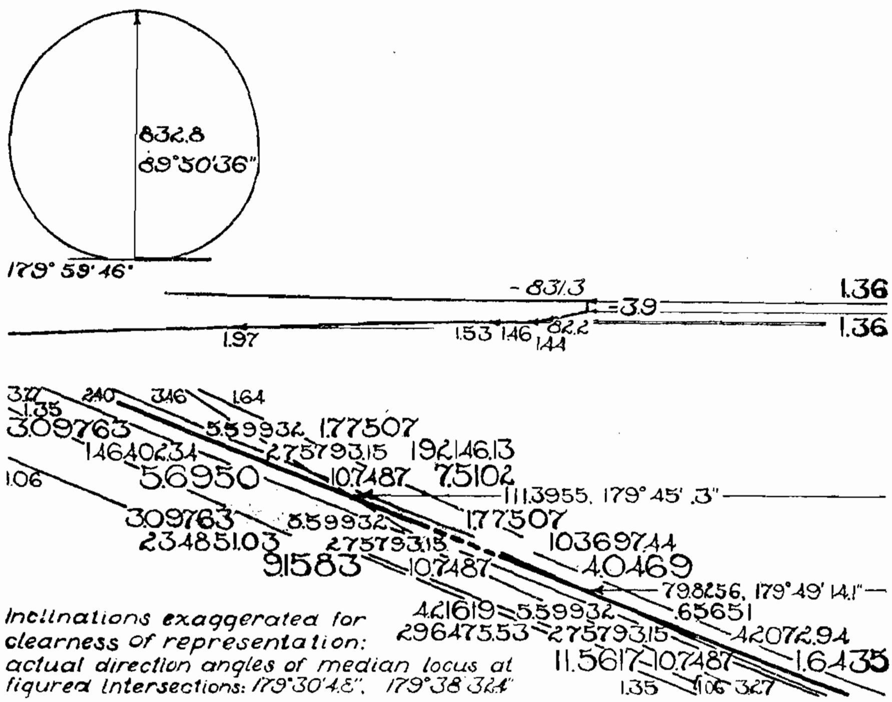


Fig. 12

and

$$\frac{\partial D_w}{\partial \alpha} = 0.00313\ (0.65651 - 4.21619) - 0.99999\ (296475.53 -$$
$$- 42072.94) + (25658.23. - 0.99999 - 0.52093.0.00313)\,(1.6435 -$$
$$- 11.5617) = 82.2\,,$$

and that of the minimum weighted absolute normal deviations sum is

$$D_{w(x,y)} = - 0.99999\,(3.09763 - 1.77507) + 0.00313\,(234851.03 -$$
$$- 103697.44) + 79.8256\,(4.0469 - 9.1583) = 1.36.$$

9. An index of the cyclical fluctuation in the employment of a division of labor in the coal-mining industry in the form of the quarterly excess or deficiency in five-thousand man-shifts worked on the surface from the average for the period (**40**) was employed by Rhodes (**17**) in his demonstration of the fitting of the parabola, $y = 32.83 - 2.07\,x - 1.76\,x^2$, with the minimum deviations sum, $D_y = 260.4$, and the least mean deviation, $D_y/n = 15.32$, by the analytic method of median loci. The solution may be confirmed by the application of the special analytic criterion to the intersection of its three median observation planes. With the direction cosines of each of the lines of intersection of the three median planes taken two at a time directed in the sense of departure from the point of intersection as parameter increments, the increment ratio is computed from the sums of the antecedent and succeeding coefficients of either member of each pair of intersecting planes for each of the three direction lines (Fig. 13). Since all three increment ratios are positive, the deviations sum increases in all directions of departure from the intersection, and the intersection is established as the minimum point. The deviations sum is computed for the intersection from the sums of the antecedent and succeeding coefficients of any of the three planes (1).

10. Power formulae to estimate the normal basal metabolism were promulgated (**41**) on the basis of previously published data (**42**) in a form equivalent to $C = 280.51\,W^{0.5}\,A^{-0.1333}$ for females, and to $C = 311.55\,W^{0.5}\,A^{-0.1333}$ for males, in which C = calories per day, W = body weight in kilograms, and A = age in years. The mode of derivation of the exponents was withheld. But the formulae were advanced with the claim of giving lesser mean percentile deviations for each sex than earlier multiple regression formulae. Upon both physiological and mathematical grounds the peculiar pattern and singular identity of the exponents in these formulae for the sexual differences in the relationship of basal metabolism to body size and age elicit immediate doubt as to their correctness. Established clinical usage and a statistical inquiry, though it utilized the standard deviation rather than the mean deviation of the errors of prediction, into their performance

---

(1) The corresponding parabola by the method of least squares is $y = 23.86 - 2.39\,x - 0.99\,x^2$ with a standard error of 22.54.

(**43**), have shown the formulae to be empirically approximate or not grossly aberrant in comparison with alternative methods for estimating the normal basal metabolism. The question as to the determination of the proper constants in power formulae based on the data to effect a minimum percentile deviation has nevertheless remained open.

On substitution of logarithms for their numerical values, the data yield for females and for males respectively series of 103 and of 109 observation equations of the form, $\log C = a + b \log W + c \log A$, as

$$3.2457 = a + 1.9713b + 1.6435c \qquad 3.2617 = a + 1.9335b + 1.3802c,$$

. . . . and . . . .

$$3.0488 = a + 1.6684b + 1.7924c \qquad 3.1504 = a + 1.7803b + 1.7853c.$$

The graphic method of median loci is obviously not feasible in the treatment of observation planes. The analytic method, though applicable, is not promising, when the independent variables are several distinct variables, and not, as in Rhodes example, different functions of a single variable, since the possibility of the selection of the appropriate trial observation equations by examination of the plotted observation points is excluded. While the graphic method is impracticable, and efficient prosecution of the analytic method is prevented by the virtual impasse in the selection of the appropriate trial observation equations, the previously advanced formulae afford first approximations which can serve as initial trial values of the required parameters.

The procedure for effecting the transition from these trial values of the parameters to the values of the parameters in the final solution for least deviations may be based upon two considerations. The deviation in the dependent variable given by the observation equation, $z_i = a + b\,x_i + c\,y_i$, is equal to the difference between the value of its first parameter, $a_{(z_i \cdot x_i, y_i)}$, and an assumed value of the first parameter, a, at the assumed values of the succeeding parameters, b, c, or to $(a_{(z_i \cdot x_i, y_i)} - a)_{b,c}$, and for a set of such observation equations the deviations sum will be least when the assumed value of the first parameter is identical with its median value, a_m, and will then be equal to $\sum_{i=1}^{n} |a_{(z_i \cdot x_i, y_i)} - a_m|_{b,c}$. Given an observation equation, $z_i = a + b\,x_i + c\,y_i$, if any arbitrary set of values of the succeeding parameters, b, c, is assigned, the value of its first parameter is determined as $a_i = z_i - b\,x_i - c\,y_i$, and the least deviations sum of a set of such observation equations with the assigned values of the succeeding parameters, $\sum_{i=1}^{n} |z_i - a - b\,x_i - c\,y_i|$, is equal to $\sum_{i=1}^{n} |a_i - a_m|_{b,c}$. For the set of observation equations, $z_i = a + b\,x_i + c\,y_i$, the least deviations sum, $\sum_{i=1}^{n} |z_i - (a + b\,x_i + c\,y_i)|$, with a given set of values of the succeeding parameters, b, c, is therefore equal to the sum of the deviations of its values of the first parameter, a_i, from their median value, a_m, or $\sum_{i=1}^{n} |a_i - a_m|_{b,c}$. The least deviations sum for a set of selected trial values of b and c may accordingly be computed if the number of observation equations is odd as $D_{z_{(b,c)}} = \sum_{i=m+1}^{n} a_i - \sum_{1}^{m-1} a_i$. By this device the least deviations sum for a given set of observation equations with any assigned set of parameter values may be determined. A second approximation to the required values of the parameters may be arrived at by interpolation between sets of trial values and extrapolation beyond their smallest deviations sum. The graph of the variation of the deviations sum as the value of the simple or the weighted median is departed from approximates to the general shape of a skew parabola. Similarly the graph of the variation of the deviations sum as the parameter point on the intersecting or coincident median loci giving the minimum deviations sum is departed from assumes the form of a parabola or a paraboloid. Parabolic formulae will accordingly serve to interpolate between successive selected values of the parameters and to extrapolate below their deviations sums in the determination of second approximations to the required values of the parameters. The transition from these second approximations to the values of the final solution is effected by application of the method of median loci and its analytic criterion.

Thus, by putting $c = 0$, and taking serial values of b, by putting $b = 0$, and taking serial values of c, and by taking sets made up of serial values of b and c, series of values of a are computed, their medians located, and their minimum deviation sums and mean deviations determined for the given sets of observation equations for females (Fig. 14) and for males (Fig. 15). Then,

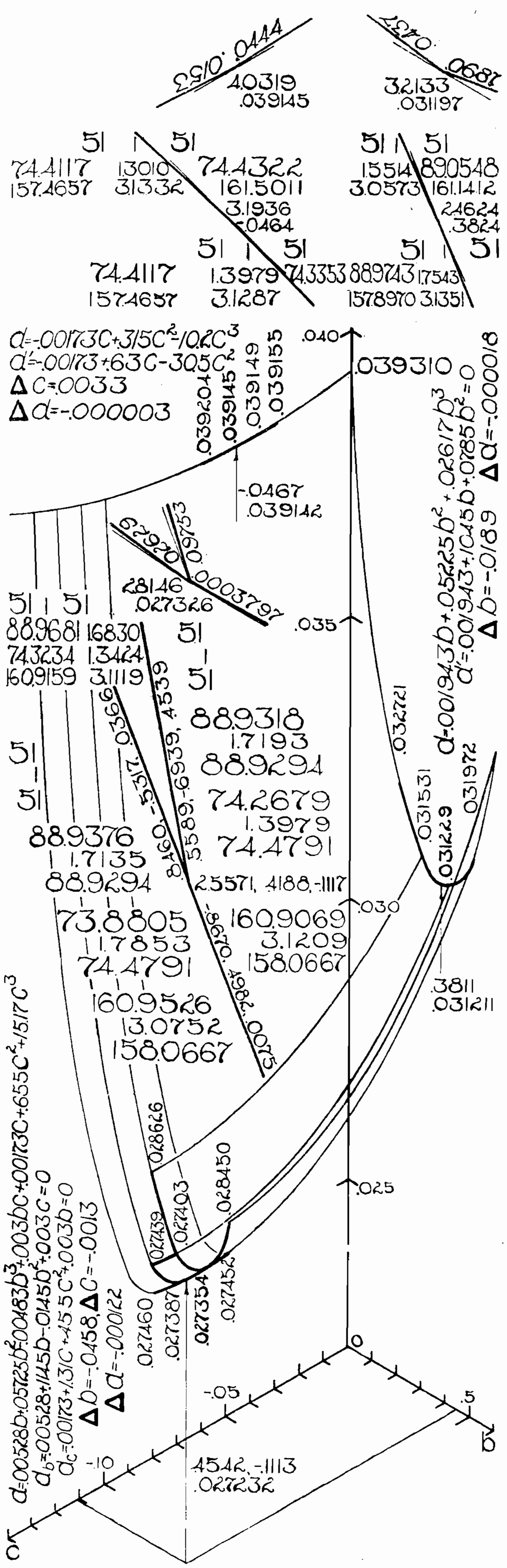


Fig. 14

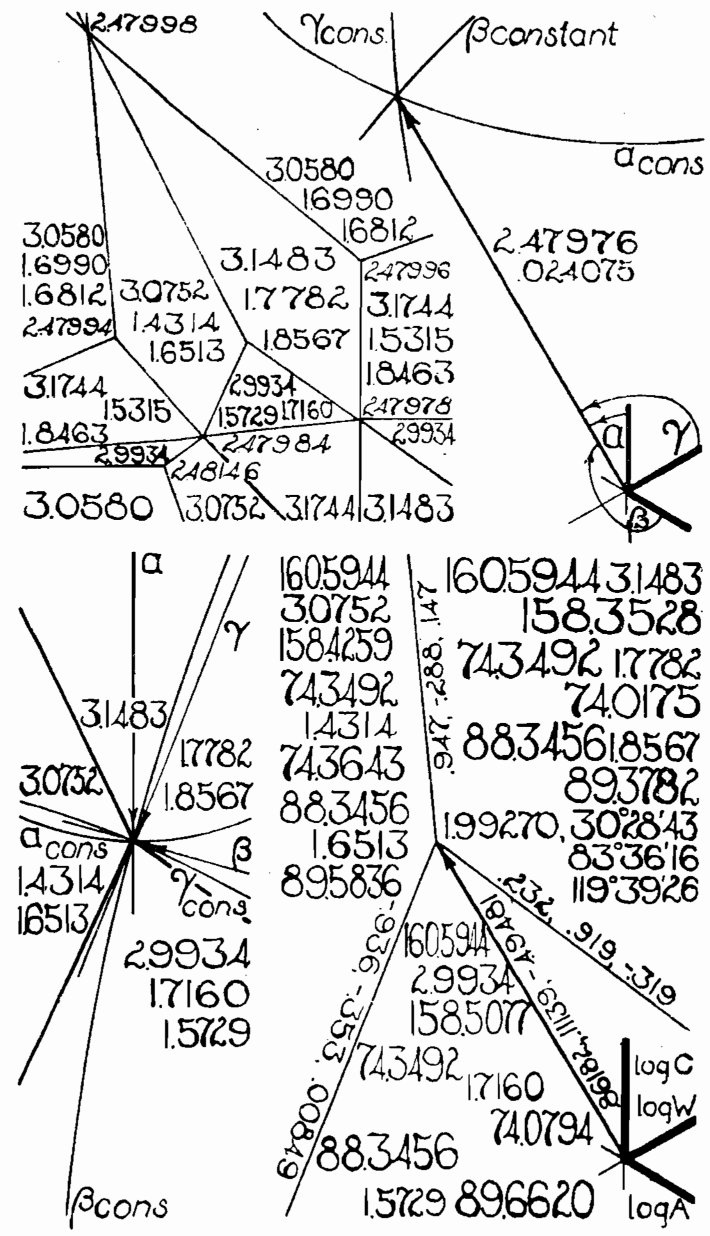


Fig. 16

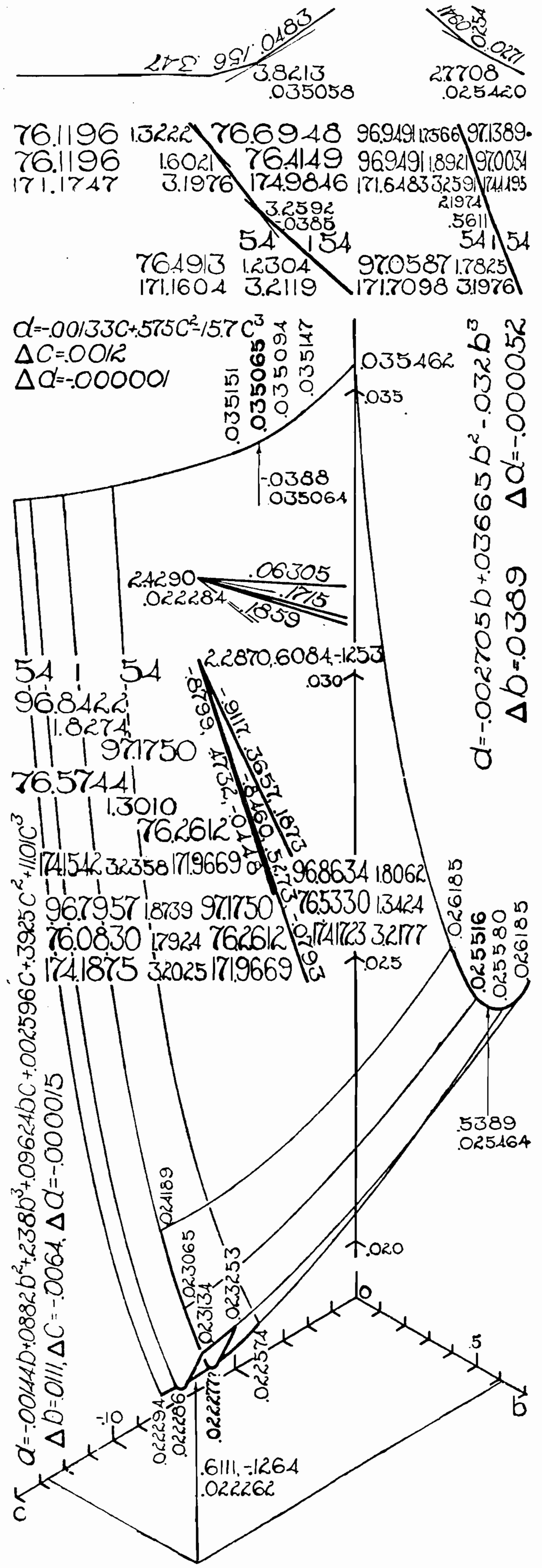


Fig. 15

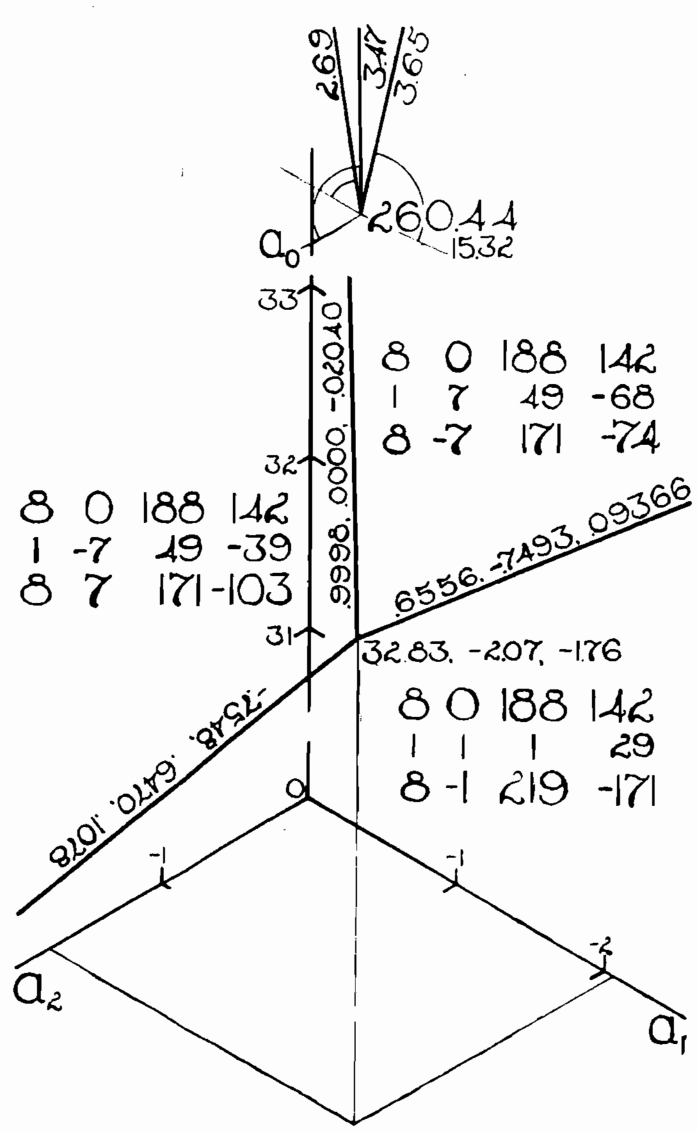


Fig. 13

taking the least mean deviation and the associated b-, c-, or (b, c)-parameter values as the origin, differences in the mean deviation as the dependent variable, and differences in the parameter values as independent variable or variables, cubic parabolas or paraboloids are fitted by the use of determinants to pass through two points on each side of, and in the case of the paraboloids through one point diagonally across from, the presumptive position of the point sought, and derivatives of the mean deviation differences with respect to the parameter differences equated to zero and solved for the parameter increments or decrements which on substitution in the equations give the greatest mean deviation decrements (Figs. 14, 15). Then, taking the resulting interpolated values of b, of c, and of b and c, as second approximations to the required values of the parameters, new median values of a are determined. The locus of the observation equation giving each of these medians is considered in its relation to those of equations giving values adjacent to the median. In the cases where the observation equations contain only two parameters of values other than zero the analytic criterion is applied to the segment of the median locus isolated for indication of the direction from the parameter point representing the second approximation to points giving lesser deviation sums. The first intersection in this direction of the isolated segment of the median locus with a contiguous segment of the median locus will be the minimum point if the sign of the increment ratio is positive for the succeeding segment. If the parameter at the required intersection should have a lesser value than the second approximation, an observation equation with a greater slope-angle than the present equation and giving a positive deviation, or with a lesser slope-angle and giving a negative deviation, is selected; if a greater value, a lesser slope-angle and a positive deviation, or a greater slope-angle and a negative deviation. In the cases where the observation equations contain three parameters, the inclinations of the adjacent planes together with their distances from the median planes determine the selection of the two planes forming an intersection with the median plane nearest the parameter point represented by the second approximation. The analytic criterion is applied to the lines of intersection of the three planes taken two at a time. The resulting minimum percentile deviation power formulae (Figs. 14, 15; Table V) may be compared with the corresponding least squares formulae derived in the course of the classical procedure of correlation analysis (Table V).

The solution for least normal percentile deviations was determined by applying trial values of the direction cosines to the set of polar observation equations, $\log C \cos \alpha + \log A \cos \beta + \log W \cos \gamma = \rho$, for females, tracing a region of the median locus, and studying the values of the partial derivatives and the deviation sums at its intersections (Fig. 16; Table VI) (1) and may be compared with the corresponding least squares formula derived by the analytic method (Table V).

Discussion.

The combination of observations in two or more variables involves a decision between the determination of a dependent variable as an explicit function of one or more independent variables, as $z = f(x, y, \ldots, u, v)$ and of the several variables as implicit functions of each other, $f(x, y, \ldots, u, v) = 0$. The alternative of the determination of one variable as an explicit function of other variables or of a mutually implicit function of several variables is not indicated by the data themselves but is dependent upon the purpose the data are to serve. And the purpose the data are to serve depends upon the use to which the function derived from them is to be put.

An explicit function derived from the present sets of complete data consisting of values of both the dependent and the independent variables in effecting their adjustment provides an instrumentality for the completion of partial sets of the present data and

---

(1) A more direct approach to the solution might have been to determine the normal deviations sums for sets of trial values of the direction cosines, approximate to their required values by interpolation, and apply the analytic criterion to the indicated intersection.

The figure representing the median locus does not attempt to delineate the actual intersections of the observation curves but only the tangent lines to the curves of intersection of the three surfaces taken two at a time at the point of common intersection the direction cosines of which are calculated from the observation equations transformed from polar to rectangular coordinates.

TABLE V.

FEMALE

Least Deviations	log C = 3.1936 − 0.04644 log A	0.039145	$C = 1562\ A^{-0.04644}$	per cent. − 8.62, + 9.43
	log C = 2.4624 + 0.38344 log W	0.031197	$C = 290.0\ W^{0.3834}$	− 6.93, + 7.45
	log C = 2.5571 + 0.41869 log W − 0.11170 log A	0.027326	$C = 360.7\ W^{0.4187}\ A^{-0.1117}$	− 6.10, + 6.49
	0.86182 log C + 0.11139 log A − 0.49481 log W − 1.99270 = 0 log C = 2.3122 + 0.57415 log W − 0.12925 log A	0.024075	$C^{0.8618}\ A^{0.1114}\ W^{-0.4948} = 98.33$ $C = 205.2\ W^{0.5741}\ A^{-0.1293}$	− 5.33, + 5.70
Least Squares	log C = 3.22288 − 0.064947 log A	0.048529	$C = 1670.6\ A^{-0.064947}$	− 10.57, + 11.82
	log C = 2.49869 + 0.361037 log W	0.039846	$C = 315.27\ W^{0.36104}$	− 8.77, + 9.61
	log C = 2.55871 + 0.427773 log W − 0.121109 log A	0.034621	$C = 362.00\ W^{0.42777}\ A^{-0.12111}$	− 7.66, + 8.30
	0.885006 log C + 0.120685 log A − 0.450205 log W − 2.159242 = 0 log C = 2.43980 + 0.508702 log W − 0.136366 log A	0.031165	$C^{0.88501}\ A^{0.12069}\ W^{-0.45021} = 144.39$ $C = 275.29\ W^{0.50870}\ A^{-0.13637}$	− 6.92, + 7.44

MALE

Least Deviations	log C = 3.2592 − 0.03847 log A	0.035058	$C = 1816\ A^{-0.03847}$	− 7.76, + 8.41
	log C = 2.1974 + 0.56113 log W	0.025420	$C = 157.5\ W^{0.5611}$	− 5.69, + 6.03
	log C = 2.2870 + 0.60844 log W − 0.12531 log A	0.022284	$C = 193.6\ W^{0.6084}\ A^{-0.1253}$	− 5.01, + 5.27
Least Squares	log C = 3.30444 − 0.070186 log A	0.047010	$C = 2015.7\ A^{-0.070186}$	− 10.26, + 11.43
	log C = 2.27832 + 0.515842 log W	0.033273	$C = 189.81\ W^{0.51584}$	− 7.38, + 7.96
	log C = 2.38867 + 0.553821 log W − 0.126260 log A	0.030095	$C = 244.72\ W^{0.55382}\ A^{-0.12626}$	− 6.70, + 7.12
	0.829631 log C + 0.119451 log A − 0.545375 log W − 1.848161 = 0 log C = 2.22768 + 0.657370 log W − 0.143981 log A	0.025597	$C^{0.84696}\ A^{0.12228}\ W^{-0.55809} = 77.834$ $C = 168.92\ W^{0.65737}\ A^{-0.14398}$	− 5.72, + 6.07

by induction of all other incomplete sets of similar data consisting of values of the independent variables alone. By means of the explicit function such sets of partial or incomplete data lacking the values of the dependent variable which are the quaesitae can be completed. An explicit function which serves to complete sets of incomplete data lacking one term by defining the values of the missing term is an instrumentality for operations of estimation or forecast. If the inquiry calls for an instrumentality for the estimation or forecast of the values of one variable when the values of other variables are given it requires the determination of an explicit function.

An implicit function derived from the present sets of complete data in effecting their adjustment generalizes the quantitative relationships subsisting between the several variables and by induction defines the relationships subsisting between the variables in all other complete sets of similar data. The implicit function which is a generalization of the present sets of complete data and an induction from the present sets of complete data to all other complete sets of similar data is a device for determining the quantitative relationships between the several variables and defining the class to which the present sets and all similar sets of the values of the variables which satisfy the relationship belong. If the inquiry calls for a definition of the quantitative relationships and of the sets of concomitant values of several variables it requires the determination of an implicits function.

The combination of observations to determine an explicit function attributes deviation exclusively to the dependent variable and denies it to the independent variables in the present sets of complete data because in the application of the function to subsequent sets of incomplete data the values of the independent variables will be given and the value of the dependent variable required. As many explicit functions are determinable for the sets of data as the number of the variables since each variable may be treated as a function of the other variables. The explicit function determined from the sets of data differs according to which variable is treated as dependent upon the other independent variables. The combination of observations to determine an implicit function attributes deviation to all the several variables, either equally as a matter of expediency, or to each in the proportion it contri-

TABLE VI.

$\sum_{m+1}^{n} x_i$	160.6116	160.6847	160.6847
x_m	3.1483	3.0752	3.0580
$\sum_{i=1}^{m-1} x_i$	158.3356	158.3356	158.3528
$\Sigma\, y_i$	74.0816	74.4284	74.4284
y_m	1.7782	1.4314	1.6990
$\Sigma\, y_i$	74.2851	74.2851	74.0175
$\Sigma\, z_i$	88.3157	88.5211	88.5211
z_m	1.8567	1.6513	1.6812
$\Sigma\, z_i$	89.4081	89.4081	89.3782
$\Phi_{\alpha(\beta)}$	− 1.958	− 0.622	0.451
$\Phi_{\alpha(\gamma)}$	− 0.184	− 0.404	− 0.409
$\Phi_{\beta(\alpha)}$	0.482	0.157	− 0.111
$\Phi_{\beta(\gamma)}$	0.446	0.055	− 0.218
$\Phi_{\gamma(\alpha)}$	− 0.184	− 0.396	− 0.418
$\Phi_{\gamma(\beta)}$	1.741	0.214	− 0.852

$$3.0580 \cos\alpha + 1.6990 \cos\beta + 1.6812 \cos\gamma = \rho$$
$$3.0752 \cos\alpha + 1.4314 \cos\beta + 1.6513 \cos\gamma = \rho$$
$$3.1483 \cos\alpha + 1.7782 \cos\beta + 1.8567 \cos\gamma = \rho$$

$\cos\alpha = 0.86254$, $\cos\beta = 0.11061$, $\cos\gamma = -0.49373$, $\rho = 1.99548$

$\angle$	$\cos\angle$	$\angle$	$\sin\angle$	$2\angle$
α	0.86254	30°23′18″	0.50598	60°46′36″
β	0.11061	83°38′58″	0.99386	167°17′56″
γ	−0.49373	119°35′10″	0.86968	239°10′20″

$\sin 2\angle$	$2\cos\angle$
0.87286	1.72508
0.21987	0.22122
−0.85871	− 0.98746

$\Sigma\, x_i$	160.6116	160.5944		160.5944
x_m	3.0580	3.0752		3.1483
$\Sigma\, x_i$	158.4259	158.4259		158.3528
$\Sigma\, y_i$	74.0816	74.3492		74.3492
y_m	1.6990	1.4314		1.7782
$\Sigma\, y_i$	74.3643	74.3643		74.0175
$\Sigma\, z_i$	88.3157	88.3456		88.3456
z_m	1.6812	1.6513		1.8537
$\Sigma\, z_i$	89.5836	89.5836		89.3782
$\Phi_{\alpha(\beta)}$	− 2.222	− 1.157		0.314
$\Phi_{\alpha(\gamma)}$	0.015	− 0.002		− 0.217
$\Phi_{\beta(\alpha)}$	0.560	0.291		− 0.044
$\Phi_{\beta(\gamma)}$	0.564	0.291	160.5944	− 0.078
$\Phi_{\gamma(\alpha)}$	0.017	− 0.005	2.9934	− 0.219
$\Phi_{\gamma(\beta)}$	2.200	1.136	158.5077	− 0.391

$$\Phi_{\alpha(\gamma)} = 0.50598\,(158.3356 - 160.6116) + \frac{0.87286}{-0.98749}(88.3157 - 89.4081) = -0.184$$

$$\Phi_{(\beta\gamma)} = 0.99386\,(74.2851 - 74.0816) + \frac{0.21987}{-0.98746}(88.3157 - 89.4081) = 0.446$$

et cetera

$$\Delta_\varrho = 0.86254\,(160.6116 - 158.3356) + 0.11061\,(74.0816 - 74.2851) - 0.49373\,(88.3157 - 89.4081) = 2.47998$$

	− 1.160	74.3492	*0.169*	
	− 0.007	1.7160	*− 0.222*	
	0.293	74.0797	*− 0.043*	
160.4952	*0.293*	88.3456	*− 0.098*	160.5683
3.1744	*− 0.008*	1.5729	*− 0.219*	3.1744
158.4259	*1.134*	89.6620	*− 0.383*	158.3528
74.2491		*− 0.0900*		74.5959
1.5315		*0.103*		1.5315
74.3643		*0.0002*		74.0175
88.1506		*0.026*		88.3560
1.8463		*0.104*		1.8463
89.5836		*0.104*		89.3782

$\angle$	$\cos\angle$	$\angle$
α	0.86182	30° 28′ 43″
β	0.11139	83° 33′ 16″
γ	−0.49481	119° 29′ 26″

$\rho = 1.99279$

$\Delta_\varrho = 2.47976$

$\Phi_{\alpha(\beta)}$	− 1.505	− 1.163	0.001	− 0.005	0.159	1.14
$\Phi_{\alpha(\gamma)}$	0.215	− 0.006	0.104	0.105	− 0.224	− 0.224
$\Phi_{\beta(\alpha)}$	0.379	0.294	− 0.0003	0.002	− 0.041	− 0.290
$\Phi_{\beta(\gamma)}$	0.434	0.291	0.026	0.028	− 0.098	− 0.348
$\Phi_{\gamma(\alpha)}$	0.212	− 0.007	0.101	0.101	− 0.219	− 0.222
$\Phi_{\gamma(\beta)}$	1.691	1.134	0.101	0.106	− 0.376	− 1.34

$\angle$	$\cos\angle$	$\angle$
α	0.86170	30° 29′ 32″
β	0.11202	83° 34′ 6″
γ	−0.49488	119° 39′ 43″

$\rho = 1.99330$

$\Sigma\, x_i$	160.4952	160.4134	160.4134	160.4134	160.5683
x_m	2.9934	3.0752	3.1744	3.1483	2.9934
$\Sigma\, x_i$	158.6069	158.6069	158.5077	158.5338	158.5338
$\Sigma\, y_i$	74.2491	74.5337	74.5337	74.5337	74.5959
y_m	1.7160	1.4314	1.5315	1.7782	1.7160
$\Sigma\, y_i$	74.1798	74.1798	74.0797	73.8330	73.8330
$\Sigma\, z_i$	88.1506	88.0722	88.0722	88.0722	88.3560
z_m	1.5729	1.6513	1.8463	1.8357	1.5719
$\Sigma\, z_i$	89.8570	89.8570	89.6 20	89.6516	89.6516

$\Delta_\varrho = 2.47978$

$\Phi_{\alpha(\beta)}$	− 0.686	0.469	0.821	0.804	1.78	1.95
$\Phi_{\alpha(\gamma)}$	0.550	0.658	0.439	0.438	0.442	0.115
$\Phi_{\beta(\alpha)}$	0.173	− 0.121	0.252	− 0.206	− 0.454	− 0.498
$\Phi_{\beta(\gamma)}$	0.312	0.046	− 0.097	− 0.095	− 0.342	− 0.470
$\Phi_{\gamma(\alpha)}$	0.551	0.651	0.431	0.430	0.434	0.111
$\Phi_{\gamma(\beta)}$	1.214	0.188	− 0.375	− 0.361	− 1.32	− 1.80

butes to the deviation of the data relative to the function derived (1). Only one implicit function is determinable for the sets of data since all the variables are accorded a like status.

The combination of observations in one variable by the determination of the mode, the median, the mean, or the midpoint of the range, or of observations in two or more variables by the determination of an explicit or implicit function by the method of the modal point, median loci (2), mean loci, or the mid-point of the least range, renders a minimum the sums of the absolute deviations raised to the zero, the first, the second, or the infinite power, characterized as most lesser-deviations, least deviations, least squares (3), and

---

(1) An implicit function in m variables the values of $m - 1$ of which were exactly known, or to be regarded as approximately so as a matter of expediency, while the values of the remaining variable were liable to error, would be equivalent to an explicit function of the variables the values of which were exactly known.

(2) The analytic solution devised by Laplace (**11, 38**) for Boscovich's conditions (**10**), that the sum of the positive corrections be equal to the sum of the negative corrections on the assumption of the same degree of probability of errors in defect and of errors in excess, and that the sum of all the corrections, positive as well as negative, be the least possible in order to reconciliate the observations as much as may be, is a method of mean and median loci. But the power-mean, x, which has the least total p^{th}-power absolute deviations sum, $\sum_{i=1}^{n} |x_i - x|^p = min.$, has equal bilateral $p - 1^{th}$-power deviations sums, $\sum_{i=1}^{k} (x - x_i)^{p-1} = \sum_{k+1}^{n} (x_i - x)^{p-1}$ (1). The arithmetic mean is the second-order power-mean which has equal bilateral first-power absolute deviations sums or a zero first-power algebraic deviations sum, while the median is the first-order power-mean giving the minimum absolute deviations sum but has equal bilaterial zero-power deviations sums (1), so that the two conditions are inconsistent. The method of mean and median loci, which has been characterized as a hybrid between the method of least squares and the method of situation (**12**), employed as a preliminary step in Rhodes' analytic method of median loci (**17**), would in the case of two variables or parameters be likely to lead directly to one of the final median observation equations, as instanced in the footnotes to numerical examples numbers 6 and 8, or at least to a near-final-median observation equation.

(3) The method of moments developed by Pearson (44) which equates the area and moments given by the function to the area and moments calculated from the data is equivalent to the method of least squares in the dependent variable for parabolas of any order with the distinction that it renders least squared deviations of all the points of some curve with the

the least greatest-deviation (1). Each method for the combination of observations involves a commitment to a determinate distribution of the relative frequencies of deviations of different magnitudes, because each method determines the best value of the average or of the function for a certain particular distribution of the relative frequencies of deviations of different magnitudes alone, and that distribution of the relative frequencies of deviations of different magnitudes requires the application of the

---

moment system determined by the observation points from the function instead of the least squared deviations of the observation points alone from the function. The fitting criterion based on the postulated equivalence of the moments of the data to the corresponding moments of the fitted valeus has also been shown by M. SASULY (*Trend Analysis of Statistics, Theory and Technique*, Washington, 1934, 14, 203-212, 220-229, 239-241) to correspond to the fitting criterion based on least squares for the n^{th} degree polynomial if its parameters are determined by postulating the equivalence of the first $n + 1$ moments.

(1) Corresponding to the non-limiting members of the succession of methods of least power-sums of the deviations for all the data are methods of least power-sums of the deviations of somewhat arbitrary parts of the data. Such is Mayer's method (45), the earliest systematic procedure for the combination of observation equations more numerous than the parameters, of simultaneous solution of a set of m composite equations, $\Sigma\,(z_i = a + x_i\, b + y_i\, c + \ldots + v_i\, m)$, in which the summations extend over the parts of the sequence of observation equations arranged in the ascending order of the values of one of the variables, delimited by the indices which are the integers nearest the values of 1 and $\frac{n}{m}$, $\frac{n}{m} + 1$ and $2\,\frac{n}{m}$, $\frac{2\,n}{m} + {} + 1$ and $3\,\frac{n}{m}$, . . . , $\frac{(m-2)\,n}{m} + 1$ and $\frac{(m-1)\,n}{m}$, $\frac{(m-1)\,n}{m} + 1$ and n, preferably in the order of the independent variable, or, in the case of more than two variables, since it is advantageous that the observation equations be so grouped that the contrasts between their several sums should be as great as possible, of that independent variable the variation of which has the most marked effect on the dependent variable, though a preliminary solution with the equations arranged in the order of the values of the dependent variable might be required to indicate the relative effects of the several independent variables on the dependent variable. Since the procedure is equivalent to passing a line or plane through m points each of which is the mean of n/m observation points, it may be designated the method of mean observation points, or, since it applies the arithmetic mean to each part of the data, the method of least squared deviations sums or of zero deviations sums. In as much as the mean observation points are the result of applying the arithmetic mean equally to each variable in each part of the data, the line or plane passing through them will more appropriately

method for the determination of the best value of the average or of the function.

The relative frequency distributions of the presumptive deviations (**48**) or the presumptive relative frequency distributions of the deviations are founded on hypotheses concerning the occurrence of deviations of different magnitudes. Different presumptive relative frequency distributions follow from different assumptions concerning the occurrence of deviations of different magnitudes. The different methods for the combination of observations

---

indicate the relationships between the several variables than it will designate the estimate of one variables to be derived from the values of the others. On this basis KOLLER (loc. cit.) has cited the line connecting the two mean observation points as a good approximation to the representation of the fundamental relationship between the variables and derived the standard error of its direction coefficient.

A variant of this method of mean observation points or of zero deviations sums occurs in Dufton's method (**46**) of dividing the plotted observation points into two groups at the median of the independent variable and drawing a line to pass through the median of the independent variable with respect to the line in each group, which may be designated the method of median observation points or of least deviations sums.

Other methods quite apart from the succession of methods of the least power-sums of the absolute deviations which tend to render the deviations a minimum without doing so in an algebraically determinate manner are Cauchy's method (**47**) of representing the dependent variable as a function of a convergent series of independent variables by the application of the theorem, that the fraction formed by the sum of all the numerators divided by the sum of all the denominators of a sequence of unequal fractions the denominators of which are all of the same sign is intermediate in magnitude between the greatest and the least of them, successively to the values of the dependent variable and the coefficients of the first parameter in the observation equations to determine the first parameter which is stet mean of the values of the dependent variable, to the corrections required on substitution of the first parameter for the values of the dependent variable and the deviations of the values of the first independent variable from their mean to determine the value of the second parameter, and to the corrections required after the substitution of all the preceding parameters and the deviations of the values of each succeeding independent variable from their mean to determine each succeeding parameter, and Goedseels' most approximative method (**28, 30**) by which, when known values of the extreme positive and negative deviations are substituted in the deviations inequalities, the limits of the smallest interval which contains the value of each parameter is determined by the successive elimination of all the other parameters, the mid-point of which interval is the most approximative value of the parameter and the semi-interval its degree of approximation.

also follow from these different assumptions. The variation of the relative frequency of the deviations with their magnitude describes a deviations distribution represented analytically by a deviations function. The presumptive deviations functions may be derived from the assumptions and the methods for the combination of observations deduced from the functions, or the methods based on the assumptions and the functions deduced from the methods (1).

---

(1) To determine what average should be taken for three observations of the same phenomenon, Laplace (**50**) postulated that an error in excess is as probable as an error in defect, that the probability of an infinite error is null, and that the total probability of all the errors is unity, referred to a theorem on the probability of the causes of events, and arrived at the reflected negative exponential function, $\Phi(x_i - x) = \frac{m}{2} e^{-m|x_i - x|}$, as representing the probability of an error as a function of the error, from which almost a century later Glaisher (**51**) in turn deduced as a direct result the self-styled remarkable conclusion, that, since the condition that the exponential probability function,

$$\prod_{i=1}^{n} \Phi(x_i - x) = \left(\frac{m}{2}\right)^n e^{-m(|x_1 - x| + |x_2 - x| + \ldots)},$$

be a maximum is that the linear error function,

$$y = |x_1 - x| + x_2 - x| + \ldots,$$

be a minimum, the definitions of the hitherto undefined simple median define the most probable value. Taking

$\log \prod_{i=1}^{n} \Phi(x_i - x) = n(\log m - \log 2) - m \sum_{i=1}^{n} |x_i - x|$, equation of the partial derivative with respect to x to zero, $\frac{\partial}{\partial x} \Pi \Phi(x_i - x) = 0$, produces the definition of the simple median for its most probable value, and equation of the partial derivative with respect to m to zero,

$\frac{\partial}{\partial m} \Pi \Phi(x_i - x) = \frac{n}{m} - \Sigma |x_i - x| = 0$, so that $m = \frac{n}{\Sigma |x_i - x|}$, the reciprocal of the mean absolute deviation (**52**). Very recently this median error-function has been derived anew from theorems on the probability of events upon the explicit basis of the first assumption alone (**53**). Conversely the error function can be deduced from the definition of the median and its minimum deviations sum property on the assumption that it is the most probable value (**49, 54**).

Similarly postulate sets can be posited as a basis for the normal error function from which the arithmetic mean can be deduced as them ost probable value (**52**), or the normal error function can be deduced from the assumption that the mean is the most probable value (**21**), but a more critical selection of assumptions and postulates and mathematical deduction pursues an intermediary course (**55**).

The best value of a variable or of a function of one or more variables is indeterminate without a deviations distribution (**49**). For each deviations distribution a certain function of the data constitutes their best value. In order that each method for the combination of observations determine their best value the data must conform to a certain deviations distribution. These deviations distributions may be represented by deviations functions. The methods for the combination of observations represent combining functions. The deviations function determines the combining function, and the combining function determines the deviations function. Probability may be defined as the relative frequency in a presumptive deviations distribution (**48**). The best value for a specific deviations distribution is then the most probable value for that deviations distribution.

The several methods for the combination of observations and their associated deviations distributions depend upon certain sets of corresponding assumptions regarding the frequencies of the occurrence of deviations of related magnitudes. The positing of such a cardinal assumption as that lesser deviations are of greater frequency, that positive and negative deviations are of equal frequency, that equal positive and negative deviations are of equal frequency, or that the greatest deviation is of finite frequency, together with the necessary supporting postulates is the foundation of a method for the combination of observations and of a deviations function. The methods for the combination of observations conforming to such sets of assumptions are respectively the methods of most lesser-deviations, of least deviations, of least squares, and of the least greatest-deviation, or the methods of the mode and the modal point, of the median, median loci, and the median locus, of the mean, mean loci, and the mean locus, and of the mid-point of the range or of the least range. The deviations distributions accordant with the same sets of assumptions are a sharp peak which would reach its limiting form in the single isolated concentrated frequency (**56**), the reflected negative exponential curve (**56**), the normal probability curve, and a limited-range curve which would reach its limiting form in the rectangle (**56**).

The method for the combination of observations to be applied in the determination of an average or of an explicit or implicit function should be implicit in the data and indicated by the data.

The method for the combination of observations assumes a certain distribution of deviations of different magnitudes and the distribution of deviations of different magnitudes indicates the appropriate method for the combination of observations. The crucial question is to match the actual deviations distribution (**48**) most closely with a presumptive deviations distribution (1). In

(1) The form of the presumptive deviations distributions ranges from the single isolated frequency through the reflected exponential curve and the normal probability curve to the rectangle. Actual deviations distributions may be classified according to their positions in relation to these presumptive deviations distribution. The method for the combination of observations appropriate to any data is the method corresponding to the presumptive deviations distribution to which is actual deviations distribution most closely conforms.

The relative values of the root mean squared deviation, the mean deviation, and the median deviation are determinate for each deviations function: foi the median deviations-function, $\Phi(x)\,dx = \frac{1}{2\eta} e^{-\frac{|x|}{\eta}}\,dx$, the ratios are $\eta : \sigma : r = 1 : 1.4141 : 0.6932$ or $\sigma : \eta : r = 1 : 0.7072 : 0.4902$, and for the mean deviations-function, $\Phi(x)\,dx = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}\,dx$, the ratios are $\sigma : \eta : r = 1 : 0.7979 : 0.6745$ or $\eta : \sigma : r = 1 : 1.2533 : 0.8453$ (53). The ratios for the actual deviations distribution may be compared with the ratios for these presumptive distributions as a test for the position of the actual deviations distribution in relation to these presumptive deviations distributions.

But the effectiveness of the comparison is somewhat impaired by the fact that for each set of data the standard deviation, mean deviation, and median deviation will differ according to the method for the combination of observations employed since in actual deviations distributions in contradistinction to presumptive distributions the mode, the median, the mean, and the mid-point of the interquartile range are not likely to coincide and the successive methods for the combination of observations are based upon these averages. Thus in numerical example number 9 where antecedent economic considerations and subsequent considerations of goodness of fit favor the solution for least deviations over the solution for least squares, the solution for least deviations gives a mean deviation of 15.32, a standard deviation of about 29.33. and a median deviation of about 10 with the ratios $\eta : \sigma : r = 1 : 1.917 : 0.653$ or $\sigma : \eta : r = 1 : 0.522 : 0.342$, which suggest a deviations distribution intermediary to the median distribution and a modal distribution, the solution for least squares gives a standard deviation of 22.54, a mean deviation of about 17.29, and a median deviation of about 12 with the ratios, $\sigma : \eta : r = 1 : 0.767 : 0.533$ or $\eta : \sigma : r = 1 : 1.304 : 0.693$, which suggest a distribution intermediary to the median and mean distripractice the more pertinent question will more often be the assumptions antecedent to which deviations function best agree with the experimental conditions, or the function resulting from the application of which method for the combination of observations will best meet the research requirements.

---

butions but perhaps nearer to the mean distribution and considered by themselves might justify a least squares solution, and a solution for least quartile deviation by the method of the mid-point of the least interquartile range would give a quartile deviation smaller than either of the preceding median deviations, standard deviations and mean deviations greater than the smaller of the preceding standard and mean deviations, and still other ratios. The ratios of the minimum standard deviation, mean deviation, and quartile deviation obtainable from the data, $\sigma : \eta : r = 22.54 : 15.32 : < 7.25 = 1 : 0.678 : < 0.322$ or $\eta : \sigma : r = 1 : 1.473 : < 0.474$, suggest approximate conformity to the median distribution. The ratios of the mean deviations and the standard deviations for corresponding formulae for least deviations and for least squares in one dependent variable, in the normals, and in the other dependent variable in Table III for numerical example number 7, which are $\sigma : \eta = 1 : 0.893$, $1 : 0.833$, and $1 : 813$ suggest approximate conformity to the mean distribution in a deviations distribution intermediary to the mean distribution and a mid-point of the range distribution. The ratios of the mean and standard deviations for corresponding formulae in Table V for numerical example number 10 which are $\sigma : : \eta = 1 : 0.807$, $1 : 783$, $1 : 0.788$, and $1 : 0.773$ for females, and $1 : 0.746$, $1 : 0.764$, and $1 : 739$ for males, suggest distributions intermediary to the median and the mean, for females approximating to the mean, and for males tending toward the median. The decisiveness of the comparison of the calculated ratios with the theoretical ratios is further vitiated by the error to which the calculated deviations are subject.

The proportionate values of the probability integral for deviations of different magnitudes are also determinate for each deviations function. Subsequent to the derivation of averages or equations by both the method of least deviations and the method of least squares, some indication of which is more properly applicable to the data may be secured by determining whether the frequency distribution of the deviations of different magnitudes given by the average or equation derived by the method of least deviations is in closer accord with the distribution defined by the least deviations probability integral, $\frac{1}{2\eta}\int_{-\upsilon\eta}^{+\upsilon\eta} e^{-\frac{|x|}{\eta}}\,dx$, than in the distribution of the deviations given by the average or equation derived by the method of least squares with the least squares probability integral, $\frac{1}{\sigma\sqrt{2\pi}}\int_{-\upsilon\sigma}^{+\upsilon\sigma} e^{-\frac{x^2}{2\sigma^2}}\,dx$, in which υ represents an arbitrary factor. If the deviations distribuzion for the least deviations average or equation conforms to the least deviations probability integral, 0.5000 of the deviations

The mode and the methods of the modal point apply to the condition of relative exactitude in a certain part of the data with gross uncertainty in the rest of the data. In the physical sciences

---

will lie within the limits of $\pm$ 0.6932 η (**53**), and 0.6321, 0.8647, 0.9502, 0.9817, 0.9933, 0.9975, 0.9991, and 0.9997 within the limits of $\pm\,\eta$, $\pm\,2\,\eta$, $\pm\,3\,\eta$, $\pm\,4\,\eta$, $\pm\,5\,\eta$, $\pm\,6\,\eta$, $\pm\,7\,\eta$, and $\pm\,8\,\eta$, and if the deviations distribution fo the least squares average or equation conforms to the least squares probability integral, 0.5000 will lie within the limits of $\pm$ 0.6745 σ, and 0.6827, 0.9545, 0.9973, and 0.9999 within the limits of $\pm\,\sigma$, $\pm\,2\,\sigma$, $\pm\,3\,\sigma$, and $\pm\,4\,\sigma$. Comparison on this basis of the actual deviations distributions for each type of formula for which equations have been derived by each method in numerical example number 10 with the theoretical distribution for the method will therefore afford some indication of the relative appropriateness of the alternative methods to the data.

TABLE VII.

Comparison of Actual and Theoretical Percentages in Deviations Distribution.

SEX	TYPE OF FORMULA	METHOD																		
		LEAST DEVIATIONS									LEAST SQUARES						
		Quartiles				Successive Multiples of Mean Deviations					Quartiles				Standard Deviations		
		25	**25**	**25**	**25**	**63.2**	**86.5**	**95.0**	**98.2**	**99.3**	**25**	**25**	**25**	**25**	**68.3**	**95.5**	**99.7**
Female	Average	27	22	27	23	59	88	98	100		24	24	31	20	65	96	100
	Age	28	22	21	29	63	91	99	100		25	26	23	25	68	96	100
	Weight	34	16	26	24	62	84	100			23	24	28	24	70	96	100
	Weight-Age	29	21	19	31	54	92	99	100		27	26	21	26	66	97	100
Male	Average	28	22	22	28	58	89	98	99	99	24	27	28	22	72	96	99
	Age	28	22	22	28	58	89	98	99	99	23	28	28	21	72	97	99
	Weight	32	18	25	25	55	89	99	100		26	23	27	24	67	97	100
	Weight-Age	29	21	21	29	55	90	98	100		25	31	28	26	69	97	100

With averages and equations derived by the method of least deviations, smaller percentages of the deviations occur within the range of the quartiles and greater percentages beyond them, and smaller percentages within the range of the first multiple and greater percentages within the ranges of the succeeding multiples of the mean deviation than are consistent with the least deviations integral. With averages and equations derived by the method of least squares the percentages of the deviations within the ranges of the quartiles and the successive multiples of the standard deviation appear to be consistent with the least squares integral. Hence the method of least squares proves to be preferentially applicable to the data.

such data subject to irregular irregularities might accrue in the course of the perfection of experimental methods and result from combining data secured by methods of very different degrees of refinement. The greater deviations will not affect the result at all. Solutions for most lesser-deviations are applicable to situations in which an approximation is required to have a negligible error as often as possible but in which when it has an appreciable error the mere occurrence of such an error is of greater importance than its precise magnitude. Such would appear to be the psychic method of combining perceptions in the learning and performance of skilled acts in which inadequacy or failure of the response to the situation is only such divergence of the response from that response which is accompanied by the accustomed feeling of its rightness as is great enough to give rise to a sense of its incompletion, a feeling of a break in its rhythm, and a need for its repetition or readjustment. The application of the method of most lesser-deviations results in the definition of a relationship or an instrumentality for an estimate or forecast which will be as exact as possible as often as possible. Manufacturers of ready-made clothing might profit from the employment of implicit functions for most lesser-deviations relating various body lengths and circumferences since it would be of advantage to produce clothes fitting the greatest number of customers with the fewest alterations and how much of an alteration was required for such customers as still required any alteration at all would be of lesser consequence. And the stocking sales-person who applies the traditional artifice of estimating the length of the foot as equal to the circuit of the fist might do better with an explicit function for most lesser-deviations and as a result have fewer items returned for exchange.

The median and the methods of median loci and the median locus apply to the condition of haphazard variation in the exactitude of the data (**53**). Such data subject to random irregularities might result from experimental methods affected by certain uncontrolled or imperfectly controlled causes of inconstancy. Since the greater deviations affect the solution only in proportion to their own magnitude (**38**), the result does not depend unduly upon the less exact data (**38, 17, 53**). When the observations are discordant by reason of the varying degrees of their uncertainty, methods based on the median are not only appropriate, but are better than methods based on the mean (**38, 12, 13**). Solutions for least deviations are applicable to situations in which an approximation is required to have the least possible mean error.

The mean and the methods of mean loci and the mean locus apply to the condition of constant exactitude throughout the data. Such data require a perfected and controlled experimental method and technique. The greater deviations may then be allowed a disproportionately greater weight in determining the result. Solutions for least squares is applicable to situations in which an approximation is required to have the least possible standard error.

The mid-point of the range and the methods of the midpoint of the least range apply to the condition that the greatest inexactitude to which the data are subject is limited. Such data affected by only finite irregularities comprise the results of any and all experimental methods. The method of the least greatest-deviation was originally employed as a crude test of the hypothesis that a given functional relation obtained between the independent and dependent variables, which was tenable if the least greatest-deviation proved to be within the limits of the error to which the observations of the dependent variable were judged to be susceptible (**27, 38**). The greatest deviations are given all the weight in determining the result. Solutions for the least greatest-deviation are applicable to situations in which an approximation is required to be exact only within the extreme limits of a stated fixed interval (**28, 29**).

The method for the combination of observations selected may finally be influenced by considerations of computational exigency or convenience. The methods of most lesser-deviations and of the least greatest-deviation are merely the limiting cases and the methods of least deviations and of least squares the simplest prototypes for odd and even finite integers of the complete sequence of all the possible methods of positive-power sums of the absolute deviations. The method of least squares itself was advocated on grounds of analytic expediency (**20**) before its first theoretical justification was advanced (**21**). Analytic facility has perhaps been coequal in influence with a special alliance to the theory of mathematical probability (**55**) in enabling it to continue after thirteen decades to be in the active process of significant development. Mean loci and their intersections and the mean locus and its critical values are completely defined analytically. The application of the method of least squares accordingly depends merely on introducing the observation numbers into an automatic algebraic mechanism. The multiplication of the number of variables alters merely the compass and length but not the feasibility of the procedure. The other methods generally depend upon the preliminary transformation of the observation numbers into observation points or observation loci and therefore require geometric plots and diagrams. Upon these diagrams the modal point must be isolated, the median loci traced, or the extreme loci distinguished. The locus which, in the case of the mean, is obtained by a single stroke of analysis, must, in the case of the median, be traced segment by segment (**12**). At the third variable such graphic methods reach their ultimate extreme, if not of practicability, then of practical resource and endurance. The methods of the modal point, which for two variables are equivalent to drawing a line through the densest region of a swarm of observation points, as presented constitutes no more than a planned approach to a sensible solution. The methods of median loci and of the median locus are provided with analytic criteria possessed of directive and definitive efficacy. The methods of selected extreme observation points, arithmetically perhaps the most simple and direct of all methods for the analytic determination of straight lines, depend on spatial perceptions which may be tested numerically. The alternative methods of the midpoint of the least range are, like the methods of median loci and the median locus, provided with analytic criteria possessed of directive and definitive efficacy. The method of median loci, as has been truly remarked, involves to some extent the practice of trial and error (**57**). In fact all such methods as require the selection of observation points or of observation equations, the designation of a particular parameter point or the tracing of particular loci, or the substitution of trial values of the parameters are by such tokens zetetic. The prosperity of the search depends not on its method and instrumentalities alone but on knowing where to seek.

REFERENCES

(1) FECHNER, G. T., *Ueber den Ausgangswertn der kleinsten Abweichungssumme, dessen Bestimmung, Verwendung und Verallgemeinerung*, « Abh. d. Kön.-Sächs. Gsllschft d. Wssnschft », 18 (Math.-Phys. Classe, 11) : 4-6, 37-50 ; 17-37 ; 11 ; 15-16 (1878).

(2) FOSTER, R. M., *Derivation of Mode as Special Case of Generalized Mean*, 1922, personal communication ; result cited in E. V. HUNTINGTON, *Mathematical Memoranda*, « Handbook of Mathematical Statistics », ed. H. L. Rietz, Boston, 1924, 7.

(3) JAÇKSON, D., *Note on the Median of a Set of Numbers*, « Bul. Amer. Math. Soc. », Ser. 2, 27 : 160-164 (1921).

(4) LAPLACE, P. S., *Théorie analytique des probabilités*, Paris, 1812, 343-345.

(5) JACKSON, D., *On the Method of Least m^{th} Powers for a Set of Simultaneous Equations*, « Ann. Math. », 25 : 185-192 (1924).

(6) LAPLACE, *op. cit.*, 2me supplement, Paris, 1818, 42-44, 47.

(7) HORST, P., *A Proof that the Point from Which the Sum of the Absolute Deviations is a Minimum is the Median*, « J. Educ. Psychol. », 22 : 463-464 (1931).

(8) TRACHTENBERG, M. I., *A Note on a Property of the Median*, « J. Roy. Stat. Soc. », 78 : 454-455 (1915).

(9) JONES, D. C., *The Mean Deviation a Minimum when measured from the Median*, « A First Course in Statistics », London, 1927, 270-271.

(10) BOSCOVICH, R. J., *Supplementa*, « Philosophiae recentioris versibus traditae », by B. Stay, Rome, 1760, II, 420-424 ; quoted in C. MAIRE and R. J. BOSCOVICH, *Voyage astronomique et geographique*, traduit du Latin, augmenté de notes, Paris : 1770, 501-506.

(11) LAPLACE, P. S., *Sur quelques pointes du systeme de monde*, « Hist. de l'Acad. d. sci., Mem. de math. et de phys. », 1789 : 32-35 ; 21-29 (1793).

(12) EDGEWORTH, F. Y., *On Observations Relating to Several Quantities*, « Hermathena », 6 : 279-283 (1887).

(13) EDGEWORTH, F. Y., *A New Method of Reducing Observations Relating to Several Quantities*, « Phil. Mag. », Ser. 5, 24 : 222-223 (1887).

(14) TURNER, H. H., *On Mr. Edgeworth's Method of Reducing Observations Relating to Several Quantities*, « Phil. Mag », Ser. 5, 24 : 466-470 (1887).

(15) EDGEWORTH, F. Y., *On a New Method of Reducing Observations Relating to Several Quantities*, « Phil. Mag. », Ser. 5, 25 : 184-191 (1888).

(16) BOWLEY, A. L., *F. Y. Edgeworth's Contributions to Mathematical Statistics*, Roy. Stat. Soc., London, 1928, 103-109 ; 56-61.

(17) RHODES, E. C., *Reducing Observations by the Method of Minimum Deviations*, « Phil. Mag. », Ser. 7, 9 : 974-992 (1930).

(18) EDGEWORTH, F. Y., *On the Use of Medians for Reducing Observations Relating to Several Quantities*, « Phil. Mag. », Ser. 6, 46 : 1074-1088 (1923).

(19) EDGEWORTH, F. Y., *Probability*, « Encyclopedia Britannica », 11th ed., 22 : 399 (1911).

(**20**) LEGENDRE, A. M., *Sur la Méthode des moindres quarrés.* « Nouvelles Methodes pour la Déterminations des Orbites des Comètes», Paris, 1805, viii, 72-75.

(**21**) GAUSS, C. F., *Theoria Motus Corporum Cœlestium in Sectionibus Conicis Solem Ambientium,* Hamburg: 1809, 205-221.

(**22**) PEARSON, K., *On Lines and Planes of Closest Fit to Systems of Points in Space,* « Phil. Mag. », Ser. 6, 2: 559-572 (1901).

(**23**) REED, L. J., *Fitting Straight Lines,* « Metron », 1 (3): 54-61 (1921).

(**24**) SNOW, E. C., *On Restricted Lines and Planes of Closest Fit to Systems of Points in any Number of Dimensions,* « Phil. Mag. », Ser. 6, 21: 367-386 (1911).

(**25**) GLAUERT, H., *The Determination of the Best Linear Relationship connecting any Number of Variables,* « Phil. Mag. », Ser. 6, 50: 205-207 (1925).

(**26**) RHODES, E. C., *On Lines and Planes of Closest Fit,* « Phil. Mag. », Ser. 7, 3: 357-364 (1927).

(**27**) LAPLACE, P. S., *Mèmoire sur la figure de la terre,* « Hist. de l'Acad. roy. d. sci., Mèm. de Math. et de Phys. », 1783: 19-23 (1786).

(**28**) GOEDSEELS, P. J. E., *Théorie des erreurs d'observation,* 3. ed., Louvain, 1909, 4-5, 13-15, 19-20, 32-35; 22-31.

(**29**) SMITH, T., *The Mid-Course Method of Fitting a Parabolic Formula of any Order to a Set of Observations,* « Proc. Phys. Soc. », 46: 560-569, 572-573 (1934).

(**30**) GOEDSEELS, P. J. E., *Simplifications de la Mèthode la plus approximative et de l'approximation minima,* « Ann. de la Soc. sci. de Bruxelles », 35^2: 351-368 (1911).

(**31**) DE LA VALLÉE-POUISSIN, C. J., *Sur la Mèthode de l'approximation minimum,* « Ann. Soc. sci. Bruxelles », 35^2: 1-16 (1911).

(**32**) MILLIKAN, R. A., and FLETCHER, H., *On the Question of Valency in Gaseous Ionization,* « Phil. Mag. », Ser. 6, 21: 753-770 (1911).

(**33**) EGGLESTON, C., *Biological Standardization of the Digitalis Bodies by the Cat Method,* « Amer. J. Pharm », 85: 99-122 (1913).

(**34**) EDGEWORTH, F. Y., and BOWLEY, A. L., *Methods of Representing Statistics of Wages and other Groups not Fulfilling the Normal Law of Error,* « J. Roy. Stat. Soc. », 65: 339-351 (1902).

(**35**) *Number of Persons Employed at Various Rates of Wages,* 1886, « General Report on the Wages of the Manual Labour Classes in the United Kingdom with Tables of the Average Rates of Wages and Hours of Labour of Persons employed in several of the principal Trades in 1886 and 1891 », London, Commercial, Labour and Statistical Department, Board of Trade, 1893, (C-6889), 470, 476.

(**36**) THACKERAY, W. G., *On the Semi-diameter of Venus,* « Mon. Not. Roy, Astron. Soc. », 46: 335-336 (1886).

(**37**) *Vertical Diameters of Venus... compared with the corresponding results of the Nautical Almanac,* « Results of the Astronomical Observations Made at the Royal Observatory, Greenwich », (Extracted from the Greenwich Observations), 1873-1884, Observer AD.

(38) LAPLACE, P. S., *Mécanique Céleste,* trans. with com., N. Bowditch, Boston, 1832, II, 434-442 ; 417-424 ; 424-434 ; 468-478 ; 443-450.

(39) WARREN, G. F., and PEARSON, F. A., *Interrelationships of Supply and Price,* Cornell Univ. Agri. Exper. Sta., Bul. 466, 116-119 (1928).

(40) *Statistical summary of output and of the costs of production, proceeds and profits of the coal mining industry for the quarter,* London, Mines Department, 1921-1925, South Wales and Monmouth.

(41) DREYER, G., *The Normal Basal Metabolism in Man, and its Relation to the Size of the Body and Age, Expressed in Simple Formulae,* « Lancet », 199 : 289-291 (1920).

(42) HARRIS, J. A., and BENEDICT, F. G., *A Biometric Study of Basal Metabolism in Man,* Carnegie. Inst. Washington, Pub. No. 279, 1919, Table *C,* 28-136 ; Table *D,* 1-103.

(43) JENKINS, R. L., *Basal Metabolism Standards : A Statistical Comparison of their Prediction Values,* « J. Nutrition », 4 : 305-321 (1931).

(44) PEARSON, K., *On the Systematic Fitting of Curves to Observations,* « Biometrika », 1 : 265-302 ; 2 : 1-23 (1902).

(45) MAYER, T., *Abhandlungen über die Umwälzung des Mondes um seine Axe,* « Kosmog. Nachr. u. Samml. a. d. Jahr, 1748 », Nürnberg, 1750, 153-156.

(46) DUFTON, A. F., *Correlation,* « Nature », 121 : 866 (1928).

(47) CAUCHY, A. L., *On a New Formula for solving the Problem of Interpolation in a Manner applicable to Physical Investigation,* « Phil. Mag.», Ser. 3, 8 : 459-468 (1836).

(48) THIELE, T. N., *Theory of Observations* London, 1903, 5-15 ; « Ann. Math. Stat. », 2 : 169-179 (1931).

(49) KEYNES, J. M., *The Principal Averages and the Laws of Error which lead to them,* « J. Roy. Stat. Soc. », 74 : 322-331 (1911) ; *A Treatise on Probability,* London, 1921, 194-205.

(50) LAPLACE, P. S., *Mèmoire sur la probabilité des causes par les événements,* « Mém. de Math. et de Phys. prés. a l'Acad. roy. d. sci. », 6 : 634-638 (1774).

(51) GLAISHER, J. W. L., *On the Law of the Facility of Errors of Observation, and on the Method of Least Squares,* « Mem. Roy. Astron. Soc »., 39 : 121-123 (1872).

(52) WHITTAKER, E. T., and ROBINSON, G., *The Calculus of Observations, A Treatise on Numerical Mathematics,* London, 1924, 187-188; 186-187.

(53) BOND, W. N., *Probability and Random Errors,* London, 1935, 57-63, 132-134 ; 41-43.

(54) BRUNT, D., *The Combination of Observations,* Combridge, 1917, 27-28, 30-31.

(55) COOLIDGE, J. L., *An Introduction to Mathematical Probability,* Oxford, 1925, 102-112, 113-117.

(56) PEARSON, K., *Second Supplement to a Memoir on Skew Variation* « Phil. Trans. Roy. Soc. London », Ser. A, 216 : 434-435, 447-448, 450 ; 436, 442-443, 451 ; 431-432, 438-440, 445-447, 448-449 (1916).

(57) YULE, G. U., *An Introduction to the Theory of Statistics,* 10th ed., London, 1932, 188.

S. S. WILKS

The Analysis of Variance and Covariance in Non-Orthogonal Data

The simple arithmetical procedure suitable for analyzing the variance in data from orthogonal and other well-designed experiments has been thoroughly discussed [1, 2, 4] and widely used. However, there are many situations in which non-orthogonal data is unavoidable. For example, in experiments on live-stock in which sex cannot be determined at the beginning of the experiment sex and treatments will not be orthogonal. Even in experiments which are designed to be orthogonal, it is not infrequent that damage or loss of some of the observations beset the orthogonality. Again, in problems arising in medical work in which all available data must be used, orthogonality with respect to each of several factors cannot in general be attained. The same type of problem arises in attempts at estimating post-parole behavior from the records of released and paroled criminals.

Use of Discrete Variates in a 2×3 Classification.

The usual procedure, which has been discussed by Yates [3] for handling problems involving non-orthogonal data consists in determining constants in linear combination by least squares, computing the sum of squared residuals, and using an appropriate test of significance for the particular hypothesis considered. In a considerable number of such problems the pertinent questions can be answered by simply applying significance tests to various sums of squared residuals computed under the different hypotheses considered. The method of determining constants and computing residuals is cumbersome at best, but there is a method of finding the sum of squared residuals without having to determine the constants, which the author proposes to discuss here. The device is simply one of introducing variates which take the values 0 or 1 and reducing the whole procedure to a problem in classical regression theory, and evaluating certain determinants.

We shall first consider the theory of the method and then illustrate it with a set of data. First, consider a two-way classification arranged for convenience in two rows and three columns, so that each observation is characterized by a continuous variate x, a particular row and a particular column. Without setting up any specific hypothesis at present we shall suppose that a general mean and any particular row or column operate additively to produce the *true* mean of the x's falling in this row and column. This can be expressed as follows

$$(1) \qquad m_i = \mu + \alpha_1 \delta_{1i} + \alpha_2 \delta_{2i} + \beta_1 \varepsilon_{1i} + \beta_2 \varepsilon_{2i} + \beta_3 \varepsilon_{3i}$$

where

$$(2) \qquad \alpha_1 + \alpha_2 = 0\,, \qquad \beta_1 + \beta_2 + \beta_3 = 0$$

and $i = 1, 2, \ldots N$, N being the number of observations. The $\delta's$ can be either 0 or 1, but for any observation, that is, any value of i, one of them must be 1 and the other 0. A similar statement holds for the $\varepsilon's$. This is due to the fact that each observation belongs to one row and one column. (1) can be regarded as a classical regression equation in which μ, the $\alpha's$ and the $\beta's$ are to be estimated subject to the restrictions (2). The independent variates are the δ's and the $\varepsilon's$. We form the sum of squares

$$(3) \qquad \varphi = \sum_{i=1}^{N} (x_i - m_i)^2$$

which must be minimized subject to (2). This is equivalent to finding the unrestricted minimum of

$$(4) \qquad \varphi' = \sum_{i=1}^{N} (x_i - m_i)^2 + \lambda_1 (\alpha_1 + \alpha_2) + \lambda_2 (\beta_1 + \beta_2 + \beta_3).$$

The normal equations are now

$$
(5)\quad
\begin{aligned}
\frac{\partial \varphi'}{\partial \mu} &= -\Sigma\,(x_i - m_i) &&= 0\\
\frac{\partial \varphi'}{\partial \alpha_1} &= -\Sigma\,\delta_{1i}\,(x_i - m_i) + \lambda_1 &&= 0\\
\frac{\partial \varphi'}{\partial \alpha_2} &= -\Sigma\,\delta_{2i}\,(x_i - m_i) + \lambda_1 &&= 0\\
\frac{\partial \varphi'}{\partial \beta_1} &= -\Sigma\,\varepsilon_{1i}\,(x_i - m_i) \qquad + \lambda_2 &&= 0\\
\frac{\partial \varphi'}{\partial \beta_2} &= -\Sigma\,\varepsilon_{2i}\,(x_i - m_i) \qquad + \lambda_2 &&= 0\\
\frac{\partial \varphi'}{\partial \beta_3} &= -\Sigma\,\varepsilon_{3i}\,(x_i - m_i) \qquad + \lambda_2 &&= 0\\
\frac{\partial \varphi'}{\partial \lambda_1} &= \alpha_1 + \alpha_2 &&= 0\\
\frac{\partial \varphi'}{\partial \lambda_2} &= \beta_1 + \beta_2 + \beta_3 &&= 0
\end{aligned}
$$

We note that $\Sigma\,\delta_{1i}\,\delta_{2i} = 0$, $\Sigma\,\delta_{1i}^2 = \Sigma\,\delta_{1i} = N_{1.}$, $\Sigma\,\delta_{2i}^2 = \Sigma\,\delta_{2i} = N_{2.}$, $\Sigma\,\delta_{1i}\,\varepsilon_{1i} = N_{11}$, $\Sigma\,\varepsilon_{1i}^2 = \Sigma\,\varepsilon_{1i} = N_{.1}$, $\Sigma\,\delta_{1i}\,\alpha_i = N_{1.}\,\overline{X}_{1.}$, etc., where $N_{1.}$ is the number and $\overline{X}_{1.}$ the mean of the x's of the individuals in the first column, with similar meanings for $N_{2.}$ and $\overline{X}_{2.}$. Similarly, $N_{.1}$ is the number and $\overline{X}_{.1}$ the mean of the x's of the individuals in the first column, with similar meanings for $N_{.2}$, $\overline{X}_{.2}$, $N_{.3}$, $\overline{X}_{.3}$. N_{11} is the number of observations in the first row and first column, with similar meanings for N_{12}, N_{21}, etc. Thus, the equations in (5) can be written as

(6)

$$
\begin{aligned}
&N\,\mu + N_{1.}\,\alpha_1 + N_{2.}\,\alpha_2 + N_{.1}\,\beta_1 + N_{.2}\,\beta_2 + N_{.3}\,\beta_3 &&= N\,\overline{X}\\
&N_{1.}\,\mu + N_{1.}\,\alpha_1 + N_{11}\,\beta_1 + N_{12}\,\beta_2 + N_{13}\,\beta_3 + \lambda_1 &&= N_{1.}\,\overline{X}_{1.}\\
&N_{2.}\,\mu + N_{2.}\,\alpha_2 + N_{21}\,\beta_1 + N_{22}\,\beta_2 + N_{23}\,\beta_3 + \lambda_1 &&= N_{2.}\,\overline{X}_{2.}\\
&N_{.1}\,\mu + N_{11}\,\alpha_1 + N_{21}\,\alpha_2 + N_{.1}\,\beta_1 + \lambda_2 &&= N_{.1}\,\overline{X}_{.1}\\
&N_{.2}\,\mu + N_{12}\,\alpha_1 + N_{22}\,\alpha_2 + N_{.2}\,\beta_2 + \lambda_2 &&= N_{.2}\,\overline{X}_{.2}\\
&N_{.3}\,\mu + N_{13}\,\alpha_1 + N_{23}\,\alpha_2 + N_{.3}\,\beta_3 + \lambda_2 &&= N_{.3}\,\overline{X}_{.3}\\
&\alpha_1 + \alpha_2 &&= 0\\
&\beta_1 + \beta_2 + \beta_3 &&= 0
\end{aligned}
$$

The determinant of the coefficients of μ, the α's, β's and λ's in these equations is not zero; therefore, the equations have a unique solution. Both λ's will be zero as can be easily seen from their values in terms of determinants. Thus, the values of μ, the α's and β's obtainable from these equations minimize the sum of squares in (3) when the conditions (2) are imposed. The value of the minimum value of φ which is the same as that of φ' is obtained by substituting the values of μ, the α's and β's from (6) into (3). If this substitution is performed, using determinants, and if the residuals $x_i - m_i$ are squared and summed over all observations, that is, with respect to the i, we get the sum of the squared residuals to be expressible as the ratio of two determinants as follows,

$$R(\mu, \alpha_r, \beta_s) = \frac{1}{\Delta} \begin{vmatrix} \Sigma X^2 & N\overline{X} & S_{1\cdot} & S_{2\cdot} & S_{\cdot 1} & S_{\cdot 2} & S_{\cdot 3} & 0 & 0 \\ N\overline{X} & N & N_{1\cdot} & N_{2\cdot} & N_{\cdot 1} & N_{\cdot 2} & N_{\cdot 3} & 0 & 0 \\ S_{1\cdot} & N_{1\cdot} & N_{1\cdot} & 0 & N_{11} & N_{12} & N_{13} & 1 & 0 \\ S_{2\cdot} & N_{2\cdot} & 0 & N_{2\cdot} & N_{21} & N_{22} & N_{23} & 1 & 0 \\ S_{\cdot 1} & N_{\cdot 1} & N_{11} & N_{21} & N_{\cdot 1} & 0 & 0 & 0 & 1 \\ S_{\cdot 2} & N_{\cdot 2} & N_{12} & N_{22} & 0 & N_{\cdot 2} & 0 & 0 & 1 \\ S_{\cdot 3} & N_{\cdot 3} & N_{13} & N_{23} & 0 & 0 & N_{\cdot 3} & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \end{vmatrix} \qquad (7)$$

where $N_{1\cdot}\overline{X}_{1\cdot} = S_{\cdot 1}$, etc., and Δ is the determinant obtained by deleting the first row and column of the determinant in the numerator of the expression for $R(\mu, \alpha_r, \beta_s)$. $R(\mu, \alpha_r, \beta_s)$ will have $N - 4$ degrees of freedom, that is, the number of observations minus the number of *independent* constants, which is 4 in this case (μ, *one* α and *two* β's). Under the assumption that the expected value of x is given by (1), and if $N_{11} = N_{12} = \ldots = N_{23}$, effects due to rows and columns are said to be orthogonal. In the orthogonal case and also other special cases including that of proportional frequencies in rows or columns the determinants in (7) are very easy to evaluate and $R(\mu, \alpha_r, \beta_s)$ reduces to the usual error sum of squares entering into the analysis of variance associated with this type of data. In the non-orthogonal case, the evaluation of the determinants becomes somewhat more tedious. If an appropriate significance test, based on the sums of squares of residuals, indicates that certain main effects are existent, then, under the assumption that μ, the α's and β's operate additively, estimates of these effects (that is, estimates of the $\alpha's$ and $\beta's$) can be obtained from equations (6) in the usual way. Equations (6) can always be set up from the elements of the determinant in the numerator of $R(\mu, \alpha_r, \beta_s)$ as given in (7). To estimate, for example, β_1 we use the ratio $\frac{\Delta_{\beta 1}}{\Delta}$ where Δ is the same as that in (7) and $\Delta_{\beta 1}$ is the determinant obtained by interchanging the first and fifth columns in the numerator of (7) and then deleting the first row and column.

Unless there are *a priori* grounds for assuming effects due to rows and columns operate additively as in (1), it is important to test for the existence of interaction between rows and columns. To do this we assume the true mean of the r-th row and s-th column to be $\mu + \overline{\alpha\beta}_{rs}$ where $\Sigma\, \overline{\alpha\beta}_{rs} = 0$ and calculate the sum of the squares of the residuals within classes, say $R(\mu, \overline{\alpha\beta}_{rs})$, which will have $N - 6$ degrees of freedom in this case. If interaction between rows and columns is non-existent, then $R(\mu, \alpha_r, \beta_s) - R(\mu, \overline{\alpha\beta}_{rs})$ and $R(\mu, \overline{\alpha\beta}_{rs})$ will be independently distributed with 2 and $N - 6$ degrees of freedom. Thus R. A. Fisher's z test can be used. The analysis of variance table for testing the interaction would take the form

(8)

	Degrees of Freedon	Sums of Squares
Interaction	2	$A - B$
Error	N—6	$R(\mu, \overline{\alpha\beta}_{rs}) = B$
Total.......	N—4	$R(\mu, \alpha_r, \beta_s) = A$

We compute $z = \frac{1}{2}\log_e \left[\frac{(N-6)(A-B)}{2\,B}\right]$ and determine significance from Fisher's tables with $n_1 = 2$, and $n_2 = N - 6$. Instead, Snedecor's F, that is $e^{\frac{1}{2}|z|}$, could be used.

If interaction is non-significant, we can then proceed to make a significance test for main effects, for example, due to columns. The procedure then is to assume the $\beta's$ to be zero in (1) and determine the sum of the squares of residuals (3) subject to the condition that $\alpha_1 + \alpha_2 = 0$. Denoting the sum of squares by

$R(\mu, \alpha_r)$ we find by the method used in determining $R(\mu, \alpha_r, \beta_s)$ that

$$(9) \qquad R(\mu, \alpha_s) = \frac{1}{\Delta'} \begin{vmatrix} \Sigma x^2 & N\overline{X} & S_1. & S_2. & 0 \\ N\overline{X} & N & N_1. & N_2. & 0 \\ S_1. & N_1. & N_1. & 0 & 1 \\ S_2. & N_2. & 0 & N_2. & 1 \\ 0 & 0 & 1 & 1 & 0 \end{vmatrix},$$

Δ' having a meaning similar to that of Δ in (7). It will be noticed that the determinants in (9) are obtained by deleting the rows and columns in the determinants in (7) associated with the β's and λ_2. The analysis of variance table for testing for effects due to columns is,

		Degrees o Freedom	Sums of Squares
	Columns	2	$A - B$
(10)	Error	$N - 4$	$R(\mu, \alpha_r, \beta_s) = B$
	Total	$N - 2$	$R(\mu, \alpha_r) = A$

Similarly, one can set up a table for testing for effect due to rows.

It should be pointed out that one can always compare variation within classes with variation between classes and test whether the estimates of the subclass means $\mu + \overline{\alpha\beta}_{rs}$ vary excessively from each other. But such a test would serve hardly more than an exploratory step, for, in the case of significance, it alone would furnish no information as to whether the variation was due to rows, to columns, or both, either additively, or in a more complex law of operation. In the event of non-significance, the test would indicate, of course, the futility of further more refined work on the same data.

Triple Classification

The extension of the above procedure to an $m \times n$ classification is immediate. In the case of an $m \times n \times p$ classification, one would set up the determinants defining the sum of squares $R(\mu, \alpha_r, \beta_s, \gamma_t)$ from which $R(\mu, \alpha_r, \beta_s)$, $R(\mu, \alpha_r)$, etc. can be obtained by deleting certain rows and columns. As an example of a triple classification consider the $2 \times 2 \times 3$ case, in which

$$(11) \qquad m_i = \mu + \alpha_1 \delta_{1i} + \alpha_2 \delta_{2i} + \beta_1 \varepsilon_{1i} + \beta_2 \varepsilon_{2i} + \gamma_1 \theta_{1i} + \gamma_2 \theta_{2i} + \gamma_3 \theta_{3i}$$

where

$$(12) \qquad \alpha_1 + \alpha_2 = 0\,, \quad \beta_1 + \beta_2 = 0\,, \quad \gamma_1 + \gamma_2 + \gamma_3 = 0\,.$$

By the same steps used in the 2×3 case we find

$$(13) \qquad R(\mu, \alpha_r, \beta_s, \gamma_t) = \frac{1}{\Delta''} \begin{vmatrix} \Sigma x^2 & N\overline{X} & S_{1\cdot\cdot} & S_{2\cdot\cdot} & S_{\cdot 1\cdot} & S_{\cdot 2\cdot} & S_{\cdot\cdot 1} & S_{\cdot\cdot 2} & S_{\cdot\cdot 3} & 0 & 0 & 0 \\ N\overline{X} & N & N_{1\cdot\cdot} & N_{2\cdot\cdot} & N_{\cdot 1\cdot} & N_{\cdot 2\cdot} & N_{\cdot\cdot 1} & N_{\cdot\cdot 2} & N_{\cdot\cdot 3} & 0 & 0 & 0 \\ S_{1\cdot\cdot} & N_{1\cdot\cdot} & N_{1\cdot\cdot} & 0 & N_{11\cdot} & N_{12\cdot} & N_{1\cdot 1} & N_{1\cdot 2} & N_{1\cdot 3} & 1 & 0 & 0 \\ S_{2\cdot\cdot} & N_{2\cdot\cdot} & 0 & N_{2\cdot\cdot} & N_{21\cdot} & N_{22\cdot} & N_{2\cdot 1} & N_{2\cdot 2} & N_{2\cdot 3} & 1 & 0 & 0 \\ S_{\cdot 1\cdot} & N_{\cdot 1\cdot} & N_{11\cdot} & N_{21\cdot} & N_{\cdot 1\cdot} & 0 & N_{\cdot 11} & N_{\cdot 12} & N_{\cdot 13} & 0 & 1 & 0 \\ S_{\cdot 2\cdot} & N_{\cdot 2\cdot} & N_{12\cdot} & N_{22\cdot} & 0 & N_{\cdot 2\cdot} & N_{\cdot 21} & N_{\cdot 22} & N_{\cdot 23} & 0 & 1 & 0 \\ S_{\cdot\cdot 1} & N_{\cdot\cdot 1} & N_{1\cdot 1} & N_{2\cdot 1} & N_{\cdot 11} & N_{\cdot 21} & N_{\cdot\cdot 1} & 0 & 0 & 0 & 0 & 1 \\ S_{\cdot\cdot 2} & N_{\cdot\cdot 2} & N_{1\cdot 2} & N_{2\cdot 2} & N_{\cdot 12} & N_{\cdot 22} & 0 & N_{\cdot\cdot 2} & 0 & 0 & 0 & 1 \\ S_{\cdot\cdot 3} & N_{\cdot\cdot 3} & N_{1\cdot 3} & N_{2\cdot 3} & N_{\cdot 13} & N_{\cdot 23} & 0 & 0 & N_{\cdot\cdot 3} & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \end{vmatrix}$$

where Δ'' has a meaning similar to Δ in (7). The meanings of the N's are similar to those in the 2×3 case. For example, $N_{1\cdot\cdot}$ is the number of observations in which α_1 is operative, $N_{2\cdot 3}$ is the number in which α_2 and γ_3 operate, etc. $S_{1\cdot\cdot} = N_{1\cdot\cdot}\overline{X}_{1\cdot\cdot}$, etc., where $\overline{X}_{1\cdot\cdot}$ is the mean of the x's on which α_1 operate. It will be noticed that due to the three sets of conditions (12) there are now three rows and columns of 0's and 1's bordering the determinant in (13). The condition for complete orthogonality for all three sets of effects under the additive hypothesis expressed by (11), is that the numbers within each of the types $N_{rs\cdot}$, $N_{r\cdot t}$, and $N_{\cdot st}$ be equal, and, of course,

$$\sum_{r,s} N_{rs\cdot} = \sum_{r,t} N_{r\cdot t} = \sum_{s,t} N_{\cdot st} = \mathrm{N}\,.$$

If there had been only two γ's, and each N with double subscripts had been equal to 1, we would then have a 2×2 Latin Square, with the α's, β's and γ's referring to rows, columns and treatments.

The determinants in the various sums of squares of residuals would then be extremely simple, and upon evaluation would reduce to the familiar expression entering into the analysis of variance for Latin Squares.

Now, suppose we wish to test the hypothesis that the effects due to the three classifications are additive, that is, that interaction is non-existent. We would calculate the sum of squared residuals within classes, say $R\,(\mu\,,\,\overline{\alpha\,\beta\,\gamma_{rst}})$ which is similar to $R\,(\mu\,,\,\overline{\alpha\,\beta_{rs}})$ for the two-way classification, and apply Fisher's z-test to the following table,

		Degrees of Freedom	Sum of Squares
	Interaction	7	$A - B$
(14)	Error	$N-12$	$R\,(\mu\,,\,\overline{\alpha\,\beta\,\gamma_{rst}}\,) = B$
	Total	$N-5$	$R\,(\mu\,,\alpha_r\,,\,\beta_s\,,\,\gamma_t) = A$

If the z-test does not yeld significance here then, of course, no interaction of any order can be assumed to exist. However, if there is interaction it may be only of first order. To test for first order interaction, as for example, that between the effects represented by the $\alpha's$ and $\beta's$ when it is assumed that the effects represented by the $\gamma's$ and the joint effects of the $\alpha's$ and $\beta's$ are additive, we determine the sum of squared residuals $R\,(\mu\,,\,\overline{\alpha\,\beta_{rs}}\,,\,\gamma_t)$ where $\sum_t \gamma_t = 0$, $\sum_{r,\,s} \overline{\alpha\,\beta_{rs}} = 0$, and use the information in the table,

		Degrees of Freedom	Sums of Squares
	Interaction $(\alpha \times \beta)$..	1	$A - B$
(15)	Error	$N-6$	$R\,(\mu\,,\,\overline{\alpha\,\beta_{rs}}\,,\,\gamma_t\,) = B$
	Total	$N-5$	$R\,(\mu\,,\,\alpha_r\,,\,\beta_s\,,\,\gamma_t) = A$

It will be noticed that $R\,(\mu\,,\overline{\alpha\,\beta_{rs}}\,,\gamma_t)$ is essentially the sum of squares of residuals for a 2×6 classification, and would be similar to (7). Tables similar to (15) can be set up for the other two types of first order interactions.

If the effects of the three classifications are additive, that is, interactions non-existent, one can then test for the significance of any one of the main effects, for example, those due to the $\gamma's$ by the following table,

		Degrees of Freedom	Sums of Squares
	Main effect (γ)	2	$A - B$
(16)	Error.	$N - 5$	$R(\mu, \alpha_r, \beta_s, \gamma_t) = B$
	Total.	$N - 3$	$R(\mu, \alpha_r, \beta_s) = A$

Although the example which we have considered in a triple classification is only the $2 \times 2 \times 3$ case, the method readily extends to the $m \times n \times p$ case. There would be m α's, n β's, and p γ's so that $\sum_r \alpha_r = \sum_s \beta_s = \sum_t \gamma_t = 0$. Analysis of variance tables can be constructed from the R's for the various hypotheses which arise. For example, the analysis of variance table corresponding to (15) would be

		Degrees of Freedom	Sums of Squares
	Interaction ($\alpha \times \beta$) ..	$(m-1)(n-1)$	$A - B$
(17)	Error.	$N-p-mn+1$	$R(\mu, \overline{\alpha\beta}_{rs}, \gamma_t) = B$
	Total.	$N-m-n-p+2$	$R(\mu, \alpha_r, \beta_s, \gamma_t) = A$

Higher Order Classifications.

The extension of the above procedure to a general k-way classification is straightforward. Theoretically, there could possibly be interactions of order as high as $k - 1$, but for practical purposes, particularly with few observations, there would be little value in considering interactions of order higher than first or second. The determinants entering into each R will each be bordered by k rows and columns of 0's and 1's; a row and column corresponding to each set of effects. Within such a given row (or column) the 1's are placed in the columns (or rows) associated with the particular set of effects to which the row (or column) corresponds, and 0's are placed elsewhere. For example, in (7) the last row corresponds to the β effects and 1's are placed in the columns associated with the β effects. A similar statement holds for the last column.

To test for interaction between the α and β effects, say, assuming the remaining effects operate additively, we use the z-test on the information in the following table,

		Degrees of Freedom	Sums of Squares
	Interaction $(\alpha \times \beta)$ ..	$(m-1)(n-1)$	$A - B$
(18)	Error.	$N-mn-p-\ldots-g+k-2$	$R(\mu, \overline{\alpha\beta_{rs}}, \gamma_t, \ldots \omega_z) = B$
	Total	$N-m-n-p-\ldots-g+k-1$	$R(\mu, \alpha_r, \beta_s, \ldots \omega_z) = A$

where there are m α's, n β's ... g ω's and k sets of effects (k-way classification).

If interaction between the α and β effects is non-significant we can test for main effects. For example, the analysis of variance constituents for testing for effects due to the α's are as follows,

		Degrees of Freedom	Sums of Squares
	Main effect (α)	$m-1$	$A - B$
(19)	Error	$N-m-n-p-\ldots-g+k-1$	$R(\mu, \alpha_r, \beta_s, \ldots \omega_z) = B$
	Total	$N-n-p-\ldots-g+k-2$	$R(\mu, \beta_s, \gamma_t, \ldots \omega_z) = A$

Analysis of Covariance

Suppose that for each observation x a value of a related continuous variate y is known. It is often important to utilize the information furnished by y to eliminate the effect due to y and to make more accurate estimates of main effects on x when the effect of y is additive. For example, in a pig feeding experiment it is important to take account of initial weight in order to test from final weights for effects due to feeding treatments. The scheme which we have used for setting up the R 's in determinantal form requires very little alteration to include a continuous variate and hence the case of analysis of covariance. Thus, if we had a y variate in the 2×3 case which operated linearly, the true mean of the i - *th* observation would now be

$$m_i' = \rho\, y_i + m_i$$

where m_i is given by (1). If we let $R(\rho\,;\mu, \alpha_r, \beta_s)$ be the minimum of the sum of the squares of the residuals $\sum_i (x_i - m_i')^2$ we get

$$R(\rho\,;\mu, \alpha_r, \beta_s) = \frac{\Delta_1}{\Delta_2}$$

whereΔ_1, is the determinant in the numerator of (7) with another row inserted between the first two rows and another column inserted between the first two columns; the elements of the inserted row and column being $\Sigma xy, \Sigma y^2, Ny, S_1'., S_2'., S._1', S._2', S._3', 0, 0$. $S_1'., S_2.', \ldots$ have the same meaning for the y's as the $S_1., S_2., \ldots$ have for the x's. Δ_2 is the determinant obtained by deleting the first row and column of Δ_1. In the case of several y variates, a corresponding number of rows and columns are inserted.

In the important case in which the true value of the regression coefficient ρ can be assumed the same in all sub-classes the problem of testing for interactions and main effects is exactly as before, except that we now have to include the extra row and column in the determinants, when making the tests. The effect of this is to reduce the number of degrees of freedom in each R by unity. The case in which ρ depends on the sub-classes, for example, is of a form similar to (1), the problem is straightforward and can be handled by the method of determinants, but we shall not enter into the details here.

Illustrative Example.

Nice and Fishman [5] reported a study on change in blood viscosity in normal and adrenalectomized rats following emotional excitement. On each of 24 normal rats they measured the viscosity of the blood, at a temperature of 26° C, drawn from the heart of the rat in quiet state. The viscosity was also measured in blood drawn after excitement (fear and pain caused by electric shocks). The same thing was done on each of 16 rats from 3 to 5 days after being adrenalectomized. The blood in some cases was venous and others arterial, it being impossible to draw either at will. It is clear that there is no orthogonalization at all in this experiment. For purposes of illustration we shall examine certain hypotheses.

We shall let x be the blood viscosity in excited state, y the viscosity in quiet state. ρ will refer to the regression coefficient of x on y. μ will refer to the general mean, α_1, α_2 to arterial and venous blood in excited state, β_1, β_2 to arterial and venous blood in quiet state, and γ_1, γ_2 to being normal and adrenalectomized. Each rat is therefore measured on x and y and classified three ways, according to the α's, β's and γ's. Thus, Nice and Fishman's data may be presented in the following form, from wich determinants for the R's for making various tests may be set up and computed:

Viscosity Ex. State	Viscosity Qu. State	Blood Ex. State		Blood Qu. State		Treatment	
x	y	$\delta_1\ (A)$	$\delta_2\ (V)$	$\varepsilon_1\ (A)$	$\varepsilon_2\ (V)$	$\theta_1(Nor.)$	$\theta_2\ (Ad.)$
3.90	3.29	1	0	1	0	1	0
5.25	3.91	1	0	1	0	1	0
4.75	4.64	1	0	1	0	1	0
4.19	3.55	1	0	1	0	1	0
5.06	3.67	1	0	1	0	1	0
5.55	4.18	1	0	1	0	1	0
5.02	3.74	1	0	1	0	1	0
5.14	4.67	1	0	1	0	1	0
4.64	3.03	0	1	1	0	1	0
5.10	4.61	0	1	1	0	1	0
5.03	3.84	0	1	1	0	1	0
6.15	4.20	0	1	0	1	1	0
4.64	3.81	0	1	0	1	1	0
5.81	4.15	0	1	0	1	1	0
5.32	4.40	0	1	0	1	1	0
5.10	4.68	1	0	0	1	1	0
5.20	4.50	1	0	0	1	1	0
5.03	4.65	1	0	0	1	1	0
5.28	4.36	1	0	0	1	1	0
4.83	4.46	1	0	0	1	1	0
5.18	4.36	1	0	0	1	1	0
5.00	4.53	1	0	0	1	1	0
5.67	4.70	1	0	0	1	1	0
5.30	4.33	1	0	0	1	1	0
3.76	3.45	1	0	1	0	0	1
5.89	4.26	1	0	1	0	0	1
6.65	4.71	1	0	1	0	0	1
3.35	3.14	1	0	1	0	0	1
4.07	3.45	1	0	1	0	0	1
6.92	5.01	1	0	1	0	0	1
5.25	4.43	1	0	1	0	0	1
5.74	4.91	1	0	1	0	0	1
5.50	4.22	1	0	1	0	0	1
5.35	4.83	0	1	1	0	0	1
4.33	3.55	0	1	1	0	0	1
5.21	3.99	1	0	0	1	0	1
4.95	4.40	1	0	0	1	0	1
4.93	4.43	1	0	0	1	0	1
6.88	4.35	0	1	0	1	0	1
5.26	4.80	0	1	0	1	0	1

Suppose, for example, we wish to test for interaction between the $\alpha's$ and $\gamma's$, that is, between type of blood and treatment in excited state. We have,

	Degrees of Freedom	Sums of Squares
Interaction $(\alpha \times \gamma)$	1	$A-B = \quad .653$
Error	36	$R(\mu, \overline{\alpha\gamma_{rt}}) = B = 20.482$
Total	37	$R(\mu, \alpha_r, \gamma_t) = A = 21.135$

We find $z = .069$, which is not significant. Therefore, the interaction $(\alpha \times \gamma)$ may be regarded as non-existent. As a matter of fact, no significant variation of viscosity in excited state can be ascribed to the $\alpha's$ and $\gamma's$, as the following table shows, in which z has the non-significant value .109,

	Degrees of Freedom	Sums of Squares
Between Classes	3	$A-B = \quad 1.373$
Within classes	36	$R(\mu . \overline{\alpha\gamma_{rt}}) = B = 20.482$
Total	39	$R(\mu) \quad = A = 21.855$

We may now test for main effects due to the $\gamma's$, that is, treatments. For this analisis, we have,

	Degrees of Freedom	Sums of Squares
Main effect (γ)	2	$A-B = \quad .287$
Error	36	$R(\mu, \alpha_r, \gamma_t) = B = 21.135$
Total	38	$R(\mu, \alpha_r) \quad = A = 21.422$

$z = .678$, which is not significant. Thus, the evidence will not support a conclusion of different main effects on blood viscosity in excited state due to treatments.

Now, suppose we make use of the information contained in y and, as far as possible, eliminate the variation due to y. We assume the regression of x on y to be the same for all sub-classes and the possible effects due to the $\alpha's$, $\beta's$ and $\gamma's$ to be additive. The constituents for the analysis of variance are,

	Degrees of Freedom	Sums of Squares
Treatment (γ)	1	$A-B = \quad .050$
Error	35	$R(\rho; \mu, \alpha_r, \beta_s, \gamma_t) = B = 11\ 223$
Total	36	$R(\rho; \mu, \alpha_r, \beta_s) \quad = A = 11.273$

The main effects due to treatments after eliminating effects due to y, the α's and β's, are clearly non-significant. In fact $z = .931$.

In a similar way various other hypotheses can be tested by using the R's, expressible as ratios of determinants.

SUMMARY

By making use of variates which can be either 0 or 1, the problem of analyzing variance and covariance in multiply-classified non-orthogonal data is reduced to a procedure of evaluating ratios of determinants.

The application of the method to the problem of testing certain hypotheses is illustrated by a numerical example.

REFERENCES

1. R. A. FISHER and J. WISHART, *The arrangement of field experiments and the statistical reduction of the results,* Imperial Bureau of Soil Science, Technical Communication (1930), No. 10.
2. F. YATES, *The Principles of orthogonality and confounding in replicated experiments,* « The Journal of Agricultural Science », Vol. 23, 1933. pp. 108-145.
3. F. YATES, *The Analysis of multiple classifications with unequal numbers in the different classes,* « Journal of the American Statistical Association », vol. 29, 1934, pp. 51-66.
4. F. YATES, *Complex Experiments,* Supplement, « Journal of the Royal Statistical Society », Vol. 2, No. 2, 1935, pp. 182-247.
5. L. B. NICE and D. FISHMAN, *Changes in the viscosity of the blood in normal, splenectomized and adrenalectomized animals following emotional excitement.* « American Journal of Physiology », vol. 107 (1934), pp. 113-119.

ARTHUR OLLIVIER

On the analysis of variation in general death rates

Comparison of the general death rates of two or more populations is usually made by studying the crude rates or by means of corrected death rates set up by the use of some form of standard population. Yule (1) has listed three methods of making such corrections and has discussed the advantages of each procedure.

The problem cannot be regarded as completely solved as soon as such corrections and comparisons have been made, for small variations in general death rates do not necessarily mean different general mortality conditions nor does the agreement of general death rates imply that the populations are subject to exactly the same mean probabilities of death. For example, five Iowa counties whose general death rates for 1925 were exactly the same, 9.2 per thousand, had death rates in 1926 which varied from 8.3 to 10.3 per thousand. Since neither the inhabitants nor the living conditions in these Iowa counties changed materially from 1925 to 1926, it does not seem reasonable to suppose them subject to the same probability of death in one year and to different probabilities in the following year. Many other similar illustrations can be given. What seems to be needed further is some criterion by which to judge the statistical significance of such variations in general death rates.

Pearson and Tocher (2) have studied several such criteria when

---

1) YULE, G. U. *On some points relating to vital statistics.* J. R. S. S., v. 97 (1934) Part I, pp. 1-84 — s. also C. GINI, *Quelques considérations au sujet de la construction des nombres indices des prix et des questions analogues.* « Metron », Vol. IV, n. 1, 1924.

2) PEARSON, KARL and TOCHER, J. F. *On criteria for the existence of differential death rates.* Biom. v. 11, (1915), pp. 159-184.

two populations only were being considered. Unfortunately, neither of the criteria considered by them to be best suited to making such tests seems capable of being used to compare more than two populations at one time, and while a complete analysis may finally demand the comparison of death rates two at a time, the beginnings of such an analysis may well be made by comparing a much larger number.

The problem to be considered in this paper may be stated as follows: Given the general death rates of several populations, is it possible to set up a criterion to test whether or not they are alike with respect to general mortality, where we mean by this last phrase, whether or not the variations in general death rates which occur could reasonably be accounted for on the hypothesis of chance fluctuations from the same basic mean probability of death? Furthermore, when a set of populations is encountered which appear to be unlike with respect to general mortality, is it possible to make such an analysis of the characteristics of the several populations as will indicate the probable reasons for the unlikeness?

The paper is divided into three sections. The first is devoted to the development of several theorems relating to the distribution of the number of deaths that would occur if the several populations under consideration were samples drawn from a composite universe by either the method of unrestricted or the method of stratified random sampling. These theorems are regarded as essential to the later development of test criteria to be applied to the problems stated above. The second section considers the use of a modified form of the squared Lexis ratio as a criterion for testing the statistical significance of variations in general death rates, while the third section consists of an application of the theory of the paper.

Section 1. In applying the criteria to be developed in the next section it will frequently be advantageous to consider what would have happened if the several populations whose death rates are being studied had been random samples drawn from the same infinite population or universe. Since, however, the human populations under consideration are composed of various groups, such as age groups, which differ widely from one another with respect to the relative frequency with which death occurs within a given time interval, it seems appropriate to suppose that the universe from which they are supposedly drawn is similarly « stratified » (1).

In drawing a sample of s individuals at random from such a stratified universe, either one of two procedures may be followed. First, individuals may be drawn at random from the universe as a whole, leaving entirely to chance the stratum from which any individual selected may be drawn. Second, a fixed number of individuals, s_i, may be drawn from the *i-th* stratum, where i ranges over all strata in the universe, and so that the sum of all s_i shall equal the sample number, s. The first method will be called the method of unrestricted random sampling and the second the method of stratified random sampling. It seems likely that neither the mean (expected) number of deaths, nor the variance in the number of deaths of repeated samples will be the same for the two methods of sampling.

Now let p_i denote the probability that death will occur to an individual in the *i-th* stratum of the universe within a given time interval. Consider the following question : What is the expected number of deaths, variance in the number of deaths, and distribution of the number of deaths, for repeated samples of s drawn by each of the two methods mentioned above ? For the method of stratified random sampling the answer is contained in Theorem 1 and its corollaries, given below.

Theorem 1. The probability that specified numbers of deaths will occur within a given time interval in an exposure of s lives, when s_i of them are drawn from the *i-th* stratum, $i=1,2,\ldots$ $\ldots, n$, and $\Sigma s_i = s$, is given by the terms of the expansion of

$$(1) \qquad \prod_{i=1}^{i=n} (p_i + q_i)^{s_i}, \quad \text{where} \quad q_i = 1 - p_i .$$

Proof : Consider a single term of the expansion of (1). It can be written as

$$(2) \qquad P_1 \cdot P_2 \cdot P_3 \ldots P_n, \quad \text{where} \quad P_i \quad \text{has the form}$$

$$(3) \qquad P_i = \frac{(s_i)\,!}{(r_i)\,!\,(s_i - r_i)\,!}\, p_i^{r_i}\, q_i^{s_i - r_i}$$

---

(1) This terminology seems to have been first used by A. L. Bowley. Cf. Neyman, J. *On two different aspects of the representative method.* J. R. S. S., v. 97, (1934), Part. IV, pp. 558 ff.

But (3) is the probability that among an exposure of s_i lives, all from the *i-th* stratum, there will occur exactly r_i deaths within the specified time interval. Since the values of P_i and P_j, $(i \neq j)$, are independent in the probability sense, the product, (2), gives the probability that in an exposure of s lives, drawn as specified, there will occur exactly Σr_i deaths within the given time interval, exactly r_i of the deaths occurring among the s_i lives exposed from the *i-th* stratum. But the complete expansion of (1) is composed of all terms similar to (2) which may be obtained by varying r_i from zero to s_i, and may therefore be considered as the probability function.

Corollary 1. The expected number of deaths among a total of s exposed lives drawn as in Theorem 1 is $E(D) = \Sigma s_i p_i$.

Corollary 2. The variance in the number of deaths for repeated samples of s exposed lives drawn as in Theorem 1 is $V = \Sigma s_i p_i q_i$. Variance is here understood to mean the expected value of the square of the deviation from the expected value.

Corollaries 1 and 2 follow directly from Theorem 1, for if the number of deaths, in the case of Corollary 1, or the square of the deviation from the expected number of deaths in the case of Corollary 2, be multiplied by the corresponding probability, given by the proper term of (1), and the whole summed for all possible samples which can arise from (1), the results will be those stated above.

It may now be noticed that stratified random sampling, as here defined, yields terms which may be described as a generalized Poisson series, for if each $s_i = 1$ we have in Theorem 1 and its corollaries exactly the results of the Poisson series.

For the case of unrestricted random sampling we have:

Theorem 2. The probability that specified numbers of deaths will occur within a given time interval among a total of s exposed lives drawn unrestrictedly at random from a stratified universe is given by the terms of the expansion of $(P + Q)^s$, where $P = \Sigma P_i p_i$, P_i designates the probability of drawing an individual from the *i-th* stratum in a single trial, and $Q = 1 - P$.

Proof: The proof of Theorem II depends upon the following lemma which is assumed.

Lemma. The probability of securing exactly s_i individuals from the *i-th* stratum $(i = 1, 2, \ldots, n)$ in a total of s drawings from a stratified universe is given by the appropriate term of the expansion of

$$(4) \qquad (P_1 + P_2 + \ldots + P_n)^s,$$

where P_i has the meaning given in the statement of Theorem 2.

Taking up the proof of Theorem 2, we note that (4) can be written in the form

$$(5) \qquad \Sigma_s \frac{s\,!}{(s_1)\,!\,(s_2)\,!\ldots(s_n)\,!} P_1^{s_1} P_2^{s_2} \ldots P_n^{s_n}$$

where the symbol Σ_s designates summation over all possible partitions of s into numbers s_i, which are positive integers or zero.

Now by Theorem 1 the probability that certain numbers of deaths will occur in repeated samples with assigned numbers, s_i, from the i - *th* stratum is given by the appropriate terms of (1). Moreover, the probability of securing a sample with such assigned numbers when the drawing is unrestrictedly at random is given by the appropriate term of (5). Since the probabilities given by these two formulas are independent of one another, the product of a term of (1) by a term of (5) gives the probability that there will occur the compound event consisting of first drawing a sample of s with the specified numbers, s_i, of individuals from the i - *th* stratum and second, of finding a specified number, r_i, of deaths occurring within the given time interval among the s_i individuals in the i - *th* stratum of the sample. Since the expression

$$(6) \quad \Sigma_s \frac{(s)\,!}{(s_1)!\ldots(s_n)!} P_1^{s_1} \ldots P_n^{s_n} \Sigma_{r_1} \frac{(s_1)\,!}{(r_1)\,!\,(s_1-r_1)\,!} p_1^{r_1} q_1^{s_1-r_1} \Sigma_{r_2} \ldots \ldots \Sigma_{r_n} \frac{(s_n)\,!}{(r_n)\,!\,(s_n-r_n)\,!} p_n^{r_n} q_n^{s_n-r_n}$$

consists of all possible terms which can arise from products of terms of (1) and (5), it may be regarded as the probability function for the distribution of the number of deaths under the conditions specified in the theorem.

To simplify (6), we note that it can be written in the form

$$(7) \quad \Sigma_s \Sigma_{r_1} \ldots \Sigma_{r_n} \frac{(s)!}{(s_1)!\ldots(s_n)!} P_1^{s_1} \ldots P_n^{s_n} \frac{(s_1)!}{(r_1)\,!\,(s_1-r_1)\,!} p_1^{r_1} q_1^{s_1-r_1} \ldots \ldots \frac{(s_n)\,!}{r_n!\,(s_n-r_n)\,!} p_n^{r_n} q_n^{s_n-r_n}$$

by the process of writing the product of the several summations in (6) in expanded form. Now in (7) we may divide the (s_i) ! which appears in a numerator by that which appears in a denominator and make other fairly obvious changes that enable us to write the result as

$$(8) \quad \Sigma^s \Sigma_{r_1} \ldots \Sigma_{r_n} \frac{(s)!}{(r_1)! \ldots (r_n)! (s_1 - r_1)! \ldots (s_n - r_n)!} (P_1 p_1)^{r_1} \ldots (P_n p_n)^{r_n}$$
$$(P_1 q_1)^{s_1 - r_1} \ldots (P_n q_n)^{s_n - r_n},$$

the number of terms in (8) remaining the same as in (7).

The expression in (8) to the right of the summation signs is one of the terms of the expansion of the multinomial

$$(9) \qquad (P_1 p_1 + P_2 p_2 + \ldots + P_n p_n + P_1 q_1 + \ldots + P_n q_n)^s$$

and since the summations in (8) extend over all integral values of r_i from zero to s_i and the first summation is over all partitions of s, it follows that (8) contains all the terms of the complete expansion of (9).

If now we put $\Sigma P_i p_i = P$, it follows that $\Sigma P_i q_i = 1 - - P = Q$, and we can write (9) in the simple form

$$(10) \qquad (P + Q)^s$$

which demonstrates the theorem.

Corollary 1. The expected number of deaths in an unrestricted random sample of s drawn from a stratified population is $E(D)_u = s \, \Sigma P_i p_i$.

Corollary 2. The variance of repeated samples of s drawn from a stratified population by the method of unrestricted random sampling is given by $V = s P (1 - P)$, where P is defined above and by the variance we mean the same as in Corollary 2, Theorem 1.

Both of these corollaries follow directly from Theorem 2. They may, however, be proved independently of that theorem by a direct application of the kind of argument already used.

It may now be stated that in drawing unrestricted random samples of individuals from a stratified universe the number of deaths occurring in the samples will tend to be distributed according to a simple Bernoulli series if we disregard the strata from which the deaths occur. It should be noted that P_i may or may not be proportional to the number of individuals in the i-th stratum of the universe and that the theorems proved above will hold in either case. It may also be pointed out that while the language in which the theorems above are stated implies that we are chiefly interested in the distribution of numbers of deaths, the results given by these theorems apply equally well to other sampling problems which satisfy the conditions imposed.

Section 2. The statistical measure which seems best suited to the consideration of the significance of the variation of several death rates simultaneously is the Lexis ratio.

The Lexis ratio is usually defined as follows (1) : Given n sets of observations with s individuals in each set, let x_i be the relative frequency of success in the i-th set. On the hypothesis that the n sets constitute a Bernoulli distribution, the standard deviation of relative frequencies is $(p\,q/s)^{1/2}$ where p is the probability of success in a single trial. The Lexis ratio is then defined to be

$$(11) \qquad L = \frac{[1/n\,\Sigma\,(x_i - p)^2]^{1/2}}{(p\,q/s)^{1/2}} = \frac{s\,[1/n\,\Sigma\,(x_i - p)^2]^{1/2}}{(s\,p\,q)^{1/2}}$$

Two modifications of (11) are usually made in practice. First, p is generally not known and is taken as the arithmetic mean, weighted or unweighted, of the relative frequencies recorded in the n sets. Secondly, in the analysis of social or economic data, it is seldom that the sets of observations contain the same number of individuals. This variation is usually neglected, the expression spq being computed from a value of s which is the arithmetic mean of the varying number of individuals in the sets. A slightly modified definition of L makes it unnecessary to neglect this variation and at the same time will give a result more suited to the consideration of our problem.

Suppose then we have n sets of observations drawn in some manner from a stratified universe. Let s_{ij} represent the number of individuals in the i-th set that belong to the j-th stratum. Let d_i be the observed number of successes which occur in the i-th set. Then if each set be supposed to have arisen by unrestricted random sampling from the same stratified universe in such a manner that the mean probability of drawing a success

(1) RIETZ, H. L., *Mathematical Statistics* (*Carus Monograph* 3). See page 152.

in any one trial is P, we have from the corollaries to Theorem 2 that $s_i P$ and $s_i P Q$ are the expected number of successes and the variance in the number of successes respectively for repeated samples of s_i. Take

$$(12) \qquad l_i = \frac{d_i - s_i P}{(s_i P Q)^{1/2}}, \quad \text{where} \quad s_i = \Sigma_j s_{ij}$$

and let

$$(13) \qquad L^2_1 = 1/n \, \Sigma_i L^2_i .$$

It may be noted that L^2_1 is the analogue of the square of the Lexis ratio defined above.

On the other hand, if each of the n sets be supposed to have arisen by stratified random sampling, we have by the corollaries to Theorem 1 that $\Sigma_j s_{ij} p_j$ and $\Sigma_j s_{ij} p_j q_j$ are the expected number of successes and the variance in the number of successes respectively for repeated samples of the i - th set. Now take

$$(14) \qquad l_i = \frac{d - \Sigma_j s_{ij} p_j}{(\Sigma_j s_{ij} p_j q_j)^{1/2}},$$

and

$$(15) \qquad L^2_2 = 1/n \, \Sigma_i l^2_i .$$

Again we may state that L^2_2 is the analogue of the squared Lexis ratio.

It seems fairly obvious that if the n sets did actually arise as unrestricted random samples, n a large number, then L^2_1 will not differ greatly from unity, and if they arose as stratified random samples then L^2_2 will be approximately one. What we need further to make these measures available as tests of the respective hypotheses is the variance and distribution of L^2_1 and L^2_2 for repeated samples of n sets. While we do not know the exact distribution of either of these two measures the following considerations yield a useful approximation.

It has been shown in Theorems 1 and 2 that for repeated samples of s_i, the number of successes is distributed according to $(P + Q)^{s_i}$ for the case of unrestricted random sampling and according to $\Pi_j (p_i + q_j)^{s_{ij}}$ for the case of stratified random sampling. It seems reasonable to assume that the distribution of relative frequencies in such composite populations as are represented by the terms of the expansions of these two expressions are nearly enough normal, as indicated by the classical theorem of Poisson (1), to justify the use of the Helmert formula (2) for the distribution of variance of the deviations from the population mean in the numerical applications which follow. Making this assumption the problem of finding the distribution of L^2_1 and L^2_2 becomes merged into the single problem of finding the distribution of

$$(16) \qquad w = \frac{1/n \, \Sigma \, (x_i - M)^2}{\sigma^2}$$

in samples of n, where the variate x is subject to the normal law

$$(17) \qquad f\,(x) = \frac{1}{\sigma \sqrt{2\,\pi}}\, e^{-\frac{(x-M)^2}{2\,\sigma^2}}$$

The distribution of w, or of functions similar to w, is well known and has been found by several writers. As has already been noted, Helmert is probably entitled to credit for being the first writer to find the distribution of this function. A detailed account of his work is given by Czuber (3). Student (4), Fisher (5), Pearson (6), and Romanovsky (7) have also studied the distribution of w and similar functions. Because of this well known character of the distribution no proof will be presented here. However, we note that

$$(18) \qquad F\,(w)\,d\,w = (n/2)^{n/2} \frac{1}{\Gamma\,(n/2)}\, w^{(n-2)/2}\, e^{-n\,w/2}\, d\,w \;,$$

gives to within infinitesimals of higher order the probability that w will lie between w and $w + d\,w$.

At this point it should be noted that in computing the Lexis ratio from observed data we do not usually know the probability

---

(1) cf. Czuber, E. *Wahrscheinlichkeitsrechnung, Erster Teil,* (1914), pp. 120-140 and 173-180.

(2) Helmert, Astron. Nach., Band 88, No. 2096-97; also ibid. Band 85, No. 2039. cf. Czuber, E., Beobachtungsfehler, pp. 147 ff.

(3) Czuber, E. *Beobachtungsfehler,* pp. 147 ff.

(4) Student, *The probable error of the mean,* Biom. v. 6, (1908-09), pp. 1-25.

(5) Fisher, R. A. *Frequency distribution of the correlation coefficient.* Biom., v. 10, (1915), pp. 507-521.

(6) Pearson, Karl, Biom., v. 10, (1915), pp. 522-529.

(7) Romanovsky, V., *On the moments of standard deviations and correlation coefficients.* « Metron », v. 5 (1926), No. IV, pp. 3-45.

of success in a single trial. We therefore generally accept the relative frequency of success in the total number of trials in the n sets of observations as an approximation to the probability, p, and are therefore dealing not with true errors, but with apparent errors, which are sometimes called « residuals » in English, or « erreurs apparentes » in French and « scheinbaren Fehler » in German. Helmert (1) has likewise given the distribution function of a sum of squares of such residuals where the original deviations follow the normal law. If we define

$$(19) \qquad w_1 = \frac{1/(n-1)\,\Sigma\,(x_i - \bar{x})^2}{\sigma^2}$$

where x is the sample mean, then w_i can be shown to be distributed exactly as is w, with the sole exception that in (18) n must be replaced by n-1.

It should be further noted that in (19) we have used the population variance for the denominator while in practice this must also be estimated from the data at hand. Strictly speaking, therefore, the distribution in which we are interested will not follow that of w or w_1, but will be more like that of Student's z or Fisher's t function. However, in the applications we wish to make of the distribution function of L^2, the number of items upon which the relative frequency of deaths, p', is based is so large that we may assume that $sp'\,q'$ relates to an infinite supply without making any material error. Then the problem of the distribution of L^2 is that of the distribution of the mean value of independent squares of residuals expressed in units of the population variance. Hence, we write $L^2{}_1 = w$ and also $L^2{}_2 = w$ and accept (18) as giving a suitable approximation to the distribution of both these measures with the condition that n be replaced by n-1 when an estimated probability is used instead of a true probability (2).

If now the transformation

$$(20) \qquad t = \sqrt{2/n}\; n\,(w - 1)/2\;,\; \alpha/2 = \sqrt{2/n}$$

be made in (18), the result will be

(1) cf. CZUBER, E., loc. cit.

(2) R. A. Fisher has stated that the squared Lexis ratio is distributed as in Chi-square. *Sec Statistical Methods for Research Workers.* Fourth Edition, p. 82. It is possible that he reached that conclusion by reasoning similar to what has just been given.

$$(21) \qquad \frac{\left(\frac{4}{\alpha^2}\right)^{\frac{4}{\alpha^2}-1/2}}{\Gamma\left(\frac{4}{\alpha^2}\right)e^{\frac{4}{\alpha^2}}}\left(1+\frac{\alpha}{2}t\right)^{\frac{4}{\alpha^2}-1}e^{-\frac{2}{\alpha}t}\,dt\,.$$

which is exactly the form of the Pearson Type III curve tabulated by L. R. Salvosa (1). By making use of this simple trasformation it is possible to employ the tables given by him to test the significance of the deviation of w, and hence of L^2_1 and L^2_2 away from its expected value of unity.

In all that has been done thus far in this section it has been assumed that composite populations such as those with which we are dealing are approximately normal. In order to throw some light on the size of the samples involved in our thinking, the value required to make almost sure a prescribed degree of approach of a normal law to a simple binomial population may be approximated as follows, if we accept the deviations of β_1, from zero and of β_2 from three as measuring the departure of the binomial from the normal. For the binomial distribution we have that

$$(22) \qquad \beta_1 = \frac{1-4pq}{spq}\,,\ \beta_2 = 3 + \frac{1}{spq} - \frac{6}{s}$$

Suppose we wish to have $\beta_1 \leqq k_1$, $|\beta_2 - 3| \leqq k_2$ where k_1 and k_2 are previously chosen positive constants. Then we must satisfy the two conditions

$$(23) \qquad k_1 \geqq \frac{1-4pq}{spq}\,,\ k_2 \geqq \left|\frac{1-6pq}{spq}\right|.$$

From these inequalities, we obtain

$$(24) \qquad s \geqq \frac{1-4pq}{k_1 pq}\,,\ s \geqq \frac{|1-6pq|}{k_2 pq}$$

respectively, and if we choose the greater of these two values of s obviously both conditions will be satisfied (2).

Before going on to section three, an example of the use of w will be given. The following data, Table I, represent a coin

(1) SALVOSA, L. R. *Tables of Pearson's Type III Function, Annals of Math. Stat.*, v. 1 (1930), pp. 191-98 and appendix.

(2) As an illustration of the above applied to death rates, suppose $p = .01$, $q = .99$ and that we wish to have $k_1 = k_2 = .01$. Solving for s gives us $s \geqq 9700$ and $s \geqq 9940$. Clearly if $s \geqq 10000$, the required degree of approximation will be attained. This is about the size of the populations used in the third section.

tossing experiment in which seven coins were tossed a total of 128 times. The probability of obtaining a head in any one toss of a single coin is assumed to be one-half.

TABLE I

Number of heads	0	1	2	3	4	5	6	7	L^2	X^2
Theoretical Frequency	1	7	21	35	35	21	7	1		
Observed, Sample 1	0	6	17	36	42	18	9	0	.857	3.244
Observed, Sample 2	0	10	25	34	28	24	7	0	1.036	3.244
Observed, Sample 3	0	10	25	30	30	24	9	0	1.271	3.244

Sample 1 gives the frequency of an actual tossing experiment, samples 2 and 3 are arbitrarily chosen frequencies which might have arisen from such an experiment and are used to illustrate a point to be made a little later. Now using $p = q = \frac{1}{2}$, we obtain $w = L^2 = .857$ for sample 1. This is then a case of subnormal dispersion although we know that it should actually represent a sample from a Bernoulli population. Transforming from w to t by means of formulas (20) we find $t = -1.14$, $\alpha/2 = .125$. Entering Salvosa's Tables with these values, we find the probability of securing a value of t less than or equal to that actually found to be $P = .128$. That is, about once in eight times we would expect to find, as a random sampling fluctuation, subnormal dispersion as great or greater than that found in this sample.

That the Chi-square test does not give exactly the same information about a sample as does the squared Lexis ratio w, is shown by considering the data for all three samples in Table I. If we follow the practice of grouping the two end classes at each end of these samples into a single class, we will get the same value of Chi-square for each of these samples, namely, $X^2 = 3.244$. Hence, judged by this test, each of the three samples is an equally good approximation to the theoretical distribution. However, the squared Lexis ratio differs widely in the three cases, for it is .857 in sample 1, 1.036 in sample 2 and 1.271 in sample 3. Sample 2 has almost normal dispersion while sample 3 has decidedly super-normal dispersion. The probability of finding a value of L^2 equal to or greater than that actually found is $P = .379$ for sample 2 and $P = .018$ for sample 3. Judged by this test, it appears that sample 2 is much more likely to appear as a random sampling fluctuation that either samples 1 or 3.

It should be noted here that the two tests, X^2 and w, are not antagonistic nor do they contradict one another. They do not give the same results when applied to a sample because they measure different characteristics of the sample. One, X^2, measures the departure of the sample frequency away from the theoretical frequency. The other measures the departure from the expected value of the variance or dispersion of the sample. It ought to be added that the X^2 test can probably be used for a wider variety of purposes than the Lexis ratio. The latter, however, seems better suited to our present purpose.

Section 3. We will consider now the application of the above theory to the analysis of the general death rates of the rural population of the state of Iowa, United States of America. Data (1) for the years 1923 to 1929 inclusive are used. This selection of data was made because it was believed from a priori considerations that the rural parts of the population of Iowa would show many characteristics of homogeneity which would be absent in other data, especially if data for both rural and urban districts were included. There is a total of ninety-nine counties in Iowa with an average rural population of about 18000, the inhabitants of all cities of ten thousand or over being not included. Data from but ninety-five counties were used; four counties were excluded because they contain state institutions such as hospitals for the insane, the deaths of the inmates of which are not separately reported. Consequently, the general death rates of these counties were excluded because we did not know how to make a correction for this factor.

Before applying the methods of the preceding section the remaining ninety-five counties were grouped into nine geographical districts; each district consists of ten or eleven counties and represents approximately one ninth of the state's area and population. These districts were numbered consecutively from south-east to north-west; Districts I, II and III form the eastern third of the state from south to north; Districts IV, V and VI run from south to north in the central third of the state; while Districts VII, VIII and IX in like manner form the western third of the state.

---

(1) These data cannot be reproduced here. They may be found in the various publications of the Bureau of the U. S. Census, Washington, D. C.. See especially the annual « Mortality Statistics », the « Birth, Still-Birth and Infant Mortality Statistics » and the « Population Reports for the 15-*th* General Census » (1930).

In Table II we have the data for these nine districts arranged to test the hypothesis that the general mortality rates of these districts differ from each other by no greater amount than can reasonably be accounted for by random sampling fluctuations. The « probability » of death upon which the theoretical number of deaths is computed for each district is the mean general mortality rate for all nine districts. Population estimates are taken from the Census of 1930 and the observed number of deaths is for 1929.

The final column of Table II is computed from formula (12).

TABLE II

District	Population	Observed Number of Deaths	Theoretical Number of Deaths	Difference	Difference Squared	spq	l_i^2
I	139341	1542	1320	222	49284	1308	37.68
II	211751	2164	2007	157	24649	1988	12.40
III	197039	1925	1867	58	3364	1850	1.82
IV	189561	1981	1796	185	34225	1779	19.23
V	246703	2145	2338	–193	37249	2316	16.09
VI	180739	1569	1713	–144	20736	1697	12.22
VIII	169361	1715	1605	110	12100	1590	7.61
VII	209683	1839	1987	–148	21904	1968	11.13
IX	160387	1273	1520	–247	61009	1505	40.52
TOTALS . . .	1704565	16153	16153	0			158.70

Its total, 158.70, divided by eight, gives the value of $L^2{}_1$ to be 19.84. Transforming this by means of formulas (20), we get the necessary entries to Salvosa's Tables to be $t = 37.68$, $\alpha/2 = .5$. This value of t is beyond the range of the tables, hence we conclude that it is highly unlikely that the general death rates of these nine districts differ by no more than random sampling fluctuations (1).

(1) The same conclusion is reached if we consider each of the ninety five counties alone as a district instead of grouping them as above. The data cannot be given here but the results are: $\alpha/2 = .146$, $t = 19.14$, which value of t is also beyond the range of the tables.

An interesting observation about the geographical distribution of death rates over the state may be made from Table II. It can be seen there that the districts having negative differences between actual and theoretical death rates are in the north and north-western part of the state, while those whose observed number of deaths exceed the theoretical number are in the south and south-east. This fact suggests that there may be a systematic variation in general death rates across the state from south-east to north-west, due perhaps to some characteristic of the population which likewise varies in this manner.

As a partial check on the theory suggested in the sentence above, the correlation coefficients between the general death rates for successive years and like-wise between the birth and death rates for the same year were computed and are given in Table III. These coefficients are too small to be of any particular value for prediction purposes but they are large enough when compared to their probable errors to be considered significant. The negative character of the correlation between birth and death rates suggests that age distribution may be the characteristic differentiating these populations since a high death rate and a low birth rate usually accompanys an aged population with the opposite true for a young population.

TABLE III

Correlation coefficients for the rural parts of 95 Iowa counties

Between death rates for successive years.		Between birth and death rates for the same year.	
Years	r	Year	r
1923-24	.63 ± .04	1924	−.26 ± .07
1924-25	.58 ± .05	1925	−.33 ± .06
1925-26	.41 ± .06	1926	−.38 ± .06
1926-27	.64 ± .04	1927	−.25 ± .07
1927-28	.50 ± .05	1928	−.15 ± .07
1928-29	.52 ± .05	1929	−.48 ± .05

The data of Table IV are arranged to test the hypothesis that the general mortality of the nine districts considered before differs significantly because the districts differ significantly with respect to the age distribution of their inhabitants. To test this hypothesis, we recognize that the population of each district is composed of individuals who fall into different age strata and that the proportional numbers of persons in each age stratum varies from district to district. After correcting the number of deaths expected to occur in each district for this age stratification, we inquire whether or not the remaining variation in general death rates between districts can reasonably be accounted for by the hypothesis of chance fluctuation.

To make the correction mentioned above, a specific probability of death, p_j, is computed from the data at hand for each age stratum into which the populations are divided in such a manner that the number of deaths expected in the state is the same as the observed number. The number of deaths expected in the *i-th* district is then $E\ (D_i) = \Sigma_j\, s_{ij}\, p_j$, where s_{ij} is the number of persons in the *j-th* stratum of the *i-th* district. In computing p_j for Table IV age strata data for the several counties were taken from the 15-*th* General Census (1930), while the deaths within each stratum were for 1929 (1).

TABLE IV

District	Observed Number of Deaths	Expected Number of Deaths	Difference	Difference squared	$s_{ij}\,p_j\,q_j$	t_i^2
I	1542	1500	42	1764	1424	1.24
II	2164	2061	103	10609	1786	5.94
III	1925	1944	-19	361	1851	.20
IV	1981	1961	20	400	1794	.22
V	2145	2250	-105	11025	2062	5.35
VI	1569	1606	-37	1369	1476	.93
VII	1715	1638	77	5929	1769	3.30
VIII	1839	1837	2	4	1689	.00
IX	1273	1356	-83	6889	1294	5.32
TOTALS . . .	16153	16153				22.50

(1) It should be mentioned here that the age strata which we were compelled to use, since no other was available, are not very favorable to our analysis especially since all persons over seventy-five years of age were grouped into one stratum. No correction could be made, therefore, for variations among the districts within this age stratum, although it was knows that such variation existed and did materially affect the number of deaths which occurred.

The last colum of Table IV was computed from formula (14). The total of this column, 22.50, divided by eight, gives us the value of $w = L^2_2 = 2.81$. Transforming this by (20), we have $\alpha/2 = 1/2$, $t = 3.62$. Referring to Salvosa's Tables with these entries we find the probability of securing a value of t equal to or greater than this by chance to be $P = .002$. The odds are strongly against the hypothesis that all variation except that due to chance fluctuation, is due to the age stratification of the population of the several districts. However, the decrease in these odds from those of Table II is considerable and we feel justified in concluding that the variation is due at least in part to age distribution (19).

Finally, we note that, in the analysis of data like those we have been considering, if there are k distinct quantitative factors, $F_1, \ldots . F_k$, (such as age strata) each of which has an effect upon the relative frequency of death, then it is possible to think of the universe from which the separate populations are presumably drawn as being stratified on the basis of each such factor. The universe would thus be distributed into cells, so that each individual in a given cell could be said to fall, on the basis of factor F_i, within the limits F_i to $F_i + d F_i$. Now if it be possible to assign to each such cell a definite relative frequency (or « probability ») of death, then it will be possible to consider the effect of stratification on the basis of one factor alone or on the basis of any combination of such factors by adding together the frequencies in the cells and by determining the mean probabilities of death for such sums. The analysis necessary for the consideration of stratification on the basis of several factors does not differ from that already considered for one factor, since the only change that needs to be made is to provide for a larger number of strata. Theoretically, therefore, it will be possible to consider the effect of stratification successively on the basis of more and more factors, until all variation except that which may reasonably be due to chance fluctuation has been eliminated. Thus it would be possible to state that a group of populations differs with respect to general death rates of its members because they differ by more than can reasonably be ascribed to chance with respect to certain of the

19) Application of the above methods to the data when the counties are not grouped into districts gives $\alpha/2 = .146$, $t = 5.48$. This value of t should be compared with that of note (17).

factors, F_1. Practically, it will nearly always be impossible to carry out such an analysis, first, because adequate data for considering the effect of stratification on any other basis than age usually does not exist, and second, because it is probably true that many of the factors affecting the relative frequency of death are not quantitatively measurable.

The University of Iowa, Iowa City, Iowa

NECROLOGIO

Harald Ludvig Westergaard 1853-1937.

La nostra rivista ha perduto quest'anno in H. L. WESTERGAARD il decano del suo comitato direttivo, uno dei suoi più venerati collaboratori.

Alla famiglia del «Metron» egli apparteneva sin dal 1923 quando il suo nome è cominciato a figurare appunto fra i membri del comitato direttivo, dal terzo volume. In «Metron» collaborò con il lucido studio *On Periods in Economic Life* (Vol. V, I-1925), nel quale tenta di ricondurre ad una estrema semplicità di metodo le ricerche intorno ai movimenti economici che particolarmente gli americani Moore e Persons pareva avessero inutilmente ingombrato con complicate elaborazioni matematiche.

Perchè il WESTERGAARD in tutta la sua attività statistica, pur rivelando un pieno possesso della matematica, mostra, forse perciò appunto, di preferire i più semplici procedimenti aritmetici quando già questi riescono a mettere in evidenza i caratteri essenziali del fenomeno studiato; ed anche nel dominio della teoria pura non trascura mai l'applicazione e l'esempio.

Tre sono le opere del WESTERGAARD che ancora oggi formano oggetto di consultazione:

Die Lehere von der Mortalität und Morbidität, pubblicata a Jena nel 1882 col motto «Observationes et numerandae et perpendendae» (una seconda edizione fu fatta nel 1901).

Die Grundzüge der Theorie der Statistik, pubblicato a Jena nel 1890 (una seconda edizione tedesca ebbe luogo nel 1928, oltre a tre edizioni in danese);

Contributions to the history of Statistics, pubblicato a Londra nel 1932.

In una prima parte dei suoi fondamenti della statistica il W. pone a base della metodologia lo schema del calcolo delle probabilità; in una seconda tratta delle applicazioni del metodo allo studio della popolazione, alla statistica economica, alla tecnica delle assicurazioni; nella terza infine traccia le prime linee di una storia della statistica, che poi si svilupparono nel succitato volume Contributions to the history of Statistics.

Nel trattato sulla mortalità e la morbilità il WESTERGAARD fa ancora precedere alle sue indagini antropologico-statistiche un'esposizione teorica probabilistica con particolare riguardo all'errore medio, del quale fa anche una larga applicazione nel campo demografico.

Ma il libro che si legge con interesse, ricco di notizie e di nomi, chiaro nell'esposizione, è la storia della statistica. In essa il W. parte dalle prime opere di Francesco Sansovino e di Giovanni Botero del secolo XVI per arrivare alle ultime decadi del secolo XIX, quando cominciarono ad affermarsi quei grandi progressi tecnici che facilitarono, più che pel passato, la raccolta accurata di vasti materiali statistici ed aprirono la via all'analisi quantitativa dei fenomeni sociali tanto in onore ai nostri giorni.

Forse l'opera del WESTERGAARD ci appare un po' troppo fredda, dove raramente si ritrova il calore del contributo personale; forse lo si sente alquanto assente dai problemi che agitano l'ora presente; egli resta tuttavia una eminente figura rappresentativa di un periodo glorioso per la nostra disciplina.

H. L. WESTERGAARD fu professore per oltre otto lustri nell'Università di Copenhagen, dopo essere stato impiegato nell'Istituto statale danese delle assicurazioni sulla vita. Membro di numerose società scientifiche ed economiche danesi ed internazionali, apparteneva dal 1902 all'Istituto Internazionale di Statistica e nel 1934 alla Sessione di Londra, circondato da grande venerazione da parte di tutti i presenti, fu calorosamente salutato da Armando Julin l'attuale Presidente dell'Istituto.

Oltre alle opere citate ricorderemo ancora di lui le seguenti:

La mortalità nelle campagne (in danese, in collaborazione con M. RUBIN) del 1886;

La statistica dei matrimoni (in danese, in collaborazione con M. RUBIN) del 1890;

L'antica economia nazionale e le sue interpretazioni sociologiche (in danese) del 1896;

L'introduzione allo studio dell'Economia Nazionale (in danese) del 1897;

L'influenza dell'alcoolismo sulla durata della vita (in danese) del 1910;

Gli scopi ed i metodi della statistica (in inglese) del 1916;

Lo sviluppo economico della Danimarca (in inglese) del 1922;

Le statistiche demografiche dei paesi scandinavi e delle repubbliche del Baltico (in inglese) del 1926.

G. PIETRA

PROF. CORRADO GINI, *Direttore responsabile.*

Roma — Tipografia Ippolito Failli.