立信統計叢書

論先進平均數

徐
華

立信會計圖書[illegible]品社出版

立信統計叢書

論先進平均數

全一冊

每冊人民幣六千五百元

著　者　徐鍾濟　華伯泉

出版者　立信會計圖書用品社
上海河南中路三三九號
重慶小什字立信大樓
天津建設路一號

總發行所　中國科技圖書聯合發行所
上海中央路二四號三〇四室

一九五一年十一月初版　（滬）

0001－2000（順）

序　言

最近東北統計局編譯之奥斯特魯莫夫教授所著之新統計學概論二冊，關於先進平均數論述甚詳。先進平均數是今日蘇聯國民經濟各部門製訂計劃定額的基礎，現在已被普遍地應用在工業和農業方面了，其重要性由此可見。

本書包含論文五篇，對於先進平均數之計算及應用，均有直接或間接之關係。前兩文係伯泉在上海財經學院教學研究之一得。第一文將奥氏所述先進平均數之計算方法提出修正。第二文將葉若夫氏等所述平均差之計算方法加以修正。雖平均差之修正公式最初已見於芮茲(1925)，今亦見於國內各教本，然作者係就二種假定作一詳細而明晰之敘述，頗便於初學之研習。

後三文係鍾濟在之江大學和浙江大學擔任高級統計學之補充教材。第三文敘述平均差的正確計算法之各種簡化公式；尤其是對中位數的平均差（記號為 δ_2），自 1943 年後國內各統計學者用不同方法獲得許多簡捷公式，這些創作可視為我國統計學者的集體貢獻，已於第五文中提及。第四文除將華氏先進平均數(記號為 $\overline{X}_a$ 或 $\overline{X}_b$) 的正確計算法化簡，並說明了先進平均數 $\overline{X}_a$ 和對算術均數的平均差(記號為 δ_1)間之密切關係。第五文建議了一個以中位數為分組標準的新的先進平均數(記號為 $\overline{X}'_a$ 或 $\overline{X}'_b$)，並指出了平均差 δ_2 之新定義，$\delta_2=(\overline{X}'_a-\overline{X}'_b)/2$。同時也可建議另一個新的差量 $\delta'_1=(\overline{X}_a-\overline{X}_b)/2$。它們均可用以測度先進平均數與落後部份的平均數二者之間的集體相差的大小。並且還於第四第五兩文中建議了兩個新的偏態量數和兩個新的峯態量數，用以測驗分配之常態性，計算簡捷，應用亦廣。惟以各文倉卒寫成，遺漏必多，尚乞全國學者不吝指教，倘能收拋磚引玉之效，則是作者們所最為感盼的。

各文曾廣泛地引用許多參考書，尤以參考奧斯特魯莫夫(1951)、薛仲三(1950)、萬鴻開(1944)、和芮茲(1925)諸氏著作之處尤多，甚至有引用原著未及註明出處的，讀者當可看出有些內容和作圖係摘自褚一飛(1943)、羅大凡譯(1948)、凱萊(1924)、賴兜(1939)、鏗道爾(1943)、克洛斯登及考登(1948)和芮茲(1927)諸書，這些都是作者們所應當聲明和感謝的。

各文稿承浙江大學和上海財經學院印成講義，之江大學楊開生君擔任繕寫，薛仲三魏宗舒兩敎授通信指正，金國寶王思立潘葆塀三敎授推薦和協助出版，均是作者們所深爲感謝的。

徐鍾濟 於 杭州之江大學 1951 年 7 月
華伯泉 於 上海財經學院

目　次

先進平均數的正確計算方法及其應用(註一)

華伯泉

一、先進平均數的意義：

先進平均數是蘇聯國民經濟各部門製訂計劃定額體系的基礎(註二)，是蘇聯的學者們創造出來的一種平均數，就是一個變數的一羣變量在其總體中先進於總體之算術平均數的那部份變量的平均數(註三)，所謂「先進」必須視所測度的變數而定；如欲測定單位時間內的產品定額時，我們當然希望單位時間內的產品愈多愈好，那末「先進於總體(註四)之算術平均數」便應體會為「大於總體之算術平均數」；如欲測定單位產品所消耗的時間，材料或電力等等有關成本的定額，我們當然希望單位產品的成本愈低愈好，那末「先進於總體之算術平均數」便應體會為「小於總體之算術平均數」，這是很顯明的。先進平均數與第三或第一、四分位數（Q_3 或 Q_1）不同；因為第三或第一、四分位數只是一個位置的量數，而先進平均數是先進於算術平均數各個變量的平均量數。

二、先進平均數的計算方法：

(一)未分組資料：

1. 計算公式： 按照定義計算先進平均數之公式為

$$\overline{X}_a=\frac{\Sigma X_a}{N_a} \quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots (I)$$

各個符號代表的意義如下：

$\overline{X}_a$： 先進平均數。

ΣX_a： 先進於算術平均數各個變量之和。

N_a： 先進於算術平均數的各個變量數。

2. 計算步驟：

(1) 求算術平均數 ($\overline{X}$)。

(2) 求先進於算術平均數各個變量之和(ΣX_a)。

(3) 代入公式(I)即可求得先進平均數($\overline{X}_a$)。

例如某車間製造某種機器零件，十個工人每日每人的平均生產量（個數）如下：

15, 16, 17, 17, 18, 19, 19, 20, 22, 25

試求其先進平均數。

解：(1) $\overline{X}=\frac{\Sigma X}{N}=\frac{188}{10}=18.8$ 個

(2) $\Sigma X_a=19+19+20+22+25=105$

(3) $\overline{X}_a=\frac{\Sigma X_a}{N_a}=\frac{105}{5}=21$ 個

答：先進平均數是 21 個。

(二)分組資料：

爲敍述有序起見，以下先從"大於總體之算術平均數"說起。

甲、普通法：

1. 計算公式：

$$\overline{X}_a=\frac{f_uX_u+\Sigma f_aX_a}{f_u+N_a} \cdots\cdots\cdots\cdots (II)$$

$$f_uX_u=\frac{(U^2-\overline{X}^2)f}{2C}$$

$$f_u=\frac{(U-\overline{X})f}{C}$$

各個符號的代表意義如下；

$\overline{X}_a$ ：先進平均數。

$\overline{X}$ ：算術平均數。

f_u ：算術平均數至算術平均數所在組之上限間所含的次數。

X_u ：算術平均數至算術平均數所在組之上限間的組中値。

U ：算術平均數所在組之上限。

f ：算術平均數所在組之次數（或稱頻數）。

C ：組距。

N_a ：大於算術平均數所在組各組次數之和。

$\Sigma f_a X_a$：大於算術平均數所在組各組次數與其組中值相乘之積的和。

2. 計算步驟：

(1) 求算術平均數 ($\overline{X}$)。

(2) 確定算術平均數所在組之次數(f)，算術平均數所在組之上限 (U) 及組距 (i)。

(3) 計算算術平均數至算術平均數所在組之上限間所含的次數 (f_u)，

(4) 計算算術平均數至算術平均數所在組之上限間的組中值 (X_u)。

(5) 求 f_u 與 X_u 之乘積 ($f_u X_u$)。

(6) 求大於算術平均數所在組各組次數與其組中值相乘之積的和 ($\Sigma f_a X_a$)。

(7) 求大於算術平均數所在組各組次數之和 (N_a)。

(8) 將所得各數代入公式(II)卽可求得先進平均數($\overline{X}_a$)。

例如某工廠中有二百個工人從事生產之統計資料如下表，試計算其每個工人每日之平均生產量，假定所計算出的定額，我們將以它作爲下期計劃的基礎。

某廠工人每日平均生產量統計表

每個工人每日平均生產量（個數）	工人數
20——30	20
30——40	40
40——50	56
50——60	44
60——70	34
70——80	6
總計	200

解：　根據題意，我們知道既然要將這個平均數作爲下期計劃的基礎，那末一定是要求先進平均數的，現在先列出其計算表如下；然後按上述計算步驟計算之。

某廠工人每日平均生產量先進平均數計算表

G	f	X	fX
20——30	20	25	500
30——40	40	35	1,400
40——50	56	45	2,520
50——60	44 } 84	55	2,420 } 5,080
60——70	34	65	2,210
70——80	6	75	450
總計	200		9,500

(1) $\overline{X}=\dfrac{\Sigma fX}{N}=\dfrac{9,500}{200}=47.5$ 個

(2) $f=56,\quad U=50,\quad C=10$

(3) $f_u=\dfrac{(U-\overline{X})f}{C}=\dfrac{(50-47.5)\times 56}{10}=\dfrac{140}{10}=14$

(4) $X_u=\dfrac{U+\overline{X}}{2}=\dfrac{50+47.5}{2}=48.75$

(5) $f_uX_u=14\times 48.75=682.5$

(6) $\Sigma f_aX_a=5,080$

(7) $N_a=84$

(8) $\overline{X}_a=\dfrac{f_uX_u+\Sigma f_aX_a}{f_u+N_a}=\dfrac{682.5+5,080}{14+84}=\dfrac{5762.5}{98}=58.8$ 個

或求得算術平均數後、直接代入下面這個公式。

$$\overline{X}_a=\frac{\dfrac{(U^2-\overline{X}^2)f}{2C}+\Sigma f_aX_a}{\dfrac{(U-\overline{X})f}{C}+N_a}\cdots\cdots\cdots\cdots(III)$$

如上例

$$\overline{X}_a=\frac{\dfrac{(50^2-47.5^2)\times 56}{2\times 10}+5080}{\dfrac{(50-47.5)\times 56}{10}+84}$$

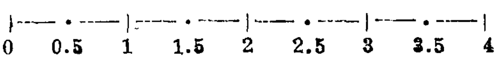
$$=\frac{\dfrac{(2500-2256.25)\times 56}{20}+5080}{14+84}$$

$$=\frac{682.5+5080}{98}=58.8\text{ 個}$$

答：　先進平均數爲58.8個。

3. 公式證明：

首先我們知道統計學上分組資料之各種特徵數的計算是建築在一種理論的假設前提上的，這就是各組的次數是平均分配在各該組距內的，事實上這種平均分配的情形正如C.奥斯特魯莫夫教授所説的是很少存在的(註五)，但是如果沒有這種理論的假設前提，那末各組的次數就不能以組中值來代表，換言之，分組資料也就無法處理，那末統計也就不能執簡馭繁了。這個理論的假設前提，蘇聯統計學者們也是承認的(註六)，所以如果一組中祇有次數 1，那末這 1 次數是假定落在組中點的，因此就以組中點的量數來代表這 1 次數，同樣假定 0—4 一組中有次數 4，那末理論上他的分配情形一定如下圖：



這四個次數的代表量數的和是：

$$0.5+1.5+2.5+3.5=8$$

而 0—4 一組的組中值是 2，2 乘 4 亦得 8，所以這四個次數都可以組中值 2 來代表，但是並不是說這四個次數都集中在 2 這一點上，而是 0—2 間有次數 2，2—4 間亦有次數 2，因爲假設的前提是次數平均分配在組距內的。

根據上面這個不移的假設和說明，所以算術平均數至算術平均數所在組之上限間在理論上必定有若干次數分配在內的，亦就是算術平

均數所在組內必定有若干次數的代表量數是大於算術平均數的，除非算術平均數確巧等於其所在組之上限，那末其所在組內之次數所代表的量數都小於算術平均數。例如上例，根據理論的假設前提，必定有十四個工人的每日平均生產量是超過 47.5 個的；也就是在 47.5—50 間尙有次數 14，如果忽略了這一點，那末所計算得到的先進平均數是不正確的；而且一定會比眞正的先進平均數估計過高。

C・奥斯特魯莫夫教授於其『新統計學概論』上册（東北計劃委員會統計局譯）中第119頁亦承認這個理論的假設前提的，他說：「……以這些組中點做爲代表各該組內各個數值總體的特徵的平均數，……因爲在確定組中點時，在組距內，是以平均分配各個標識的數值爲前提的」。但是該書第130頁所舉的例子中（本文所引用的例子卽錄自此書）計算所得的先進平均數確忽略了這一點，因此

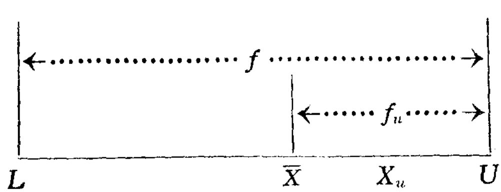
$$\overline{X}_a = \frac{5080}{84} = 60.5 \text{ 個}$$

而正確的先進平均數應該是 58.8 個，C・奥斯特魯莫夫教授對於求先進平均數所應用的公式雖然沒有寫出；但是根據其所舉的例子，可以寫成

$$\overline{X}_a = \frac{\Sigma f_a X_a}{N_a}$$

這明確的是忽略了算術平均數與其所在組之上限間在理論上有若干次數分配存在的；而且極大多數的事實也必定是如此的。

至於求算術平均數與其所在組之上限間的次數可以作圖如下以明之。



假定上圖 LU 代表算術平均數所在組，L 是下限，U 是上限，組距

爲 C，即 $LU=C$。

$\overline{X}-U$ 表示算術平均數與其所在組之上限間的距離。

f 代表 LU 間共有的次數。

f_u 代表 $\overline{X}-U$ 間所含有的次數。

根據理論的假設前提 f 是平均分配在 LU 間的，所以 $\overline{X}-U$ 間所含有的次數可以用比例求得之，即

$$LU : \overline{X}U = f : fu$$

$$\therefore \qquad f_u=\frac{\overline{X}Uf}{LU}$$

又因爲　$LU=C$，　$\overline{X}U=U-\overline{X}$

故　$$f_u=\frac{(U-\overline{X})f}{C}$$

同時 f_u 的每一次數可以 $\overline{X}U$ 間的組中值爲其代表量數。今假定其組中值用符號 X_u 來代表，則

$$X_u=\overline{X}+\frac{U-\overline{X}}{2}=\frac{U+\overline{X}}{2}$$

所以算術平均數與其所在組之上限間的次數之總量數爲 f_u 乘 X_u。即

$$f_uX_u=\frac{(U-\overline{X})f}{C}\cdot\frac{(U+\overline{X})}{2}=\frac{(U^2-\overline{X}^2)f}{2C}$$

同樣大於算術平均數各組之次數均可以各該組之組中值爲代表，而求得各組次數之總量數，即各個 f_aX_a。

然後根據算術平均數的定義，將各組次數之總量數加總，再以各組次數之和除之，即可直接得到先進平均數的公式如下：

$$\overline{X}_a=\frac{f_uX_u+\Sigma f_aX_a}{f_u+N_a}\text{……………………此即公式}(II)$$

若以　$f_u=\dfrac{(U-\overline{X})f}{C}$　及　$f_uX_u=\dfrac{(U^2-\overline{X}^2)f}{2C}$

代入上式，即可得公式 (III)。

乙、簡捷法：

如果算術平均數以上有許多組，而且變量甚大時 那末求大於算術平均數各組之次數與各該組組中值之乘積便極爲費時與繁複，則可應用下列簡捷方法計算之。

1. 計算公式：

$$\overline{X}_a=\frac{f_uX_u+N_aX_o+C\Sigma f_ax'}{f_u+N_a}\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots(IV)$$

或

$$\overline{X}_a=\frac{\dfrac{(U^2-\overline{X}^2)f}{2C}+N_aX_0+C\Sigma f_ax'}{\dfrac{(U-\overline{X})f}{C}+N_a}\cdots\cdots\cdots\cdots\cdots(V)$$

$$x'=\frac{X-X_0}{C}$$

各個符號的代表意義如下：

X_0：　假定算術平均數。

x'：　組距單位離差。

其餘各個符號的代表意義同前。

2. 計算步驟：　同普通法，可是要先確定一假定算術平均數，普通是以次數較多，且其上與其下各組次數之和是近乎相等之一組的組中值爲假定算術平均數；但是亦可以選擇任何一組之組中值爲假定算術平均數，然後據此求各組之組距單位離差。

例如上例可用簡捷法計算之，先列計算表如下：

某廠工人每日平均生產量先進平均數計算表

G	f	x'	fx'
20——30	20	—2	—40
30——40	40	—1	—40
40——50	56	0	0
50——60	44	1	44
60——70	34 }84	2	68 }130
70——80	6	3	18
總　　計	200		50

(1) $X_0=45$

$$\overline{X}=X_0+\frac{\Sigma fx'}{\Sigma f}\times C=45+\frac{50}{200}\times 10=47.5\text{ 個}$$

(2) $f=56,\quad U=50,\quad C=10,$

(3) $f_u=\frac{(U-\overline{X})f}{C}=\frac{(50-47.5)\times 56}{10}=\frac{140}{10}=14$

(4) $X_u=\frac{U+\overline{X}}{2}=\frac{50+47.5}{2}=48.75$

(5) $f_uX_u=14\times 48.75=682.5$

(6) $\Sigma f_ax'=130$

(7) $N_a=84$

(8) $\overline{X}_a=\frac{f_uX_u+N_aX_0+C\Sigma f_ax'}{f_u+N_a}$

$$=\frac{682.5+84\times 45+10\times 130}{14+84}$$

$$=\frac{682.5+3780+1300}{98}=\frac{5762.5}{98}=58.8\text{ 個}$$

或求得算術平均數後，直接代入公式(V)，即得

$$X_a=\frac{\frac{(50^2-47.5^2)\times 56}{2\times 10}+84\times 45+10\times 130}{\frac{(50-47.5)\times 56}{10}+84}$$

$$=\frac{682.5+3780+1300}{14+84}=\frac{5762.5}{98}=58.8\text{ 個}$$

3. 公式證明：

因為 $x'=\frac{X-X_0}{C}$

所以 $X=X_0+x'C$

故 $\Sigma f_aX_a=\Sigma f_a(X_0+x'C)=N_aX_0+C\Sigma f_ax'$。

以此代入公式(II)便得

$$\overline{X}_a=\frac{f_uX_u+N_aX_0+C\Sigma f_ax'}{f_u+N_a}\text{…………此即公式}(IV)$$

同理將 $\Sigma f_a X_a = N_a X_0 + C\Sigma f_a x'$ 代入公式 (III)，即得公式 (V)。

至於測定有關成本的先進平均數，便需計算小於總體之算術平均數的那部份變量的平均數，設以 $\overline{X}_b$, X_b, f_b, N_b 及 X_l, f_l 依次代替以上的 $\overline{X}_a$, X_a, f_a, N_a 及 X_u, f_u；所有符號下面添加的小註 b 或 l 都是表示「小於總體之算術平均數」之意。那末以上公式 (II) 及 (IV) 依舊可以適用，但是其中

$$f_l = \frac{(\overline{X} - L)f}{C},$$

$$f_l X_l = \frac{(\overline{X}^2 - L^2)f}{2C}。$$

L 代表算術平均數所在組之下限，因此公式 (III) 變成

$$\overline{X}_b = \frac{\dfrac{(\overline{X}^2 - L^2)f}{2C} + \Sigma f_b X_b}{\dfrac{(\overline{X} - L)f}{C} + N_b} \cdots\cdots\cdots\cdots\cdots\cdots (III')$$

公式 (V) 變成

$$\overline{X}_b = \frac{\dfrac{(\overline{X}^2 - L^2)f}{2C} + N_b X_0 + C\Sigma f_b x'}{\dfrac{(\overline{X} - L)f}{C} + N_b} \cdots\cdots\cdots\cdots (V')$$

上述公式的證明與前面的證法完全相似，所以不再重複敍述。

三、先進平均數的應用：

在蘇聯，統計學是建設社會主義和建設共產主義的武器，是國民經濟計劃工作的武器，是技術改進的武器，蘇聯自十月革命至今三十三年間在經濟建設上有如此輝煌的成績，無可否認的統計學是起了一定的作用，這與偉大的列寧和斯大林底卓越與正確的領導是分不開的，列寧這樣指示說：「一系列關係着現代國家的經濟情況及其發展的根本問題，在從前是憑一般的推測和近似的資料求得解決，而今天如果不根據統一的提綱去搜集全國各地的大量材料，並由統計專家加以整理，便難期解決問題。」(註七) 又說：「統計學應當用數字表示出已完全顯現或

正在顯現在生活中的被研究現象的各種社會類型。」(註八)「我們要把這個統計學(馬克思主義的經濟統計理論統計教科書)帶到羣衆中去，推廣普及，以便勞動者漸漸學會自己判斷，應當怎樣做工作，作多少；應當怎樣休息。休息多少。」(註九) 斯大林這樣指示說：「缺了統計，核算便不能向前進；而缺了核算，一切建設工作，一切國家工作，一切計劃工作都是不可想像的。」(註十) 斯大林在第一次全蘇聯斯塔哈諾夫工作者會議上說「斯塔哈諾夫運動這一男女工人運動的目的，是要打破現今的技術定額，打破現行的生產標準，打破現行的生產計劃和對照表，……沒有技術定額便無法進行計劃經濟。除此而外，其所以需要技術定額，是爲了督促落後羣衆來趕上先進份子，技術定額是一種巨大的調節力量，它能在生產中把廣泛工人羣衆組織在工人階級先進份子周圍，所以，我們需要有技術定額，但我們所需要的不是現行的定額，而是較高的定額，……有些人說，技術定額是需要的，可是現在就必須把它提高到斯塔哈諾夫、布塞根、兩位維諾格拉多娃及其他等人所已經達到的高度，這也是不對的，這樣一種定額在目前這個時候是不實際的，因爲那些比斯塔哈諾夫和布塞根等人底技術素養低些的男工和女工，就會不能完成這一種定額。我們現在所需要的技術定額，應當是一個介乎現行技術定額與斯塔諾夫和布塞根等人所達到的標準間的定額，……很明顯的，我們必須規定一個介乎現行技術定額和斯塔哈諾夫同志所達到的標準間的定額」(註十一)。由於這許許多多的正確指示，蘇聯由重視統計而實行計劃經濟，由舊的計劃定額而新的計劃定額，即是由算術平均定額而採用先進平均定額，蘇聯部長會議當批准一九四七年度的計劃時，曾通令各部局規定計劃的數字不應以算術平均定額爲基礎，而應以先進平均定額爲基礎；就是說，應該向先進的工業企業看齊，計劃定額的製訂方式及其計算方法，不論是在直接的計劃製訂過程當中，或者是在計劃執行的鬥爭當中，或者是在爲發掘並利用國內資源的鬥爭當中，都具有決定的意義。所以蘇聯國民經濟的任何一個部門，在審議計劃定額時所發生的最重要的問題，就是計劃定額的如何製訂方式和計算方