138545

立信統計叢書

高級統計學

鄒依仁著

（一九五一年初版）

2,0

立信會計圖書用品社出版

立信統計叢書

高級統計學

鄒依仁著

（一九五一年初版）

立信會計圖書用品社出版

立信統計叢書

高級統計學

全一冊

版權所有
不准翻印

每冊人民幣一萬四千二百元

著者 鄒依仁

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

一九五一年七月初版
一九五一年十一月再版 (滬)

3001—6000 (東)

說　　明

1. 本書係著者在復旦大學，重慶大學，中央大學（現稱南京大學），上海財經學院等校連續講授高等統計、統計原理十餘年的講稿，經多次修改而編成。如將本書作爲高級統計或統計原理的教本或主要參考書，最爲適宜。

2. 凡曾修習普通統計學與微積分入門者，即可閱讀本書的教材，而無任何困難。

3. 本書內容極力避免數理，並求深入淺出，多舉實例來說明最近的統計學理與方法。至於牽涉虛玄而不實用的理論，則皆已删除不用。

4. 本書所附各種統計表格，皆適合於實用性質。其查閱方法已詳見各實例中。

高級統計學

目　錄

试读结束：需要全本请在线购买： www.ertongbook.com

第一章

動差方法與統計特徵數

(Method of Moments & Statistics)

第一節　動差與統計特徵數的意義

在物理靜力學槓桿定律中，力乘力支距稱爲動差 (Moment)，若支點左右之力率相等，則槓桿成平衡。在統計次數分配中，以各組之次數（相當於力）乘各量數與某一定值 A（相當於支點）之差（相當於力支距），求其總和，再以總次數除之，如 $\frac{1}{N}\Sigma f(x-A)$，其商亦稱爲對 A 之動差。若此定值恰爲算術平均數，則此動差爲零（相當於平衡）即 $\frac{1}{N}\Sigma f(x-M)=0$。

統計學家利用此種動差的意義並擴展至高次，卽作出動差的基本定義如次：

以量數絕對數值零爲起點的各次動差爲

第一次動差 $V'_1=\frac{\Sigma xf}{N}$　　V 讀作 nu。

第二次動差 $V'_2=\frac{\Sigma x^2f}{N}$

…………………………

第 n 次動差 $V'_n=\frac{\Sigma x^nf}{N}$

V'_n 爲動差的基本定義。

如以量數的平均數 M 爲起點的各次動差爲

$$V_1=\frac{\Sigma(x-M)f}{N}$$

$$V_2=\frac{\Sigma(x-M)^2f}{N}$$

....

$$V_n=\frac{\Sigma(x-M)^nf}{N}。$$

凡對統計資料應用分析方法計算出之各種代表量數統稱爲統計特徵數。例如集中趨勢的各種平均數，離中趨勢的各種離中量數，偏斜度的各種偏斜係數，峯度內各種峯度，以及相關度的各種相關量數等等皆是也。所謂主要統計特徵數者爲各類量數中最通用與最有代表性者，如集中趨勢的算術平均數，離中趨勢之標準差，偏斜度之偏斜係數 a_3，峯度中之 a_4，以及相關度之簡單相關係數等是也。

如採用動差方法來計算上述各主要特徵數，則非特其方法爲有系統的進行，且可避免死背統計公式之苦也。

第二節　動差展開式

由動差方法可以作有系統的計算各主要特徵數。對主要特徵數所用的當以算術平均數作爲起點的各次動差。但在實際計算時，以算術平均數作爲起點的動差，因計算太煩而變爲不實用。一般所用者當以零爲起點之各次動差是也。爲適應此種實際情況起見，吾人必須發展一套使以算術平均數作爲起點的各次動差而化成以零爲起點的各次動差的展開式。其展開方式甚易。今且按步討論之。

在前節中已述及以零爲起點的第一次動差 $V'_1=\frac{\Sigma xf}{N}$ 卽代表算術平均數之意義，亦卽 $M=V'_1$。至於以算術平均數爲起點之第一次動差爲

$$V_1=\frac{\Sigma(x-M)f}{N}=\frac{\Sigma xf}{N}-M\frac{\Sigma f}{N}=V'_1-M=M-M=0,$$

由此知以算術平均數爲起點的第一次動差應爲零。

現將推論至二次，三次，四次，以及 n 次。

$$V_2=\frac{\Sigma(x-M)^2f}{N}=\frac{\Sigma(x^2-2Mx+M^2)f}{N}$$

$$=\frac{\Sigma x^2f}{N}-\frac{2M\Sigma xf}{N}+\frac{M^2\Sigma f}{N}$$

$$=V'_2-2MV'_1+M^2=V'_2-2M^2+M^2$$

$$=V'_2-M^2。$$

但吾人已知以算術平均數作爲起點的第二次動差之意義即爲標準差自乘方，亦即

$$V_2=\sigma^2。$$

此種標準差自乘方普通稱爲差異或即二次動差 (Variance)。由其展開式，可得

$$\sigma=\sqrt{V_2}=\sqrt{V'_2-M^2}$$

進而討論至三次與四次。

$$V_3=\frac{\Sigma(x-M)^3f}{N}=\frac{\Sigma(x^3-3x^2M+3xM^2-M^3)f}{N}$$

$$=V'_3-3MV'_2+3M^2V'_1-M^3$$

$$=V'_3-3MV'_2+2M^3。$$

吾人已知 V_3 代表偏斜度的意義。爲避免單位起見，偏斜係數 a_3 爲

$$a_3=\frac{V_3}{\sigma^3}。$$

同理，
$$V_4=\frac{\Sigma(x-M)^4f}{N}=V'_4-4MV'_3+6M^2V'_2-3M^4。$$

吾人已知 V_4 代表峯度的意義。與 a_3 相同，爲避免單位起見，峯度 a_4 爲

$$a_4=\frac{V_4}{\sigma^4}。$$

如擴展 V 至 n 次，則

$$V_n=\frac{\Sigma(x-M)^nf}{N}=V'_n-NV'_{n-1}M+\frac{n(n-1)}{2}V'_{n-2}M^2-\cdots。$$

當 n 在四次以上並無相當的特徵數作爲對照，故並無實際意義可言。由

此知一般發展至四次已經足够。

第三節　動差還元式

在實際應用而　由 V'_n 計算所得 V_n 的數值時，計算中可能發生錯誤而無法覺察。爲求出並校正此等錯誤計，可以發展一套動差還元式以測驗之。所謂動差還元式者即反動差發展式而以 V_n 的數值求出 V'_n 的數值是也。

先使各量數 x 對算術平均數 M 的離中差 $x-M$ 爲 X，即

$$X=x-M$$

或即　$$x=X+M,$$

因此，　$$V'_n=\frac{\Sigma x^n f}{N}=\frac{\Sigma(X+M)^n f}{N}$$

$$=\frac{\Sigma\left(X^n+nX^{n-1}M+\frac{n(n-1)}{2!}X^{n-2}M^2+\cdots\cdots+M^n\right)f}{N}$$

$$=\frac{\Sigma X^n f}{N}+\frac{nM\Sigma X^{n-1}f}{N}+\cdots\cdots+\frac{\Sigma M^n f}{N}$$

$$=V_n+nMV_{n-1}+\cdots\cdots+M^n。$$

故　$$V'_2=V_2+M^2$$

$$V'_3=V_3+3V_2M+M^3$$

$$V'_4=V_4+4V_3M+6V_2M^2+M^4$$

由還元結果如所得 V'_2, V'_3, V'_4 之數值相同於以前所計算者，則知在計算過程中並無錯誤存在。反之，V'_2, V'_3, V'_4 等數值如與以前所計算者顯有不同時，則知兩者間定有錯誤發生而必須加以校正之是也。

第四節　應用動差方法計算主要特徵數的實例

設已知下列次數分配，而應用動差方法計算其主要特徵數，M, σ, α_3, α_4 等。

试读结束：需要全本请在线购买： www.ertongbook.com

組中值 x'	次數 f
47	5
48	2
49	13
50	23
51	58
52	96
53	134
54	127
55	111
56	74
57	37
58	16
59	4
60	2
61	1
	$N=703$

爲簡化計算起見,取“54”爲假定起點卽假定原點如次:

x'	x	f	xf	x^2f	x^3f	x^4f	$(x+1)^4f$
47	−7	5	−35	245	−1715	12005	6480
48	−6	2	−12	72	−432	2592	1250
49	−5	13	−65	325	−1625	8125	3328
50	−4	23	−62	368	−1472	5888	1863
51	−3	58	−172	522	−1566	4698	928
52	−2	96	−192	384	−768	1536	96
53	−1	134	−134	134	−138	134	0
54	0	127	0	0	0	0	127
55	1	111	111	111	111	111	1776
56	2	74	148	296	592	1184	5994
57	3	37	111	333	999	2997	9472
58	4	16	64	256	1024	4096	10000
59	5	4	20	100	500	2500	5184
60	6	2	12	72	432	2592	4802
61	7	1	7	49	343	2401	4096
		$N=703$	−704	3267	−7712	50859	55396
			+473		+4001		
			−231		−3711		

由於 Σxf, Σx^2f, Σx^3f, Σx^4f 等數值計算 V'_1, V'_2, V'_3, V'_4 等數值,然後由 V'_1, V'_2, V'_3, V'_4 等數值應用動差展開式計算M, V_2, V_3, V_4 等數值，最後計算 σ, a_3, a_4 等數值。最後一列 $\Sigma(x+1)^4f$ 卽所謂 Charlier 氏還元法以觀察表格上諸數值內是否有所差誤。

$$V'_1=M'=\frac{-231}{703}=-0.3286 \qquad V_1=0, \quad M=54-0.3285=53.6714$$

$$V'_2=\frac{3267}{703}=4.6472 \qquad V_2=V'_2-V'_1=4.6472-0.1079=4.5393$$

$$V'_3=\frac{-3711}{703}=-5.2788,\ V_3=V'_3-3V'_2V'_1+2V'^3_1$$

$$=-5.2788+4.5798-0.0708=-0.7698$$

$$V'_4=\frac{50859}{703}=72.3456\quad V_4=V'_4-4V'_3V'_1+6V'_2V'^2_1-3V'^4_1$$

$$=72.3456-6.9363+3.0089-0.0349$$

$$=68.3833$$

由此等 V_2, V_3, V_4 等數值計算得

$$\sigma=2.1305,\quad \alpha_3=\frac{V_3}{\sigma^3}=\frac{-0.7696}{9.6704}=-0.0796$$

$$\alpha_4=\frac{V_4}{\sigma^4}=\frac{68.3833}{20.6052}=3.3187$$

今且歸納實際計算的步驟如此:

1. 先確定一假定原點卽假定起算點。

2. 計算各 xf, x^2f, x^3f, x^4f, 與 $(x+1)^4f$ 與其總和。

3. 應用 Charlier 的還元法展開

$$\Sigma(x+1)^4f=\Sigma x^4f+4\Sigma x^3f+6\Sigma x^2f+4\Sigma xf+\Sigma f$$

而觀察 xf, x^2f, x^3f, x^4f 等數值有無差誤。

4. 計算以假定原點零爲起算的動差 V'_1, V'_2, V'_3, V'_4。

5. 由 V'_1, V'_2, V'_3, V'_4, 計算以算術平均數 M 爲原點的動差 V_2, V_3, V_4。

6. 應用動差還元式觀察 V_2, V_3, V_4 等數值是否爲正確。

7. 如次數分配爲連續性時,應用薛伯氏 (Sheppard) 校正法校正各次動差。

8. 以校正以後的動差 μ'_1, μ'_2, μ'_3, μ'_4 等數值計算 M, σ, α_3, α_4。

應用動差方法應注意三點:

1. 非但對已歸類資料可以應用動差方法來作有系統的計算各主要特徵數,卽對未歸類資料(卽資料未作成次數分配者)亦可同樣的應用動差方法。當資料爲未歸類時,所有次數皆爲 1, 卽

$$f_1=f_2=f_3=\cdots\cdots=1,$$

而動差公式原爲

$$V'_n=\frac{\Sigma x^n f}{N},\quad V_n=\frac{\Sigma(x-M)^n f}{N},$$

在未歸類資料時爲

$$V'_n=\frac{\Sigma x^n}{N},\quad V_n=\frac{\Sigma(x-M)^n}{N}。$$

2. 組中值間的單位問題。如組中值間或組距不爲一個單位時，則計算所得的 M 與 σ 皆應乘以組中值間或組距的原單位。但對於 a_3 與 a_4 因不受單位的任何影響，故不必再乘以單位。

3. 如次數分配的量數爲連續性時，因係根據次數集中在組中值的假定下而計算各次動差 V_n，故必須應用薛伯氏校正法：

$$\mu_1=V_1=0$$

$$\mu_2=V_2-\frac{h^2}{12}$$

$$\mu_3=V_3$$

$$\mu_4=V_4-\frac{h^2}{2}V_2+\frac{7}{240}h^4$$

μ 爲校正後的動差記號。如組距爲一個單位時，則 $h=1$，校正式爲

$$\mu_1=V_1=0$$

$$\mu_2=V_2-\frac{1}{12}$$

$$\mu_3=V_3$$

$$\mu_4=V_4-\frac{1}{2}V_2+\frac{7}{240}。$$

第五節　動差與簡單相關係數

以上四節所述者皆限於一個資料或一個變數的動差方法。如擴充至二個資料或二個變數時，則動差即稱爲積差（Product-moment）。其意義與一個資料相同。

以零爲起點的第 nk 次積差公式爲

$$V'_{nk}=\frac{\Sigma x^n y^k f_{xy}}{N}。$$

以算術平均數爲起點的第 nk 次積差公式爲

$$V_{nk}=\frac{\Sigma(x-M_x)^n(y-M_y)^k f_{xy}}{N}。$$

如以算術平均數爲起點的第一次論

$$V_{11}=\frac{\Sigma(x-M_x)(y-M_y)f_{xy}}{N},$$

而簡單相關係數之定義卽爲以算術平均數爲起點的第一次積差除以標準差單位是也。

$$r=\frac{V_{11}}{\sigma_x\sigma_y}。$$

上式所以除以 σ_x，σ_y 之原因，係避免單位之不同。由此知簡單相關係數 r 當不受單位之影響也。

第二章
二項理論次數分配
(Theoretical Binomial Frequency Distribution)

第一節　二項理論次數分配的意義

二項理論次數分配簡稱二項分配。其來源為代數中二項展開式。如以機率的形式為討論出發點，則卽發展成所謂二項次數分配。吾人可以一銅幣為例。一般銅幣共有二面，一卽正面，一卽反面。若銅幣之組織很為均勻時，則將其旋轉之可能性，出現正面(H)或"字"與反面(T)或"背"有同等之機會。亦可謂旋轉一個銅幣一次時，可能得正面 H 或可能得反面 T，兩者中必得其一，則為確定之事實。以記號表示之為

$$H+T=1$$

或
$$\frac{1}{2}+\frac{1}{2}=1。$$

如旋轉兩個銅幣一次或一個銅幣二次時，則出現結果稍為繁複，其可能性為

$$HH,\ HT,\ TH,\ TT$$

亦卽
$$H^2,\quad 2HT,\quad T^2,$$

以機會表示之為

$$\left(\frac{1}{2}\right)\left(\frac{1}{2}\right),\quad \frac{1}{2},\quad \left(\frac{1}{2}\right)\left(\frac{1}{2}\right)$$

亦卽
$$\frac{1}{4}\qquad \frac{1}{2}\qquad \frac{1}{4}$$

而其總和爲一，其意義爲表示確定性，亦卽在三種可能性中必定出現一種是也。

如擴充此種情况至旋轉 n 個銅幣一次或一個銅幣 n 次時，則其結果合於二項展開式

$$(H+T)^n=H^n+nH^{n-1}\cdot T+\frac{n(n-1)}{2!}H^{n-2}\cdot T^2+\cdots\cdots$$
$$\cdots\cdots+nH\cdot T^{n-1}+T^n=1,$$

以機率表示之爲

$$\left(\frac{1}{2}+\frac{1}{2}\right)^n=\left(\frac{1}{2}\right)^n+n\left(\frac{1}{2}\right)^{n-1}\cdot\left(\frac{1}{2}\right)+\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\cdots\cdots+n\left(\frac{1}{2}\right)\cdot\left(\frac{1}{2}\right)^{n-1}+\left(\frac{1}{2}\right)^n=1。$$

若以得正面 H 爲成功作爲 p，得反面 T 爲失敗作爲 q，則可表示爲

$$(p+q)^n=p^n+np^{n-1}\cdot q+\frac{n(n-1)}{2!}p^{n-2}\cdot q^2+\cdots\cdots+q^n=1。$$

在旋轉銅幣時，$p=q$，在其他情况如投擲骰子等等，則 p 可能不等於 q，但其展開形式當屬相同。因統計實際資料中常發現正與反的二方向，例如工業生產品統計中個別的生產品合於某種標準或不合於某種標準，生命統計中生存與死亡兩種可能性，當合於二項分配的條件，故可應用此種理論次數分配而無疑義。

第二節　二項理論次數分配的普通項與極大項

二項分配之各項已如上述。在 n 爲不大時，甚易求出此種分配之任意一項。在 n 爲相當大時，吾人恆欲求出其普通項作爲計算之標準。此種普通項實際上卽爲二項展開式之普通項，其形式爲

$$y=f_x={}_nC_x\cdot p^x\cdot q^{n-x},$$

亦可寫作爲　$$f_x=\binom{n}{x}\cdot p^x\cdot q^{n-x}=\frac{n!}{x!(n-x)!}p^x\cdot q^{n-x},$$

內 c 爲組合之符號是也。

若 n 爲極大時，則此種分配之項數極多，如欲確定何項爲最大，甚屬困難，但吾人可以下列簡易方法求出之。

试读结束：需要全本请在线购买： www.ertongbook.com