統計學概要

朱君毅 編著

正中書局印行

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,亦隨之猋趨邁進,訓練統計專才,政府視爲急務;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,冀爲訓練之教材,期作初學之讀物,倘祈國內專家,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,細爲繪製,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

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，分析變爲單純，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，爲研究社會科學與自然科學者之必要工具．

統計學之定義　統計學爲討論如何綜合，分析，及比較數字材料之學．

統計方法之特性

1.統計方法　爲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　爲客觀者，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**　以一種品質爲標準,具此品質者爲一類,不具者又爲一類,兩者截爲二物,各不相關,如'有'與'無','是'與'非','智'與'愚','善'與'惡','明'與'瞽','聰'與'聾'等是.

2. **變數統計**　舉一種品質,而爲之不斷分級,由少而多,或由淺而深,如將溫度或智慧,分爲無窮等級是.

II. 統計數列

將統計材料,爲之依序排列,即爲統計數列,或統計分配.依材料數字之大小排列,所得者爲次數數列;依材料發現時間之先後排列,所得者爲時間數列;依材料所在位置之次序排列,所得者爲空間數列.

次數數列,又可分爲下列二種:

連續數列　材料數字,有無窮之値者,爲連續數列,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,均爲自然整數,其單位不可細分者,爲不連續數列,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, 卽爲全距.

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,如有一數字爲 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,不便用圖表示. C 種意義極明,歸類亦不致錯誤,然各組間仍乏連續性,不便用圖表示. 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,最爲合理,而應用亦最爲普徧.

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,不論其原値如何,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,故其値可以中點代表之.

2.　每組之數量,均均勻分配於該組之內,故其中點値亦足以代表之.

(三) **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.

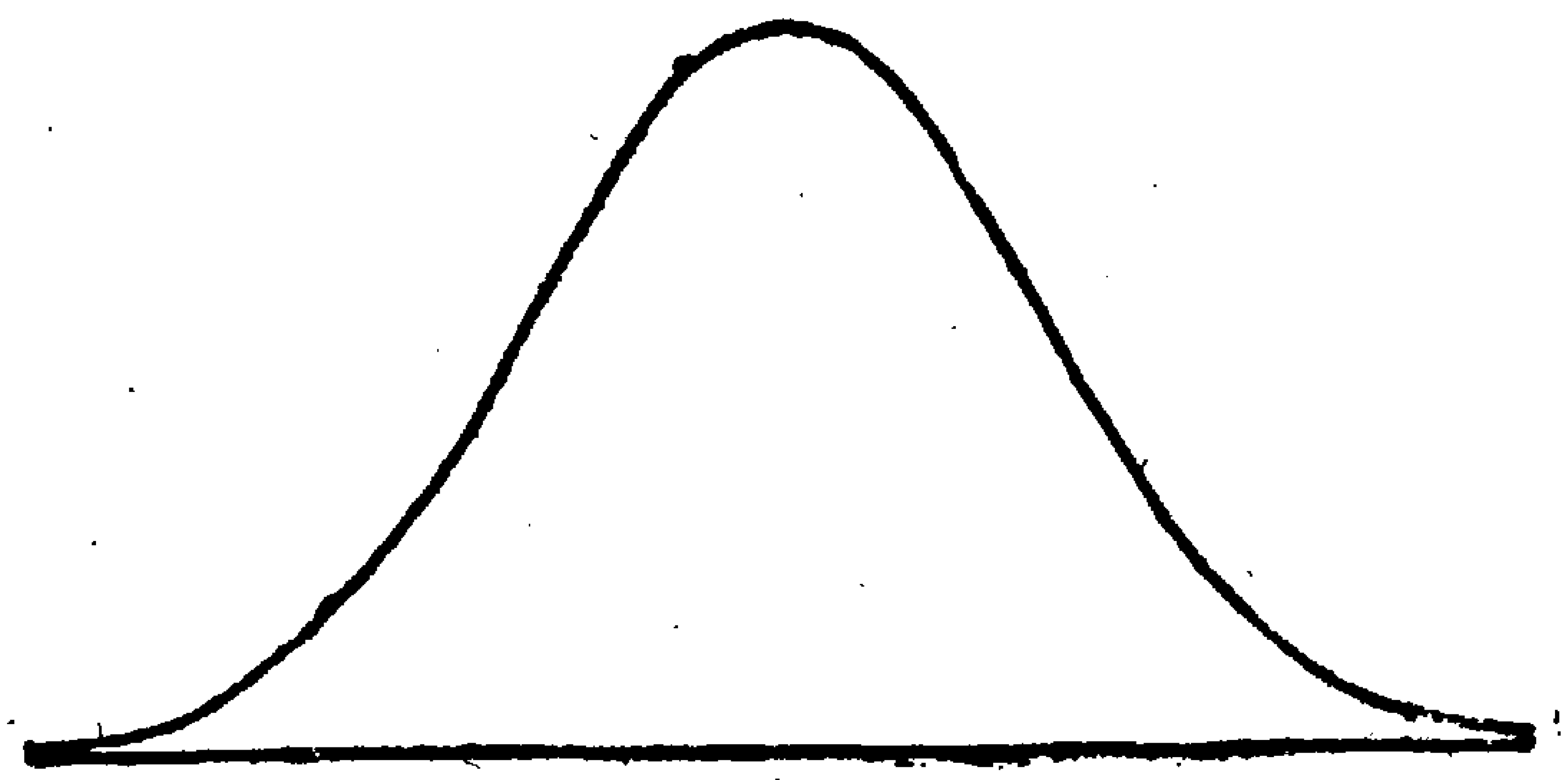


圖 1.　常態分配

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

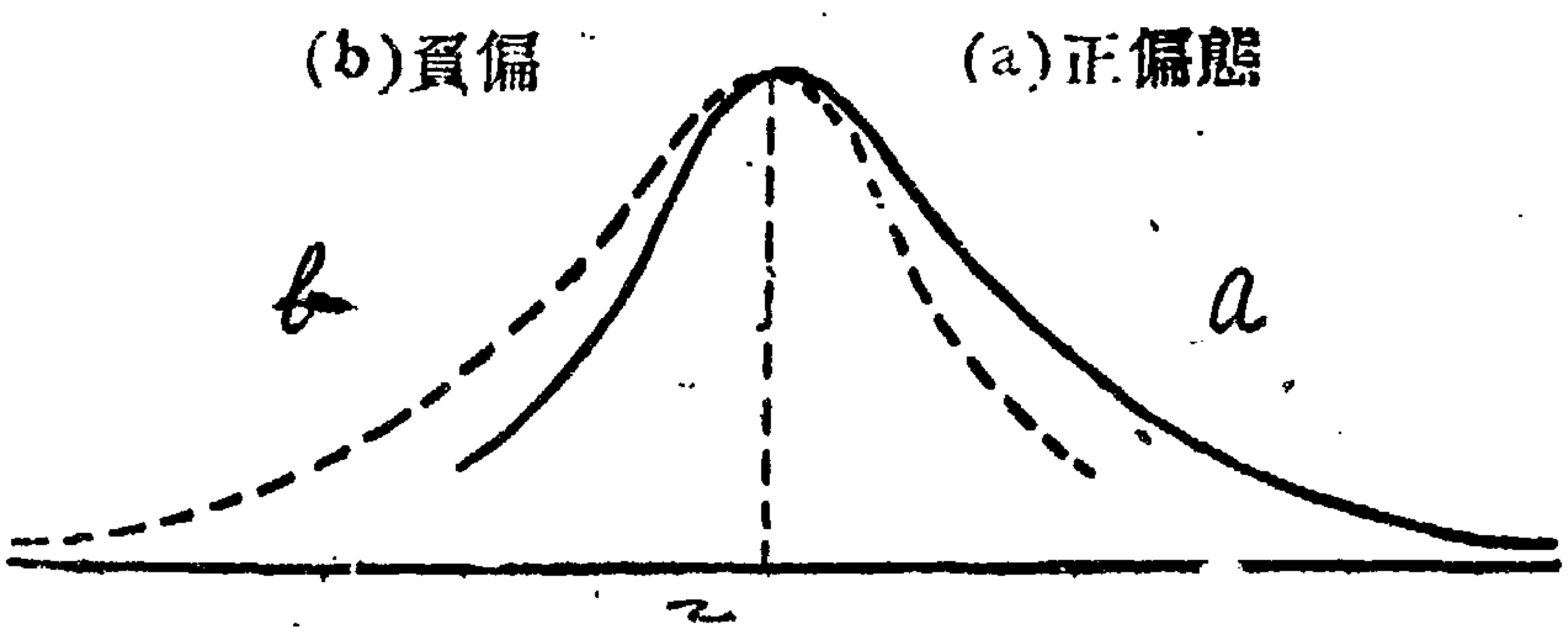


圖 2.　偏態分配

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.

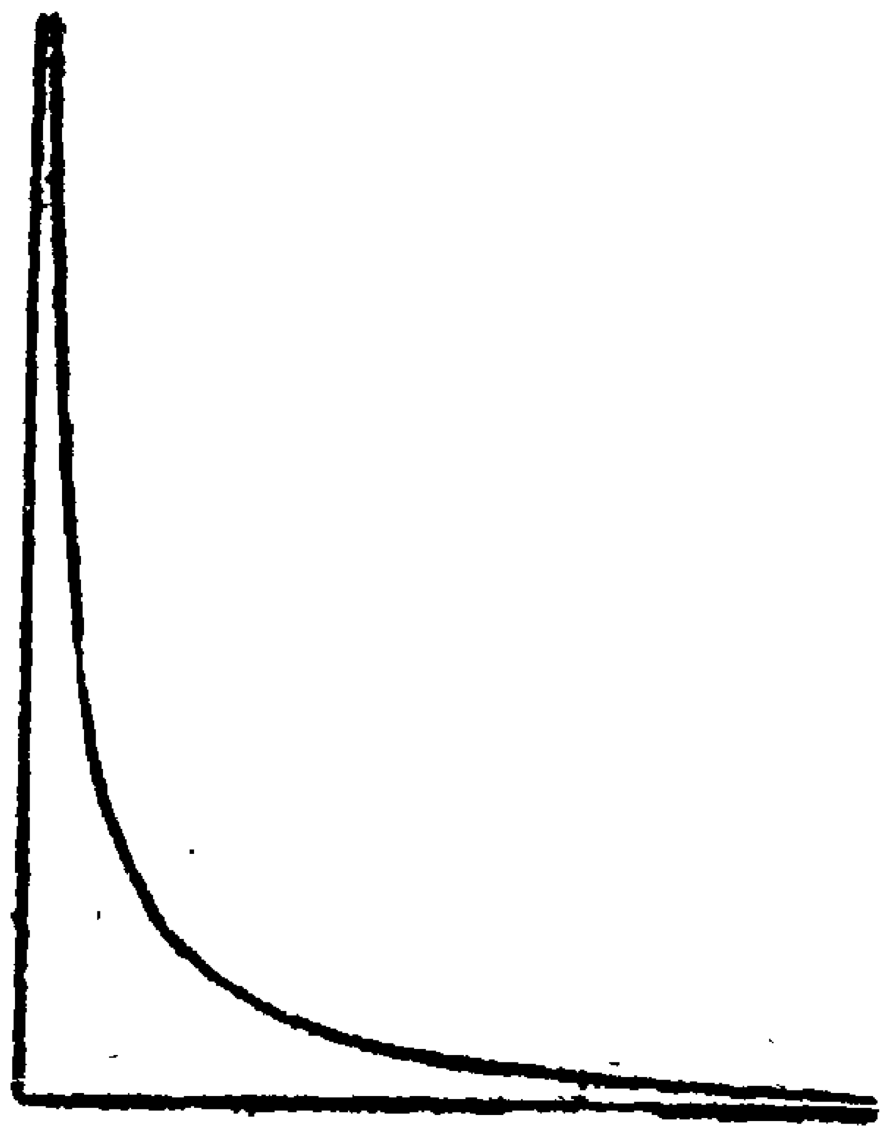


圖 3. J 形分配

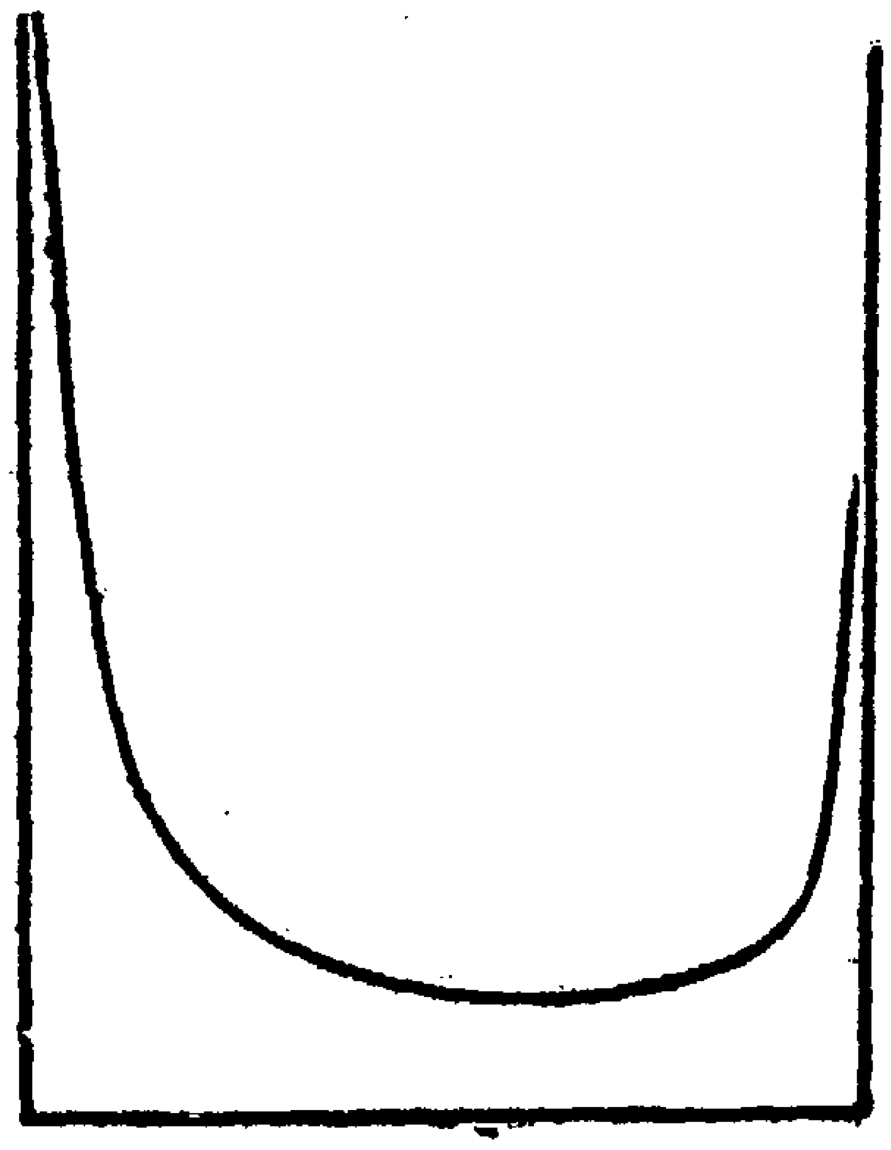


圖 4. U 形分配

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.

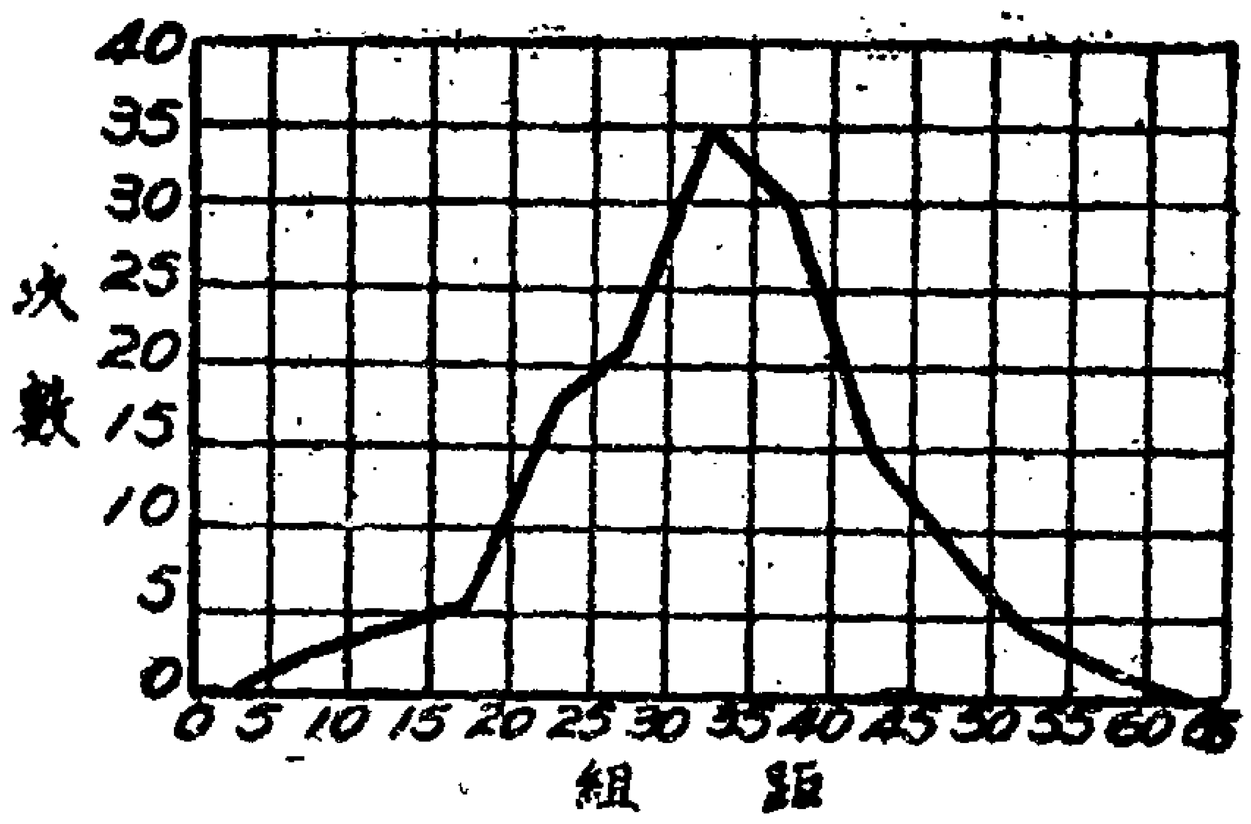


圖 5. 次數多邊圖

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.

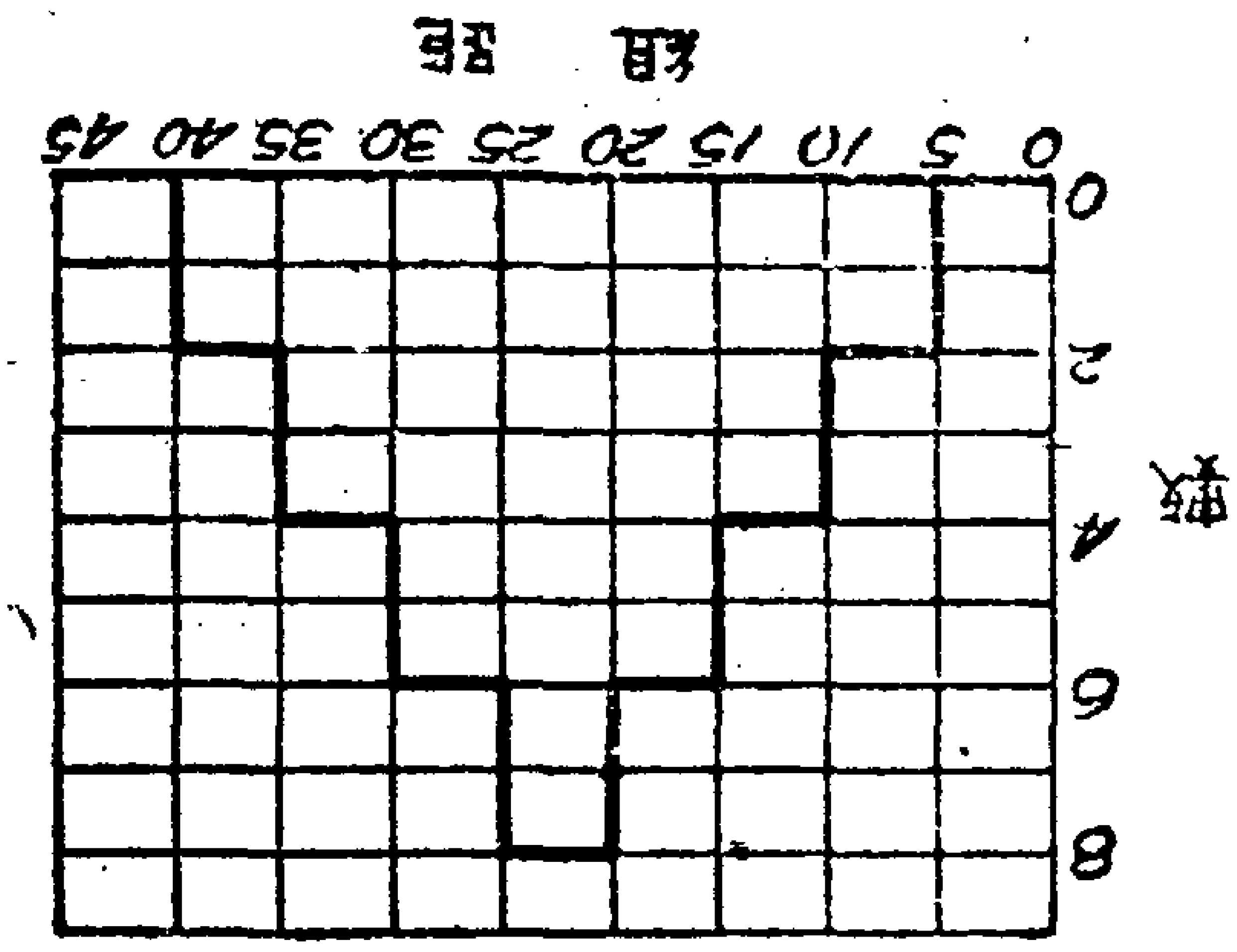


圖 6. 直方圖

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.)

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		

比相當組最高之值小

'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.

大

'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.

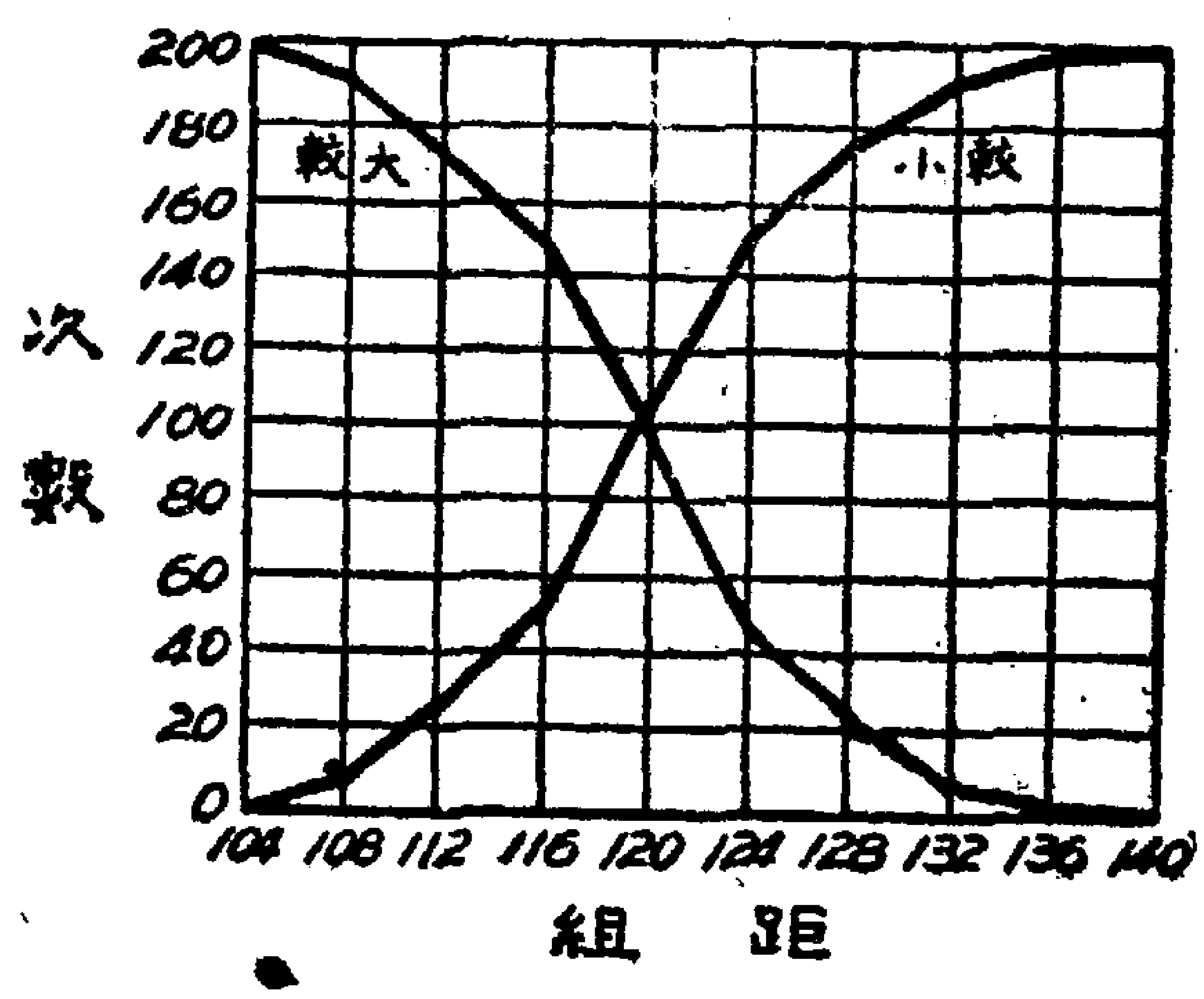


圖 7. 累積次數圖

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.

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,所得卽爲算術平均數.

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).

X
75
76
77
78

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

$$=\frac{711}{9}$$

79

80

81

82

83

711

$$=79$$

方法二:

公式 $$M=\frac{\Sigma(f\cdot X)}{N\ \Sigma f}.$$

在以上公式內:

f=個別數值之次數,

$\cdot$ =乘號.

X	f	$f\cdot X$
75	3	225
76	4	304
77	5	385
78	6	468
79	4	316
80	3	240
81	2	162

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

$$=\frac{2347}{30}$$

$$=78.23.$$

82	2	164
83	1	83
	30	2347

方法三:

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

在以上公式內:

$$M-p=\text{組中點}.$$

X	$M-p$	f	$f\cdot M-p.$
30—35	32.5	2	65
35—40	37.5	1	37.5
40—45	42.5	10	425
45—50	47.5	17	807.5
50—55	52.5	18	945
55—60	57.5	19	1092.5
60—65	62.5	18	1125
65—70	67.5	10	675
70—75	72.5	2	145
		97	5317.5

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

$$=\frac{5317.5}{97}$$

$$=54.8$$

(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
77	$-1/-6$
78	0
79	1
80	2
81	3
82	4
83	5/15

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

$$=78+\frac{15-6}{9}$$

$$=78+1$$

$$=79.$$

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	$-4/-10$

77	5	0	0
78	6	1	6
79	4	2	8
80	3	3	9
81	2	4	8
82	2	5	10
83	1	6	6/47

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

$$= 77 + \frac{47 - 10}{30}$$

$$= 77 + 1.23$$

$$= 78.23.$$

方法三:

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

在此公式內:

i＝組距之大小.

X	f	d'	f·d'
5—10	2	−4	− 8
10—15	4	−3	−12

15—20	5	−2	−10
20—25	6	−1	−6/−36
25—30	7	0	0
30—35	6	1	6
35—40	4	2	8
40—45	3	3	9
45—50	1	4	4/27
	38		

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

$$M=27.5+\frac{27-36}{38}\times 5$$

$$=27.5-1.184=26.316.$$

C. 算術平均數之性質

(1) 算術平均數之值,係從全分配之各個項目而決定.此爲由計算而得之平均數.

(2) 各個別數值,對於算術平均數之差數,其和爲零.

(3) 算術平均數,爲一個固定數值.

(4) 算術平均數爲一最通用之平均數.

(5) 爲最易了解.

(6) 其計算法比較簡單.

(7) 欲計算其值,祗須具有全分配之總值,與各項目之總

次數.

(8) 可用代數法處理之.例如全分配各小組之平均數,可以求得,則此平均數可再予平均,以求得到各組合併之平均數.倘各小組之次數不同,則宜求加權平均數.假定第一小組 13 數之平均數爲 10,第二小組 42 數之平均數爲 16,則 55 數之合併之平均數應爲

$$M=\frac{N_1M_1+N_2M_2}{N_1+N_2}=\frac{13(10)+42(16)}{13+42}=\frac{802}{55}=14.6.$$

(9) 其數值常受極端量數之影響,而生極大之增減,因之設分配有極端量數存在時,算術平均數不能爲一代表之數.

D. 算術平均數之應用　在以下各種情形時,應採用算術平均數:

(1) 在決定集中趨勢時,欲使每個數值,發生均等之重量.

(2) 欲求得最高之可靠性.

(3) 在分析統計資料過程中,倘 計算標準差,積差相關與可靠性之必要.

2. 中位數

A. 中位數之定義　項目照大小排列時,其中間一項之數值卽爲中位數.若項目爲偶數,則居中二數之平均,卽爲中位數.

B. 中位數之計算法

(1) 不連續數列

a. 項目爲奇數時,則中位數爲該組項目中之第 $\frac{N+1}{2}$ 個項目之數值.

例: 2, 4, 6, 8, 10 各項目之中位數爲 6.

b. 項目爲偶數時,則中位數爲該組最中二項目,即第 $\frac{N}{2}$ 及 $\frac{N}{2}+1$ 次之平均.

例: 2, 4, 6, 8, 10, 12 各項目之中位數爲 7.

(2) 連續數列

公式:
$$Md_n = l + \frac{\frac{N}{2} - N_b}{f_m} i,$$

或
$$Md_n = u - \frac{\frac{N}{2} - N_a}{f_m} i.$$

在以上二公式內:

Md_n = 中位數,

l = 含有中位數之組之下限,

u = 含有中位數之組之上限,

f_m = 含有中位數之組之次數,

i = 組距,

N = 次數之總和,

N_b = 少於中位數組各組之次數之和,

N_a = 大於中位數組各組之次數之和.

X	f	計　算
0—5	2	$\frac{N}{2}=\frac{47}{2}=23.5.$
5—10	4	次數 f 欄由上向下遞加:得
10—15	5	$N_b=2+4+5+6=17.$
15—20	6	中位數組之次數為 8,組距為 5,
20—25	8	$23.5-17=6.5;\ \frac{6.5}{8}\times 5=4.06.$
25—30	7	中位數組之下限為 20;
30—35	6	$\therefore\ Md_n=20+4.06=24.06.$
35—40	5	或由下向上遞加:得
40—45	3	$N_a=1+3+5+6+7=22;$
45—50	1	$23.5-22=1.5;\ \frac{1.5}{8}\times 5=.9375.$
	47	中位數組之上限為 25,
		$\therefore\ Md_n=25-.9375=24.06.$

若用公式代入:

$$\frac{N}{2}=23.5,\qquad N_b=2+4+5+6=17,$$

$$l=20,\qquad f_m=8,\qquad i=5.$$

$$Md_n=l+\frac{\frac{N}{2}-N_b}{f_m}i=20+\frac{\frac{47}{2}-17}{8}\times 5=24.06.$$

或

$$N_a=1+3+5+6+7=22,\qquad u=25,$$

$$Md_n=u-\frac{\frac{N}{2}-N_a}{f_m}i=25-\frac{\frac{47}{2}-22}{8}\times 5=24.06.$$

C, 中位數之性質

(1) 中位數爲一位置排列上之平均數.

(2) 中位數受項目多少之影響,而不受極端量數之影響.

例:

	X	X'	X''
	45	45	45
	46	46	46
	47	47	47
	48	48	48
	49	100	49
			100
中位數	47	47	47.5
算術平均數	47	57.2	55.8

(3) 數列中間之項目密集時,中位數最富有代表性.

(4) 中位數計算容易.

(5) 其值不爲特殊項目所搖動.

(6) 數列兩端項目爲無限制時,中位數亦可求得.

(7) 中位數不若算術平均數之爲一般人所認識.

(8) 欲求中位數,項目必依次排列.

(9) 中位數不能用代數式處理之. 例如全分配各小組之中位數,不能爲之平均而求得全分配之中位數.

D. 中位數之應用 在以下各種情形時,應採用中位數:

(1) 當有特殊量數,對於平均數發生過大之影響時.

(2) 數列中有一部分材料,不能嚴格決定其數值時.

(3) 在分析統計資料過程中,倘有計算四分位點及其他百分位數之必要時.

3. 衆數

A. 衆數之定義 衆數爲數列中發現最多之數值.若分配爲常態,則衆數爲最高縱線之値.

B. 衆數之計算法 衆數之嚴格數理決定,殊不易爲.統計學家多用各種方法求得衆數之近似數.含有衆數之組之中點値,不能一概作爲衆數之値,因組距之大小一變,衆數之値亦變.故計算衆數,通常採用以下二法:

(1) 方法一:——估計法

公式: $M_0=l+\frac{f_2}{f_2+f_1}i$, 或 $M_0=u-\frac{f_1}{f_2+f_1}i$.

在以上公式內:

M_0=衆數,

l=衆數組之下限,

u=衆數組之上限,

f_1=少於 l 之鄰組之次數,

f_2=大於 u 之鄰組之次數,

i＝組距之大小.

X	f	計 算
10—20	10	衆數組爲 30—40，以 60 爲最大次數
20—30	20	20 30
30—40	60	$M_0=30+\frac{30}{30+20}\times 10=30+6=36$,
40—50	30	20 20
50—60	10	或 $M_0=40-\frac{20}{30+20}\times 10=40-4=36$.

「若兩鄰組之次數相等，則衆數適爲衆數組之中點. 若較大鄰組之次數爲大，則衆數值有向衆數組中點以上移動之趨勢；反之其理亦同.」

(2) 方法二:——改正估計法 上法未將衆數組之次數列入公式，故衆數組之次數 $f_0=60$ 或其他數時，衆數均不變更. 此公式有欠靈敏，改正如下:

公式: $$M_0=l+\frac{\Delta_1}{\Delta_1+\Delta_2}i, \text{或 } M_0=u-\frac{\Delta_2}{\Delta_1+\Delta_2}i.$$

在以上公式內，M_0, l, u, f_1, f_2, i 等號均如前:

f_0＝衆數組之次數,

$\Delta_1=f_0-f_1$, 衆數組次數與下一組次數之差,

$\Delta_2=f_0-f_2$, 衆數組次數與上一組次數之差.

仍用上例，則

$\Delta_1=f_0-f_1=60-20=40$,

$\Delta_2=f_0-f_2=60-30=30$,

+0.7

故 $$M_0=30+\frac{40}{40+30}\times 10=30+5.7=35.7,$$

或 $$M_0=40-\frac{30}{40+30}\times 10=40-4.3=35.7.$$

(3)　方法三:——經驗法

公式: $$M_0=M-3(M-Md_n).$$

若分配爲微偏態，則根據經驗，算術平均數(M)與中位數(Md_n)間之距離，約佔算術平均數與衆數間之距離之三分之一.故引用以上公式，亦可求得衆數之近似值.玆用圖說明三種平均數之關係如下:

算術平均數，中位數，衆數之關係

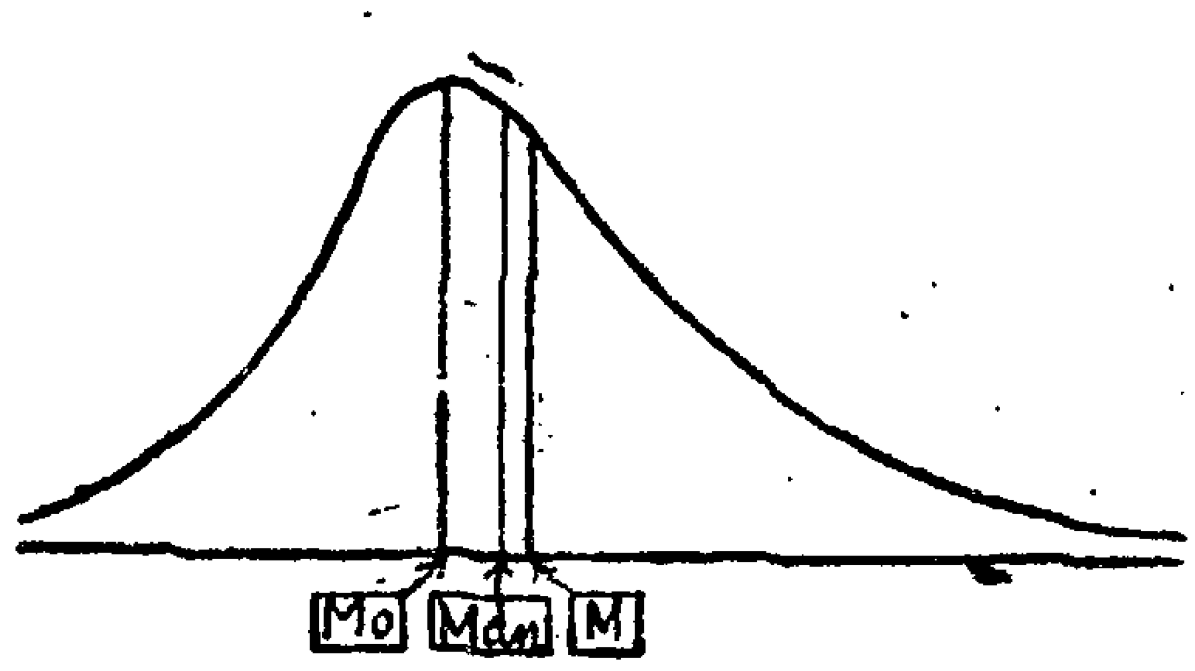


圖 8.　正偏態曲線

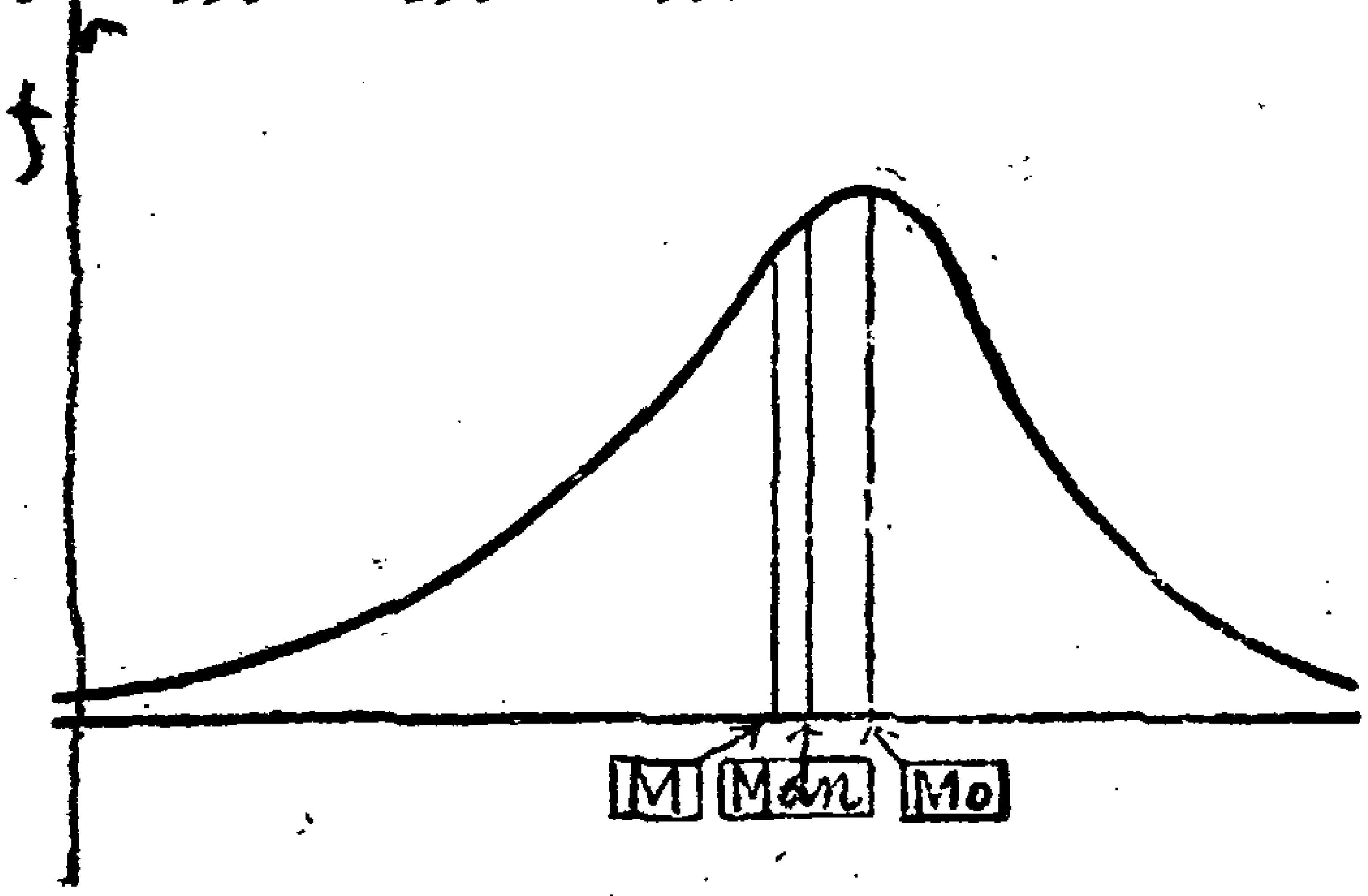


圖 9.　負偏態曲線

(4)　其他方法:　此外尚有其他方法，足以計算衆數者，如重新歸類法，修勻次數分配法，移動平均數法，數理曲線等.

C. 衆數之性質

(1) 衆數爲各平均數中之最富有代表性者.

(2) 衆數值不受極端量數之影響.

(3) 衆數爲一位置排列上之平均數.

(4) 項目少時,最易用觀察法指定之.

(5) 若無大量量數,衆數之意義極微.

(6) 數量少時,有時並無衆數存在;因其中各量數,均無重複者.

D. 衆數之應用 在以下各種情形時,應採用衆數:

(1) 比較數個時期之工資:若第二期之平均工資較第一期增加時,算術平均數不能指出此種增加,是由於低級工資之全體提高,抑由於幾個高級工資之提高;工資減少時,其困難亦同.衆數可以指出大部分工資之變動.

(2) 研究人口統計:結婚或死亡之最慣常年齡(衆數),實比平均年齡爲較有意義.且研究死亡常有二種衆數,一爲嬰兒死亡,一爲老年死亡,若用算術平均數,則不能表出.

(3) 欲速求集中趨勢之大概,可用衆數,以其計算較易,且有時可由觀察而得.

4. 幾何平均數

A. 幾何平均數之定義 N量數之幾何平均數,爲N數量乘積之N次根.

B. 幾何平均數之計算法

公式: $G=\sqrt{X_1\cdot X_2\cdot X_3\cdots\cdots X_N}$.

例: $G=\sqrt[3]{4\times 8\times 16}=8$.

但計算幾何平均數，普通應用對數，以節省時力，其公式如下:

$$\log G=\frac{\log X_1+\log X_2+\log X_3+\cdots\cdots+\log X_N}{N}.$$

例: 求以下各數之幾何平均數: 784, 563, 649, 283, 764.

$$\begin{aligned}
\log 784&=2.8943\\
\log 563&=2.7505\\
\log 649&=2.8123\\
\log 283&=2.4518\\
\log 764&=2.8831\\
\hline
&13.7920(5\\
\hline
\log G&=2.7584 \qquad \therefore\ G=573.2.
\end{aligned}$$

C. 幾何平均數之性質

(1) 幾何平均數爲一由計算而得之値，且受各量數大小之影響.

(2) 倘數列中有一量數爲零，則幾何平均數必爲零.

(3) 倘數列中有一量數爲負數，則幾何平均數不能決定.

(4) 幾何平均數可用代數方法處理之.

(5) 幾何平均數不爲一般人所認識.

(6) 幾何平均數計算不易.

D. 幾何平均數之應用 在以下各種情形時,應採用幾何平均數:

(1) 計算指數.

(2) 計算人口之增加率及二個時期間之人口數.

(3) 計算複利.

(4) 計算進步率.

5. 調和平均數

A. 調和平均數之定義 調和平均數, 又稱倒數平均數, 爲各數值倒數之平均數之倒數.

B. 調和平均數之計算法 計算調和平均數之方法, 可用以下公式說明之:

$$H=\frac{1}{\frac{1}{N}\Sigma\frac{1}{X}}.$$

在以上公式內, H 爲調和平均數之符號.

C. 調和平均數之應用

例 1: 以工作爲例, 設某人步行上山時之速率爲每小時十二公里, 下山時步行每小時之速率爲八公里, 則某人徒返步行, 每小時之平均速率爲多少公里?

$$H=\frac{1}{\frac{1}{2}\left(\frac{1}{8}+\frac{1}{12}\right)}=9.6\text{公里.}$$

工作速度，可用兩種形式表示，一爲每單位時間之工作分量；一爲每單位工作所費之時間，二者互爲倒數．例1，每小時步行里數，係以第一種形式表示．在例1中，欲決定應用算術平均數，抑應用調和平均數，應先決定何者爲不變之數，(1)工作之分量，抑(2)時間之長短．若工作之分量不變，則應用調和平均數；若時間之長短不變，則應用算術平均數．例1所示之工作分量，卽某段山路之距離爲不變，應採用調和平均數．玆將一般規則列表如下：

速度形式	不變因子	
	工作分量(d)	時間(t)
每單位時間之工作分量(d/t)	(1) 調和平均數	(2) 算術平均數
每單位工作之時間長短(t/d)	(3) 算術平均數	(4) 調和平均數

例1：　以物價爲例，貨物之時價亦有兩種形式．一爲每單位物品所值貨幣數量，一爲每單位貨幣所購物品多少．二者互爲逆數．設鷄蛋之時價，爲甲種大洋每元四隻，乙種每元五隻，丙種每元二十隻，若每種各購買100隻，則計算平均價格，應用調和平均數，卽每元大洋，所購得之鷄蛋爲；

$$H=\frac{1}{\frac{1}{3}\left(\frac{1}{4}+\frac{1}{5}+\frac{1}{20}\right)}=6\text{個}.$$

例 2 係以平均貨幣購得物品多少形式表示，此時欲決定應用算術平均數，抑應用調和平均數，應先決定不變之因子爲：(1)所費之貨幣，抑(2)所購物品之量．若所費之貨幣不變，而所購之物品變更，則應用算術平均數；若物品之量不變，而所費之貨幣變更，則應用調和平均數．例 2 所購之物品各100隻爲不變，故應採用調和平均數．今將一般規則列表如下：

時價單位之形式	不變因子	
	所費之貨幣(M)	所購物品之量(C)
每單位物品所值貨幣數量(M/C)	(1) 調和平均數	(2) 算術平均數
每單位貨幣所購得之物品數量(C/M)	(3) 算術平均數	(4) 調和平均數

D. 調和平均數之性質

(1) 調和平均數用於時間速率之平均及物價之平均上．

(2) 調和平均數，可用代數公式計算．

(3) 一個數列之調和平均數，較其幾何平均數爲小；其幾何平均數又較其算術平均數爲小．

IV. 離差

離差之定義　有若干數量，若祇知各量數之代表數（即平均數），而不知各量數之差異之程度，則意義甚少．因各量數對於代表數之差異若大，則此代表數即失去其代表之性質而為用極微．故差異程度之測量，或離差之計算，實甚重要．

因之比較兩種或多種分配，不特須知其代表數，更要知其離差，以觀其同異之點．兩種分配，可有以下三種之不同情形：

1．集中數（即平均數或代表數）不同，而離中程度（即離差）相同，如圖 A.

2．集中點相同，而離中程度不同，如圖 B.

3．集中點與離中程度均不相同，如圖 C.

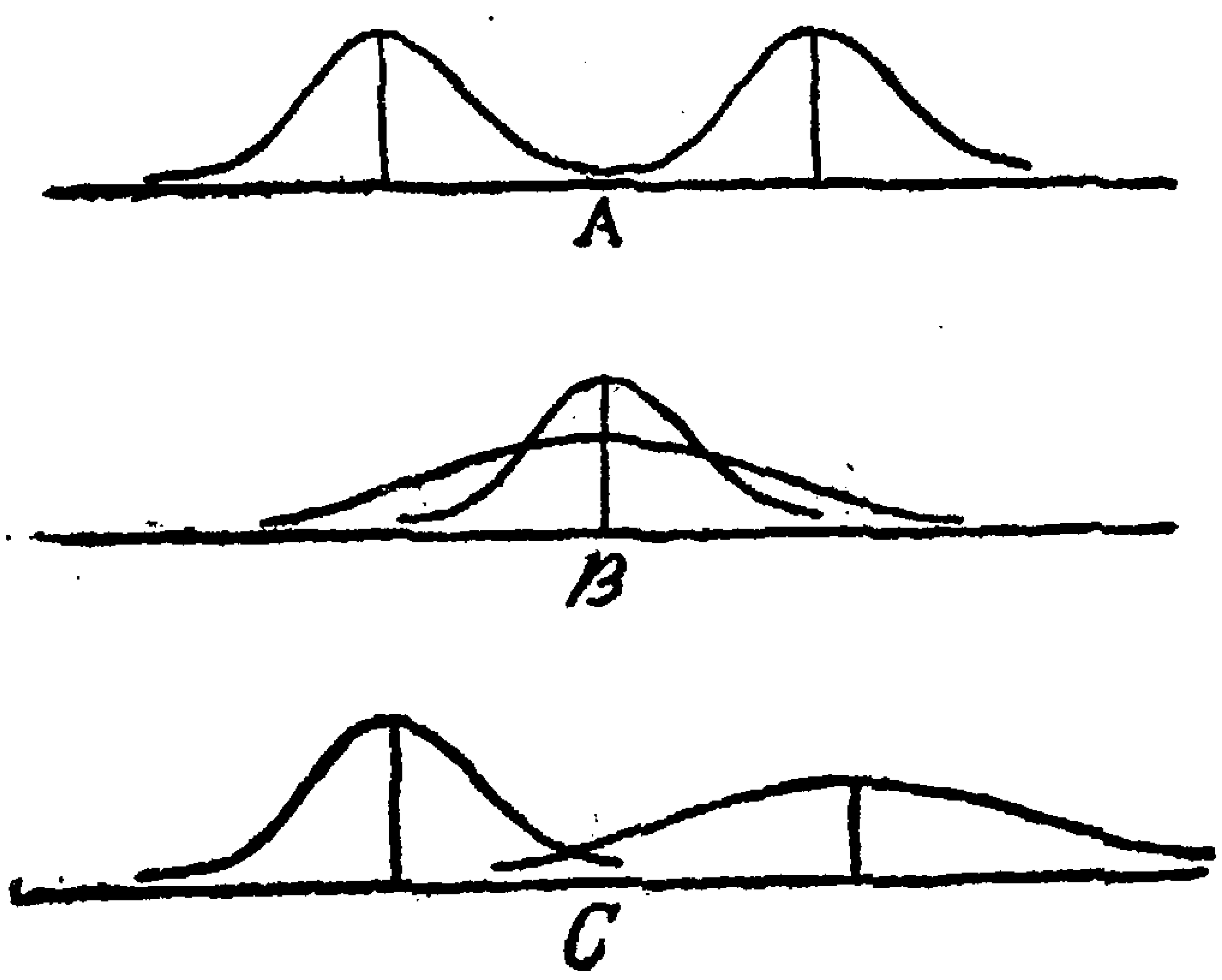


圖 10．離差圖

離差之種類

1．全距（又稱兩極距），

2．四分位差，

3．標準差，

4．平均差，

5．相對離差，

6．偏斜度．

1．全距

A. 全距之定義 全距爲數列中最大數與最小數之差.例如一數列中之最大數爲 85,最小數爲 15,則全距爲 70.

B. 全距之短處 計算離差時,普通不用全距之值,因其有以下二種缺點:

(1) 數列中若有極端量數,足使全距重受影響而大變其值.舉例如下:

X	f	
25	1	此數列之全距爲 30－25＝5.
26	2	
27	3	
28	4	
29	3	
30	1	
⋮	⋮	
50	1	若加一極端量數50,則全距爲50－25＝25.

(2) 全距之值,不能顯示分配之形狀;即不能顯示次數究爲密集於中心,抑大致均勻分配於數列上.舉例如下:

X	f_1	f_2
25	4	1
26	4	1
27	5	2
28	5	18

29	5	10
30	5	2
31	4	1
32	4	1
	36	36

在f_1項下，次數大致相同而均分配於數列上；在f_2項下，各次數均密集於 28 及 29 二個中間數量上．二分配形完全不同，但此數列之最大數值與最小數值則仍相同；若求全距之值，顯然不能表出此二分配之不同．

2． **四分位差**

A． **四分位差之定義** 四分位差爲第一四分位點與第三四分位點相差數之一半．

若將全部量數分爲相等四部，則第一部分終了時，爲第一四分位點，第二部分終了時，爲第二四分位點或中位數，第三部分終了時，爲第三四分位點．

B． **四分位差之計算法**

公式：
$$Q=\frac{Q_3-Q_1}{2}.$$

在以上公式內：

Q＝四分位差，

Q_3＝第三四分位點，

Q_1＝第一四分位點.

X	f	計算
0—5	4	$\frac{N}{4}=16.5$
5—10	5	Q_1組以下各組次數之和,$N_b=4+5+7=16$.
10—15	7	$\frac{N}{4}-N_b=16.5-16=0.5$,
15—20	9	Q_1組之下限 $l=15$; Q_1組之次數,$fq_1=9$,
20—25	10	
25—30	12	$Q_1=l+\frac{\frac{N}{4}-N_b}{fq_1}\times i=15+\frac{0.5}{9}\times 5=15.28$.
30—35	8	Q_3組以上各組次數之和,$N_a=1+2+3+5=11$.
35—40	5	$\frac{N}{4}-N_a=16.5-11=5.5$,
40—45	3	Q_3組之上限,$u=35$; Q_3組之次數,$fq_3=8$,
45—50	2	$Q_3=u-\frac{\frac{N}{4}-N_a}{fq_3}\times i=35-\frac{5.5}{8}\times 5=31.56$,
50—55	1/66	$Q=\frac{Q_3-Q_1}{2}=\frac{31.56-15.28}{2}=8.14$.

C. **四分位差之意義**

(1) $\pm Q$ 包括全分配量數之一半或 50%.

(2) 若分配爲常態時,則第一四分位點及第三四分位點,各與中位數之距離相等.

(3) Q 爲橫坐標上之一距離,用以表示 Q_1 與 Q_3 間各量數與中位數相差之程度.若分配爲常態,則在一個正四分位差($+Q$)或一個負四分位差($-Q$),與中位數之間,任何量數與中位數之差,不超出四分位差數之外.

3. 標準差

A. 標準差之定義　標準差爲差數平方之平均數之平方根.

B. 標準差之計算法

(1) 通常法:

公式: $\sigma=\sqrt{\frac{\Sigma x^2}{N}}$; 或 $\sigma=\sqrt{\frac{\Sigma fx^2}{N}}$.

在以上公式內:

σ=標準差(σ 讀 sigma, 亦可寫爲 S.D.)

x=量數與算術平均數之差,

N=次數之總和.

X	x	x^2	計　算
35	−5	25	
36	−4	16	
37	−3	9	$\sigma=\sqrt{\frac{\Sigma x^2}{N}}$
38	−2	4	
39	−1	1	
40	0		$=\sqrt{\frac{110}{11}}$
41	1	1	
42	2	4	$=\sqrt{10}$
43	3	9	
44	4	16	$=3.16$.
45	5	25	
$M=40$		110	

(2) 簡捷法:

公式: $\sigma=\sqrt{\frac{\Sigma fd'^2}{N}-c^2}$; 或 $\sigma=\sqrt{\frac{\Sigma fd'^2}{N}-c^2}\times i$.

在後一公式內:

$d'=\frac{X-A}{i}$ 卽各量數與假定平均數之差,以組距爲單位.

$c=\frac{\Sigma fd'}{N}=\frac{M-A}{i}$ 卽校正數,以組距爲單位.

若組距爲 1,則得前一公式.

X	f	d'	fd'	$f\cdot d^2$
30—35	2	−4	−8	32
35—40	3	−3	−9	27
40—45	5	−2	−10	20
45—50	7	−1	−7	7
50—55	8	0	−34	
55—60	6	1	6	6
60—65	5	2	10	20
65—70	3	3	9	27
70—75	2	4	8	32
75—80	2	5	10	50
80—85	1	6	6	36
	44		49	257

$\Sigma f\cdot d^2=257$, $c^2=\left[\frac{49-34}{44}\right]^2=0.34^2=0.1156$,

$$\therefore\quad \sigma=\sqrt{\frac{257}{44}-0.1156}\times5=11.96.$$

C.　標準差之意義

(1)　若分配爲常態，在一個標準差距離，與曲線下之面積，包括全面積之34.13%；在正負兩個標準差($\pm\sigma$)距離與曲線下之面積，包括全面積之68.26%. 玆列表如下：

$1\sigma=34.13\%$,　　$\pm1\sigma=68.26\%$,

$2\sigma=47.73\%$,　　$\pm2\sigma=95.46\%$,

$3\sigma=49.86\%$,　　$\pm3\sigma=99.72\%$,

$4\sigma=49.999\%$,　　$\pm4\sigma=99.998\%$.

玆再用圖表出之：

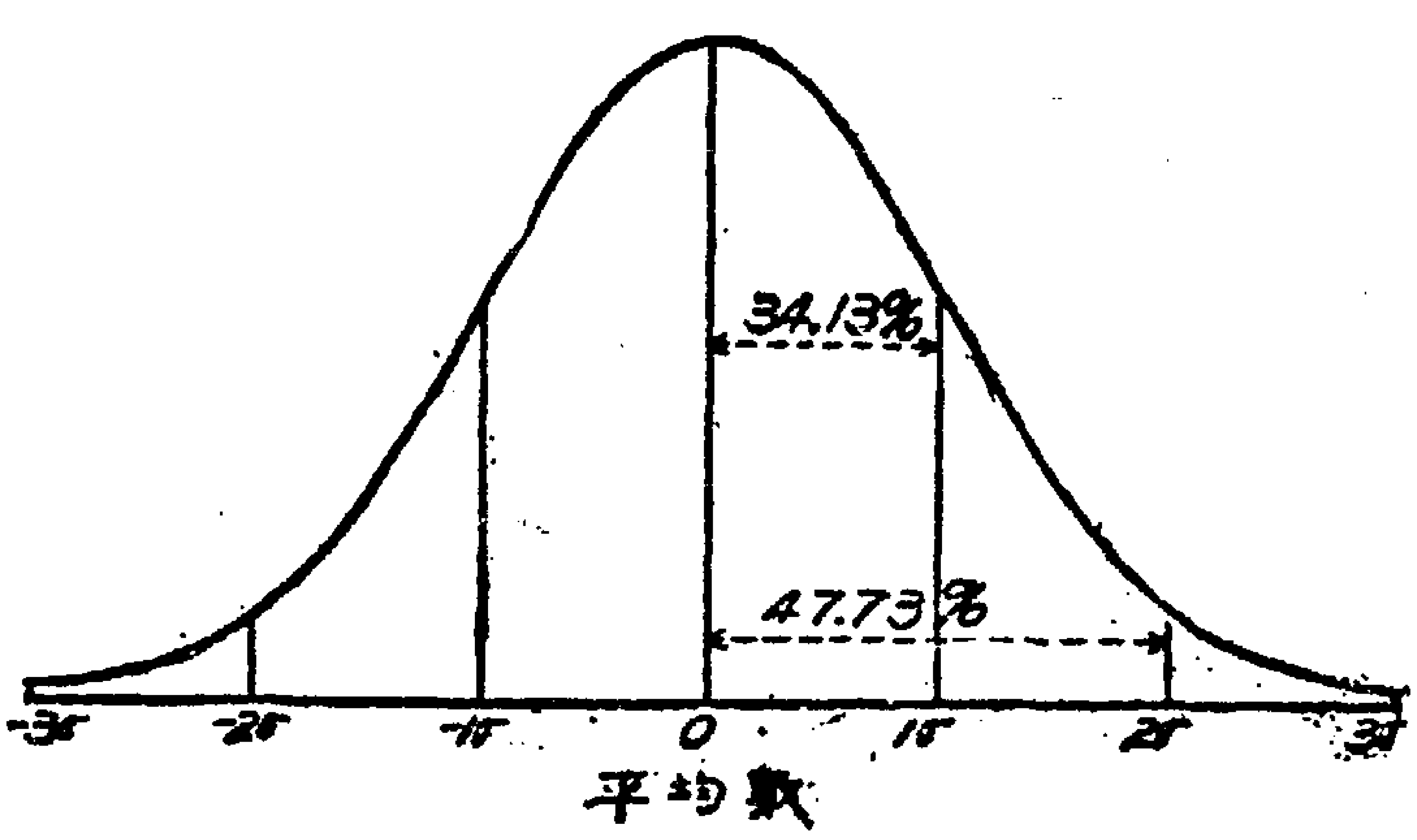


圖 11.　常態分配下之標準差

(2)　若分配爲常態，共含有百分之68.26量數中，任何量數與平均數之差，不超出標準差之數值.

(3)　六乘標準差，包括橫坐標與線下全面積百分之99.72.

4.　平均差　　2×3σ

A. 平均差之定義 平均差為各數值對於算術平均數或中位數之差數(各差數不計符號,即為正號)之算術平均數.

B. 平均差之計算法

(1) 通常法(以不連續數列為例,並以算術平均數為平均數)

公式: $$A.D.=\frac{\Sigma|X|}{N}.$$

在以上公式內:

A.D.=平均差.

$|X|$=各數值對於眞正算術平均數之離差,不計符號.

例:

X	$\|X\|$ = X − M
3	6
5	4
7	2
9	0
11	2
13	4
15	6
$M=9$	$\Sigma\|X\|=24$

$\Sigma X = 63$

$M = \frac{\Sigma X}{N} = \frac{63}{7} = 9$

$$\therefore \quad \text{A.D.}=\frac{\Sigma|X|}{N}=\frac{24}{7}=3.4.$$

(2) 簡捷法(以連續數列爲例,並以中位數爲平均數).

a. 簡捷法一(假設各組次數集中於組中點)

公式: $$\text{A.D.}=\frac{\Sigma f|d'|+(N_b-N_a)c+f_m|c|}{N}\times i.$$

在以上公式內:

A＝中位數組之中點(稱爲假定中位數),

$d'=\dfrac{x-A}{i}$, 各組中點與假定平均數之差, 以組距爲單位,

$c=\dfrac{Md_n-A}{i}$, 眞正中位數與假定中位數之差(卽校正數), 以組距爲單位,

$\Sigma f|d'|$＝各數值對於假定中位數之差之和, 而不計符號,

$|c|$＝校正數之絕對值(卽不計符號),

N_b＝小於 l (中位數組之下限)各組次數之和,

N_a＝大於 u (中位數組之上限)各組次數之和,

f_m＝中位數組之次數.

例:

X	組中點	f	$\lvert d' \rvert$	$f\lvert d'\rvert$
104—108	106	7	3	21
108—112	110	18	2	66
112—116	114	27	1	27
116—120	118	49	0	
120—124	122	52	1	52
124—128	126	23	2	46
128—132	130	16	3	48
132—136	134	5	4	20
136—140	138	3	5	15
		200		265

$N_b=7+18+27=52$,　　　　　眞正中位數$=119.92$,

$N_a=3+5+16+26+52=99$,　假定中位數$=118$,

$f_m=52$,

$$c=\frac{119.92-118}{4}=0.48.$$

$$\therefore\quad \text{A.D.}=\frac{\Sigma f|d'|+(N_b-N_a)c+f_m|c|}{N}\times i$$

$$=\frac{265+0.48(52-99)+52(0.48)}{200}\times 4$$

$$=1.33\times 4=5.32.$$

b. 簡捷法二(假設各組次數平均分配於組內)

公式:　$$\text{A.D.}=\frac{\Sigma f|d'|+(N_b-N_a)c+f_m\left(c^2+\frac{1}{4}\right)}{N}\times i.$$

在以上公式:

$|d'|$, N_b, N_a, f_m A, c, 之符號與簡捷法一公式內所用之符號相同.

仍由上表之計算,

$$N_b=52, \qquad N_a=99, \qquad f_m=52,$$

$$Md=119.92, \qquad A=118, \qquad \Sigma f|d'|=265.$$

$$c=\frac{119.92-118}{2}=0.48, \qquad c^2=0.2304.$$

$$\text{A.D.}=\frac{\Sigma f|d'|+(N_b-N_a)c+f_m\left(c^2+\frac{1}{4}\right)}{N}\times i$$

$$=\frac{265+0.48(52-99)+52(0.2304+0.25)}{200}\times 4$$

$$=1.33\times 4=5.32.$$

校正數$|c|$之數值可由0至0.5,當c等於或近於0.5時,簡捷法一與簡捷法二結果集於相等,但當c近於零時,簡捷法二之結果大於簡捷法一,簡捷法二之公式亦稱校正公式.

5. 相對離差

A. 相對離差之意義　以上所述,係爲絕對離差.若兩個分配,其平均數之大小懸殊,或單位不同,則此二種離差不能比較.

(1) 平均數之大小懸殊

離差係由平均數而生，故平均數之大小，影響離差之大小. 例如鼠身之平均長爲四吋，人體之平均長爲六十四吋. 今若鼠身之離差爲半吋，而人體之離差亦爲半吋，則此二種半吋，意義不同，不能視爲相等. 因四吋之身，而有半吋之離差，其差甚大；六十四吋之體，而祇有半吋之離差，其差甚微. 又例如小學教員之平均薪水爲四十元，其離差爲十元；大學教授平均薪水二百五十元，而其離差亦爲十元，則此二種十元，亦不能視爲相同.

(2) 平均數之單位不同

量身之高用吋，量體之重用磅. 若平均身高爲六十五吋，離差爲五吋；平均體重爲一百二十磅，離差爲九磅，則此九之離差，是否較五爲大，亦不易言，因之，若有數種離差時，必須將離差以其相當平均數除之，如是，則上述二種困難，均可解決.

B. 相對離差之計算法

(1) 公式一：求標準差之離差

$$V_\sigma=\frac{\sigma}{M};\quad 或\quad V_\sigma=\frac{\sigma}{M}\times100.$$

(2) 公式二：求四分位差之離差

$$V_Q=\frac{\dfrac{Q_3-Q_1}{2}}{\dfrac{Q_3+Q_1}{2}}=\frac{Q_3-Q_1}{Q_3+Q_1};\ 或\ V_Q=\frac{Q_3-Q_1}{Q_3+Q_1}\times100.$$

在以上公式內:

V_σ＝標準差之相對離差,

V_Q＝四分位數之相對離差.

相對離差爲百分數或爲純數,與數量之單位,如斤,斗,尺,元等無關.因之,如有兩個離差,其中一以磅爲單位,一以吋爲單位,仍可將其換爲相對離差,彼此比較,以觀二者孰大孰小.

6. 偏斜度

A. **偏斜度之意義** 偏斜度爲表示次數分配離開常態(或不能對稱)之程度之一個名詞.分配完全合乎常態或完全對稱時,則算術平均數,中位數及衆數,均在一點上.分配不能合乎常態或變爲偏態時,則各種平均數彼此分離.算術平均數,以受極端量數之影響最大,其離開衆數之距離亦最大.故偏斜愈大,則算術平均數與衆數之距離愈遠.

因之,算術平均數與衆數間距離之大小,可用作偏斜度之量數.但偏斜之測量,其目的原在比較.若差數之大小懸殊,或單位不同,則比較又復發生困難.故偏斜量數,應以差數除之,以解決此困難.

B. 偏斜度之計算法

(1) 公式一: 以 M 與 M_0 間之距離爲準,

$$S_k=\frac{M-M_0}{\sigma},$$

但 $$M_0 = M - 3(M - Md_n),$$

$$S_k = \frac{M - [M - 3(M - Md_n)]}{\sigma}$$

$$= \frac{3(M - Md_n)}{\sigma}.$$

若分配爲常態，則 $M = M_0$，偏斜度爲零.

若分配爲正偏態，則 $M > M_0$，偏斜度爲正.

若分配爲負偏態，則 $M < M_0$，偏斜度爲負.

(2) 公式二: 以四分位點之位置爲準，

$$S_k = \frac{(Q_3 - Md_n) - (Md_n - Q_1)}{(Q_3 - Md_n) + (Md_n - Q_1)} = \frac{Q_1 + Q_3 - 2Md_n}{Q_3 - Q_1}.$$

若分配爲常態，則 $Q_3 - Md_n = Md_n - Q_1$，偏斜度爲零.

若分配爲正偏態，則$Q_3 - Md_n > Md_n - Q_1$，偏斜度爲正.

若分配爲負偏態，則$Q_3 - Md_n < Md_n - Q_1$，偏斜度爲負.

此與相對離差相同，所得之數係爲純數，不受所用測量單位之影響，而可彼此比較.

V. 常態曲線

常態曲線之繪法 常態曲線又稱機率誤差曲線，係基於常態分配所繪成之曲線. 繪製此種曲線圖須根據下列公式:

$$Y = Y_0 e^{-\frac{x^2}{2\sigma^2}},$$

其中

x＝任何 X 與其平均數 M 之差，

y＝圖內任何縱線，

y_0＝平均數上之縱線(即縱線之最高者)，

e＝2.71828，即自然對數之底數，

σ＝標準差＝$\sqrt{\frac{\Sigma x^2}{N}}$.

因 $y_\sigma=\frac{N}{\sigma\sqrt{2\pi}}$，　　故 $y=\frac{N}{\sigma\sqrt{2\pi}}e^{\frac{-x^2}{2\sigma^2}}$

其中 N, σ, π, e 均爲常數，故 x 變動，y 亦隨之變動．算學家且將 x 以 σ 爲單位$\left(\text{即}\frac{x}{\sigma}\right)$時，位於 x 點上之縱線與平均點上縱線之比$\left(\frac{y}{y_0}\right)$求出，另詳於附表．例如

$x=0$，$y=1.00000$(或謂之 100,000 亦可)；

$x=0.1$ 時，　$y=0.99501$；等等.

已知上述公式之應用，可按下法繪製曲線圖.

1．用精細方格，依紙之大小，分配縱橫線，務使此圖高闊適中.

2．橫坐標可分爲六段，自 -3σ 至 3σ，中點爲 0，此點上作一最高縱線(y_σ)，其長宜爲 1，或 10，或 100 等.

3．自 0 至 1σ (或 -1σ)，再分十小段，爲 $0.1\sigma, 0.2\sigma$ 等.

4.　參閱附表一，$x=0.1\sigma$，$y=0.99501$，故從 0.1σ 上作一縱線，其長$=0.99501$ 而於上端作一點；餘類推．玆將各點列表如下：

$\frac{x}{\sigma}$	$\frac{y}{y_0}$	$\frac{x}{\sigma}$	$\frac{y}{y_0}$	$\frac{x}{\sigma}$	$\frac{y}{y_0}$
0.0	100,000	1.0	60,653	2.0	13,534
0.1	99,501	1.1	54,607	2.1	11,025
0.2	98,020	1.2	48,675	2.2	08,892
0.3	95,600	1.3	42,956	2.3	07,100
0.4	92,312	1.4	37,531	2.4	05,614
0.5	88,250	1.5	32,465	2.5	04,394
0.6	83,527	1.6	27,804	2.6	03,405
0.7	78,270	1.7	23,575	2.7	02,612
0.8	72,615	1.8	19,790	2.8	01,984
0.9	66,689	1.9	16,448	2.9	01,492
				3.0	01,111

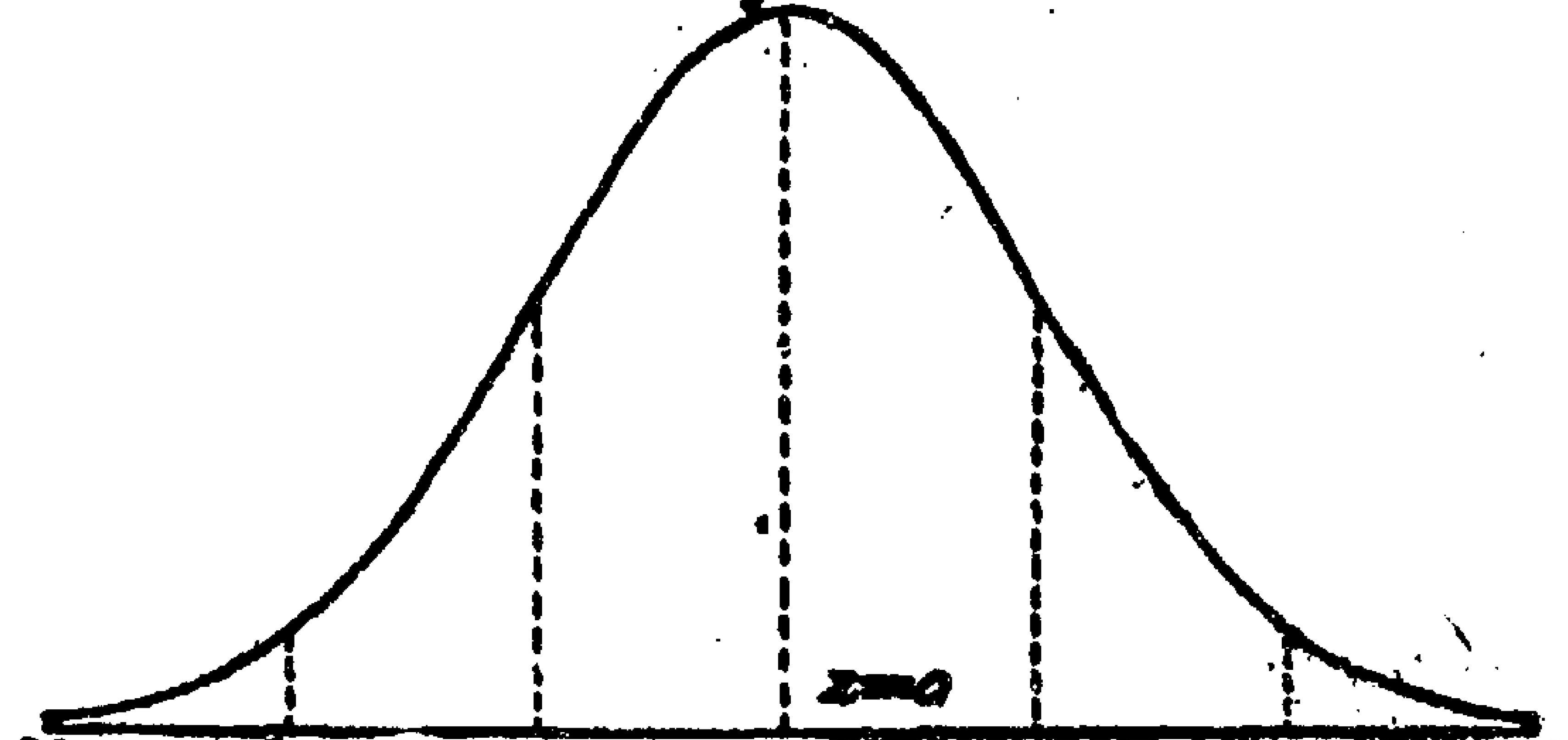


圖 12　常態曲線圖

5. 將各點用線連好，即爲常態曲線圖如上.

常態曲線下之面積及機率 在常態曲線之橫坐標上，若以平均點爲中點，視其値爲 0，則此點左右可以 σ 之距離爲單位，任分橫坐標爲若干段，而以 $\left|\frac{x}{\sigma}\right|$ 表示之. 在任何 $\left|\frac{x}{\sigma}\right|$ 點上，立一縱線 y，則 y 與 y_0（平均點上之縱線）之間曲線下，所含面積多少可由附表二查出. 該表以常態曲線之總面積爲單位，即總面積爲 1，各段面積爲總面積之分數. 例如 0 與 0.4σ 之間，其面積爲 0.1554，即該面積爲總面積百分之 15.54. 0 與 1.12σ 之間，其面積爲 0.3686. 因常態曲線兩邊對稱，故 0 與 -0.4σ 之間，其面積亦爲 0.1554. 同樣 0 與 -1.12σ 之間，其面積亦爲 0.3686.

常態曲線下各段之面積，乃代表常態分配時，x 在該段之相對次數，爲各段次數與總次數之比例. 此種相對次數稱爲機率. 各段次數不能小於 0，亦不能大於總數，故相對次數僅可由 0 至 1，而機率必爲在 0 至 1 間之數. 例如常態分配時，x 與 0 與 0.4σ 間之機率，即可由常態曲線之面積表查出爲 0.1554，則 x 在 0 與 -0.4σ 間之機率亦爲 0.1554，在 -0.4σ 與 $+0.4\sigma$ 間之機率爲 $0.1554\times 2=0.3108$，在 -0.4σ 與 $+\infty$ 間之面積爲 $0.5-0.1554=0.3446$；在 $\pm 0.4\sigma$ 以外之機率爲 $1-0.3108=0.6892$. 故若已知由 0 至任何 x 點間之常態曲線面積，則四種機率，即 0 至 x（或 0 至 $-x$），x 至 ∞（或 $-x$ 至 $-\infty$），$\pm x$ 以內，

$\pm x$ 以外之機率同時可以求得. 茲將常態分配中相當於各標準 x 之機率,及各指定機率之分別表如下:

表 1. 常態分配中相當於各標準距離 x 之機率

x	o	1P.E.(或.6745σ)	1σ	2σ	3σ
由0至 x	0	0.25	0.3413	0.4773	0.4986
由 x 至∞	0.5	0.25	0.1587	0.0227	0.0014
$\pm x$ 以內	0	0.50	0.6826	0.9545	0.9973
$\pm x$ 以外	1.0	0.50	0.3174	0.0455	0.0027

表 2. 常態分配中相當於各指定機率之距離(x/σ)

機率由0至 x	0.25	0.45	0.475	0.49	0.495
由 x 至∞	0.25	0.05	0.025	0.01	0.005
$\pm x$ 以內	0.50	0.90	0.95	0.98	0.99
$\pm x$ 以外	0.50	0.10	0.05	0.02	0.01
距離 x/σ	0.6745	1.6449	1.9600	2.3263	2.5758

在常態分配中,普通除用標準差 σ 之距離爲 x 之單位外,亦有時用四分位差或 QD 之距離爲單位. 因 x 在 $\pm 1\,QD$ 以內或以外之機率相等(各等於 0.50),故此 QD 之單位距離,通常稱爲機誤(probable error 或 P.E.). 由面積表,

$$\text{P.E.}=0.6745\sigma.$$

故在分配爲常態或近於常態時,四分位差或機誤約爲標準差三分之二. 然吾人應注意設分配爲偏態時,x 在正負 P.E. 以內之

機率未必即 0.50，甚且中位數兩旁各一個 P.E. 距離內之機率（即曲線之面積）不同，如圖 13 $A \neq B$. 故在非常態分配時，機誤並非一適宜之單位.

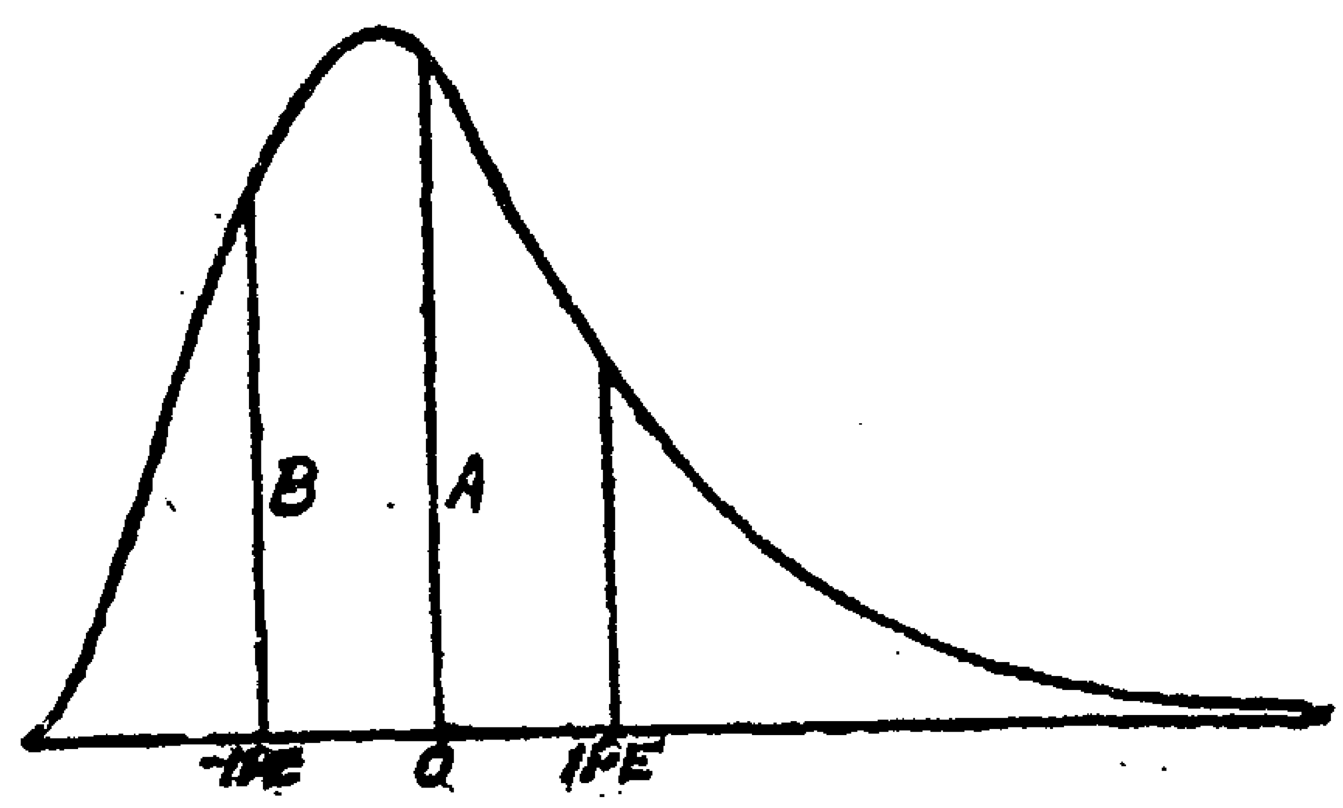


圖 13. 偏態曲線圖

VI. 可靠性

可靠性之意義 吾人若在無限量(或大量)之事物中，取其有限之若干量數而求其平均數，苟非偶爾凑巧者，决不能與眞正平均數相同. 此無限量（或大量）之事物稱爲全體，被取出之若干量數，稱爲樣本. 眞正平均數即全體之平均數，今以 M_p 表示之. 實得之平均數，或 M，即樣本之平均數. 平常實得之平均數，不過眞正平均數之近似數耳. 實得數 M 與眞正數 M_p 相較，其差數愈小，則實得數之可靠性愈大，差數愈大，則實得數之不可靠性愈大. 故 M 與 M_p 之差數與 '可靠性' 相背，而與 '不可靠性' 相並行. 故吾人必須明瞭各實得統計量數之可靠程度，並利用常態曲線之概念，爲研究本問題之開端.

算術平均數之可靠性 設吾人有大量資料，譬如有大量輪胎(大小相同，牌號相同，由同樣車輛使用)之數字資料，記載各輪胎所行駛之里程，吾人可從全體資料中研究，亦可隨機抽取

之件之多小

一個或數個樣本研究。所謂隨機抽樣者，卽全體資料中每一份子均有相等被抽取為樣本之機會.譬如抽取一個200項之樣本，則從樣本資料卽可獲得許可有用之報告.例如樣本平均數或與全體平均數相差不大.且樣本愈大，則樣本平均數亦愈接近於全體平均數.

若抽取另一個200項之樣本，則第二樣本未必與第一樣本相同.設可能抽取1000個項數相同(卽200項)之樣本，並計算各樣本之平均數時，則1000個樣本平均數成一次數分配，稱為樣本平均數之分配，吾人亦可求1000個樣本平均數之標準差如下式:

$$\sigma_M=\sqrt{\frac{(M_1-M_p)^2+(M_2-M_p)^2+\cdots+(M_k-M_p)^2}{k}}$$

公式中　σ_M 稱為樣本平均數之標準誤差，

$M_1, M_2, \cdots\cdots$ 為各樣本之平均數，

M_p 為全體資料之總平均數，

k 為平均數之個數(卽抽樣之個數).

上述盡量抽取許多樣本資料之方法並不常有.普通全體資料中僅有一，二，或少數樣本.若全體中各數字之分配為常態曲線，則各隨機樣本之平均數之分配，亦為常態曲線，以全體平均點 M_p 為中點.因此甚易由有限項之樣本計算平均數之性質，推論無限量或大量之全體平均數.若全體之分配非常態，則隨機

樣本之平均數，當樣本愈大時，亦愈趨近於常態（小樣本計算之各量數當另行討論）.

圖 14 表示由大量之全體（972 個工人之工資）抽取 100 個樣本（各樣本爲 10 項）之結果，實線表示全體分配，爲一偏態曲線. 然樣本平均數（卽10個工人之平均工資）之分配曲線（虛線）則近於對稱，且其分布範圍亦更有限制. 若各樣本之項數愈多，則 100 樣本平均數分配之曲線分布亦愈小（代表分布之量數爲標準差，平均數之標準誤差與樣本大小 N 之方根成反比例，其公式見下節）.

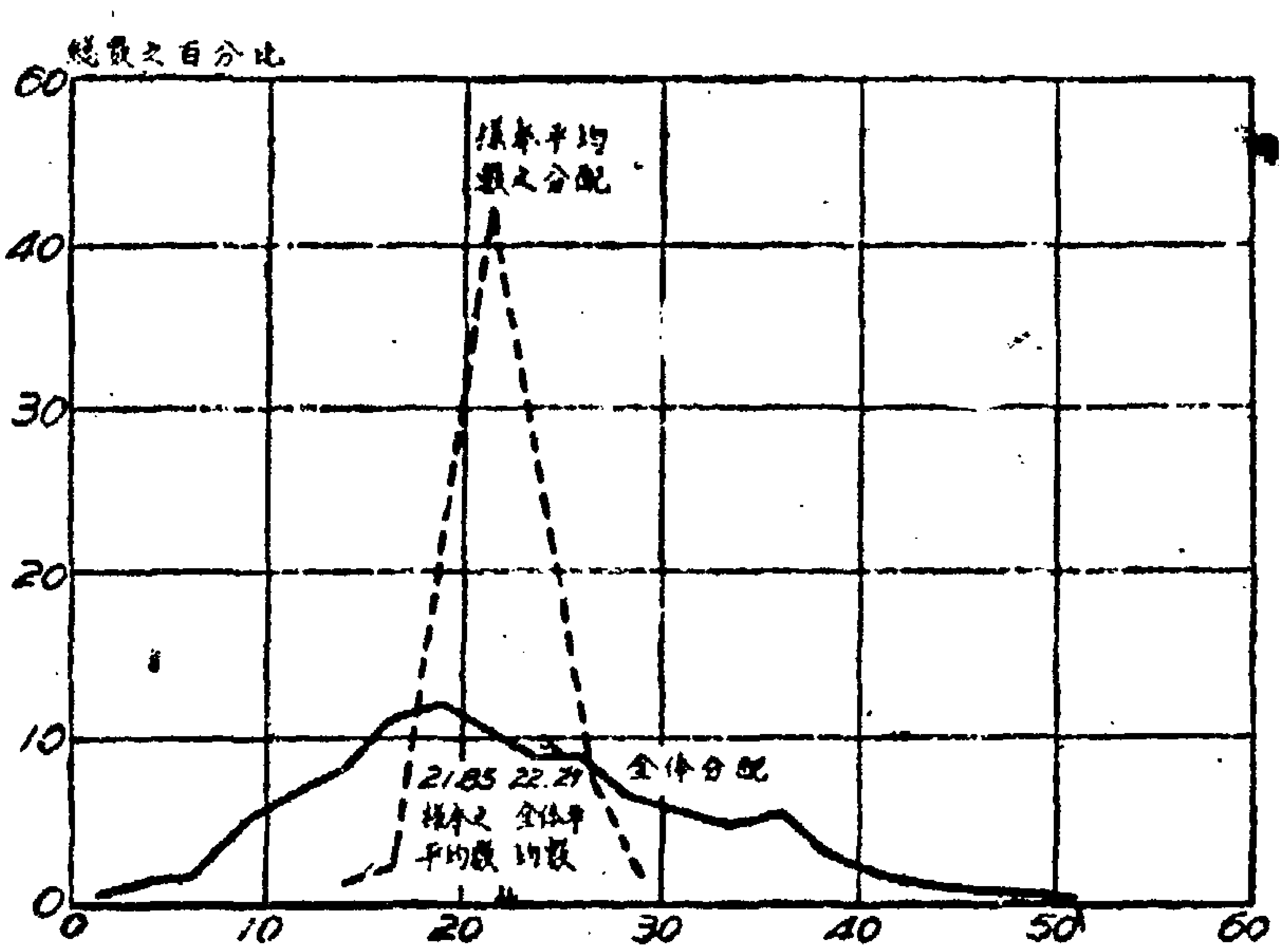


圖 14. 工 資 元 數 分 配 圖

樣本算術平均數之標準誤差 算術平均數之可靠性通常以標準誤差爲單位. 設已知全體分配之標準差爲 σ_p，則樣本算術平均數之標準誤差之公式爲:

$$\sigma_M=\frac{\sigma_p}{\sqrt{N}}.$$

假設前述大量輪胎所行駛之平均里程爲 $M_p=15,200$ 哩，標準差爲 $\sigma_p=1,248$ 哩．今取出 900 隻輪胎爲試驗樣本，其結果爲 $M=15,223$ 哩，$\sigma=1230$ 哩．顯然此樣本之算術平均數若非偶爾湊巧，必不能與眞正平均數 M_p 相脗合，其不可靠可知．其不可靠或可靠之程度，可由資料用公式求得，

$$\sigma_M=\frac{\sigma_p}{\sqrt{N}}=\frac{1,248}{\sqrt{900}}=\frac{1,248}{30}=41.6\text{ 哩}.$$

其意義卽指由全體中重複隨機抽取同樣大小之樣本，將有百分之 68.26 之樣本，此樣本平均數在眞正平均數 $15,200\pm41.6$ 哩以內，卽在 15,241.6 及 15,158.4 哩內者，其機會將爲百分之 68.26．在 15.200 ± 83.2 哩（卽 $M_p\pm2\sigma_M$）以內者，將有百分之 95.45 機會．在 $15,200\pm124.8$ 哩（卽 $M_p\pm3\sigma_M$）以內者，將有百分之 99.73 機會．

因樣本平均數之分配爲常態，有時爲便利起見，亦採用機誤 P.E. 爲單位．算術平均數之機誤公式爲：

$$\text{P.E.}_M=0.6745\frac{\sigma_p}{\sqrt{N}}$$

$$=0.6745\times41.6=28.1.$$

在 $15,200\pm28.1$ 哩（卽 $M_p\pm\text{P.E.}_M$）內之機會，有百分之 50，

本例樣本平均數 $M=15.223$ 適在此界限內.

VII. 相 關

相關之意義 宇宙事物,紛呈雜列,驟視之若千頭萬緒,淆亂無章; 細剖之,則脈絡貫通,彼此呼應. 例如雨量之多寡與收穫之豐嗇, 智慧之高下與造詣之深淺, 醫藥之興弛與死亡之增減,貨幣之貴賤與物價之漲落,莫不此消彼長或消長齊趨. 吾人若能明其關係,則執此可以知彼,根據已往可以逆料未來,以之駕馭自然之環境,控制社會之現象,實有莫大之功效.

相關之法, 卽係觀察數種現象之變化, 因而尋求其相互之關係,並用一種係數,以確定其關係之有無與深淺,以作預測之根據焉. 玆將相關之定義述之如下:

若有二種變量,彼此發生如是之關係,卽:一種變量發生變化時,而他種變量亦隨之發生變化,卽一種變量之增減,常伴以他種變量之增減, 且一種之變化程度愈大, 他種之變化程度亦隨之愈大,此之謂二種變量之相關.

若一種變量增加,他種變量亦有增加之趨勢,則相關爲正; 若一種變量增加,他種變量反有減少之趨勢,則相關爲負; 若一種變量增加或減少時,他種變量不受影響,則相關爲零.

相關之種類甚多, 有分爲直線相關與曲線相關者, 此係由相關形式之不同而別; 有分爲複相關與純相關者, 此係由相關

所用變量之多寡而定；有分品質相關與變量相關，此係由相關材料質量之區別而言. 最普通者爲直線相關. 而用皮而生氏之積差公式計算之方法，茲舉例如下：

相關之計算法　皮而生氏積差相關之通常求法.

公式：
$$r=\frac{\Sigma x\cdot y}{\sqrt{\Sigma x^2}\cdot\sqrt{\Sigma y^2}}.$$

在以上公式內：

r＝皮而生氏積差相關係數，

x＝各 x 量數與其平均數之差，

y＝各 y 量數與其平均數之差.

個人	x 量數	y 量數	x	y	x^2	y^2	$+xy$	$-xy$
A	50	22	−12	−8.4	144	70.56	100.8	
B	53	25	−9	−5.4	81	29.16	8.6	
C	56	31	−6	3.6	36	12.96		−21.60
D	58	28	−4	−2.4	16	5.76	9.60	
E	60	26	−2	4.4	4	19.36	8.80	
F	61	30	−1	−0.4	1	0.16	0.10	
G	61	32	−1	1.6	1	2.56		−1.60
H	64	30	2	−1.4	4	0.16		−0.80
I	67	28	5	−2.4	25	5.76		−12.00
J	70	34	8	3.6	64	12.96	28.80	
K	71	36	9	5.6	81	31.36	50.4	
L	73	40	11	9.6	121	92.16	105.6	
	$M_1=62$	$M_2=3)$			578	282.92	353	−36

$$r=\frac{\Sigma x\cdot y}{\sqrt{\Sigma x^2}\cdot\sqrt{y^2}}$$

$$=\frac{353-36}{\sqrt{578}\cdot\sqrt{282.92}}=0.78.$$

相關係數之可靠性及其意義如下：

$$\sigma_r=\frac{1-r^2}{\sqrt{N-2}}=\frac{1-0.78^2}{\sqrt{12-2}}=0.12.$$

$$\therefore\quad r=0.78\pm0.12.$$

若全體眞正之相關係數爲 $+0.78$，則由 12 對隨機樣本計算之相關，將有百分之 68.26 在 0.66 與 0.80 間，然此僅爲一粗疏而欠適宜之公式。因 0.78 並非全體之係數，卽使全體之相關 $r_p=0.78$，樣本 r 並非依照常態分配於 r_p 之兩旁，上式僅當 N 爲大而 r_p 爲小時，樣本 r 方近於常態也。

皮而生氏積差相關之簡捷求法：

公式：
$$r=\frac{\frac{\Sigma x'y'}{N}-c_xc_y}{\sqrt{\frac{\Sigma x^2}{N}-c_x^2}\sqrt{\frac{\Sigma y^2}{N}-c_y^2}};$$

或
$$r=\frac{\Sigma x'y'-\frac{(\Sigma x')(\Sigma y')}{N}}{\sqrt{\Sigma x^2-\frac{(\Sigma x')^2}{N}}\sqrt{\Sigma y^2-\frac{(\Sigma y')^2}{N}}}.$$

在以上公式內：

$x'=x-x_0$，即各 x 量數與其假定平均數(x_0)之差，

$y'=y-y_0$，即各 y 量數與其假定平均數(y_0)之差，

$c_x=\frac{\Sigma x'}{N}$ 即假定平均數 x_0 之校正數，

$c_y=\frac{\Sigma y'}{N}$ 即假定平均數 y_0 之校正數.

個人	量數 x	量數 y	x'	y'	x'^2	y'^2	$+x'y'$	$-x'y'$
A	50	22	-10	-8	100	64	80	
B	53	25	-7	-5	49	25	35	
C	56	34	-4	-4	16	16		-16
D	58	28	-2	-2	4	4	4	
E	60	26	0	-4		16		
F	61	30	1	0	1			
G	61	32	1	2	1	4	2	
H	64	30	4	0	16			
I	67	28	7	-2	49	4		-14
J	70	34	10	4	100	16	40	
K	71	36	11	6	121	36	66	
L	73	40	13	10	169	100	130	
					626	285	357	-30

$$c_x=\frac{24}{12}=2, \qquad \therefore \quad c_x{}^2=4.$$

$$c_y=\frac{5}{12}=0.417, \qquad \therefore \quad c_y{}^2=0.173889.$$

$$\frac{\Sigma x'^2}{N}=\frac{626}{12}, \qquad \frac{\Sigma y'^2}{N}=\frac{285}{12},$$

$$\frac{\Sigma x'y'}{N}=\frac{357-30}{12}, \qquad c_x c_y=2\times 0.417$$

$$r=\frac{\frac{\Sigma x'y'}{N}-c_x c_y}{\sqrt{\frac{\Sigma x^2}{N}-c_x{}^2}\sqrt{\frac{\Sigma y^2}{N}-c_y{}^2}}$$

$$=\frac{\frac{357-30}{12}-2\times 0.417}{\sqrt{\frac{626}{12}-4}\sqrt{\frac{285}{12}-0.173889}}$$

$$=\frac{26.416}{33.659}=0.78.$$

$$\sigma_r=\frac{1-r^2}{\sqrt{N-2}}=\frac{1-0.78^2}{\sqrt{12-2}}=0.12, \qquad r=0.78\pm 0.12.$$

皮而生氏積差相關用相關表求法:

公式:

$$r=\frac{\frac{\Sigma fx'y'}{N}-c_x c_y}{\sqrt{\frac{\Sigma fx'^2}{N}-c_x{}^2}\sqrt{\frac{\Sigma fy'^2}{N}-c_y{}^2}},$$

或

$$r=\frac{\Sigma fx'y'-\frac{(\Sigma fx')(\Sigma fy')}{N}}{\sqrt{\Sigma fx'^2-\frac{(\Sigma fx')^2}{N}}\sqrt{\Sigma fy'^2-\frac{(\Sigma fy')^2}{N}}}$$

在以上公式內，先設各 x, y 爲各組之中點，h, k 依次爲 x 及 y 之組距，並依次選任一組之中點 x_0, y_0 爲假定平均數；再設

$x' = \dfrac{x-x_0}{h}$，各 x 組中點與假定平均數之差，以組距 h 爲單位；

$y' = \dfrac{y-y_0}{k}$，各 y 組中點與假定平均數之差，以組距 k 爲單位，

$c_x = \dfrac{\Sigma fx'}{N} = \dfrac{x-x_0}{h}$，假定平均數 x_0 之校正數，以組距 h 爲單位，

$c_y = \dfrac{\Sigma fy'}{N} = \dfrac{y-y_0}{k}$，假定平均數 y_0 之校正數，以組距 k 爲單位.

$\dfrac{1}{N}\Sigma f_{xy}x'y'$ 稱爲平均積差 (mean product) 或互變異量 (covariance)積差. $\Sigma f_{xy}x'y'$ 之計算方法甚多，茲用兩種方法計算如下:

1. 先求同一 y' 行各 f_{xy} 依次與 x' 之乘積，並求其和 $S_x' = \Sigma f_{xy}x'$，再求各 y' 依次與 S_x' 之乘積，其和即爲 $\sum\limits_{y} y'S_x' = \sum\limits_{x,y} f_{xy}x'y$. 例如下表:

第一列　$y'=-6$，僅 $x'=-3$ 時，次數爲 1 卽 $f_{xy}=1$，故 $S_x'=1(-3)=-3$，$y'S_x'=(-6)(-3)=18$.

第二列　$y'=-5$，次數無，故 $S_x'=0$，$y'S_x'=0$.

第三列　$y'=-4$，當 $x'=-6,-4,-3$ 及 -1 時次數爲 1，故 $S_x'=(-6)+(-4)+(-3)+(-1)=-14$，$y'S_x'=(-4)(-14)=56$.

第四列　$y'=0$，當 x' 爲 $-3,-2,-1,0,1,2,3$ 時，f_{xy} 依次爲 $1,2,4,8,5,0,1$，故 $S_x'=1(-3)+2(-2)+4(-1)+8(0)+5(1)+0(2)+1(3)=-3$，$y'S_x'=0(-3)=0$.

其他 S_x' 及 $y'S_x'$ 亦可依此計算，惟須注意 $\sum_y S_x'=35=\sum f_x x'$ 及 $\sum_y y'S_x'=286=\Sigma f_{xy}x'y'$.

2. 先求同一 x' 行各 f_{xy} 依次與 y' 之乘積，其和爲 $S_y'=\sum_y f_{xy}y'$，再求各 x' 依次與 S_y' 之乘積，其和卽爲 $\sum_x x'S_y'=\sum_{x,y} f_{xy}x'y'$. 例如下表:

第一行　$x'=-6$，僅 $y'=-4$ 時，次數爲 1，卽 $f_{xy}=1$，故 $S_y'=1(-4)=-4$，$x'S_y'=(-6)(-4)=24$.

第四行　$x'=-3$，當 $y'=-6,-5,-4,-3,-2,-1,0$ 時，f_{xy} 依次爲 $1,0,1,2,1,2,1$，故 $S_y'=1(-6)+0(-5)+1(-4)+2(-3)+1(-2)+2(-1)$

表 3. 94 個學生在張氏及陸氏智力測驗上分數

學號	張氏測驗分數	陸氏測驗分數	學號	張氏測驗分數	陸氏測驗分數	學號	張氏測驗分數	陸氏測驗分數
1	40	63	33	47	61	65	48	68
2	43	52	34	45	47	66	17	39
3	30	53	35	43	62	67	40	56
4	57	79	36	36	61	68	33	53
5	23	41	37	52	55	69	38	57
6	22	53	38	44	61	70	44	57
7	38	59	39	31	38	71	36	59
8	36	53	40	37	71	72	34	56
9	40	66	41	39	40	73	47	60
10	52	71	42	43	56	74	53	74
11	42	49	43	23	51	75	26	49
12	25	42	44	45	63	76	52	61
13	51	64	45	39	61	77	56	65
14	51	61	46	27	47	78	23	41
15	32	46	47	45	50	79	33	53
16	37	65	48	35	47	80	31	65
17	45	60	49	47	49	81	33	58
18	50	73	50	31	41	82	30	53
19	36	56	51	42	55	83	46	65
20	13	47	52	31	47	84	28	53
21	61	79	53	23	46	85	23	25
22	30	65	54	8	37	86	38	55
23	49	64	55	39	57	87	59	70
24	54	54	56	35	54	88	25	56
25	45	68	57	35	48	89	46	60
26	27	55	58	21	38	90	51	71
27	39	63	59	32	55	91	22	55
28	42	55	60	49	60	92	50	54
29	30	47	61	32	50	93	53	65
30	39	59	62	58	82	94	52	51
31	45	48	63	48	71			
32	57	65	64	32	56			

茲將表3之數字，轉入分佈圖內如下：

X　　X 張氏智力測驗總分數

Y　　Y 陸氏智力測驗總分數

	5—	10—	15—	20—	25—	30—	35—	40—	45—	50—	55—	60—	fy	y'	fyy'	fy'^2	S_x'	$y'S_x'$
25—				1									1	−6	−6	36	−3	18
30—													0	−5	0	0		0
35—	1		1	1		1							4	−4	−16	64	−14	56
40—				2	1	1	1						5	−3	−15	45	−9	27
45—		1		1	2	3	2	1	3				13	−2	−26	52	−8	16
50—				2	1	5	1	1	1	3			14	−1	−14	14	−1	1
55—				1	2	4	8	5		1			21	0	−77	0	−3	
60—							3	3	7	3			16	1	16	16	26	26
65—						2	1	1	3	1	2		10	2	20	40	16	32
70—							1		1	4	1		7	3	21	63	18	54
75—											1	1	2	4	8	32	9	36
80—											1		1	5	5	25	4	20
f_x	1	1	1	8	6	16	17	11	15	12	5	1	94		70	387	35	286
x'	−6	−5	−4	−3	−2	−1	0	1	2	3	4	5						
$f_x x'^1$	−6	−5	−4	−4	−12	−16	−67	11	30	36	20	5	102					
$f_x x'^2$	36	25	16	72	24	16	0	11	60	108	80	25	173					
$S'y$	−4	−2	−4	−20	−8	−14	0	2	9	14	16	4	−7					
$x'Sy'$	24	10	16	60	16	14	0	2	18	42	64	20	286					

$+1(0)=-20$, $x'S_y'=(-3)(-20)=60$.

第七行 $x'=0$, 當 $y'=-3,-2,-1,0,1,2,3$ 時, f_{xy} 依次爲1, 2, 1, 8, 3, 1, 1, 故 $S_y'=1(-3)+2(-2)+1(-1)+8(0)+3(1)+1(2)+1(3)=-8+8=0$, $x'S_y'=0$.

其他 S_y' 及 $x'S_y'$ 亦可仿此計算，惟須注意 $\sum_x S_y'=-7=\Sigma f_y y'$ 及 $\Sigma x'S_y'=286=\Sigma f_{xy}x'y'$, 此可作複核之用.

由上表,

$$N=94.$$

$$x_0=37.5, \qquad y_0=57.5.$$

$$h=5, \qquad k=5.$$

$$x'=\frac{x-37.5}{5}, \qquad y'=\frac{y-37.5}{5}.$$

$$c_x=\frac{-67+102}{94}=+0.3723,$$

$$c_y=\frac{-77+70}{94}=-0.07447,$$

$$c_{x^2}=0.1386, \qquad c_{y^2}=0.0055,$$

$$c_xc_y=-0.0277.$$

$$\frac{\Sigma fx'^2}{N}=\frac{473}{94}=5.0319, \qquad \frac{\Sigma fy'^2}{N}=\frac{387}{94}=4.1170.$$

$$\frac{\Sigma f x' y'}{N}=\frac{286}{94}=3.0426.$$

$$r=\frac{\frac{\Sigma f x' y'}{N}-c_x c_y}{\sqrt{\frac{\Sigma f x'^2}{N}-c_x{}^2}\sqrt{\frac{\Sigma f y'^2}{N}-c_y{}^2}}$$

$$=\frac{3.0426-(-0.0277)}{\sqrt{5.0319-0.1386}\sqrt{4.1170-0.0055}}$$

$$=\frac{3.0703}{\sqrt{4.8933}\sqrt{4.1115}}=\frac{3.0703}{2.2121\times 2.0277}$$

$$=\frac{3.0703}{4.4855}=0.6845. \qquad \therefore \quad r=0.684\pm 0.550.$$

$$\sigma_r=\frac{1-r^2}{\sqrt{N-2}}=\frac{1-0.6845^2}{\sqrt{94-2}}=\frac{0.535}{\sqrt{92}}=0.0554.$$

相關係數之意義　兩種量數爲完全相關時，其相關係數爲 $+1.00$；兩種量數爲完全負相關時，其相關係數爲 -1.00；兩種量數爲完全不相關時，其相關係數爲 0. 自 -1.00 至 $+1.00$ 可得相關數之各種等級. 就事實言，$+1.00$ 與 -1.00 不可多得. 故重要問題，爲何種相關係數表示高相關，何種表示低相關，例如 ± 0.65 之相關係數，是否應視爲高；± 35 之相關係數，是否應視爲低，此與事實之性質有關. 有時 ± 0.65 之相關係數，在某二種事實上，應視爲高，而在其他二種事實上，則視爲低. 茲將

麥柯爾(Wm. A. McCall)氏所定之標準錄下:

0 至 ±0.4	表示低度相關,
±0.4 至 ±0.7	表示切實相關,
±0.7 至 ±1.00	表示高度相關.

相關係數之可靠性 但相關係數之高低,亦須視量數之多少而定. 量數愈多,則相關係數之可靠性愈大; 若量數太少,則相關係數雖高,亦不甚可靠. 茲將表示可靠性之各公式分述如下:

1. r 之標準誤差(σ_r)與平均數之可靠性適相似. 設已知全體之相關係數爲 r_p 時,其公式爲

$$\sigma_r=\frac{1-r_p^2}{\sqrt{N-1}}.$$

然應用上式時,有兩種限制: (1)當全體相關係數爲零時,r 之抽樣分配,方爲常態,當全體係數爲正時,r 之抽樣分配爲負偏態,若 r_p 爲 0.80,則不同抽樣之 r 僅能較高 0.20,但抽樣係數最低可能爲 -1.00(即可能較低 -1.80),當 r_p 近於零時,則 r 之分配近似常態,何時不能應用上列 σ_r 之公式,雖無一定之界限,但鐵貝特(Tippett)氏認爲 r_p 近於 0.80 時,r 分配之偏態甚爲重要,不能應用上列公式; (2)即使 r_p 爲零或近於零時,小樣本相關係數之分配亦非常態,但當 N 增加時,r 之分配漸趨於常態.

若吾人之目的在測驗相關之顯著與否，卽測驗在全體中，是否可假設二量數無相關．若假設不能證實，則相關認爲顯著，故測驗此假設時，必須以 $r_p=0$ 代入 σ_r 式中，卽

$$\sigma_r=\frac{1}{\sqrt{N-1}}.$$

然當 N 爲少數時，此式亦不可用，譬如 94 個學生之張氏及陸氏測驗分數，用相關求得 $r=0.675$．應用公式：

$$\sigma_r=\frac{1}{\sqrt{N-1}}=\frac{1}{\sqrt{12-1}}=\frac{1}{\sqrt{11}}=0.3015.$$

當測驗 r 是否顯著時，卽以其標準誤差除 r 與 0 差，或

$$\frac{r}{\sigma_r}=\frac{0.675}{0.0567}=11.905.$$

若 r 大於 σ_r 之三倍時，則統計學家認爲其相關存在，並非由於機會，顯然本例相關爲顯著．設相關指示不顯著時，吾人不妨增大樣本，再行測驗．

2. 公式：

$$\sigma_r=\frac{1-r^2}{\sqrt{N-2}}$$

有時可用以測驗直線相關 r 之可靠性，上例 $N=12$，$r=0.675$，

$$\sigma_r=\frac{1-r^2}{\sqrt{N-2}}$$

$$=\frac{1-0.675^2}{\sqrt{94-2}}=(0.544)(0.1043)$$

$$=0.0568.$$

則 $r=0.675\pm0.057$. 若全體之相關係數爲 0.675, 則由 94 對隨樣本計算之相關係數, 將有百分之 68.26 在 0.618 與 0.732 之間. 然此亦爲粗疏而非適宜之手續, 卽使 $r_p=0.675$, 樣本 r 並非常態分配於 r_p 之兩旁, 蓋僅當 N 爲大而 r 爲小時, 樣本 r 之分配方近於常態也.

VIII. 表列

表列之意義 將統計材料在縱橫兩行內作有秩序之排列, 使便於閱讀, 稱爲表列.

表列之級次 表列可依其項目之多寡, 而分爲一級, 二級, 三級等表.

1. 一級表 若縱標目或橫標目祇有一層, 則稱爲一級表. 如表 4 之橫標目祇有 '人口數' 一層.
2. 二級表 若縱標目或橫標目下再分一層, 則稱爲二級表.

   如表5之橫標目 '人口數' 爲一層, 縱標目及 '男' '女' 又爲一層.

3. 三級表　若縱標目.或橫標目與其下各層,共爲三層,則稱爲三級表.

如表 6 之橫標目計有'人口數'一層,其下又分'普通戶''其他各類戶'一層,再下又分'男''女'一層.餘依類推.

表 4. 四川省彭縣雙流崇寧三縣之人口總數

民國三十一年

縣別	人口數
彭縣	369,055
雙流	156,085
崇寧	94,331

表 5. 四川省彭縣雙流崇寧三縣之人口總數

民國三十一年

縣別	人口數	
	男	女
彭縣	192,771	176,284
雙流	80,110	75,975
崇寧	48,966	45,365

表 6. 四川省彭縣雙流崇寧三縣之人口總數

民國三十一年

縣別	人口數			
	普通戶		其他各類戶	
	男	女	男	女
彭縣	184,626	174,845	8,145	1,439
雙流	72 081	74,931	8,029	1.044
崇寧	46,501	44,903	2,465	462

製表之規則

1. 關於表之標題者

   A. 表之標題當力求顯明易解,並須切合表之內容.

   B. 標題之位置應在表之頂端,可分層臚列.

   C. 標題中所舉各點,其次序應與表中所列之項目一致.

   D. 表之內容有地域及時間區別時,標題中亦應註明.

   E. 如表之內容甚長,佔數頁地位時,各頁均須註明標題,除第一頁外,餘頁並須於標題後註明'續前'二字.

2. 關於表之項目者

   A. 表中項目之次序按下列各種標準排列:

   (1) 重要之程度.

   (2) 等級之高低.

   (3) 時間之先後.

(4)　數量之大小.

(5)　筆畫之簡繁.

(6)　地方區域之位置，如有規定次序可資遵循時，務須依照規定.

B.　表中項目一律由左至右橫寫.

C.　大項目可分細目，惟須低一格，並用線格分開.

D.　對於某一項目特別注重時，可用較重之字體顯明之.

E.　凡橫幅過長之表，左部所註項目不便閱覽時可於右端重列一次.

F.　為便於檢查或引用起見，項目較多之表，可於項目上面或旁面用羅馬字標明次序.

G.　表上項目過多，使表身左右太長，可以過表續排，謂之橫跨；若使表身上下太長，可以過頁續排，謂之縱跨.

3.　關於表之線格者

A.　表中各縱行間須以直線畫分之，細目間用細縱線，大項目間則用較粗之縱線，橫行間則無須用線.

B.　表之左右兩邊無須用邊線，惟橫跨之表，右邊須畫細邊線.

C.　表之頂底均畫粗橫線，惟縱跨之表，則底線應為細橫線.

D. 表中上部項目與數字之間，須用橫線畫分之.

E. 表中排列數字處不可用橫線分開，以免障礙視線.

4. 關於數字之排列者

A. 表中數字，須一律用阿拉伯字，以其整齊且可節省篇幅.

B. 表中各縱行數字之位置須上下相對，以便加減比較. 如有小數，則小數點尤須列在一垂直線上，以免計算錯誤.

C. 總計須寫在表之上左角上，並用重體字排；百分數之數字，可用斜體字排.

D. 數字多至四位以上，須用分逗點，普通每三位分作一段.

E. 空格內實際上缺乏數字者用三點(…)表示之；有數字而未經報告者用短橫線(一)表示之.

5. 關於表中其他事項者

A. 表中所列材料來源須於表之下端詳細註明.

B. 表中所列事項有須加以解釋者，應於表之下端一一說明.

IX. 圖示

圖示之意義 用圖形表示數量之事實，以揭出材料之顯著

特性，而隱藏其細微部分，稱為圖示．圖示適宜，則閱者不加思索，一覽而興趣勃發，意義著明，對於圖中之事實，卽瞭如指掌．

圖示之種類 圖示之種類，各家之分法不一，兹概分爲以下五種，並各爲舉例如下：

1. 線圖 例如：
   - A. 次數分配曲線圖．
   - B. 累積次數曲線圖．
   - C. 盈虧圖．
   - D. 帶形圖．
   - E. 羅倫氏曲線．
   - F. Z形圖．
2. 條圖 例如：
   - A. 橫條圖．
   - B. 縱條圖．
   - C. 甘氏進度圖．
3. 面積圖 例如：
   - A. 三角形圖．
   - B. 方形圖．
   - C. 圓形圖．
4. 體積圖 例如：
   - A. 立方圖．

B.　球體圖.

C.　圓柱圖.

D.　二變數次數立體圖.

5.　統計地圖　例如:

A.　陰影地圖.

B.　交叉線地圖.

C.　加點地圖.

D.　着色地圖.

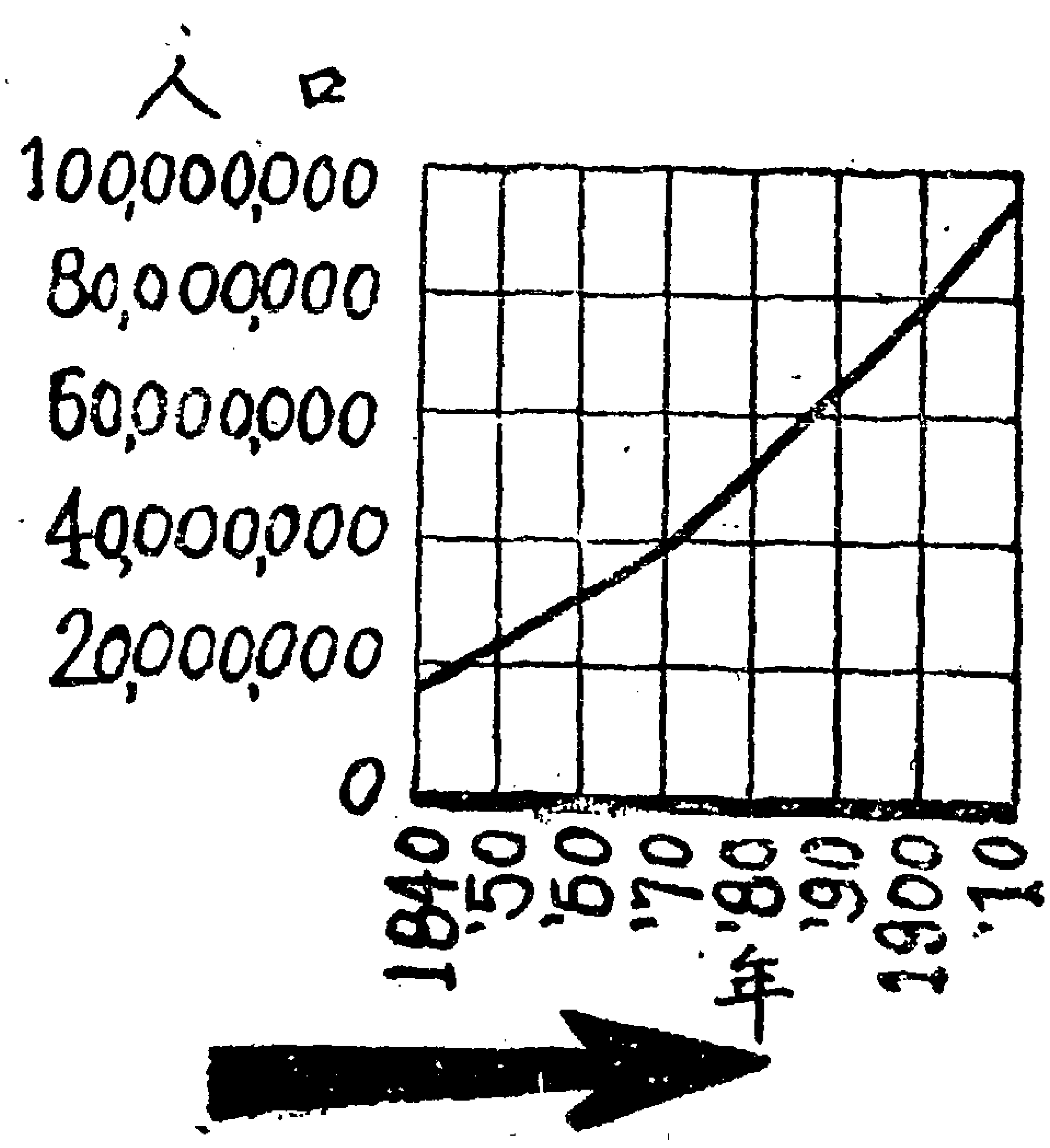
E.　插針地圖.

F.　插旗地圖.

圖示法之通則及示例

(美國圖示標準委員會規定)

1.　圖之一般排法宜自左至右.



圖例 1.

2. 能用線代表量數最好，蓋面積與體積均易誤解.

年	噸
1900	270,588
1914	555,091

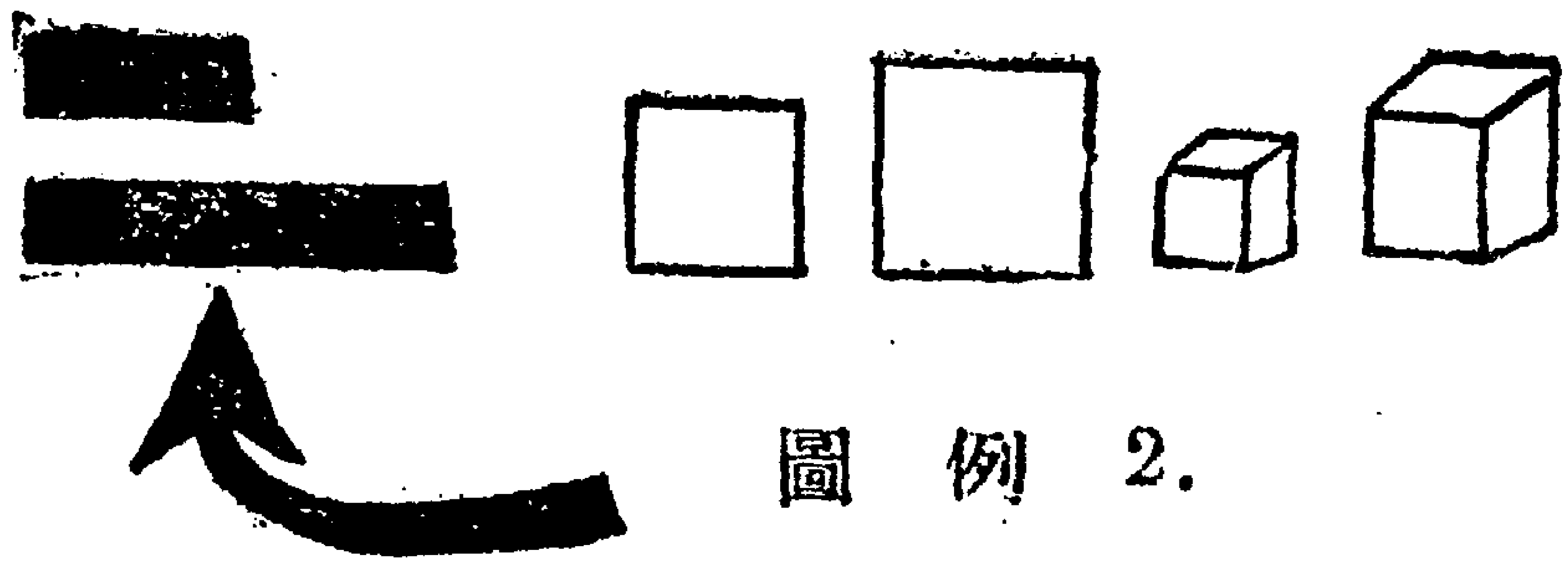


圖　例　2.

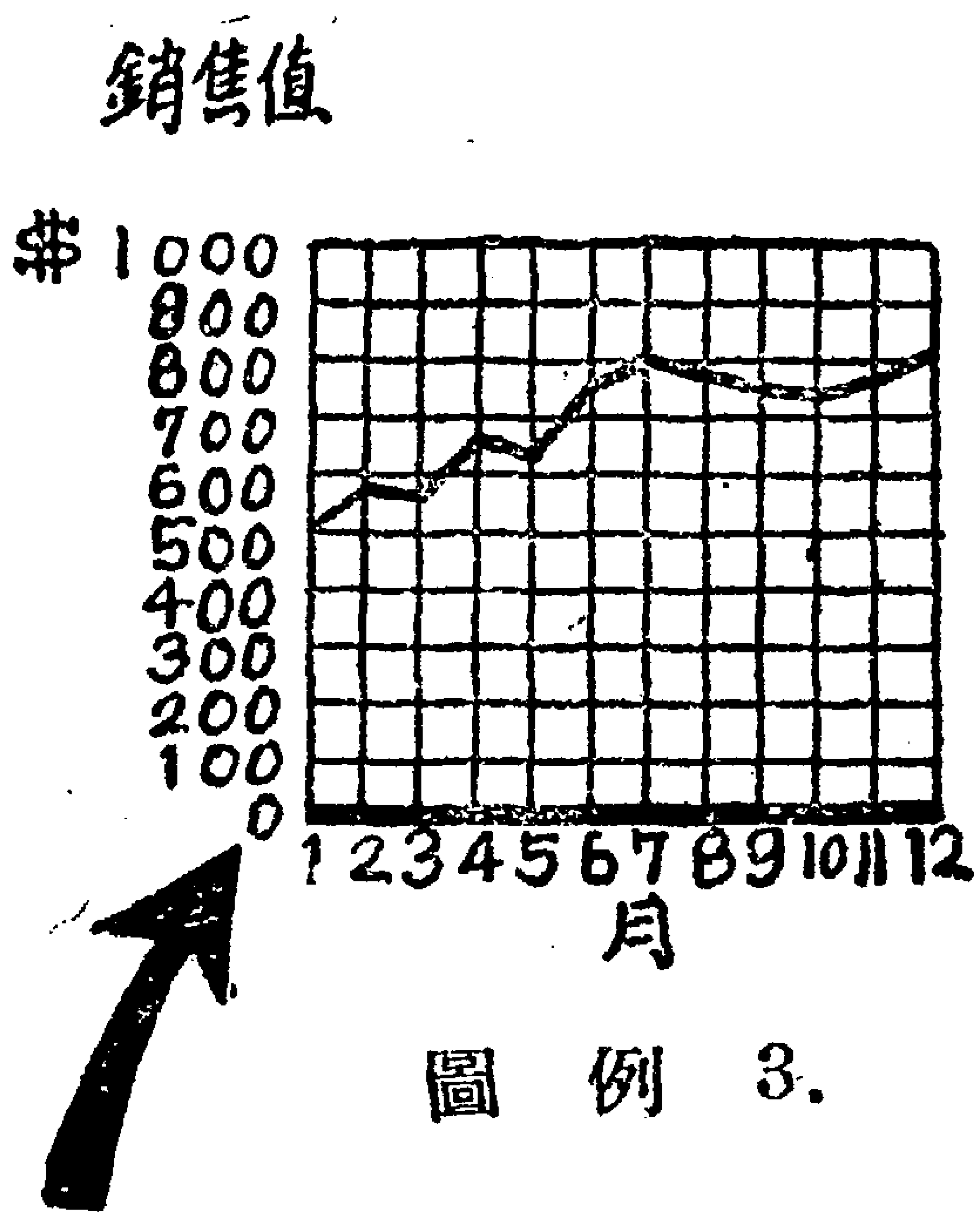


圖　例　3.

3. 曲線圖上最好能將縱坐標零度畫出.

4. 倘縱坐標零度照常不致用到，則圖上宜用一種波線，而仍保存縱坐標零度之位置.

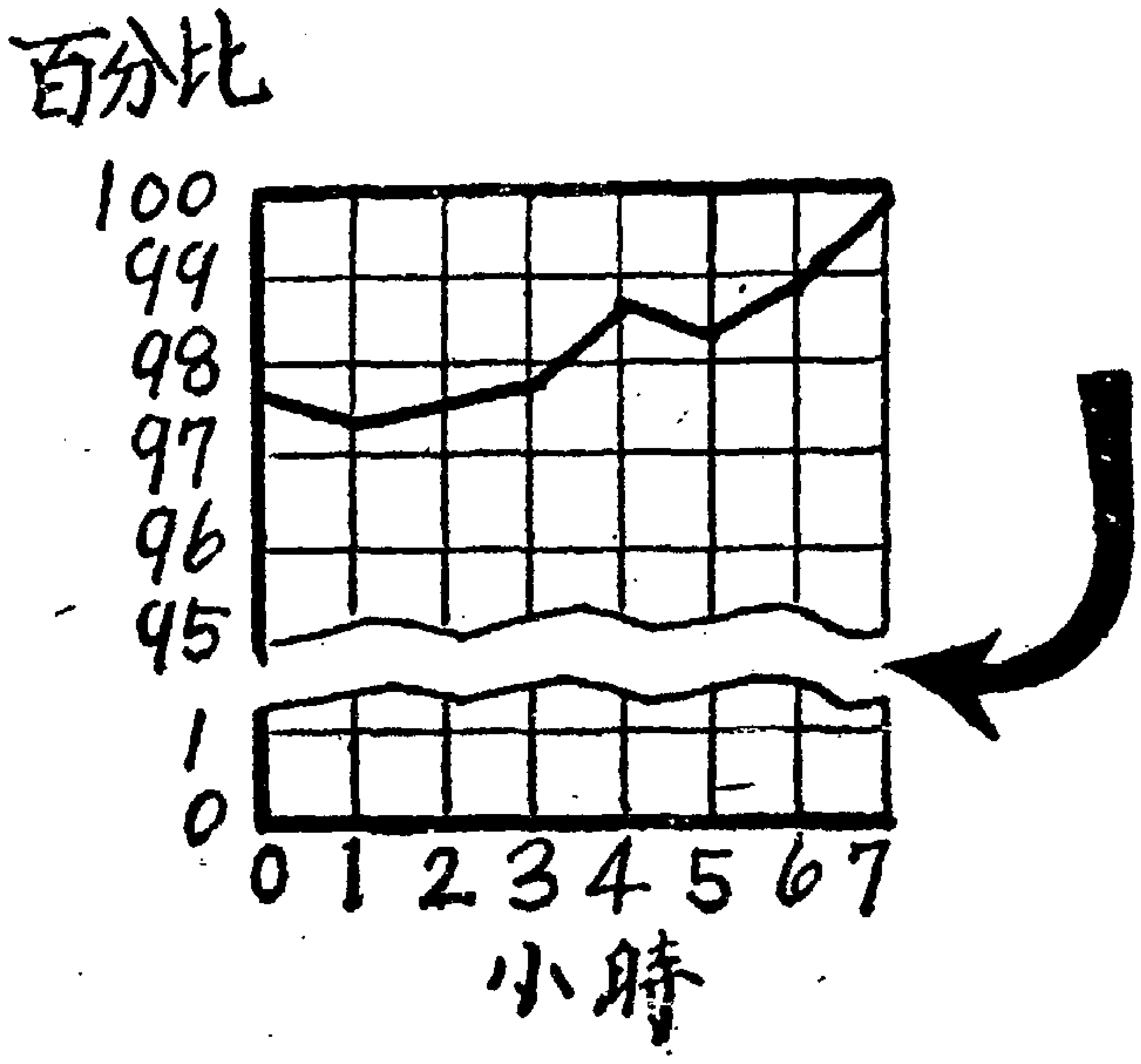
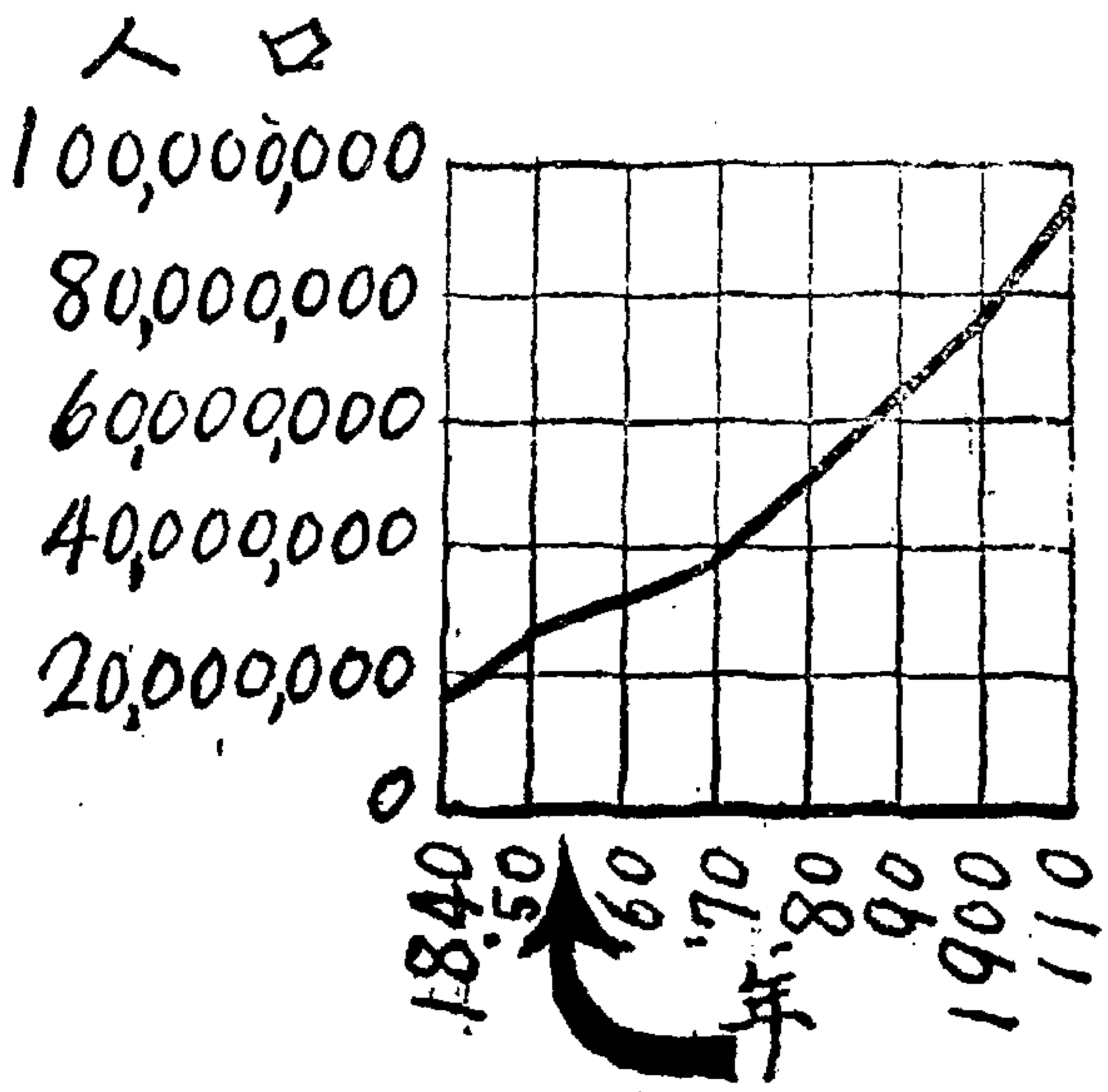
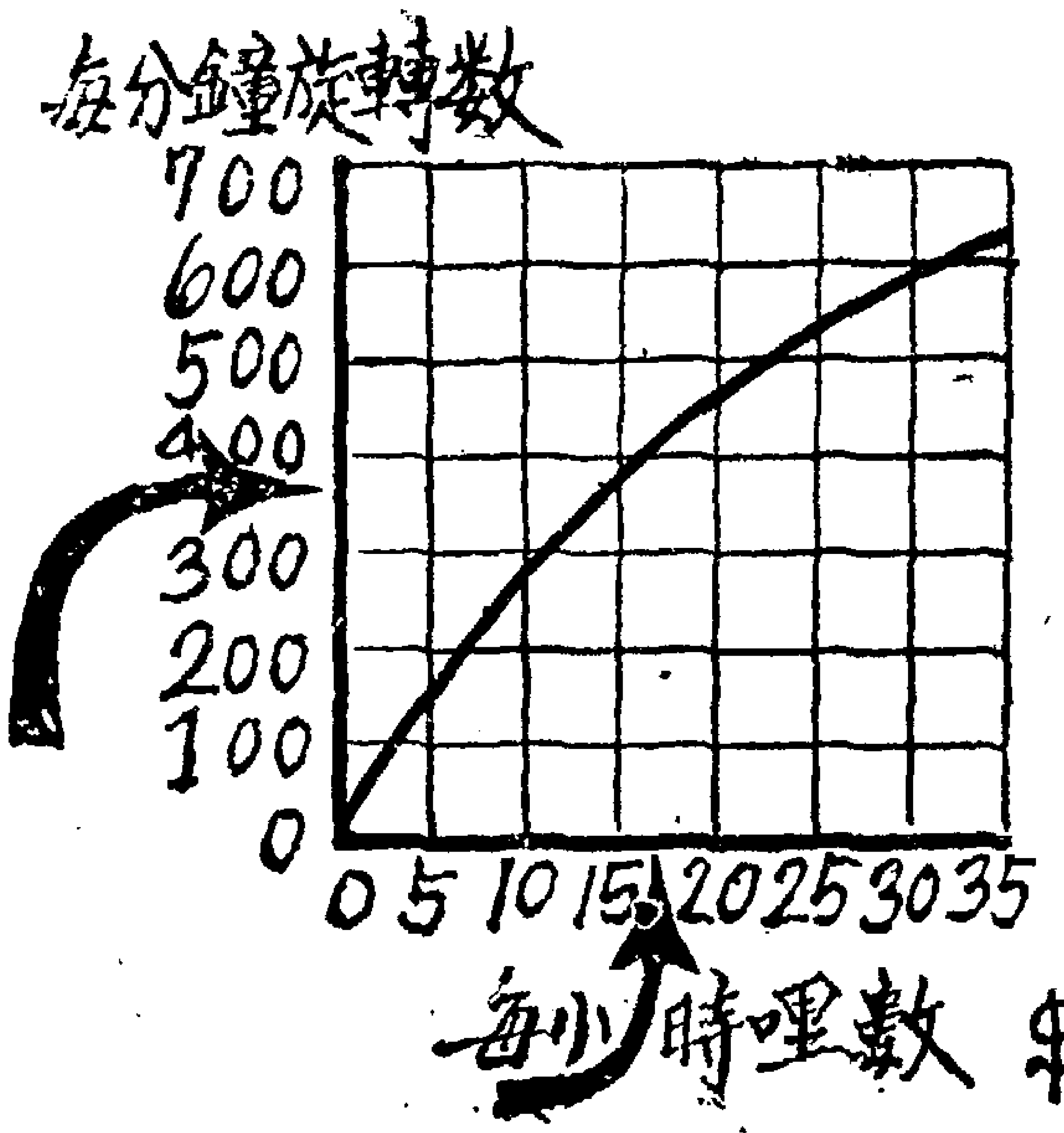


圖　例　4.

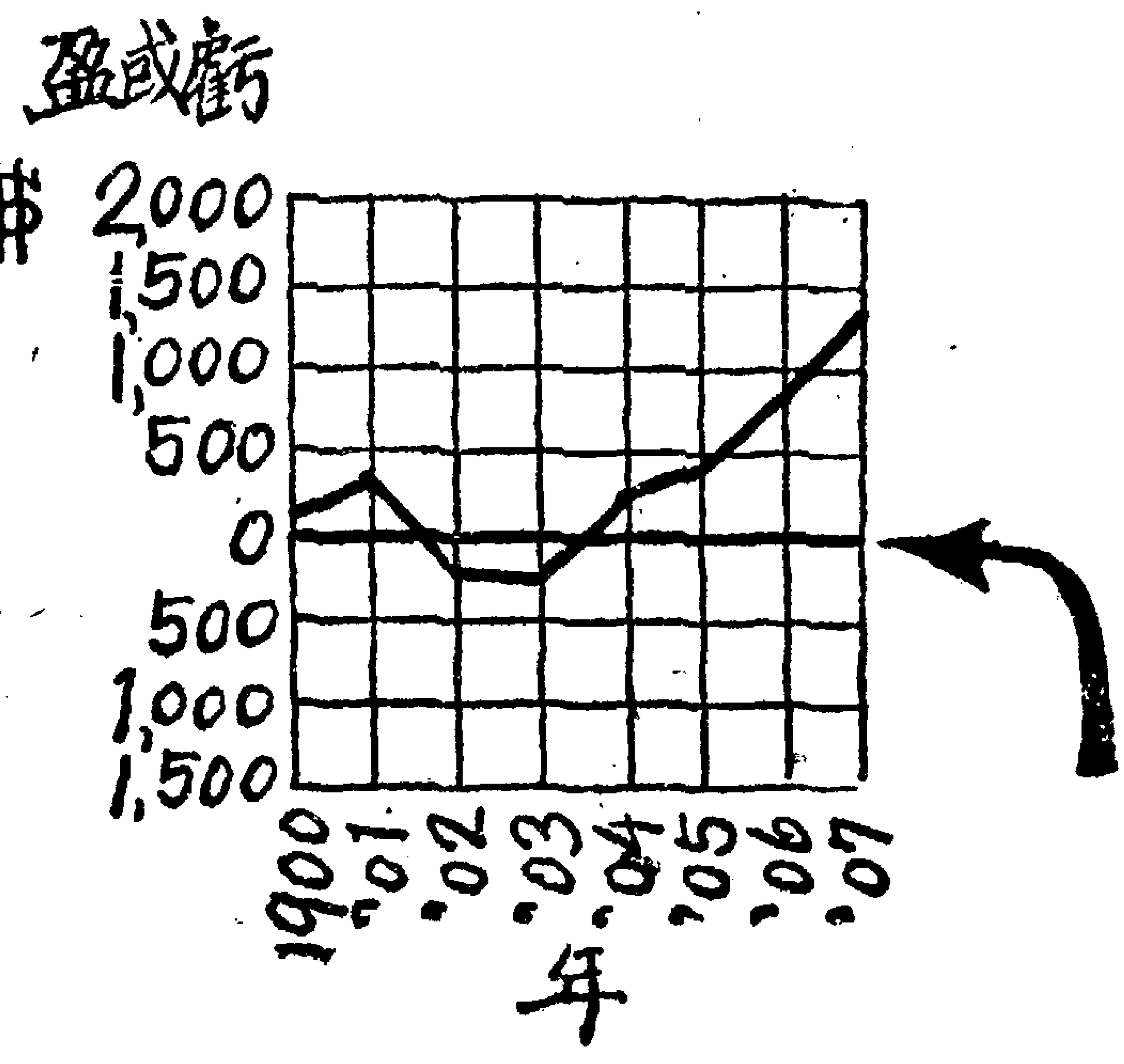
5. 縱橫坐標之零度線，應較粗大，以別於其他縱橫線.



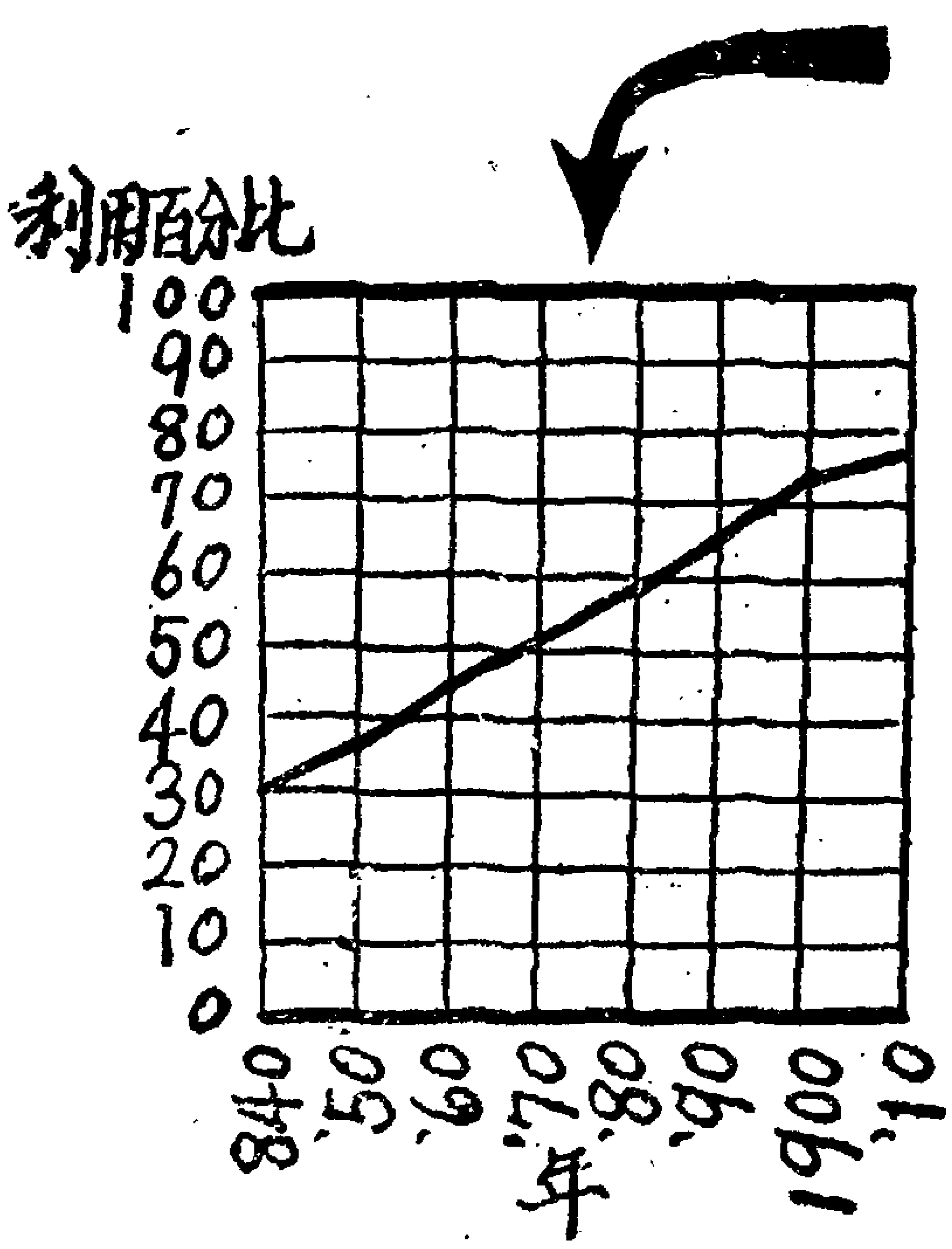
圖例 5.(A)



圖例 5.(B)

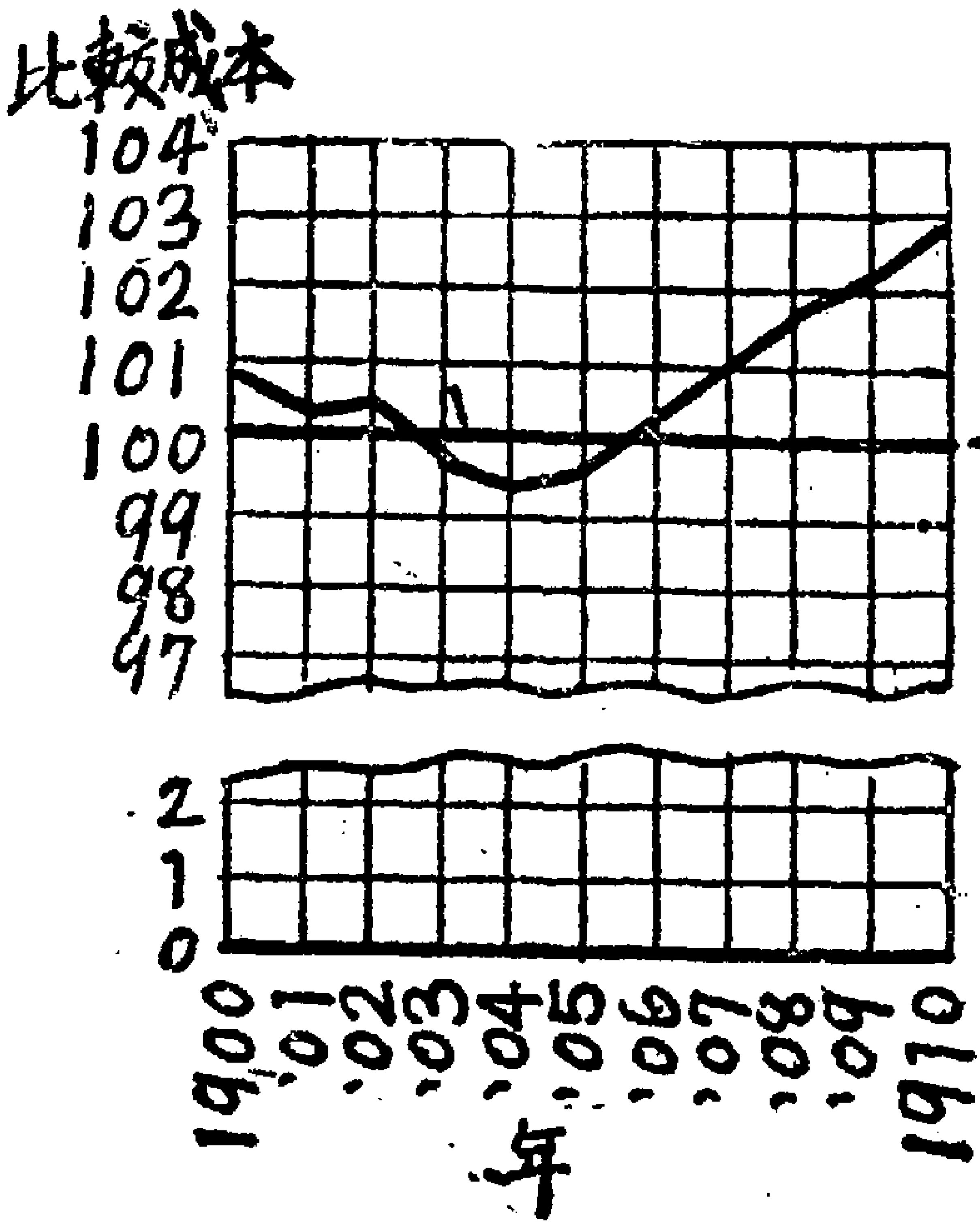


圖例 5.(C)

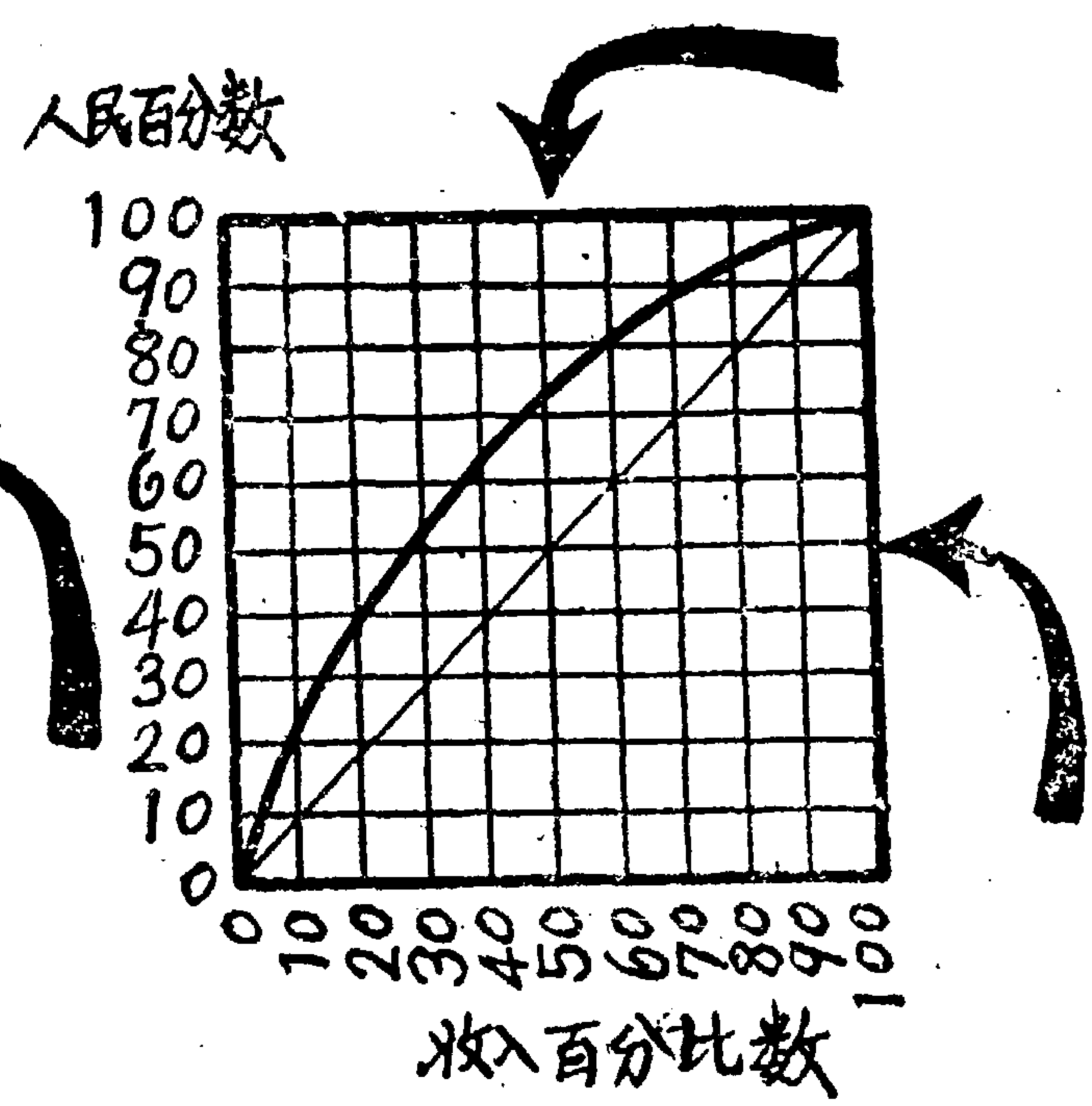


圖例 6.(A)

6. 若曲線圖代表百分比者，則百分線宜粗大以示區別. 其他用作比較基線亦宜較粗大.

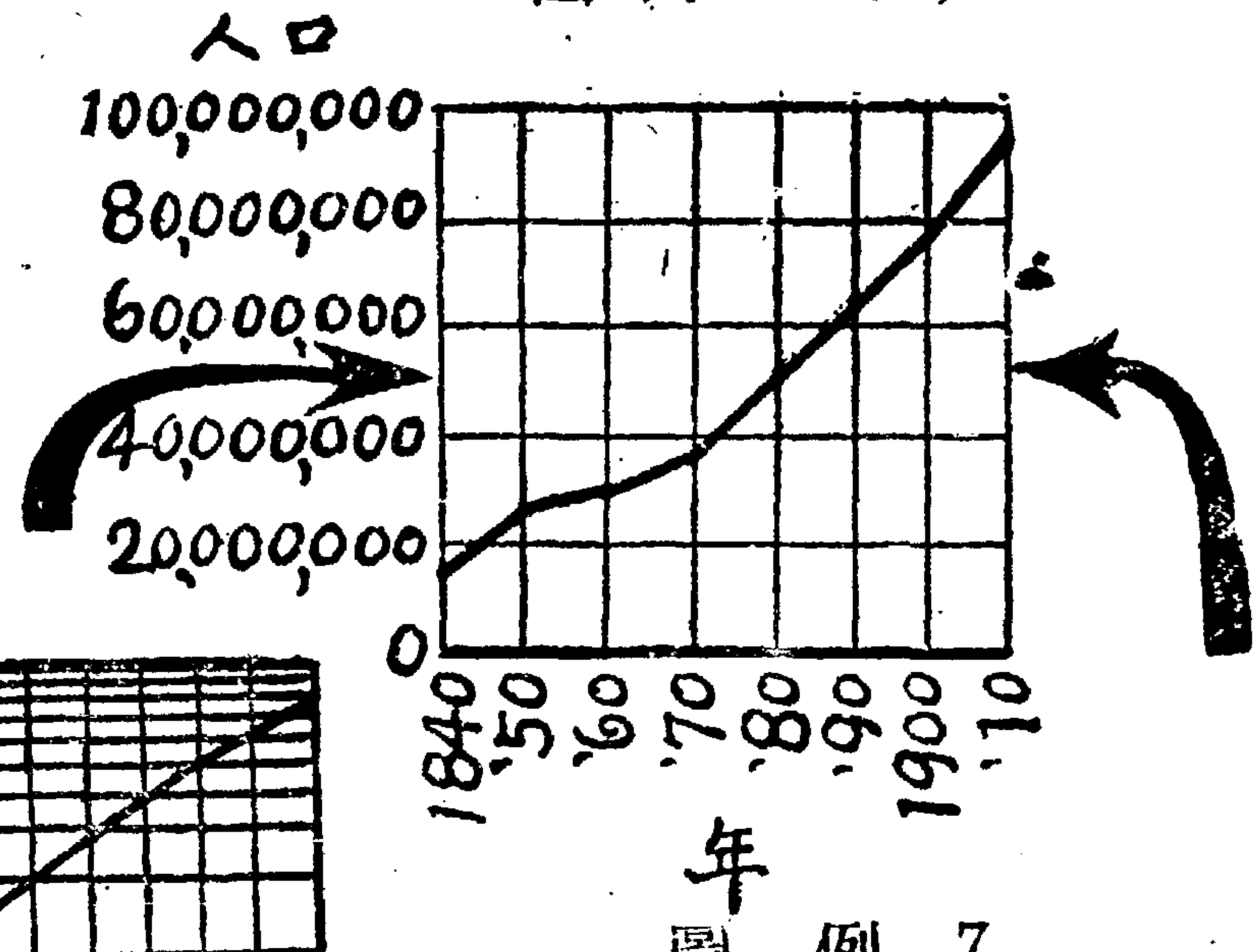


圖例 6. (B)

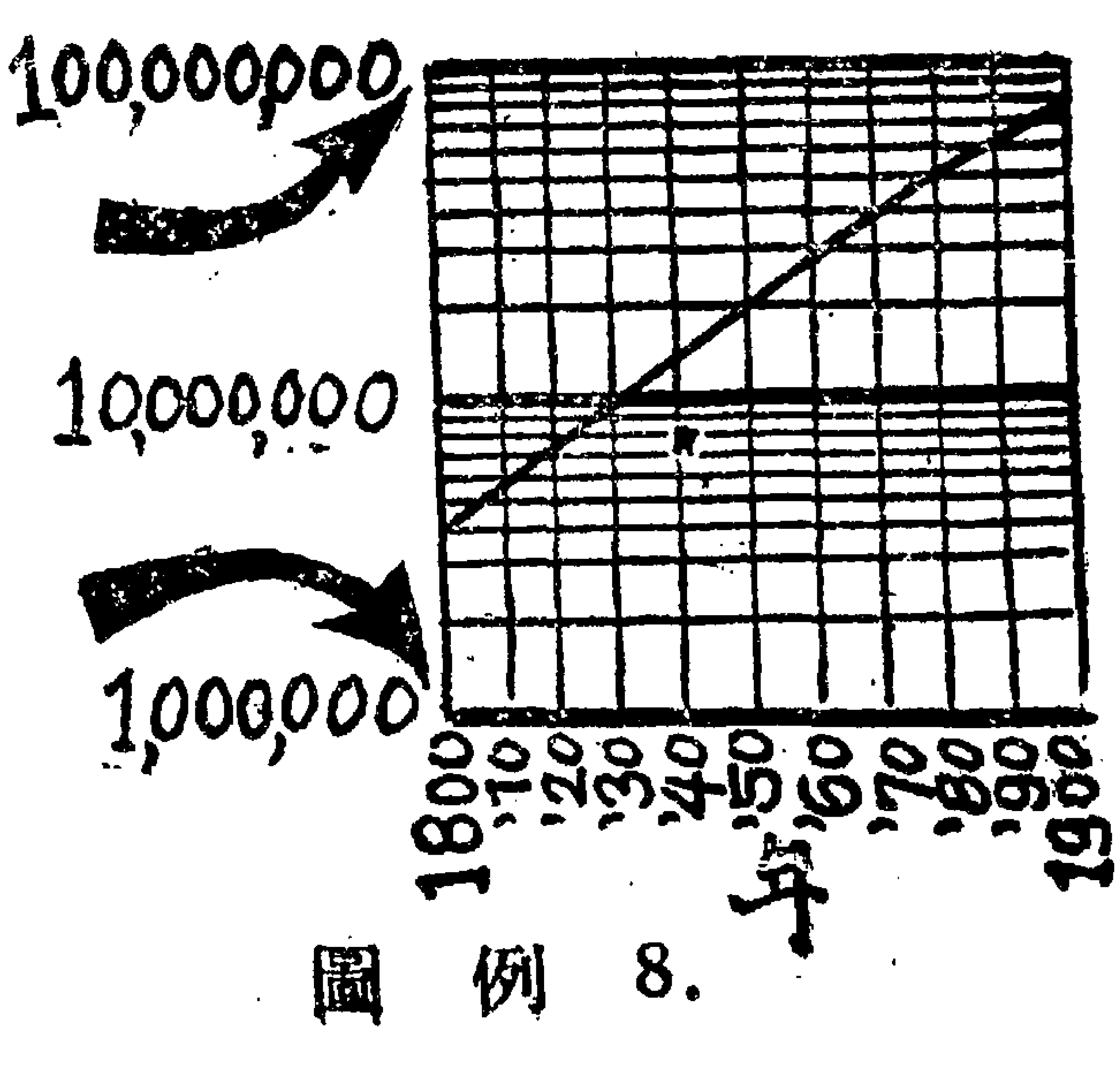


圖例 6. (C)

7. 若圖表示時間者，則兩旁縱界線不宜粗大，以示時間之起終，不能加以限止。



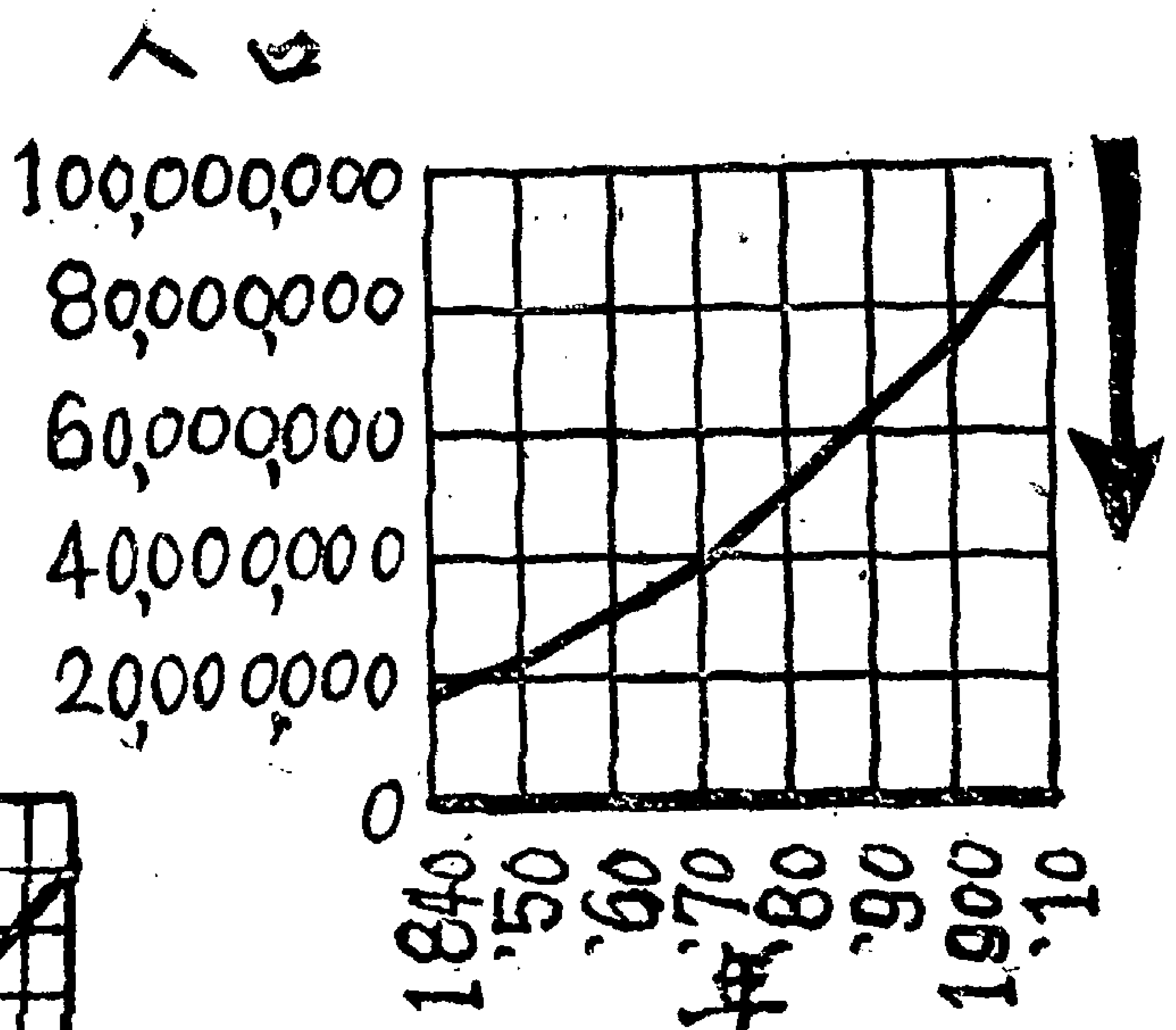
圖例 7.



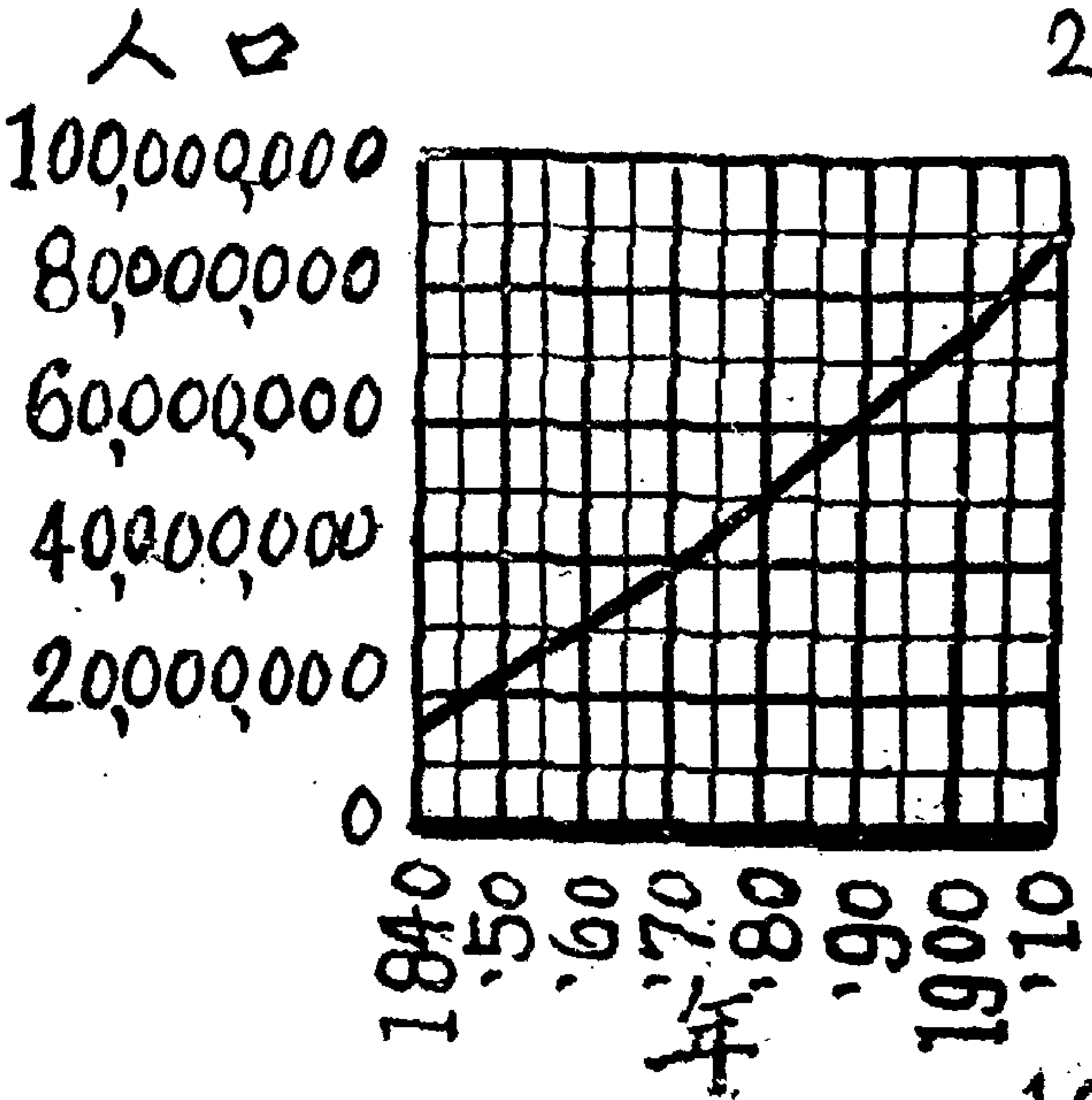
圖例 8.

8. 若曲線畫在對數縱線上，則圖之上下兩界線，應各畫在對數級上十數之某次方。

9. 縱橫線除輔助閱圖所必需外，不宜太多.



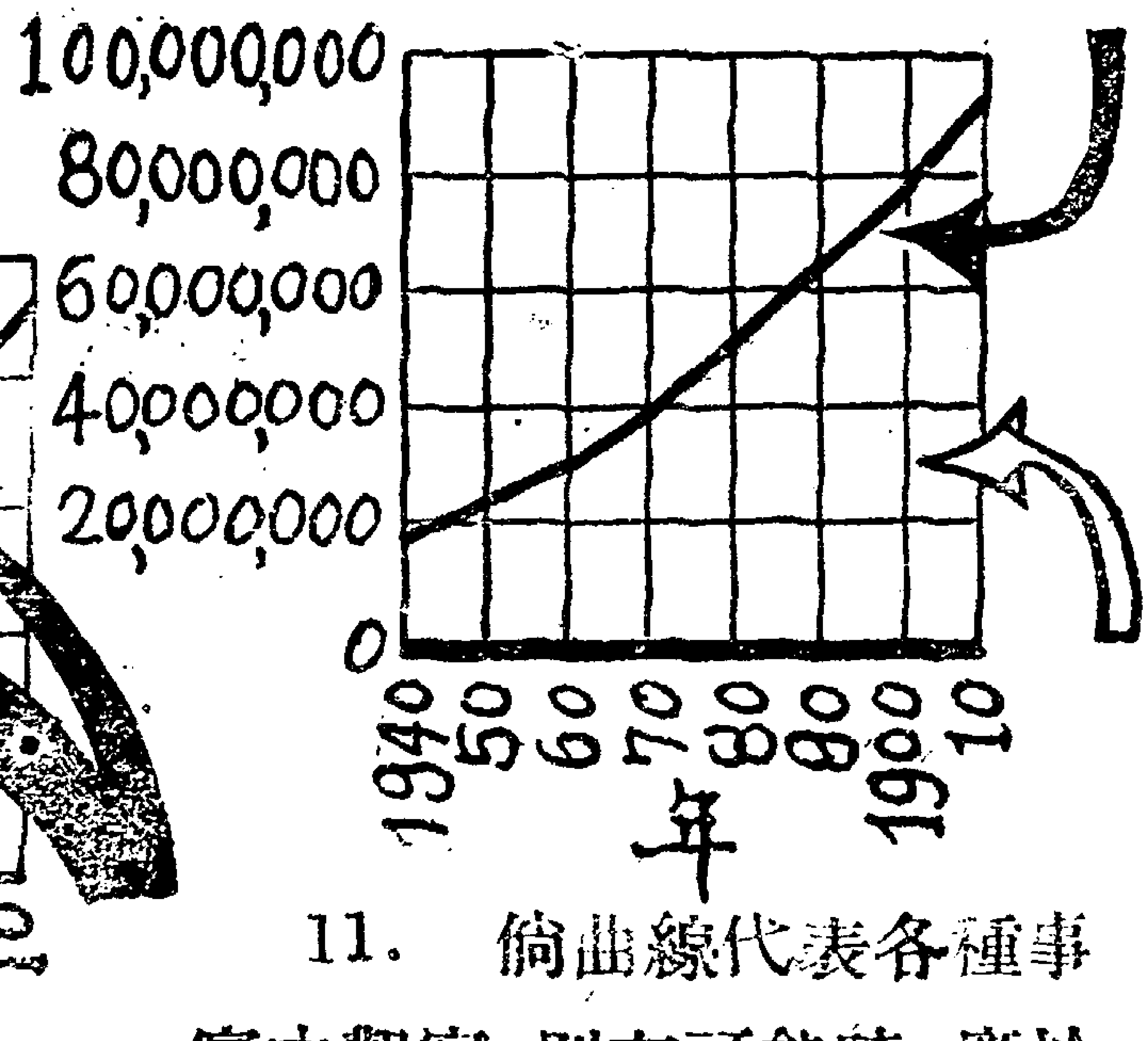
圖例 9. (A)

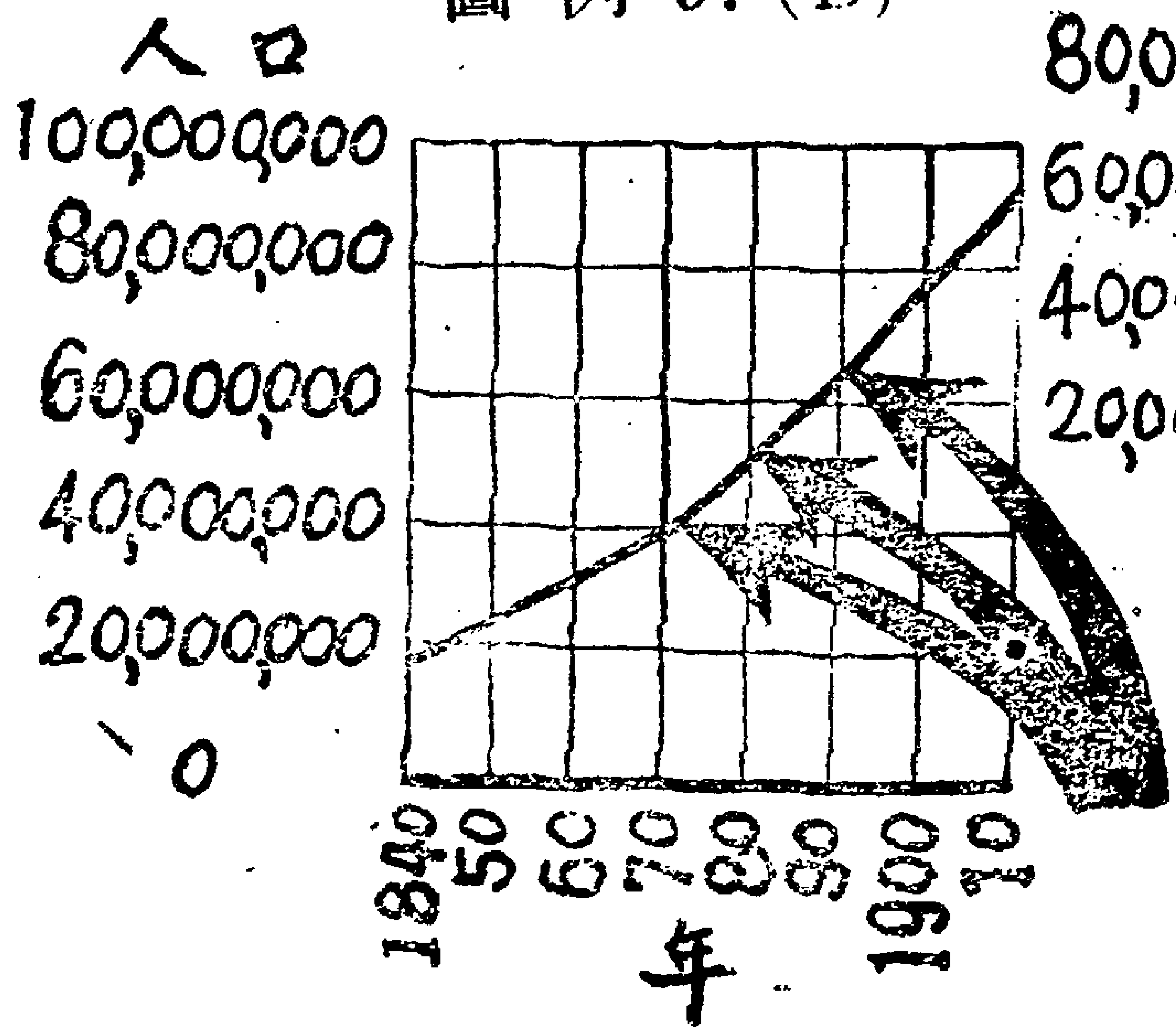


圖例 9. (B)

10. 圖上曲線宜與格線有所區別.

圖例 10.





圖例 11. (A)

11. 倘曲線代表各種事實之觀察，則在可能時，應於曲線上表明此類觀察之點.

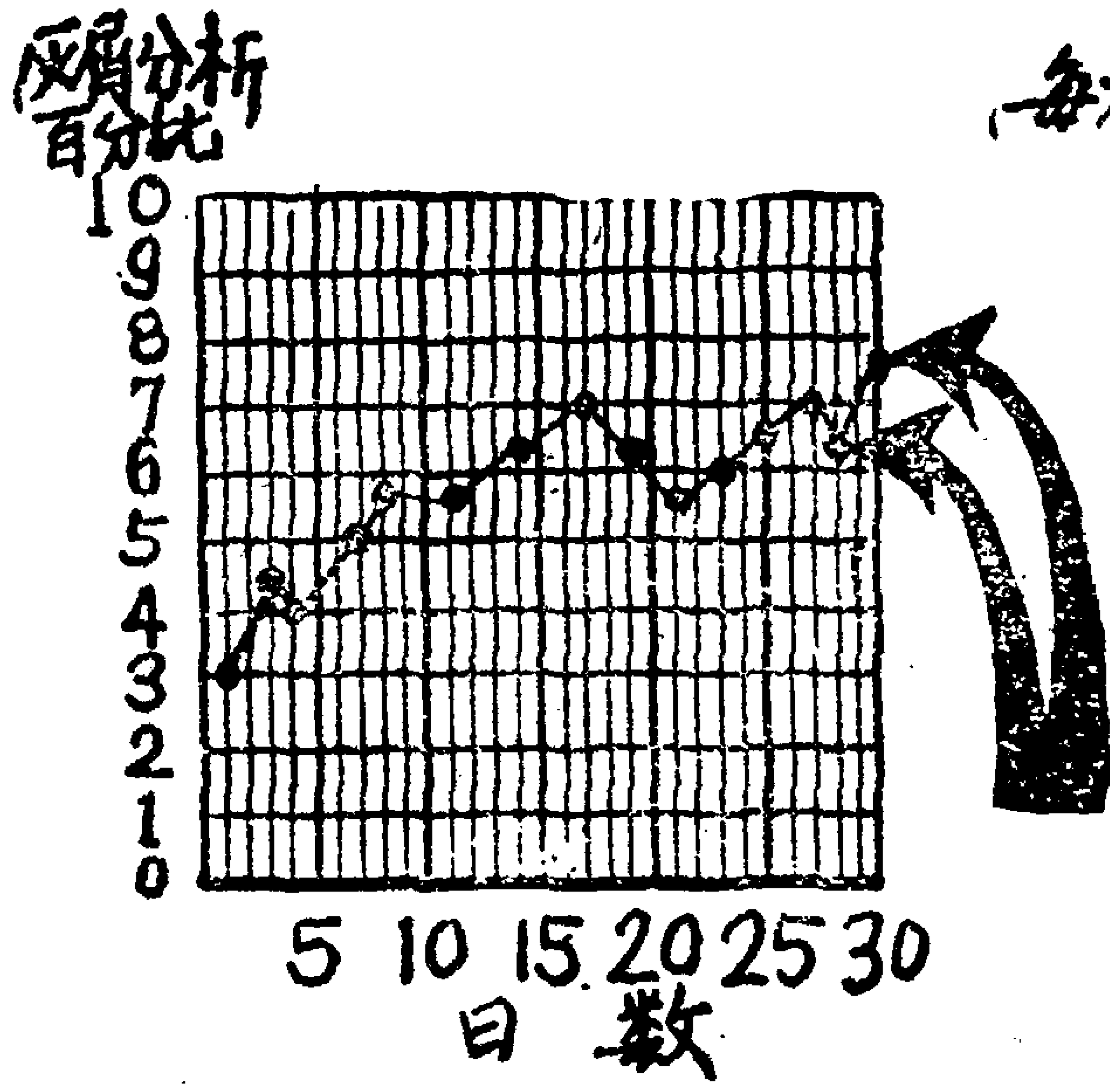


圖 例 11. (B)

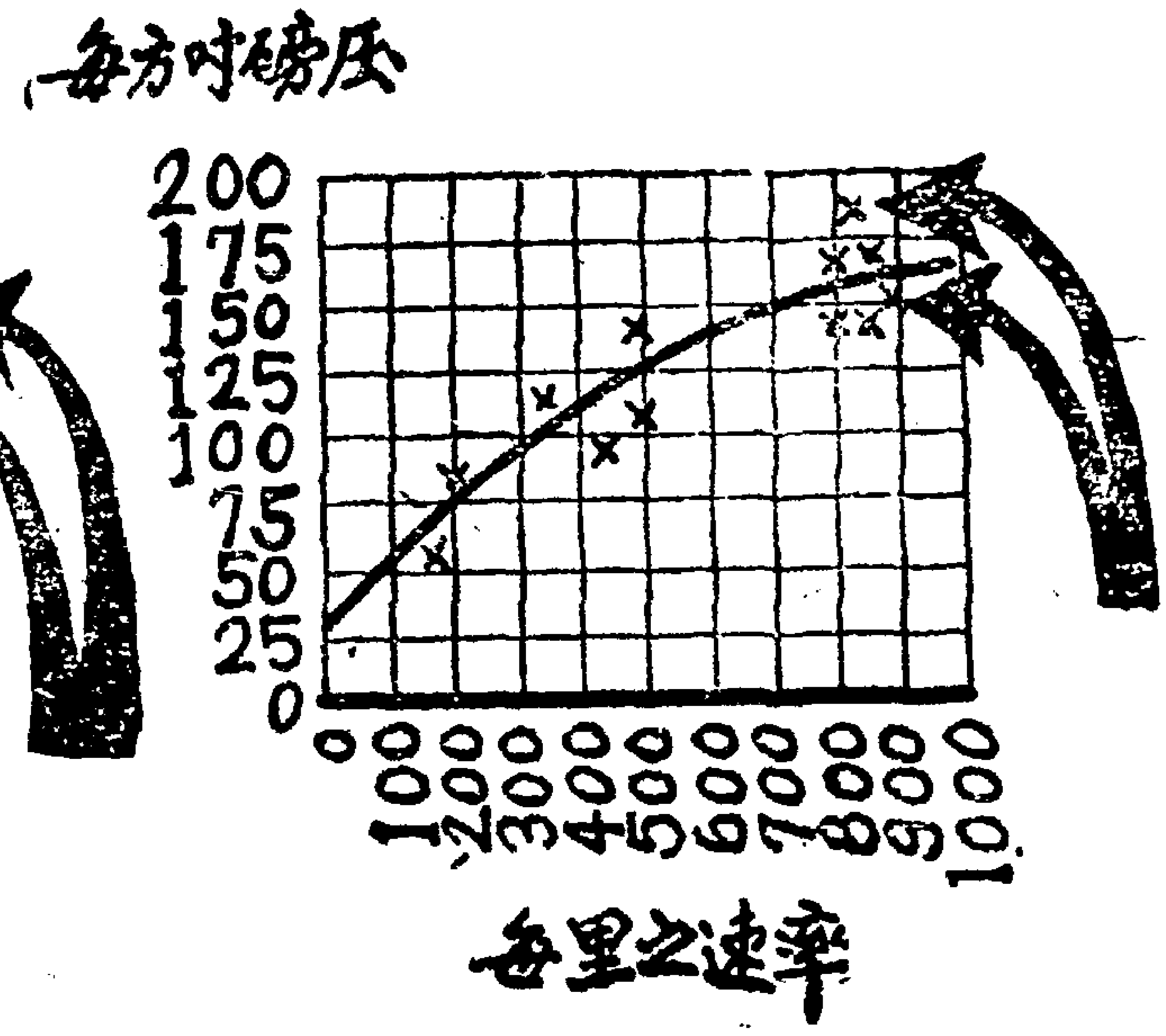


圖 例 11. (C)

12. 圖上橫量表之讀法，宜自左至右，縱量表之讀法自下而上.

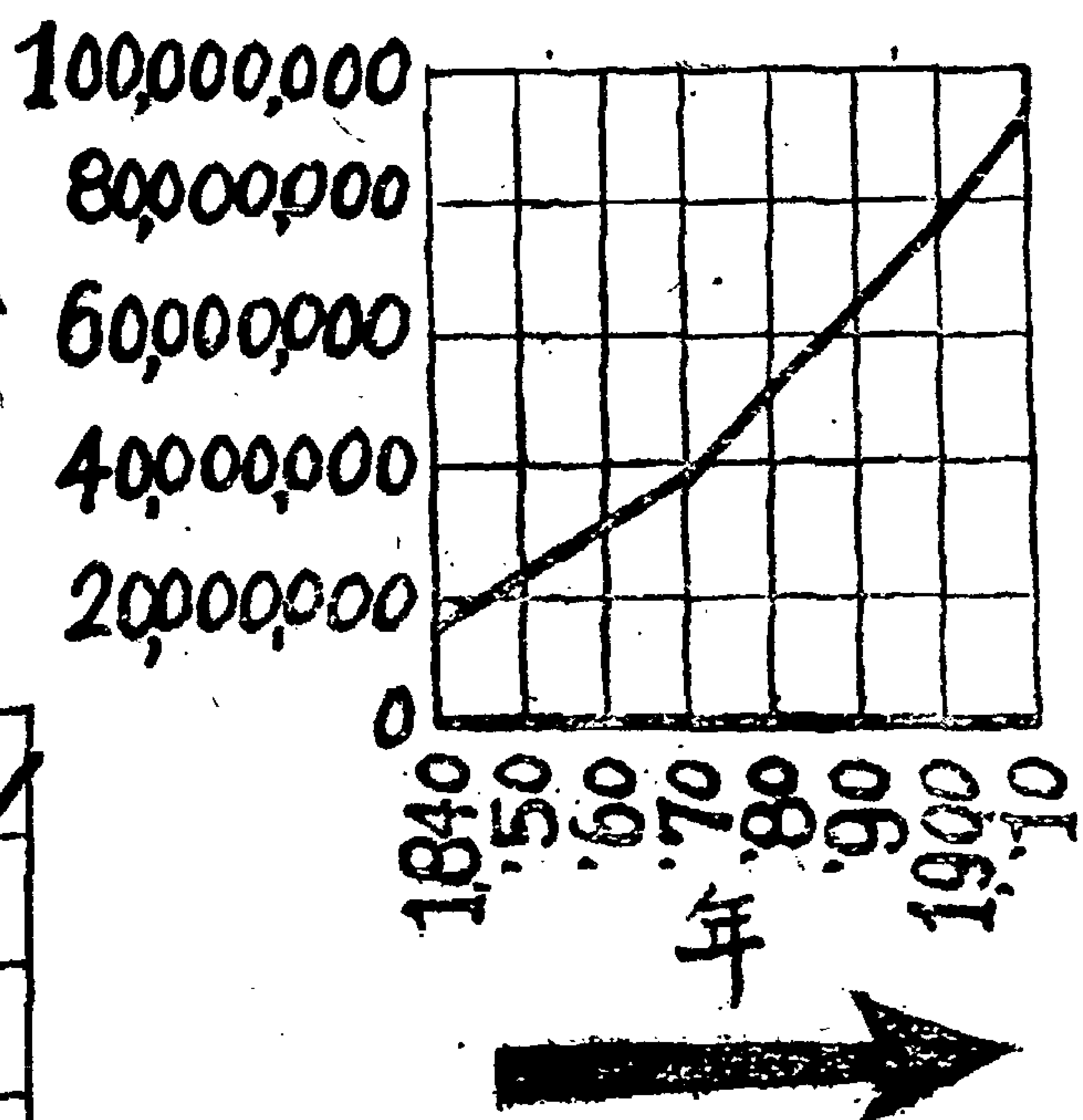


圖 例 12.

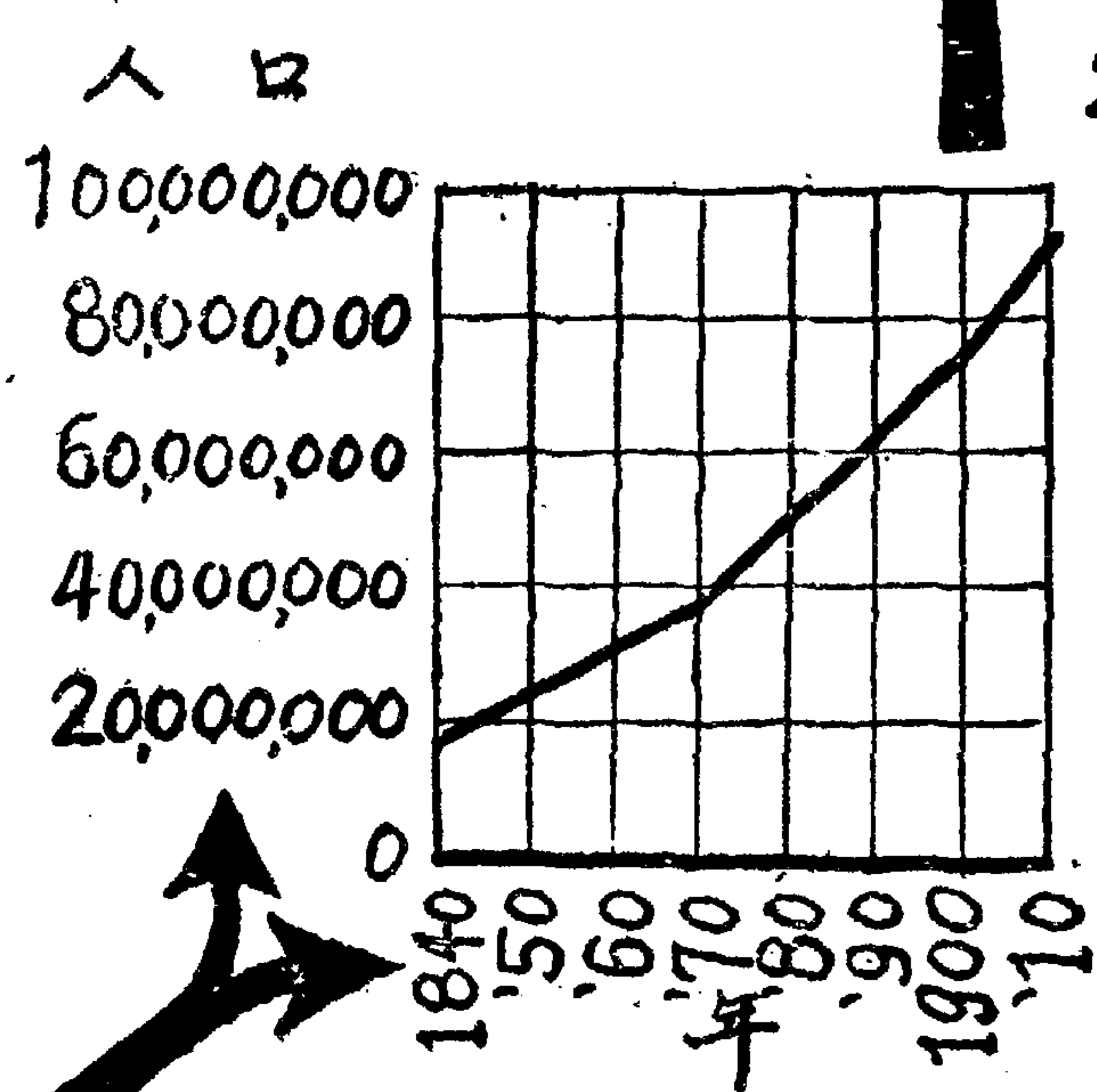
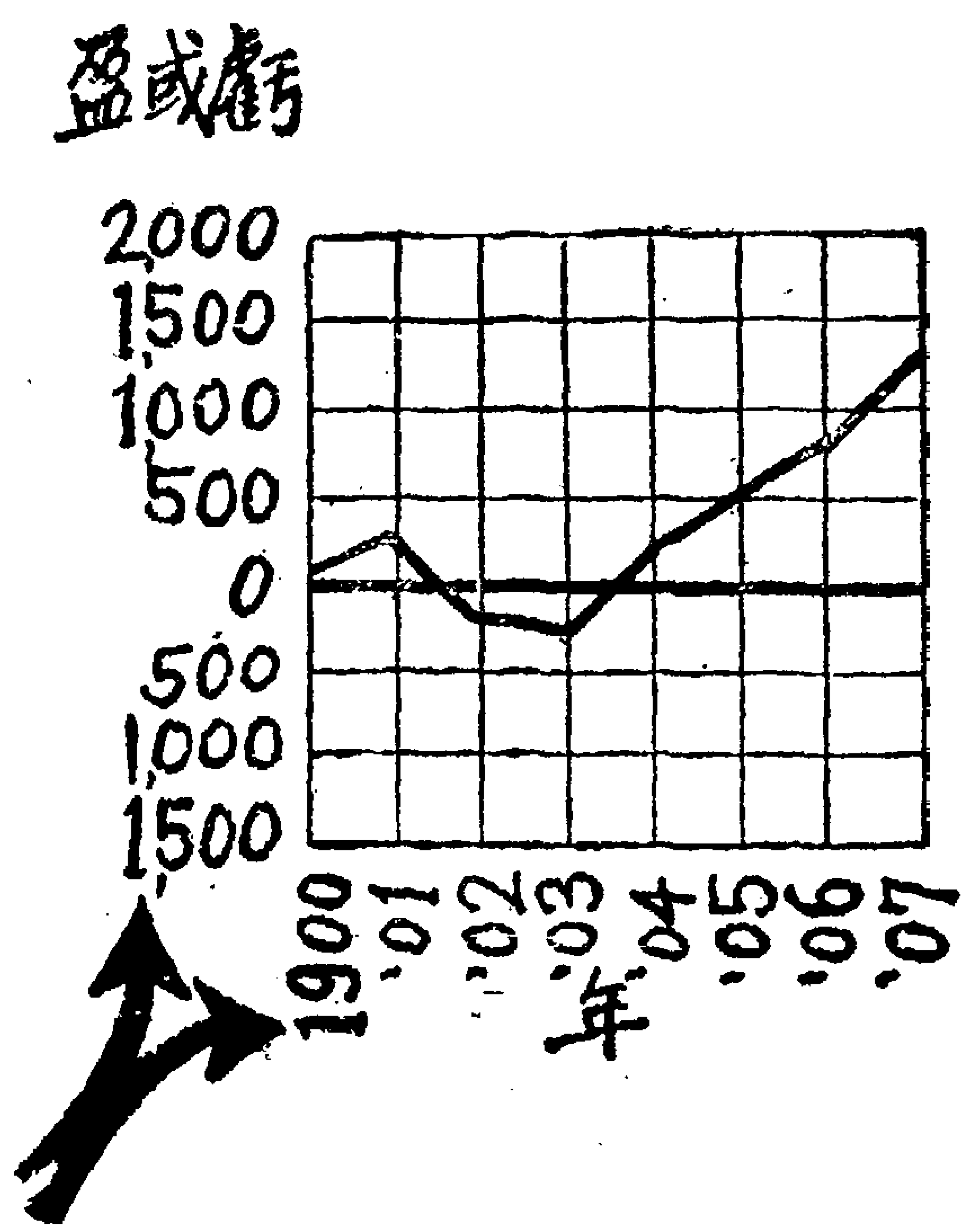
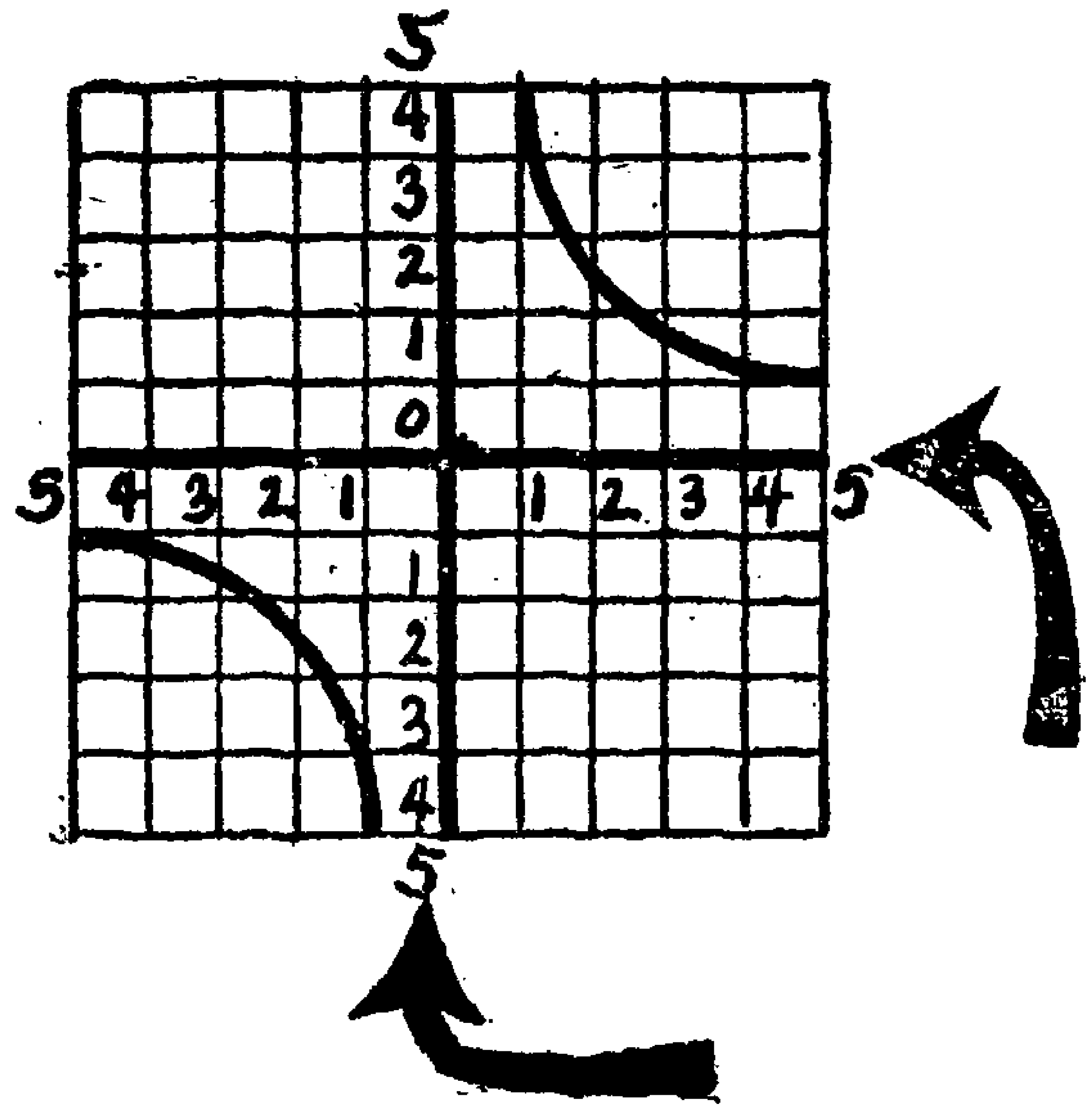


圖 例 13. (A)

13. 量表上之數字，宜放在縱坐標之左，橫坐標之下，或沿縱橫軸上.

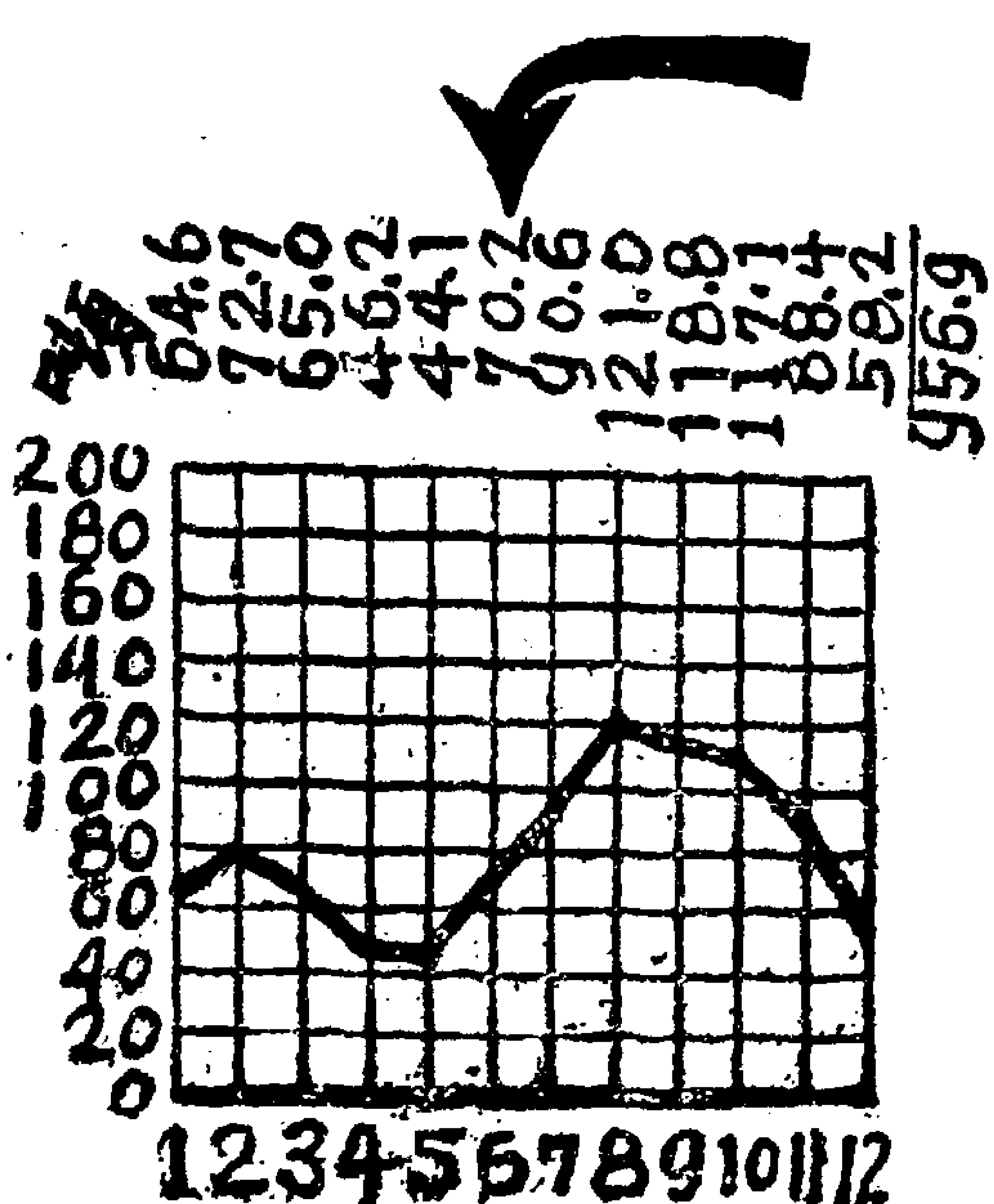


圖例 13. (B)

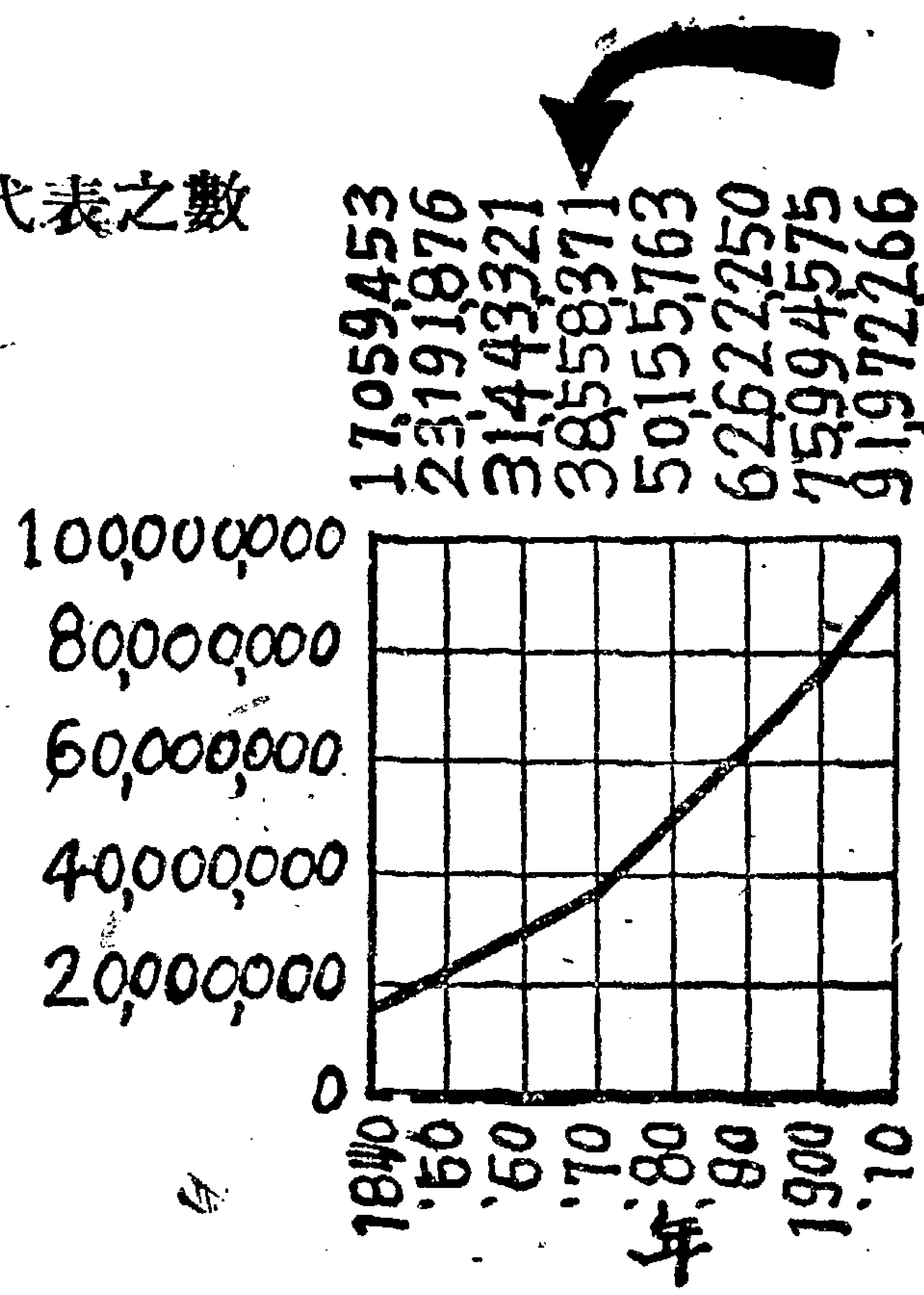


圖例 13. (C)

14. 圖上有時應載所代表之數字的事實或公式.

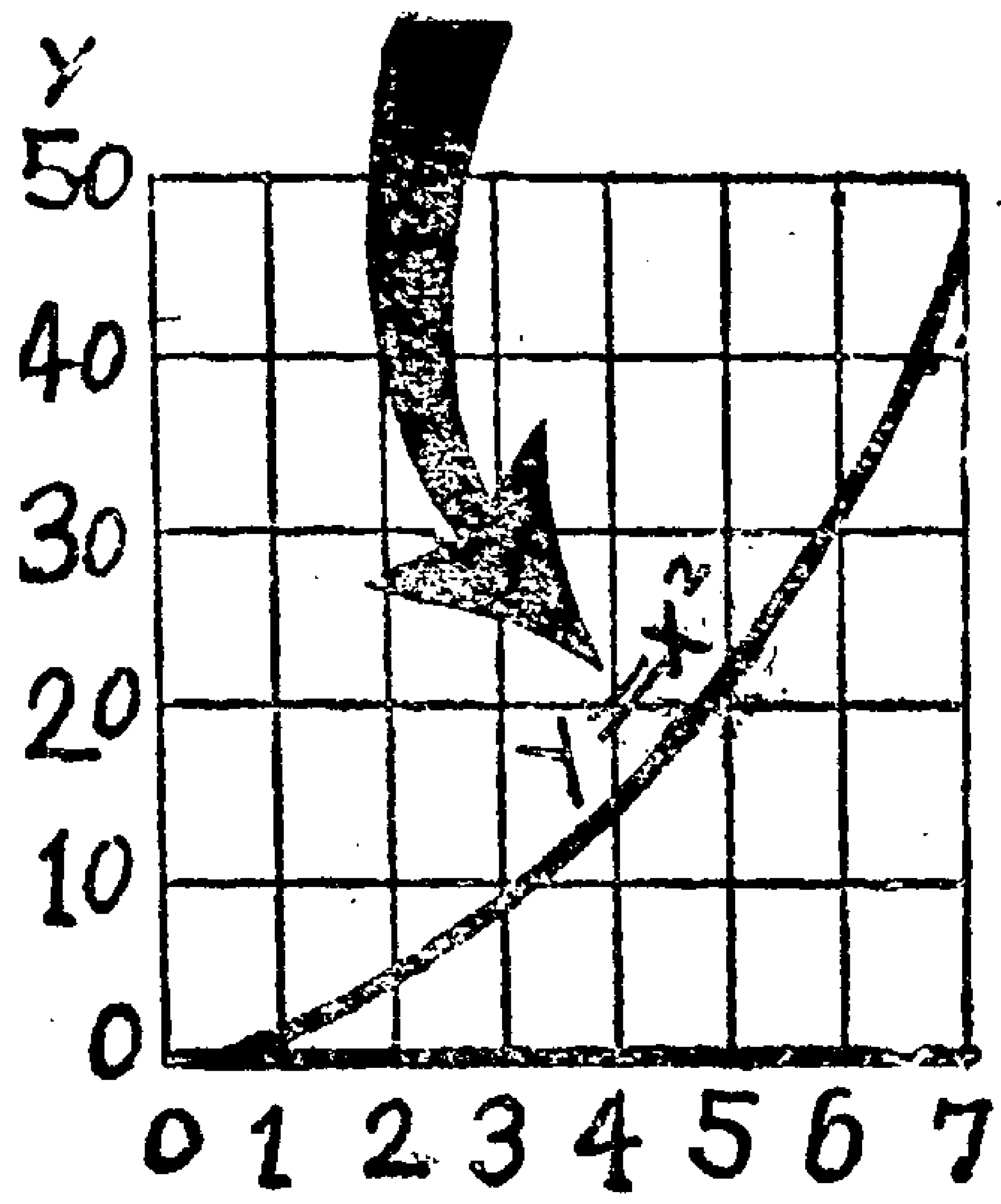


圖例 14. (B)



圖例 14. (A)

15.　若數字之事實，未在圖上表出，可另載一表以伴之.



圖·例 15.（A）

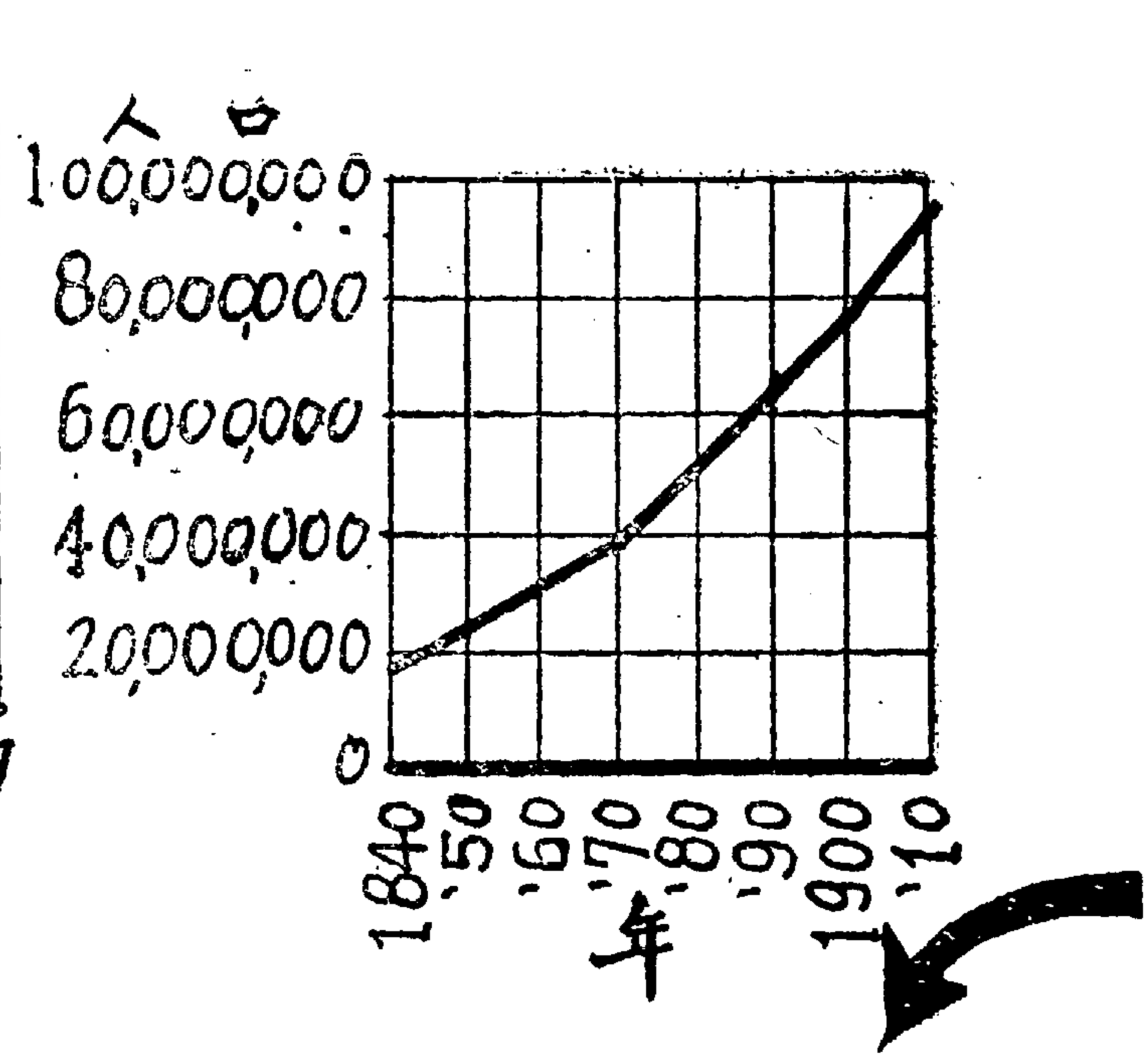


圖 例 15.（B）

年	人口
1840	17,069,453
1850	23,191,876
1860	31,443,321
1870	38,558,371
1880	50,155,783
1890	62,822,250
1900	75,994,575
1910	91,972,266

16. 文字與數字之排列宜使以橫坐標爲底，而閱讀順目；或以右邊縱界線爲底，而閱讀順目.

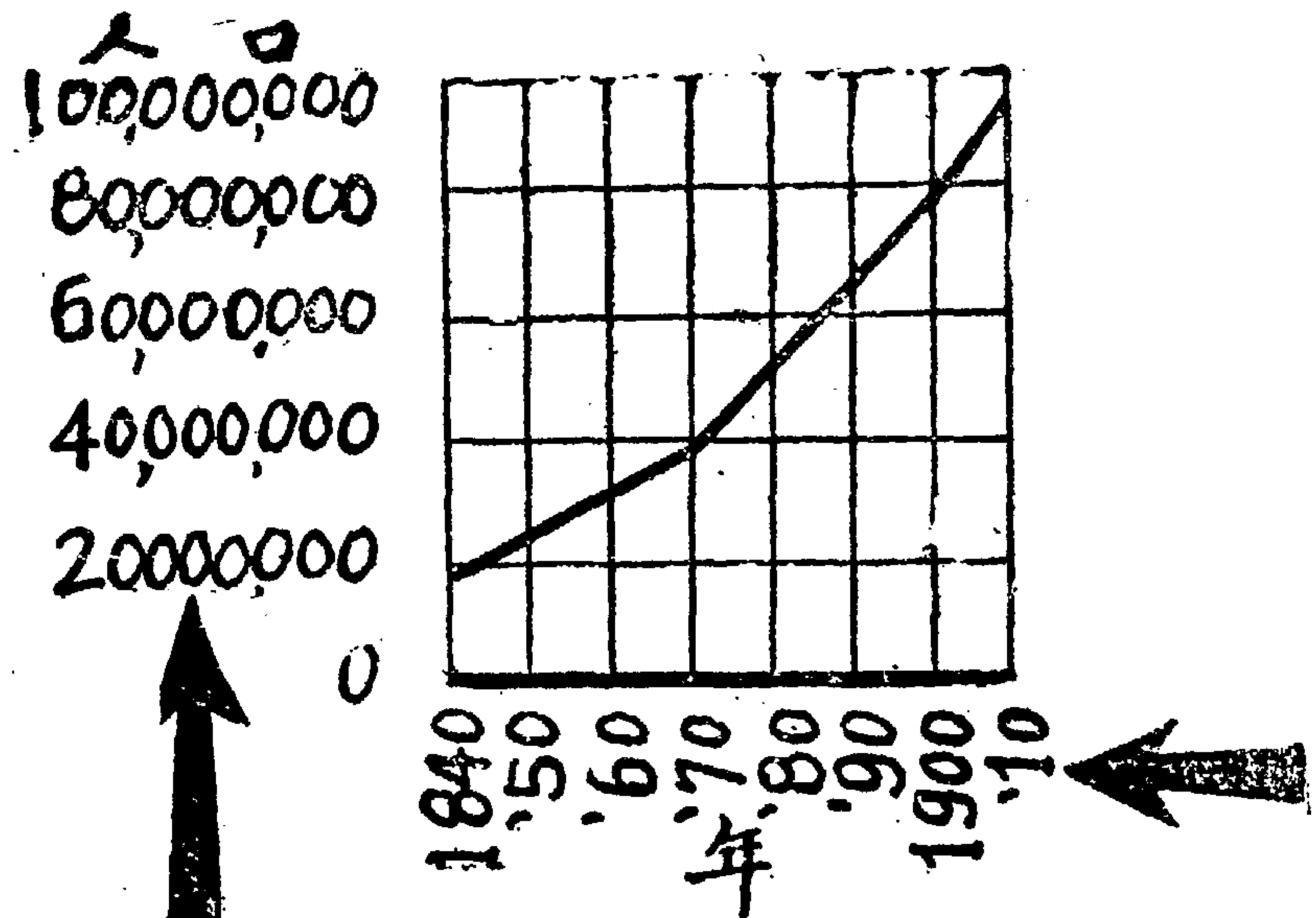


圖 例 16.

17. 圖之題目，宜詳備明晰，遇必要時不妨多加子目或說明使其更爲清晰.

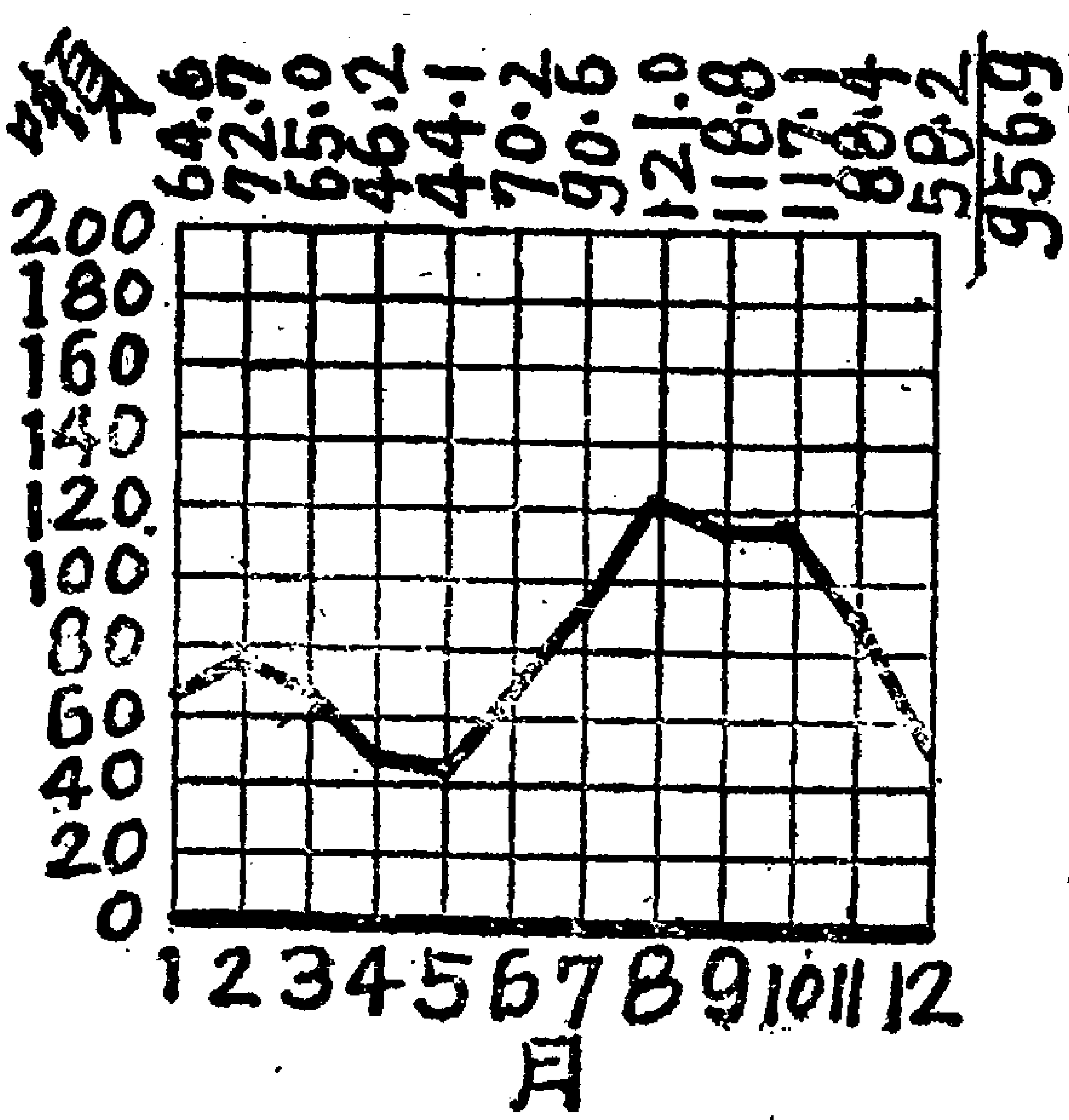


圖 例 17.

第二工廠在一九一四年按月之鋁質鑄造物產額.產額以短噸計算.碎鋁之銷售不計在內.

附 錄 一

主要進修與參考中西文統計書目

王仲武:	統計學原理及應用	商務印書館
王仲武:	統計公式及例解	商務印書館
王書林:	教育統計學	商務印書館
艾　偉:	高級統計學	商務印書館
朱君毅:	心理與教育之統計法	商務印書館
朱君毅:	教育統計學	商務印書館
朱君毅:	教育統計學綱要	商務印書館
朱君毅:	統計方法大綱	正中書局
朱君毅:	教育測驗與統計	商務印書館
李植泉:	統計學原理	商務印書館
芮寶公:	統計學概論	中華書局
金國寶:	統計學大綱	商務印書館
金國寶:	統計學	商務印書館
唐啓賢:	統計學	黎明書局
陳善林:	統計學	中華書局
寧恩承:	統計方法	商務印書館
褚一飛:	統計學概論	立信會計圖書用品社

鄭堯柈: 統計學 商務印書館
羅大凡: 數理統計 北新書局

Camp, Burton H., Mathematical Part of Elementary Statistics, D. C. Heath and Co., New York, 1931.

Chaddock, Robert E., Principles and Methods of Statistics, Houghton Mifflin Co., New York, 1925.

Croxton, F. F., and Cowden, D. J., Practical Business Statistics, Prentice-Hall Inc., New York, 1934.

Crum, William L. and Patton Alston C., An Introduction to Economic Statistics, Mcgraw-Hill Book Co., New York, 1925.

Davies, George R. and Crowder, Walter F., Methods of Statistical Analysis in Social Sciences, John Wiley and Sons Inc., New York, 1933.

Day, Edmund E., Statistical Analysis, Macmillan Co., New York, 1930.

Garrelt, Henry E., Statistics in Psychology and Education, Longmans Green and Company, New York, 1926.

Harper, F., Elements of Practical Statistics, Macmillan Co., New York, 1930.

Holzinger, Karl J., Statistical Methods for Students in Education, Ginn and Company, New York, 1928.

Jerome, Harry, Statistical Method, Harper and Bros., New York, 1924.

Kelley, Truman L., Statistical Method, Macmillan Co., New York, 1923.

Mills, Frederick C., Statistical Methods, Henry Holt and Co., New York, 1924.

Odell, Co., W., Statistical Method in Education, D. Appleton-Century Co., New York, 1935.

Richardson, C. H., An Introduction to Statistical Analysis, Harcourt, Brace and Co., New York, 1931.

Riggleman, John R., and Frisbee, Ira N., Business Statistics, McGraw-Hill and Co., New York, 1932.

Rugg, Harold O., Statistical Methods Applied to Education, Houghton Mifflin Co., New York, 1917.

Thurstone, L. L., Fundamentals of Statistics, Macmillan Co., New York, 1925.

Yule, G. Udny, An Introduction to the Theory of Statistics, Charles Griffen & Co. Ltd., London, 1929.

註：西文書目，係根據 Herbert Arkin and Raymond R. Colton 著 An Outline of Statistical Methods 書中所列之十八種標準統計教科書.

附錄二

本書統計名詞漢英對照表

橫坐標	Abscissa
材料之分析	Analysis of data
算術平均數	Arithmetic average (or mean)
假定平均數	Assumed average (or mean)
不對稱分配	Asymmetrical (or Skewed) distribution
平均數	Average
平均差	Average (or mean) deviation
鐘形分配	Bell-shaped distribution
破裂紋	Break
集中趨勢	Central tendency
機遇	Chance
機變	Chance variation
圖	Chart; diagram; graph
材料之分配	Classification of data
組距	Class interval
相關係數	Coefficient of correlation
材料之蒐集	Collection of data
縱行	Column

連續數列	Continuous series
校正數	Correction
相關	Correlation
品質相關	Correlation of attributes
累積次數	Cumulative frequency
累積次數表	Cumulative frequency table
曲線	Curve
離差	Deviation
不連續數列	Discrete series
經驗法	Experimental method
極端量數	Extreme measure
極不對稱分配	Extremely asymmetrical distribution
偶數	Even number
第一四分位點	First quartile
次數分配	Frequency distribution
次數多邊圖	Frequency polygon
次數數列	Frequency series
甘氏進度圖	Gantt progress chart
幾何平均數	Geometric mean
圖示法	Graphic method
圖示	Graphic presentation

歸類	Grouping
調和平均數	Harmonic mean
高度相關	High correlation
直方圖	Histogram (or Bar chart)
指數	Index number
項目	Item
J形分配	J-shaped distribution
較小	'Less than'
直線相關	Linear correlation
對數尺數	Logarithmic scale
通常法	Long method
羅倫氏曲線	Lorenz curve
低度相關	Low correlation
下限	Lower limit
數理曲線	Mathematical curve
中位數	Median
估計法	Method of estimate
組中點	Mid-point of class interval
衆數	Mode
微偏態	Moderately asymmetrical
較大	'More than'
負偏態	Negative skewness

常態分配	Normal distribution
觀察法	Observation method
實得相關	Obtained correlation
奇數	Odd number
縱坐標	Ordinate
皮而生氏積差相關	Pearson's product moment correlation
百分位數	Percentile
正偏態	Positive skewness
材料之顯示	Presentation of data
機誤	Probable error
乘積	Product
預測	Prognosis
純相關	Pure correlation
數量之材料	Quantitative data
四分位差,或二十五分差	Quartile deviation, or Semi-interquartile range
全距,或兩極距	Range
相對離差	Relative deviation
可靠性	Reliability
代表數	Representative measure
橫行	Row

分布圖	Scalter diagram
簡捷法	Short method
偏態分配	Skewed distribution
偏斜度	Skewness
標準差	Standard deviation
統計資料	Statistical data
統計數列	Statistical series
統計技術	Statistical technique
統計學	Statistics
品質統計	Statistics of attributes
變數統計	Statistics of variables
切實相關	Substantial correlation
總和	Summation
對稱分配	Symmetrical distribution
製表法	Tabulation method
第三四分位點	Third quartile
標題	Title
眞正相關	True correlation
U形分配	U-shaped distribution
上限	Upper-limit
加權平均數	Weighted average
零線	Zero line

附表一

常態曲線之縱線表

本表爲常態曲線圖內，各橫坐標距離上之縱線長度，各縱線以中間之最高縱線

$$y_0=\frac{N}{\sigma\sqrt{2\pi}}=\frac{N}{2.5066\sigma}$$

爲單位；橫坐標距離以標準差 σ 爲單位，設各縱線之長度爲 y；各橫坐標之距離爲 x，且以中間之 $x=0$.

本表之用法如下，設橫坐標距離 $x=0.3\sigma$ 時，則縱線 $y=0.95600y_0$；橫坐標距離 $x=1.4\sigma$ 時，則縱線 $y=0.37531y_0$. 餘類推.

x/σ	0	2	2	3	4	5	6	7	8	9
0.0	10000	99995	99[illegible]80	99[illegible]55	99920	99375	99820	99755	9[illegible]685	99[illegible]93
0.1	99501	99396	99283	99158	99025	98881	98723	99565	98393	98211
0.2	98020	978[illegible]9	97609	97390	97161	96923	96676	95420	96156	958[illegible]2
0.3	95600	95309	95010	94702	94387	94355	93723	93382	93024	92677
0.4	92312	91399	91558	91169	9[illegible]774	90371	89961	89543	89119	88688
0.5	88250	87805	87353	86896	86432	85962	85488	85006	84519	8406[illegible]
0.6	83527	83023	82514	82010	81481	80957	8[illegible]409	79896	79359	78817
0.7	78270	77721	77167	76610	76048	75484	74916	74[illegible]42	73769	73193
0.8	72615	72033	71448	70861	7[illegible]272	69681	69087	68493	67896	67298
0.9	66689	66097	65494	64891	64287	63683	63077	62472	618[illegible]5	61259
1.0	60[illegible]53	60047	59440	58831	58222	57623	57017	56414	55810	55209
1.1	54607	54[illegible]07	53409	5[illegible]812	52214	51620	51027	50437	4[illegible]848	49260
1.2	48675	48092	47511	46933	46357	45783	45212	44644	44078	43516
1.3	42956	423[illegible]9	41845	41294	4[illegible]747	40202	39661	39123	38569	38058
1.4	37531	37[illegible]07	36487	35971	35459	34950	34445	33944	33447	32954

1.5	32465	31930	31500	31023	30550	30082	29618	29158	28702	28251
1.6	27804	27361	26923	26489	26059	25634	25213	24797	24385	23978
1.7	23575	23176	22782	22392	22008	21627	21251	20879	20511	20148
1.8	19790	19436	19086	18741	18400	18064	17732	17404	17081	16762
1.9	16448	16137	15831	15530	15282	14939	14650	14364	14083	13806
2.0	13534	13265	13000	12740	12483	12230	11981	11737	11496	11259
2.1	11025	10795	10570	10347	10129	09914	09702	09495	09290	09090
2.2	08392	08698	08507	08320	08136	07956	07778	07604	07433	07265
2.3	07100	06939	06780	06624	06471	06321	06174	06029	05888	05750
2.4	05614	05481	05350	05222	05096	04973	04852	04734	04618	04505
2.5	04394	04285	04179	04074	03972	03873	08775	03680	03586	03494
2.6	03405	03317	03232	03148	03066	02986	02908	02831	02757	02684
2.7	02612	02542	02474	02408	04343	02280	02218	02157	02098	02040
2.8	01984	01929	01876	01823	01772	01723	01674	01627	01581	01536
2.9	01492	01449	01408	01367	01328	01288	01252	01215	01179	01145
3.0	01111	00819	00598	00432	00309	00219	00153	00106	00073	00050
4.0	00034	00022	00015	00010	00006	00004	00003	00002	00001	00001
5.0	00000									

附表二

常態曲線下之面積表

本表為由常態曲線中間(即 $x=0$)至各指定橫坐標距離內常態曲線之面積，各段面積以全面積為單位，橫坐標距離以標準差為單位.

本表之用法如下：設 x 由 0 至 0.3σ 時，則曲線所包括之面積＝全面積之 $\frac{1179}{10000}$ 或 0.1179；x 由 0 至 1.4σ 時，則曲線所包括之面積＝全面積之 $\frac{4192}{10000}$ 或 0.4192，餘類推.

x/σ	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	0000	0040	0080	0120	0159	0199	0239	0279	0319	0359
0.1	0398	0438	0478	0517	0557	0596	0636	0675	0714	0753
0.2	0793	0832	0871	0910	0948	0987	1026	1064	1103	1141
0.3	1179	1217	1255	1293	1331	1368	1406	1443	1480	1517
0.4	1554	1591	1628	1664	1700	1736	1772	1808	1844	1879
0.5	1915	1950	1985	2019	2054	2088	2123	2157	2190	2224
0.6	2257	2291	2324	2357	2389	2422	2454	2486	2518	2549
0.7	2580	2612	2642	2673	2704	2734	2764	2794	2823	2852
0.8	2881	2910	2939	2967	2995	3023	3051	3078	3106	3133
0.9	3159	3186	3212	3238	3264	3289	3315	3340	3365	3389
1.0	3413	3438	3461	3485	3508	3531	3554	3577	3599	3621
1.1	3643	3665	3686	3718	3729	3749	3770	3790	3810	3830
1.2	3849	3869	3888	3907	3925	3944	3962	3980	3997	4015
1.3	4032	4049	4066	4083	4099	4115	4131	4147	4162	4177
1.4	4192	4207	4222	4236	4251	4265	4279	4292	4306	4319
1.5	4332	4345	4357	4370	4382	4394	4406	4418	4430	4441
1.6	4452	4463	4474	4485	4495	4505	4515	4525	4535	4545
1.7	4554	4564	4573	4582	4591	4599	4608	4616	4625	4633
1.8	4641	4649	4656	4664	4671	4678	4686	4693	4699	4706
1.9	4713	4719	4726	4832	4738	4744	4750	4758	4762	4767
2.0	4773	4778	4783	4783	4793	4798	4803	4808	4812	4817
2.1	4821	4826	4830	4834	4838	4842	4846	4850	4854	4857
2.2	4861	4865	4868	4871	4875	4878	4881	4884	4887	4890
2.3	4893	4896	4898	4901	4904	4906	4909	4911	4913	4916
2.4	4918	4920	4922	4925	4927	4929	4931	4932	4934	4936
2.5	4938	4940	4941	4943	4945	4946	4948	4949	4951	4952
2.6	4953	4955	4956	4957	4959	4960	4961	4962	4963	4964
2.7	4965	4966	4967	4968	4969	4970	4971	4972	4973	4974
2.8	4974	4975	4976	4977	4977	4978	4979	4980	4980	4981
2.9	4981	4982	4983	4984	4984	4984	4985	4985	4986	4986
2.0	4986	4987	4987	4988	4988	4988	4989	4989	4989	4990

中華民國三十七年三月初版

統計學概要

全一册　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）

整：書
校：

滬・本　　2/1—0.15