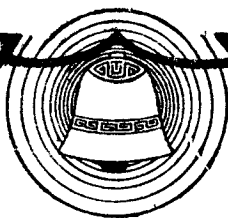


統計學概要

朱君毅 編著

正中書局印行



版權所有
翻印必究

中華民國三十七年三月初版

統計學概要

全一冊 定價國幣三元六角

(外埠酌加運費匯費)

編 著 者 朱 君 毅

發 行 人 蔣 志 澄

印 刷 所 正 中 書 局

發 行 所 正 中 書 局

(2279)

校
：
整
：
書

滬 - 本

2/1-0.15

序

抗戰以前，我國統計圖書，並時迭出。政府西遷，研讀受阻，著譯不絕如縷。邇者建設伊始，百廢俱興，統計事業，亦隨之疾趨邁進，訓練統計專才，政府視為急務；傳播統計常識，學人責無旁貸。作者鑒於坊間缺乏淺易統計書籍，不揣庸陋，公餘冗次，編輯是書，提綱挈領，趨近就顯，冀為訓練之教材，期作初學之讀物，尙祈國內專家，有以教之。稿成，全書內容，承徐鍾濟博士，余介石教授，詳加校閱；書內各圖，復承趙峻山，鍾伯元二君，細為繪製，均甚佩感，並此誌謝。

朱君毅 三十六年六月八日於南京

目次

1. 緒言	2. 偏態分配 5
統計學之定義 1	3. J形或極不對稱分配 ... 6
統計方法之特性 1	4. U形分配 6
統計技術之要素 1	次數分配之圖示法 6
統計之大別 2	1. 次數多邊圖之繪法 ... 6
1. 品質統計 2	2. 直方圖之繪法 7
2. 變數統計 2	3. 累積次數 7
II. 統計數列	III. 平均數
連續數列 2	平均數之定義 8
不連續數列 2	平均數之種類 8
全距 3	1. 算術平均數 9
組距 3	2. 中位數 15
組限 3	3. 衆數 19
應用組距之假定 4	4. 幾何平均數 22
次數分配之種類 5	5. 調和平均數 24
1. 常態分配 5	IV. 離差

離差之定義… … … 27	相關之計算法… … … 50
離差之種類… … … 27	相關係數之意義… … … 59
1. 全距… … … 27	相關係數之可靠性… … … 60
2. 四分位差… … … 29	VIII. 表列
3. 標準差… … … 31	表列之意義… … … 62
4. 平均差… … … 33	表列之級次… … … 62
5. 相對離差… … … 37	製表之規則… … … 64
6. 偏斜度… … … 39	IX. 圖示
V. 常態曲線	圖示之意義… … … 66
常態曲線之繪法… … … 40	圖示之種類… … … 67
常態曲線下之面積及機率… 43	圖示法之通則及示例… … 68
VI. 可靠性	附 錄
可靠性之意義… … … 45	主要進修與參考中西文統
算術平均數之可靠性… … 45	計書目… … … 77
樣本算術平均數之標準誤	本書統計名詞漢英對照表… 80
差… … … 47	附 表
VII. 相關	常態曲線之縱線表… … 85
相關之意義… … … 49	常態曲線下之面積表… … 86

1. 緒 言

社會情態，變異萬端，因果複雜，若無綜合化簡之術，條分縷析之方，則無從執簡以馭繁，溯因而測變；即自然現象，便於實驗，易加控制，可使因子減少錯綜，分析變為單純，然觀察或生偏誤，測量或未盡美；例如理化實驗，其波動氣壓，寒燠晴雨，稍有變異，輒生影響，欲求推算完善，仍有賴於統計。故統計方法，為研究社會科學與自然科學者之必要工具。

✓統計學之定義 統計學為討論如何綜合，分析，及比較數字材料之學。

✓統計方法之特性

1. 統計方法 為處理大量數字材料之唯一方法。
2. 統計方法 祇適用於可以化成數量之材料。
3. 統計方法 為客觀者，但結果難免受主觀解釋之影響。
4. 統計方法 適用於社會科學，同時適用於自然科學。

✓統計技術之要素

1. 材料之搜羅與綜合。
2. 材料之分類與縮短。

3. 材料之分析與比較。

4. 材料之報告，用(1)文字式，(2)表列式，(3)圖示式。

✓ 統計之大別

1. 品質統計 以一種品質為標準，具此品質者為一類，不具者又為一類，兩者截為二物，各不相關，如‘有’與‘無’，‘是’與‘非’，‘智’與‘愚’，‘善’與‘惡’，‘明’與‘瞽’，‘聰’與‘聵’等是。

2. 變數統計 舉一種品質，而為之不斷分級，由少而多，或由淺而深，如將溫度或智慧，分為無窮等級是。

II. 統計數列

將統計材料，為之依序排列，即為統計數列，或統計分配。依材料數字之大小排列，所得者為次數數列；依材料發現時間之先後排列，所得者為時間數列；依材料所在位置之次序排列，所得者為空間數列。

次數數列，又可分為下列二種：

連續數列 材料數字，有無窮之值者，為連續數列，如人之身高，體重，智慧，其測量之數字，均可至小極微，而使之連續不斷。

不連續數列 材料數字，均為自然整數，其單位不可細分者，為不連續數列，如車輛或房間之多少，不能有小數數字之表示。

全距 最大量數與最小量數間之距離爲全距。如有八十個分數，最高者爲 80，最低者爲 50，則 $80 - 50 = 30$ ，即爲全距。

組距 組距爲每組上限與下限之距離，選定組距時須注意以下各點：

1. 決定組距之大小時，須能使分配中之特點顯出。
2. 組距不可太大，大則分類必近於粗率，不可太小，小則不易於處理。普徧言之，組距之數目，應在十五組與三十組之間，但不可少於六組。

3. 各組應爲等距，以便可用圖示表出數列；又在計算算術平均數時，以一組爲一單位，較爲簡捷；且與其他分配，亦可比較。

4. 每組應有上下組限，凡‘以上’‘不滿’等式樣，如‘不滿 50’‘85 以上’，均不宜採用。

組限 組限爲每組上限與下限之數字表示，爲便利計算與表列，組限應爲整數。組距中點最好亦應使之爲整數。欲達此二種目的，應使組之單位數爲奇數。五個單位，或其倍數，成爲一組，最爲理想。組限不可含糊，使數字不易歸類。如數量爲不連續數列，則組限不難決定，例如工人分爲 1—100, 101—200, 201—300，則下限爲 1, 101, 201，上限爲 100, 200, 300。如量數爲連續數列，則有以下分類方式：

A. 25 與不滿 30, B. 25—29, C. 25—29.99,

30 與不滿 35,	30—34,	30—34.99
35 與不滿 40,	35—39,	35—39.99,
40 與不滿 45,	40—44,	40—44.99,
45 與不滿 50.	45—49.	45—49.99.

D. 27.50,	E. 25—30,
32.50,	30—35,
37.50,	35—40,
42.50,	40—45,
47.50.	45—50.

以上各法, A 種用字太多, 書寫費時. B 種組限含糊, 如有一數字為 29.87, 雖應歸入 25—29 一組, 而實際上已超過 25—29 一組上限, 且不達到 30—34 一組之下限, 況數列無連續性, 不使用圖表示. C 種意義極明, 歸類亦不致錯誤, 然各組間仍乏連續性, 不使用圖表示. D 種以組中點表示一組, 未能顯出下限與上限, 意義不易明瞭. E 種顯示各組間之連續性, 可用圖表示. 惟須注意者, 即每組之上限含有‘將達到而不包括’之意. 例如 25—30 一組, 意即任何一數, 自 25 起直至 29.99 ……均應歸入此組, 一滿 30, 即應歸入 30—35 之一組. 故 E 種方式, 最為合理, 而應用亦最為普遍.

應用組距之假定 將若干量數, 歸入其相當之組, 則該組各量數之值, 即以該組中點代表之. 例如設 50—55 一組之次數

爲3,意即該組共有三個分數,而每個分數,不論其原值如何,計算時悉以該組之中點 52.50 視之。故應用組距,以便利計算時,吾人應切實明瞭二種假定:

1. 每組之量數,均集中於該組之中數,故其值可 中點代表之。

2. 每組之數量,均均勻分配於該組之內,故其中點值亦足以代表之。

✓ 次數分配之種類 次數分配,可分爲四種如下:

✓ 1. 常態分配「又稱爲對稱分配,或鐘形分配。若在此分配上,求得一中點,則次數兩面開展,情形相同。例如機遇變化,人類測量,心理測量等數字,多屬於此類分配,如圖 1。

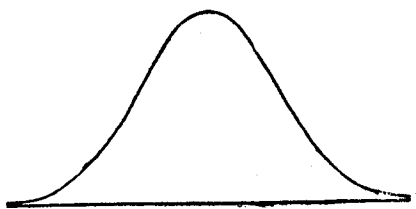


圖 1. 常態分配

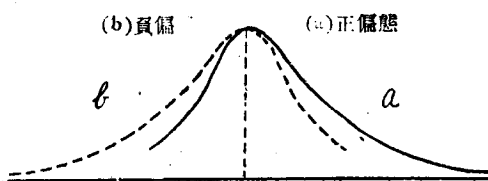


圖 2. 偏態分配

✓ 2. 偏態分配
次數分配上,其一方開展較他一方面爲多,爲偏態分配。偏態分配,又可

分爲兩種: (a) 右(正)偏態分配及 (b) 左(負)偏態分配。大多數

次數分配均屬之，如圖 2。

✓ 3. J形或極不對稱分配 其分配形狀如 J 形，故稱之。例如一國之財富，或機關職員之薪水，其分配類多似之，如圖 3。

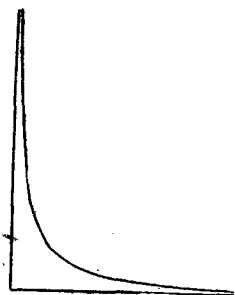


圖 3. J 形分配

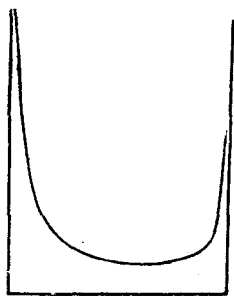


圖 4. U 形分配

✓ 4. U 形分配 其形如 U 字，故稱之。例如人口死亡率，其分配多屬此類，如圖 4。

✓ 次數分配之圖示法

✓ 1. 次數多邊圖之繪法

- A. 以橫坐標代表量表。
- B. 以縱坐標代表每組量數之次數。

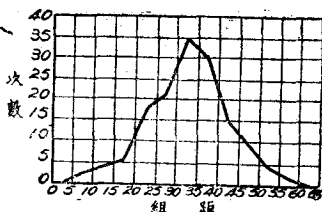


圖 5. 次數多邊圖

C. 依照次數，在各組中點中，用點標出高度。

D. 將各點連接之，惟左右最後各一點，應接到其鄰組

之組中點爲止。

✓2. 直方圖之繪法

- A. 以橫坐標代表量表。
 B. 當組距相等時，以縱坐標代表每組量數之次數。(若組距不等時，以每組直方形面積代表次數。)

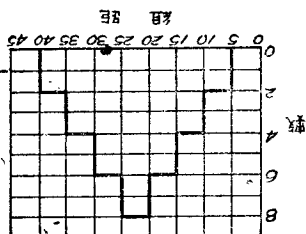


圖 6. 直方圖

- C. 以組距爲底，以每組之次數爲高，作直方形。

3. 累積次數

- A. 累積次數表之意義

組 距	次 數	累積次數 (較小)	累積次數 (較大)
104—108	7	7	200
108—112	18	25	193
112—116	27	52	175
116—120	49	101	148
120—124	52	153	99
124—128	23	176	47
128—132	16	192	24
132—136	5	197	8
136—140	3	200	3
	200		

‘較小’之意，為較相當組最高值為少，例如第三欄次數 7，表示該組最高值為 108，共有 7 個量數，其各值均較小於 108。同理在第三欄 7 與 18 之和為 25 個量數，其值均較小於 112。

‘較大’之意，為較相當組最低值為多，例如第四欄次數 3，表示該組最低值為 136，共有三個量數，其值均較大於 136。同理第四欄 3 與 5 之和為 8 個量數，其值均較大於 132。

- B. 累積次數圖之繪法
在‘較小’曲線上，將各組上限各點連接之。
在‘較大’曲線上，將各組下限各點連接之。

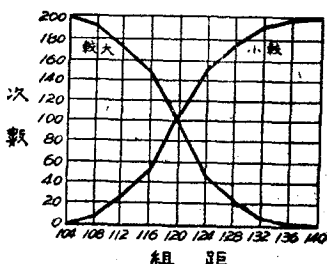


圖 7. 累積次數圖

III. 平均數

平均數之定義 平均數為一代表值，用以總括或形容一羣之材料。平均數亦為確定集中趨勢之位置之一量數。

平均數之種類 其主要者五種如下：

1. 算術平均數，
2. 中位數，

3. 衆數,
4. 幾何平均數,
5. 調和平均數.

✓ 1. 算術平均數

A. 算術平均數之定義 各數字項目相加,以其總次數除之,所得即為算術平均數.

B. 算術平均數之計算法

✓ (a) 通常法

方法一:

公式,
$$M = \frac{\Sigma(X)}{N}.$$

在以上公式內:

M = 算術平均數,

N = 次數之和,

X = 數值,

Σ = 和號(Σ 讀 sigma).

$\frac{X}{75}$

75

76

77

78

$$M = \frac{\Sigma(X)}{N}$$

$$\begin{array}{r}
 79 \\
 80 \\
 81 \\
 82 \\
 \underline{83} \\
 711
 \end{array}
 \qquad
 \begin{array}{r}
 = \frac{711}{9} \\
 \\
 = 79
 \end{array}$$

方法二：

公式
$$M = \frac{\sum(f \cdot X)}{N}$$

在以上公式內：

f = 個別數值之次數，

$\cdot$ = 乘號。

<u>X</u>	<u>f</u>	<u>f · X</u>	
75	3	225	
76	4	304	$M = \frac{\sum(f \cdot X)}{N}$
77	5	385	
78	6	468	
79	4	316	$= \frac{2347}{30}$
80	3	240	
81	2	162	$= 78.23$

82	2	164
83	1	83
	<u>30</u>	<u>2347</u>

方法三：

公式
$$M = \frac{\Sigma(f \cdot M - p)}{N}$$

在以上公式內：

$M - p =$ 組中點。

<u>X</u>	<u>M - p</u>	<u>f</u>	<u>f · M - p</u>	
30—35	32.5	2	65	
35—40	37.5	1	37.5	
40—45	42.5	10	425	$M = \frac{\Sigma(f \cdot M - p)}{N}$
45—50	47.5	17	807.5	
50—55	52.5	18	945	
55—60	57.5	19	1092.5	$= \frac{5317.5}{97}$
60—65	62.5	18	1125	
65—70	67.5	10	675	$= 54.8$
70—75	72.5	2	145	
		<u>97</u>	<u>5317.5</u>	

✓ (b) 簡捷法

方法一：

公式
$$M = A + \frac{\Sigma(d')}{N}$$

在以上公式內：

A = 假定平均數，

d' = 個別數值與假定平均數之差。

X	d'	
75	-3	
76	-2	$M = A + \frac{\Sigma(d')}{N}$
77	-1/-6	
78	0	$= 78 + \frac{15-6}{9}$
79	1	
80	2	$= 78 + 1$
81	3	$= 79.$
82	4	
83	5/15	

方法二：

公式
$$M = A + \frac{\Sigma(f \cdot d')}{N}$$

X	f	d'	$f \cdot d'$
75	3	-2	-6
76	4	-1	-4/-10