高級統計學

Advanced Statistics

萬青館存

章校長教正

編者 [illegible]曾 卅六、六、

高級統計学 P.1.

Advanced Statistics

緒論

宇宙現象不論自然的或~~变动的~~社會的都是变动不居而這个变化在我们居住的這个时代来得特别激烈，不管是在生產方法中，在生產関係中，在社會組織形態中，最後在我们的理論的認識觀念中，這激烈的变动仍然是前撲後継，不断的发生着，社會发展動盪的結果，已把正常的東西，安定的東西，神的東西，宿命的東西，以及依常態曲線而不散的東西，等々觀念顯露出不对称的集团，非典型的平均及变化系統的現实世界要求統計学者对于這些複雜而不規則的現象的测定方法的特性加以研究。

一切現象是受许多複雜原因直接或間接的影响错综複

雜，互相排斥互相關联着，有其因果規律的存在，並非決定於"盲目的神灵"或偶然的意志，這是必然的中心觀念與思想，亦即科學研究的主要原則。吾人研究事物即根據此原則常欲除去許多紛雜的微小原因，而集中於幾个主因；但紛雜的原因常不易除去而仍存在着，為便利研究起見，吾人把原因析為恒常原因與偶然原因兩部份。恒常原因是主要的基本的，如果將這部原因除掉則结果便會变质。偶然原因是表面的附随的，這部份原因無论其有無与否，对於结果不发生本质的影响。這兩种原因交互错综地作用到一切現象上去。無论那种現象都受這兩种原因的影响。

有些現象恒常原因之作用強，偶然原因之作用弱，則其各个结之间差异不大，是謂典型現象。反之若偶然原因之作用強，恒常原因被蒙蔽，其各个结果之间差异很大，是謂之非典型現

种研究方法就是大量观察法也就是统计方法.对这种由大量
个体集合而成的集团现象施以研究时,则在个体现象中,所不
能看到的现象性状,个体现象中所不能显出的规律,即可在这
个集团现象中显现出来,因为个体现象既然主要係受着偶然
原因的作用,则这些偶然原因作用,在个体上是呈现着种种不
同方向不同强度的形態,而当把这些个体大量集合起来观察
时这些不同方向不同强度的偶然原因作用,即相互抵銷而减
弱,於是在个体上被隐蔽的恒常原因作用,此时得以显现,这样
显现出来的恒常原因作用,不外就是这个集团中所存的规律
性,统计方法就是用大量观察从非典型现象的集团中得求规
律性,自然界中有许多现象属於典型的,故大多用实验方法,但
其中有实验方法未能尽量研究者亦復不少,如氣象遗传等是
也,至一般社会现象大都是非典型现象,且不能实验,则惟用统

高级统计学 p.2.

象,因为典型现象既然都是主要受着恒常原因的作用,偶然原因不过只是附随的次要的影响,所以对受着同种恒常原因作用的典型现象,只须究其个体研究,便可看出这种现象中所有的恒常原因作用,此时所用的实在研究方法,与该时典型现象的个体拿来,施一种人工的装置,把个体中的偶然作用的原因除掉,析出其恒常原因作用,这个方法就是"实验",但是对于研究非典型现象实验方法便没有效果,因为非典型现象都是主要受着偶然原因作用的影响,所以各呈个个殊异的状态,只就这些现象个体的研究,不能看出其中所有的恒常原因作用,此时须另想他法,这种方法是将许多同具有某种特质的非典型现象之个体汇集起来,看做一个大量的集团,对这种集团加以研究,以求具有同一特质的现象中所存的恒常原因作用,这

計法去而研究.

簡明確切為科学必備条件,而经济的研究方法,又為科学之基本要求,簡明確切莫过於数量,晚近科学研究都向着数量化的大道邁進,实為今日科学長足進展的由来,科学必備的精神重在实証,原理与法則的確立,须據真实証明,而可靠的統計資料,即為最佳的实証.統計中之抽樣方法,是最经济的科学方法,此法近来得诸先輩的发揚光大,而愈精深而合实用,而致用尤廣.此近来科学界研究運用統計方法研究事物之獲得極大的效果,故統計的数字的分析在科学範圍內有這樣的深入,這樣的在很多科学部门內占着重要的地位,然統計学者的企求並不以此為满足,他们对於統計正如同其他科学一樣,雖然主張其獨立活動的权利,却对各种科学的问题,企圖僅依統計学之

助而得解决，未代替科学研究法而单独的利器，即在如何利用取得之材料，以确定具有科学性之一般关系。

一般关系之确立，与已越出统计学之范围而步入某种科学的领域。统计方法与可施於任何科学范围，其因此求得一般关系乃属该门科学本身之问题，而并非统计上之一般关系。统计方法不仅为一种如何整理材料之工具，使研究者能藉此作实际的调查，以便观察所得情形，能与期待的理论相符，或因调查所得而发现某种新的理论，此乃为建立一切科学上理论之必经步骤。例如以经济学言，其所有理论，或先由学者之思想所构成而后经事实证明者，或经事实之启迪，而後经学者发现者，要之经济理论之确立概为经济学者的工作，而非偏於统计学者的范围，可以概见。

统计的方法上知识因近世的问题与需要，成为从事该学

高級統計學 上 P.4

不可少的工具。工欲善其事，必先利其器。處現在科學研究，且趨數量化的時代，吾人怎能不徹底了解把握這科學的利器，換言之，要明瞭統計學的原理，以確立統計學的基礎。

統計學是描寫一群數量分配狀況的方法，最簡單者為分組，分劃次相差等，由分組而得次數分配，依次數分配的性狀，而用集中量數、離中量數、偏態、峰度等，指出其分配之特徵，曲線與曲線修勻以描寫之。對時間數列，則將長期趨勢、季節變動、循環變動、偶然變動等以研究之。此外對一切隨機事件之相對變動或差異的概括，則常使用指數。研究事象協變之法，則用聯立程式，則有單相關、複相關、偏相關等，而其相互變動乾諸之趨勢又可分為直線相關與非直線相關。此皆為普通統計均已修習，唯因篇幅之限制，未加詳述，並有許多高深方法也完全未能

研习故当有高级统计学来补救。

高级统计学中所研究的是应用机率原理注重实例，使能承上——统计学一接下——数理统计学，以为统计方法及理论研究的张本，兼推演主要公式之来源，继以实例为出发点。以统计学既然是物质的计算，就不能应用纯粹数学方法之推论，导引以解决实际上所遭遇的问题。故本书内容先研究次数分配及相关，为上篇，下篇完全讨论抽样问题，上下共二十

上篇内容：先研究描写次数问题，讨论动矩之计算、机率概念、二项分配及配合法之概要，以导引常态曲线 $y=\frac{N}{\sqrt{2\pi}\sigma}e^{-\frac{x^2}{2\sigma^2}}$

即所谓Laplace-Gauss曲线，凡社会或自然现象的次数分配用此曲线可以描写，大致统计学上好多理论皆建设在此曲线上。次及用数学描写非常态即偏态曲线式之分配问题。此项研究在十九世纪末曾分为三方面：其一为共 J. P. Gram (1879) T. N.

高級統計學 P.5.

Thiele (1889) 及 C. V. L. Charlier (1905) 二氏 相同联者称為斯堪提纳维亚系. 此系之曲線是用两种函级数代表之為

Type A: $y = A_0\psi(x) + A_1\psi'(x) + A_2\psi''(x) \cdots A_n\psi^{n}(x) \cdots \left[\psi(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2\sigma^2}}\right]$

Type B: $y = B_0\phi(x) + B_1\Delta\phi(x) + B_2\Delta^2\phi(x) \cdots B_n\phi(x)\Delta^n \cdots \left[\phi x = e^{-n}\frac{(n)^x}{x!}\right]$

其二為 K. Pearson 從事此式之研究, 以其系亦得名稱波而生曲線系. 此係之次為曲線由於積分下列方程式得之.

$$\frac{dy}{dx} = \frac{(x-a)y}{b_0 + b_1x + b_2x^2}$$ 分為十三型, 主型三: 即

$$y = y_0\left(1+\frac{x}{a_1}\right)^{m_1}\left(1-\frac{x}{a_2}\right)^{m_2}, \qquad y = y_0\left(1+\frac{x^2}{a^2}\right)^{-m} e^{-v\tan^{-1}\frac{x}{a}},$$

$$y = y_0(x-a)^{q_2} x^{-q_1}$$ 以為其他十變型之代替;

其第三方面為 Edgeworth (1900) Kapteyn (1903) (1916) 及其他諸人 以其研究之方法而稱之為換置法或變数換置法.

Method of transformation variable 此法之大意为:为某人之身长分配为常態,并以体重与身长的三次方成正比例,则体重之分配必非常態为偏態.

本书对于相关特以绝大的注意,期能对于相关度及相关函数之原理及应用加以優量精確的說明,故有相关度之测定原理,两个变数相关的原理及计算,内分为直线,非直线,别联合,详述其原理及其应用,介绍最新简便方法,关于相关函数分述直线迴歸,非直线迴歸之原理及应用,最后讨论常態相关面之原理及应用及其与各相关度之关係.

下篇全部研究抽樣,以抽樣方法为研究科学的人省了多少劳力且又得多少知识,其实抽樣是否可靠,可靠之程度为何,此为研究科学者不可不知也,按今生论抽样问题为统计学中最基本的问题.年来统计学者对抽样问题理论阐明很多,实对科

高級統計学上 P.6.

学有相当的貢献。抽樣原理中最主要的是研究樣本常数的分布及其与母体常数之相差。抽樣理論上方法分成兩部分，即大樣本与小樣本。大樣本理論多為皮尔生氏所倡導，小樣本理論多為R.A. Fisher氏与其他学者晚近所研究。当樣本相当大時，樣本常数之分配合于常態，此時之统計方法與常態曲線不过談到用不同技術使常態曲線一般化。

宇宙現象既然係受許多複雜原因之影响，於原因之分析實為了解事象必要之手續。统計学在這方面創了不少方法，變異分析法之主旨在說明變異乃由多数因子所集成，有些因子為已知因子，其對於試驗之影响，可設法逐步析出，將其剩下部分歸之於試驗誤差，本书中亦評論之。

以上所述就是本书的内容上的意義，茲之如之。

民國卅三年九月

李蕃序於國立復旦大學

高級統計學 七

上篇 次數分配及相關

第一章 動差 moment

第一節 次數分析與動差

次數分配之分析，可大別為兩種：一為普通目的求得分配之各種度量，即足以敘述分配之性質，其最主要而常用者為全體分配之集中量數與離中量數等。再進一步的分析法，即為曲線之配合，至次數分配成系統者，要不外下列兩類：一為皮爾生氏曲線系 Pearson's systems of frequency curves；其他為常態曲線之種種變化，例如 Thiele, Gram, Edgworth 等次數函數是。

後者一類之次數函數，由機率的假定的假設推求而得，從理論以觀察，自較皮爾生之系統為滿足，惟其長處亦即其缺點。蓋一般之實現象之次數分配，每有相當的偏度（尤其在經濟社會生物等現象中），而此等函數未能適應偏度稍大之分配，即使勉強適應偏態之函數，亦一種受限制之分配（社會經濟教育心理等現象中甚多）。或丁形、U形分配，不能用此種函數配合。皮氏數曲線則無此缺點。

配合曲線之方法，通常用最小二乘法，然若曲線之方程式非多項式而為超越函數 Transcendental function，例如皮爾生氏之曲線型，則最小二乘法殊不切實用，或竟為不可能者。關於此點，皮氏乃發明以動差方法以代二乘法最小，其實偏次數曲線

求之項式(拋物線派)此法與最小二乘法完全一樣

用動差配合曲線，其論點理論上或實際上並非唯一的方法。皮氏自認此法在理論上不能視為最佳之法，蓋所謂最佳者，其含糊的定義至於事實上，特殊的曲線往往有特殊的便利的配合法，即如人壽保險學上之實例是。然此種特殊方法不能施之於一般曲線，故不能視為普通有系統的配合方法，若論普通的有系統的配合法，則除動差法外，尚未有此發現，故動差之計算，仍為次數分配分析的基礎。

皮式曲線配合法有賴動差之能否準確求得，倘用不準確的動差配合曲線，則失却配合之本意。且動差之應用，並不僅為配合，凡次數分配之重要度量，除百分位數外，幾乎皆可利用動差表示之，蓋其算術平均數即第一次動差，標準差即為第二動差之平方根，偏態為$=\frac{\mu_3}{\mu_2^{3/2}}$，峯度為$\beta_2=\mu_4/\mu_2^2$。倘分配為常態，則平均差為標準差之$\sqrt{\frac{2}{\pi}}$倍，故亦可用動差表示之。總之動差之計算實為次數分配分析之基礎問題。

第二節 定義及符號

動差為力學上通用之名詞，用以測量一動力，此動力有成旋轉之趨勢，而根據於動力之某中點與原點之距離。此在統計學上之重要意義亦頗相同，組次數即可認為動力，各組之中點與原點之距離乃計算動差最重要之一點。值一原點之選擇，配之動差可以計算，以每

组次数对于X軸上原点之距離之某次方之和，再以總次数除之，
為第一動差用X距離之一次方，第四動差用X距離之四次方
設有一数列其 x_1' x_2' x_3' …… x_n' 為自某点起之离中差，並設 f_1
f_2，f_3 …… f_n 為相应对之觀察次数，動差之定義為：

$$m_r = \frac{\sum_{i=1}^{N} x_i'^r f_i}{\sum_{i=1}^{N} f_i} \quad 或 \quad m_r = \frac{\sum x'^r f_x}{\sum f_x} \qquad x = X - \bar{X}$$

$\bar{X} = A + h m_1'$，式中A為任意原点，及h為单位，使 $x' = (X-A)/h$
Σf = 次数總和，通常以N記之。

$\frac{\Sigma x' f}{N}$ = 任意原点以组距為单位之第一次動差（未校正）記以 m_1'

$\frac{\Sigma x'^2 f}{N}$ = 〃 〃 〃 〃 （ 〃 ） 〃 m_2'

$\frac{\Sigma x'^r f}{N}$ = 〃 〃 〃 〃 （ 〃 ） 〃 m_r'

不帶撇（'）之 m_r 代表以算術平均数為原点，以组距為单位之第r級動差，即次数分配之動差，但與曲線之動差不同，其式如下：
$\int y\,dx = N$ = 次数總和
$\int xy\,dx / N$ = 以平均数為原点由曲線而得之第一級動差，以 μ_1 記之
之 之記之

三. $\Sigma x^r y\, dx/N =$ 〃 〃 $=$ 〃 〃 M_2之三

$\Sigma x^r y\, dx/N =$ 〃 r 〃 M_r之三

上計述即先求以任意點為心並從轉化以平均數為心之動差。此係本生教授於其動差論(moment)首推論之、第一、第二、第三、乃至高級之動差，即為離中差之一次、二次、三次乃至更高次方之平均離中差，為隨任何一點起而計算亦例得之動差，不過備僅算之便利耳。

茲推論 m'_r(輔助動差)與 m_r(主要動差)之關係述之如次：

由定義則

$$m_r = \Sigma(X-\bar{x})^r f/N \quad \text{其中} \quad X = a + h x', \quad \bar{x} = \Sigma X f/N$$

$$\Sigma X f = \Sigma a f + \Sigma h X f = Na + h\Sigma X f \quad \text{除以}N\text{得}\ \bar{x} = a + h m'_1$$

$$m_r = \Sigma(X-\bar{x})^r f/N = \Sigma[(a+hx') - (a+hm'_1)]^r f/N$$

$$= \Sigma[h(x'-m'_1)]^r f/N = \frac{h^r}{N}\Sigma(x'-m'_1)^r f$$

$$= \frac{h^r}{N}\Sigma\left[x'^r - r x'^{r-1} m'_1 + \frac{r(r-1)}{2!}x'^{r-2} m'^2_1 - \frac{r(r-1)(r-2)}{3!}x'^{r-3} m'^3_1 + \cdots\right] f$$

$$= h^r\left[\frac{\Sigma x'^r f}{N} - r\frac{\Sigma x'^{r-1} f}{N} m'_1 + \frac{r(r-1)}{2!}\frac{\Sigma x'^{r-2} f}{N} m'^2_1 - \right.$$

$$\left.\frac{r(r-1)(r-2)}{3!}\frac{\Sigma x'^{r-3}}{N} f m'^3_1 + \cdots\cdots\right]$$

$$= k^{r}\left[m'_r - r m'_{r-1} m'_1 + \frac{r(r-1)}{2!} m'_{r-2} m'^2_1 - \frac{r(r-1)(r-2)}{3!} m'_{r-3} m'^3_1 + \cdots\cdots\right]$$

如果 $k=1$ $r=0, 1, 2, 3, 4$, 则有

$k=1$ $r=0$, $m_0 = m'_0 = 1$

$k=1$ $r=1$ $m_1 = m'_1 - \frac{1}{1} m'_0 m'_1 = m'_1 - m'_1 = 0$

$k=1$ $r=2$ $m_2 = m'_2 - 2m'_1 m'_1 + \frac{2}{2} m'_0 m'^2_1 = m'_2 - 2m'^2_1 + m'^2_1 = m'_2 - m'^2_1$

$k=1$ $r=3$ $m_3 = m'_3 - 3m'_2 m'_1 + 3m'_1 m'^2_1 - m'^3_1 = m'_3 - 3m'_2 m'_1 + 2m'^3_1$

$k=1$ $r=4$ $m_4 = m'_4 - 4m'_3 m'_1 + 6m'_2 m'^2_1 - 4m'_1 m'^3_1 + m'_0 m'^4_1$

$= m'_4 - 4m'_3 m'_1 + 6m'_2 m'^2_1 - 3m'^4_1$

例:

X	f	x'	$x'f$	$x'^2 f$	$x'^3 f$	$x'^4 f$	$(x'+1)^4 f$
57	29	-4	116	464	1856	7424	2349
62	23	-3	69	207	621	1863	368
67	51	-2	162	324	618	1296	51
72	151	-1	151	151	151	151	0
77	192	0	498	0	3276	0	192
82	239	1	239	239	239	239	3824
87	157	2	314	628	1256	2512	12717
92	93	3	279	837	2511	7533	23808
							15125

97	29	4	116	464	1856	7424	7776
102	6	5	30	150	750	3750	
	1,000		978	3,464	6612	33192	
			480		3.336		

$m_1' = \frac{\Sigma x' f}{N} = \frac{480}{1000} = .480$ $\qquad m_3' = \frac{\Sigma x'^3 f}{N} = \frac{3336}{1000} = 3.336$

$m_2' = \frac{\Sigma x'^2 f}{N} = \frac{3464}{1000} = 3.464$ $\qquad m_4' = \frac{\Sigma x'^4 f}{N} = \frac{33192}{1000} = 33.192$.

$m_0 = m_0' = 1 \qquad m_1 = 0$

$m_2 = m_2' - m_1'^2 = 3.464 - (.480)^2 = 3.2336$

$m_3 = m_3' - 3m_2' m_1' - 2m_1'^3 = 3.336 - 3 \times 3.464 \times (.480) - 2(.480)^3 = 1.43097$

$m_4 = m_4' - 4m_3' m_1' + 6m_2' m_1'^2 - 3m_1'^4$

$= 33.192 - 4 \times 3.3336 \times .480 + 6 \times 3.464 \times (.480)^2 - 3 \times (.480)^4 = 30.4164$

上例系用動差公式求得 $m_1'\ m_2'\ m_3'\ m_4'$ 及 $m_1,\ m_2\ m_3\ m_4$ 因而更进一步求得 $\bar{X}$, σ, β_1, β_2 K_1 及 K_2 等。对於表中 fx', fx'^2, fx'^3, fx'^4 以及総和数目是否有误可用 Charlier check 针正法校验之

因 $\Sigma(x'+1)^4 f = \Sigma x'^4 f + 4\Sigma x'^3 f + 6\Sigma x'^2 f + 4\Sigma f x' + \Sigma f$.

第三節 総和法 Summation method

在前例中已说明求$\Sigma x^{y}f$之方法，但求之组距之数目相当多，且之各个数值相当大时，普通法之手续可谓尽麻烦之能事，而其结果又易错误，在先进国家皆以計算機代替之，而我国则不易齐備，故能應用総和法，則简便多矣

(1) 総和法公式之来源

項目	$f(x)$	S_0	S_1	S_2
1	$f(1)$	$f(1)+f(2)+\cdots f(n)$	$f(1)+2f(2)+\cdots n\,f(n)$	$f(1)+3f(2)+6f(3)+\cdots\frac{n(n+1)}{2}f(n)$
2	$f(2)$	$f(2)+f(3)+\cdots f(n)$	$f(2)+2f(3)+\cdots(n-1)f(n)$	$f(2)+3f(3)+\cdots\frac{(n-1)n}{2}f(n)$
3	$f(3)$	$f(3)+f(4)+\cdots f(n)$	$f(3)+2f(4)+\cdots(n-2)f(n)$	$f(3)+3f(4)+\cdots\frac{(n-2)(n-1)}{2}f(n)$
4	$f(4)$	$f(4)+f(5)+\cdots f(n)$	$f(4)+2f(5)+\cdots(n-3)f(n)$	$f(4)+3f(5)+\cdots\frac{(n-3)(n-2)}{2}f(n)$
$n-2$	$f(n-2)$	$f(n-2)+f(n-1)+f(n)$	$f(n-2)+2f(n-1)+3f(n)$	$f(n-2)+3f(n-1)+6f(n)$
$n-1$	$f(n-1)$	$f(n-1)+f(n)$	$f(n-1)+2f(n)$	$f(n-1)+3f(n)$
n	$f(n)$	$f(n)$	$f(n)$	$f(n)$
		$\sum_{n=1}^{n} f(n)=m_0'$	$\sum_{n=1}^{n} nf(n)=m_1'$	$\sum_{n=1}^{n}\frac{n^2+n}{2}f(n)=\frac{1}{2}(m_2'+m_1')$

S_3

$$f(1)+4f(2)+\cdots\frac{n(n+1)(n+2)}{3!}f(n)$$
$$f(2)+4f(3)+\cdots\frac{(n-1)n(n+1)}{3!}f(n)$$
$$f(3)+4f(4)+\cdots\frac{(n-2)(n-1)n}{3!}f(n)$$
$$\vdots$$
$$f(n-2)+4f(n-1)+10f(n)$$
$$f(n-1)+4f(n)$$
$$f(n)$$
$$\sum_{n=1}^{n}\frac{n^3+3n^2+2n}{6}=\frac{1}{6}\left[m_3'+3m_2'+2m_1'\right]$$

S_4

$$f(1)+5f(2)+\cdots\frac{n(n+1)(n+2)(n+3)}{4!}f(n)$$
$$f(2)+5f(3)+\cdots\frac{(n-1)n(n+1)(n+2)}{4!}f(n)$$
$$f(3)+5f(4)+\cdots\frac{(n-2)(n-1)n(n+1)}{4!}f(n)$$
$$\vdots$$
$$f(n-2)+5f(n-1)+15f(n).$$
$$f(n-1)+5f(n)$$
$$f(n)$$
$$\sum_{n=1}^{n}\frac{n^4+6n^3+11n^2+6n}{24}=\frac{1}{24}\left[m_4'+6m_3'+11m_2'+6m_1'\right]$$

此式為以任意点為心之動差,但应用所需者為以平均数為原点之動差,今应用由輔助動差化為主要動差之公式且令 $S_2=\alpha$

則有 $m_2=m_2'-m_1'^2=(2S_3-\alpha)-\alpha^2=2S_3-\alpha(1+\alpha)$

同理得 $m_3=6S_4-3m_2(1+\alpha)-\alpha(1+\alpha)(2+6)$

$$m_4=24S_5-2m_3\{2(1+\alpha)+1\}-m_2\{6(1+6)(2+6)-1\}-$$
$$\alpha(1+\alpha)(2+\alpha)(3+\alpha).$$

証:

次数	Σ	Σ^2	Σ^3	Σ^4
29	1000	5480	19,372	54506
23	971	4480	13,892	35136
81	948	3509	9,412	21244
151	867	2561	5,903	11832
192	716	1694	3,342	5929
237	534	978	1,648	2557
157	285	444	670	939
93	128	169	216	266
29	35	41	47	83
6	6	6	6	53
1000	5480	19,372	54508	135506

$S_1=0$

$S_2=5.48$

$S_3=19.372$

$S_4=54.508$

$S_5=137.503$

[illegible] 十一

代入之式

$S_2 = d = 5.48$

$m_2 = 2S_3 - d(1+d) = 2\times 19.372 - 5.48\times 6.48 = 3.2336$

$m_3 = 6S_4 - 3m_2(1+d) - d(1+d)(2+d) = 6\times 54.508 - 3\times 3.2336\times 6.48$

$5.48\times 6.48\times 7.48 = -1.43077$

$m_4 = 24S_5 - 2m_3\{2(1+d)+1\} - m_2\{6(1+d)(2+d)-1\} - d(1+d)(2+d)(3+d)$

$= 24\times 132.503 - 2\times(-1.43077)\times 6.48 + 1 - 3.2336\times 6\times 6.48\times 7.48 - 1 -$

$5.48\times 6.48\times 7.48\times 8.48 = 30.4164$

又前表中 x^r 代表 $x(x-1)(x-2)(x-3)\cdots(x-r+1)$ 共有 r 个因子，故

$x^4 = x(x-1)(x-2)(x-3)$

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

$x^2 = x(x-1)$

$x' = x$

在 $S_{(3)}$ 直线下之 $n(n-1)(n-2)/3!\,fn$ 之来源如次 因其为

$[1+3+6+\cdots\cdots+n(n-1)/2]fn$ 故在增量学 calculus of finite

difference 之総和级数法则

$$\sum_{n=2}^{n=1}\frac{n(m-1)}{2} = \sum_{2}^{n=1}\frac{(m)^2}{2} = \frac{m^{(3)}}{3!}\Big]_2^n = \frac{n^{(3)}}{3!}$$

同理 $S_{(4)}$ 直行 下之 $[n(n-1)(n-2)(n-3)/4!]f_n$ 項 亦由此法而得.

使 $S_r = \Sigma x^r f_n$

$S_{(r)} = \dfrac{\Sigma x^r f_x}{r!}$

$S(0) = S_0 = \Sigma f_x$

$S(1) = \Sigma x f_x = S_1$

$S_{(2)} = \dfrac{\Sigma x(x-1)f_x}{2!} = \dfrac{S_2 - S_1}{2}$

$S_{(3)} = \dfrac{\Sigma x(x-1)(x-2)f_x}{3!} = \dfrac{S_3 - 3S_2 + 2S_1}{6}$

$S_{(4)} = \dfrac{\Sigma x(x-1)(x-2)(x-3)f_x}{4!} = \dfrac{S_4 - 6S_3 + 11S_2 - 6S_1}{24}$

設 $\sigma_{(r)} = \Sigma x^r f_x / \Sigma f_r = \dfrac{r!\,S(r)}{S(0)}$.

由上式得 $\sigma_{(0)} = 1$

$\sigma_{(1)} = V_1'$

$\sigma_{(2)} = \dfrac{V_2' - V_1'}{2}$

$\sigma_{(3)} = \dfrac{V_3' - 3V_2' + 2V_1'}{6}$

$\sigma_{(4)} = \dfrac{V_4' - 6V_3' + 11\,8V_2' - 6V_1'}{24}$

若以 $S_{(r)}$ 表 S_r 则得

$$S_0 = S(0)$$
$$S_1 = S(1)$$
$$S_2 = 2S(2) + S(1)$$
$$S_3 = 6S(3) + 6S(2) + S(1)$$
$$S_4 = 24S(4) + 36S(3) + 14S(2) + S(1)$$

若取[illegible]及[illegible]x之中間列对於[illegible]可[illegible]其[illegible]如次:

x	f	$S_{(1)}$	$S_{(2)}$	$S_{(3)}$	$S_{(4)}$
$-n$					
$-(n-1)$					
⋮					
-4					$\bar{S}(4)$
-3				$\bar{S}(3)$	
-2			$\bar{S}(2)$		
-1	$\bar{S}(0)$	$\bar{S}(1)$			
0	f_0				
1	$\overset{+}{S}(0)$	$\overset{+}{S}(1)$			
2			$\overset{+}{S}(2)$		

x	fx	$S(1)$	$S(2)$	$S(3)$	$S(4)$
3				+ $S(3)$	
4					+ $S(4)$ - - - - - - - -

第五節 薛伯氏校正數 Sheppard's correction

計算分組次數表之動差，求得統計常數 Parameters 如平均數、標準差、偏斜度、峯度等，皆假定各項集中於組中点，例如某一組女生在 98 lbs–5 103 lbs 之間，故此組之重量作為 100.5 lbs。但實際上各組重量不同，且當次數分配完全對稱或偏態不大時，在小於平均數各組中較小數值之次數較多於較大數值之次數。故若以中點代表全組中各項之數值，則由是求得之第二第三第四級動差將與實際動差稍有相差，故欲得正確數必需校正。

按 W. F. Sheppard 校正法，其應用之數列需合於下列之三条件：

(1) 曲線兩端須低降與 x 軸相切者，用數學的術語述之，即假定之次數與數的各級導微在兩端皆為零。然此种条件經之次

次分佈之形式不合 第Pearson之各种次数曲線型而言 其第一二,三,八,九,十,十一,十二 各种曲線往々一端与x之軸相遇而並不系切然含一级導数于一端為零,而二级三级導数並不等於零或為J字式或為U形,蓋伯氏之校正自不符之不適用自不待言,故此种校正之非依任分配皆可應用者,倘不顧此种校正之假定而用之,所得結果反不如不校者

(二)读者連續变量之身長等,如几形曲線间断之徵亦不勉強用之。

(三)读由分組材料求出者,亦自原始材料求出者之读校正。

校正数公式來源:

設 $\varphi(x)$ 為一机率函数, h 為单位, $x = X - \bar{X}$, 則

$$\int_a^b x^r \varphi^{(s)}(x)\,dx = h_j \Sigma x_j^r \varphi^{(s)}(x_j)$$

設此分配在 $b-a$ 之距离内,將此距离分為 n 組

$(x-\frac{1}{2}h,\ x+\frac{1}{2}h)$ 即 $b-a=n\omega$ 第 j 組内之机率当為

$$p_i = \int_{x_j-\frac{1}{2}h}^{x_j+\frac{1}{2}h} \varphi(x)\,dx$$

設 y_x 之形式為一多項式如下

$$y = a + bx + cx^2 + dx^3 + ex^4$$

今設觀察值為 $y_{-2}, y_{-1}, y_0, y_1, y_2$ 試求 $\int_{-\frac{1}{2}}^{\frac{1}{2}} y\,dx$ 之值以此等觀察值表之

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} y\,dx = \int_{-\frac{1}{2}}^{\frac{1}{2}} (a + bx + cx^2 + dx^3 + ex^4)\,dx = \left| ax + \frac{b}{2}x^2 + \frac{c}{3}x^3 + \frac{d}{4}x^4 + \frac{e}{5}x^5 \right|_{-\frac{1}{2}}^{\frac{1}{2}}$$

$$= a + \frac{c}{12} + \frac{e}{80}$$

令

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} y\,dx = h y_0 + k(y_{-1} + y_1) + l(y_{-2} + y_2)$$

$$y_0 = a$$

$$y_{-1} + y_1 = 2(a + c + e)$$

$$y_{-2} + y_2 = 2(a + 4c + 16e)$$

則

$$a + \frac{c}{12} + \frac{e}{80} = h y_0 + k(y_{-1} + y_1) + l(y_{-2} + y_2) = ha + 2k(a + c + e)$$

$$+ 2l(a + 4c + 16e) = ha + 2ka + 2kc + 2ke + 2la$$

$$+ 8lc + 32le = (h + 2k + 2l)a + (2k + 8l)c + (2k + 32l)e$$

用比較係數法得

$$\left.\begin{aligned} 1 &= h + 2k + 2l \\ \frac{1}{12} &= 2k + 8l \\ \frac{1}{80} &= 2k + 32l \end{aligned}\right\}$$

即

$$\left.\begin{aligned} h + 2k + 2l &= 1 \\ 24k + 96l &= 1 \\ 160k + 2560l &= 1 \end{aligned}\right\}$$

用行列式解之得:

$$h = \frac{1}{D}\begin{vmatrix} 1 & 2 & 2 \\ 1 & 24 & 960 \\ 1 & 160 & 2560 \end{vmatrix} = \frac{5178}{5760}$$

$$k = \frac{1}{D}\begin{vmatrix} 1 & 1 & 2 \\ 0 & 1 & 960 \\ 0 & 1 & 2560 \end{vmatrix} = \frac{308}{5760}$$

$$c=\frac{1}{D}\begin{vmatrix}1&2&1\\0&24&1\\0&160&1\end{vmatrix}=\frac{17}{5760}\qquad D=\begin{vmatrix}1&2&2\\0&24&96\\0&160&2560\end{vmatrix}=5760$$

即

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} y\,dx=\frac{1}{5760}\{5178y_0+308(y_1+y_{-1})-17(y_2+y_{-2})\} \qquad (6)$$

苟更以 y_0 y_1 y_2 y_3 表示 $\int_{-\frac{1}{2}}^{\frac{1}{2}} y\,dx$ 之值，同樣用上法進行運算得

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} y\,dx=\frac{1}{24}(27y_0+17y_1+5y_2-y_3) \qquad (7)$$

按此方旋轉矩之定義，求r次矩之修正值

$$\mu_r=\frac{1}{N}\sum_{i=0}^{n-1} X_i^r y_i$$

$$\begin{aligned}\mu_r&=\frac{1}{N}\left[X^r y_0+(X+1)^r y_1+\cdots\cdots+(X+\overline{N-1})^r y_{n-1}\right]\\&=\frac{1}{N}\left[X^r\int_{-\frac{1}{2}}^{\frac{1}{2}} y\,dx+(X+1)^r\int_{\frac{1}{2}}^{1+\frac{1}{2}} y\,dx+\cdots\cdots+(X+(\overline{N-1}))^r\int_{N-1\frac{1}{2}}^{N-\frac{1}{2}} y\,dx\right]\\&=\frac{1}{5760}\{X^r[5178y_0+308(y_1+y_{-1})-17(y_2+y_{-2})]+(X+1)^r[5178y_1\\&\quad+308(y_0+y_2)-17(y_{-1}+y_3)]+(X+2)^r[5178y_1+308(y_1+y_3)\\&\quad-17(y_0+y_4)]+\cdots\cdots\cdots+(X+\overline{N-1})^r[5178y_{n-1}+308(y_{n-2}+y_n)\\&\quad-17(y_{n-3}+y_{n+1})]\}\end{aligned}$$

y_x 之係數亦為

$$(\xi+\iota)^r 5178+308(\xi+\iota+1)^r+308(\xi+\iota-1)^r-17(\xi+\iota-2)^r$$
$$-17(\xi+\iota+2)^r.$$

令 $\xi+\iota=h$ 代入上式 新 y_x 之係數為

$$5178h^r+308[(h-1)^r+(h+1)^r]-17[(h-2)^r+(h+2)^r]$$

$$\therefore\ \mu_r=\frac{1}{5760N}\sum_{x=0}^{\infty}y_x\{[5178h^r]+308[(h-1)^r+(h+1)^r]$$
$$-17[(h-2)^r+(h+2)^r]\}$$

$$=\frac{1}{5760N}\sum y_x\{5178h^r+308[h^r-rh^{r-1}+\frac{r(r-1)}{2!}h^{r-2}\cdots\cdots$$
$$+h^r+rh^{r-1}+\frac{r(r-1)}{2!}h^{r-2}+\cdots\cdots]-17[h^r-2rh^{r-1}$$
$$+4\frac{r(r-1)}{2!}h^{r-2}+\cdots\cdots+h^r+2h^{r-1}r+4\frac{r(r-1)}{2!}h^{r-2}\cdots\cdots]$$

$$=\frac{1}{5760N}\sum y_x\{5178h^r+240r(r-1)h^{r-2}+3r(r-1)(r-2)(r-3)h^{r-4}\cdots\}$$
$$\cdots\cdots(8)$$

依次令 $r=1,2,3,4$ 代入式(8)得

$$r=1\quad \mu_1=\frac{1}{5760N}\sum y_x\times5760h'=\frac{1}{N}\sum y_x h=\mu_1'$$

$$r=2\quad \mu_2=\frac{1}{5760N}\sum y_x(5760h^2+240\times2)=\frac{1}{N}\sum y_x h^2+\frac{480}{5760}$$

$$= \mu_2' + \frac{1}{12}$$

$$r=3,\quad \mu_3 = \frac{1}{5760N}\Sigma y_x(5760 h^3 + 240\times 3\times 2h)$$
$$= \frac{1}{N}\left(\Sigma y_x h^3 + \frac{1}{4}\Sigma y_x h\right) = \mu_3' + \frac{1}{4}\mu_1' = \mu_1$$

$$r=4,\quad \mu_4 = \frac{1}{5760N}\Sigma y_x(5760h^4 + 240\times 4\times 3h^2 + 3\times 4\times 3\times 2)$$
$$= \frac{1}{N}\Sigma y_x h^4 + \frac{1}{2}\Sigma y_x\left(h^2 + \frac{1}{80}\right) = \mu_4' + \frac{1}{2}\mu_2' + \frac{1}{80}.$$

以上諸式中 μ_r 为加厘正之積矩，μ_r' 为已厘正者，因是可得

$$\mu_1' = \mu_1 \qquad \mu_2' = \mu_2 - \frac{1}{12}$$
$$\mu_3' = \mu_3 \qquad \mu_4' = \mu_4 - \frac{1}{2}\mu_2 - \frac{1}{80}.$$

6. 計算積矩之步驟

总計算手續简便起見，吾人实际計算積矩時，可根据以下之步驟，为适合皮尔生考數曲線之目的，故关於計算皮氏各型曲線所共同需要之事項也：β_1，β_2，K_1 及 K_2 之意义及計算，特此依附带說明。

$$\beta_1 = \mu_3^2 / \mu_2^3 \qquad \beta_2 = \mu_4 / \mu_2^2$$

$$K_1 = 2\beta_2 - 3\beta_1 - 6 \qquad K_2 = \frac{(\beta_2+3)^2\beta_1}{(8K_1+12\beta_1+48)K_1}$$

第二章　機率概念

第一節　機率之意義

何謂機率：簡言之，乃即某種可能發生之比例也。而此種現象在吾人日常生活中固無時無地不在存在也。例如今天天氣晴朗，而晴朗之度數甚高，則吾人於出門時問回來時是否下雨？當答下雨之机會甚微。又如某人某甲身体甚健，試問某人今日身体健康否？當答健康之机會甚大。上述各例中吾人所以作此答者，以吾人因有相當之經驗，知某人身体健康之理由，但不能因此而下結論曰：今日必晴，某人必身体健康，亦僅能認為相當可能之事实耳。某種事实之現狀及其因此所得之結論二者之間有一種關係，吾人固不能由此種關係而能獲得確定之結論，但此種關係於統計上稱謂机率關係。

基上所言，某種事实之現狀對於机率之影响甚大。例如以銀幣言擲地後向上者為背或為面，銀面本身現狀所施之影响至大，故在未將銀幣擲地以前，應將銀幣本身詳加審察，背面是否相同？是否相互对称？輕重是否相等？如銀幣確能合於上述理想條件，銀幣擲地後所得之結果，其為背為面應有絕对同樣之可能。反之如發現幣之面背重量对称，以及相等均有缺點時，則此种对於銀幣本身現狀之補充認識，对於机率之影响至大。如銀幣背方較重，則擲地之結果面向上之机會較多，此種情形如試驗之次數較少，擲下之結果未必向上者必為幣面，不過加以多次之試驗，同時將每次之結果加以記錄，則幣面向上之總次數，當較幣背向上之次數為多。由此可見現狀之補充認識，可以影响机率之結果。如現狀之補充認識為有利者，則其所施

之影响，當可增加机率之程度。

因是之故，欲研究某种事实之机率，必須其現狀有深刻之認識。下述為便於研究起見，一般人均以抽象之符號表示所求之机率。例如某種事實发生之可能机会為 h，其实現之机会為 a，則此实現机会(a)，當為上述可能机会(h)所限制。

例：(h) 一袋内含有黑白球各一

(a) 由袋内抽出一球色白

故(a)在(h)條件之下，當為可能实現之事实，然非必然实現之事实。試□袋内含有黑白球各一，則由袋内抽出之球，黑白二者均同屬可能之事实。故如二者有一发生之情形時，則吾人可認為(a)与(h)間確有机率之关係，而以符號 a/h 代表取得白球之机率。

推廣論之：若一事件共有 h 個不同的情形，而在此等不同之情形中，共有 a 個為成功事件，則失敗事件則有($h-a$)個，亦以 p 代表成功之机率，q 代表失敗之机率，則

$$p = a/h$$
$$q = (h-a)/h$$
$$p+q = \frac{a+h-a}{h} = h/h = 1$$

故在一單独試驗中，可不確定或為成功或為失敗，而確定成功与失敗之机率總和為一單位。此當合理。在 K 次試驗中，其成功的可能数目當為 KP，失敗当為 Kq，且因

若為n個試驗,以平常言之,當為(n)或為[illegible](n-α)或[illegible]個失敗.

第二節 · 事件之區別選擇數

[illegible]能合方法總數,[illegible]各事件所有之選擇[illegible]如[illegible]選擇數[illegible],因此欲知各[illegible]之計算[illegible]選擇數[illegible]不同

[illegible]事件選擇數之[illegible]方法[illegible]性質之不同,[illegible]成下列三類:

1.互相排斥事件 如:袋內有白球五個黑球七個,由袋內任取一球,是/白球或黑球二事件,即為相互排斥之事件,因一[illegible]白球出現後,黑球即不能出現,反之亦然.計算之法,[illegible]所得[illegible]事件[illegible]相加而得之.如上述之黑白球選擇為五數[illegible]之一球,如第一白球之選擇數為一,故任何一白球之選擇數為五,以任何一球之選擇數為12

2.独立之複合事件 如甲袋備白球五個,乙袋除黑球七個,如同時自甲乙二袋各取一球,成為第一白球與第一黑球之二事件,則為独立之複合事件,蓋同時取出黑白二球為複合事件,而取白球與取黑球為各別之事件,不能互相影响,故又為独立之事件,此種事件之選擇數由諸独立之事件之各別選擇數相乘而得,如甲乙二袋中之黑白球,於甲袋中取出白球之選擇數為5,於乙袋中取出黑球之選

擇數為7.今同時各取一球,其選擇數為5×7＝35.因選取任何一白球之時,其同時所有之黑球選擇數均為7次,今白球可以選取5次,則其共有之次數自然為35矣,故得其同時聯合選擇之方法公式為 mn.

此種之可以應用於三數四數以至任何數各自獨立之事件聯合選擇數,如擲幣三枚,其面向上之形式可得八種:

$$(正_1+背_1)(正_2正_3+正_2背_3+背_2正_3+背_2背_3)=2\times4=8$$

3.關聯之複合事件. 如一袋中含五白球七黑球,則由其中取出二球即成為關聯之複合事件,此種選擇數可依其合法分為二種:

(甲)如二球先後取出,則先取白球次取黑球之二事件即有互相影響,因先將一球取出後,其全部之球數即為減少,而摸出他一球之機會即增加,則袋中諸球之全體數為n.第一次取出一球,其得以選擇者為袋中所有諸球之全體,故其選擇數為n.第二次取出一球,其得以選擇者,為袋中所有諸球之全體除去第一次取出之一球所餘下之各球,故其選擇數為(n−1).第三次取出一球之~~數~~選擇數為(n−2).同理至第(r)次取出一球之選擇數為(n−r+1),於是得其排列選擇數之計算公式如下:

$$_nP_r = n(n-1)\cdots\cdots(n-r+1)$$

(二) 組合 [illegible]

[illegible]

例如 設有 a b c 三球 [illegible]

[illegible]

[illegible] 排列 [illegible] $_3P_2 = 3 \times 2 = 6$ [illegible]

abc 三球每次二球之排列法存數 $_3P_2$ 為為

$$(ab+bc+ac)\times 2! = 3\times 2\times 1 = 6$$

今此 $(ab+ac+bc)=3$ 為 abc 三球每次二球之組合之數 $_3C_2$ 故是 $_3P_2 = {}_3C_2\cdot 2!$ 其推廣之則

$$_nP_r = {}_nC_r\cdot r!$$

將上式移項可得

$$_nC_r = \frac{_nP_r}{r!}$$

又知 $_nP_r = n(n-1)\cdots\cdots(n-r+1)$ 代入上式得

$$_nC_r = \frac{n(n-1)\cdots\cdots(n-r+1)}{r!}$$

再將分子分母各乘$(n-r)!$即$(n-r)(n-r-1)\cdots\cdots 2\cdot 1$
遂得組合選擇數之簡單公式如下

$$_nC_r = \frac{n!}{r!(n-r)!}$$

第三節
機率之度量

設袋內含有相同之五球,白色者二個,黑色者三個
a_1, a_2 代表抽出之第一白球和第二白球
b_1, b_2, b_3 代表抽出之第一黑球第二黑球和第三黑球
如以a代表抽出之白球(不論為第一白球或第二白球)則
a等於a_1和a_2;如以
b代表抽出之黑球(不論抽出之黑球為第一黑球第二黑球或第三黑球)則b等於b_1或b_2或b_3. 而依機率相等之原則

$$\frac{a_1}{n} = \frac{a_2}{n} = \frac{b_1}{n} = \frac{b_2}{n} = \frac{b_3}{n}$$

是則以上五者之機率均相等,而其中之二有利於一個白球之抽出,其中之三有利於一個黑球之抽出,故可以有利機會與可能機會之比例作為衡量機率之標準,試就上例言,設以P代表抽得白球之機率,q代表抽得黑球之機率,則

$$P=\frac{a}{n}=\frac{2}{5}\qquad q=\frac{b}{n}=\frac{3}{5}$$

而抽得一白球或一黑球之機率均等於一,$(P+q=1)$

由此可見研究任何機率問題概可分為二部

(A) 直接對其機率所趨相等之判斷

例如 $\frac{a_1}{n}=\frac{a_2}{n}=\frac{b_1}{n}=\frac{b_2}{n}=\frac{b_3}{n}$

(B) 如何由有利數與可能數之比例度量其機率

$$P=\frac{a}{n}=\frac{2}{5}\qquad q=\frac{b}{n}=\frac{3}{5}$$

設以可能數n與a, b, c……等事實為例

以$\bar{a}$代表a事實之不發生。例如袋內含有白球二個黑球三個,設(a)代表抽出白球之一,則$\bar{a}$既代表a事實之不發生,當為抽出紅球或黑球之一.

設以

$a+b$ 代表 a 或 b。設以一般擲骰，而以 a 代表所得之點為積數，以 b 代表所得者為三者之積數，則 $a+b$ 為代表得此 3 或 4 或 6。

設以 ab 代表 a，b 同時實現，例如 6 点同時為[illegible]及三[illegible]之積數

設以可能事件[illegible] a，b 二事件為例

$$\frac{ab}{h},\quad \frac{a+b}{h},\quad \frac{a}{bh},\quad \frac{b}{ah}$$

以上四種不同之机率，按以上所言，各有不同之意[illegible]，[illegible]之[illegible]如下：

(1) $\frac{ab}{h}$ a，b 二事件同時出現之机率

(2) $\frac{(a+b)}{h}$ a 事或 b 事實現之机率

(3) $\frac{a}{bh}$ 除可能事實者外，在 b 事實現之条件下 a 事實現之机率

(4) $\frac{b}{ah}$ 除可能事實者外，在 a 事實現之条件下 b 事實現之机率。

以上四種不同机率之意义，說明者以 a b 併為一意，則可得下列四種不同之事實：

$ab,\quad a\bar{b},\quad \bar{a}b,\quad \bar{a}\bar{b}$

以上四种不同合併事实，实为ab併為一起後，應有之可能事件其实现之机会当相等 故可以 $\alpha, \beta, \gamma, \delta$ 表之

ab 同時实现之方式有 α 個

a 出现 b 不出现之方式有 β 個

b 出现 a 不出现之方式有 γ 個

ab 均不出现之方式有 δ 個

設是上述之 $\alpha+\beta+\gamma+\delta$ 個方式之可能性相等 則

a 出现之机率為 $P_1 = \dfrac{\alpha+\beta}{\alpha+\beta+\gamma+\delta}$

b 出现之机率為 $P_2 = \dfrac{\alpha+\gamma}{\alpha+\beta+\gamma+\delta}$

a及b 至少有一出现之机率為 $P_3 = \dfrac{\alpha+\beta+\gamma}{\alpha+\beta+\gamma+\delta}$

a及b同時出现之机率為 $P_{12} = \dfrac{\alpha}{\alpha+\beta+\gamma+\delta}$

a 已出现 b出现之机率為 $P = \dfrac{\alpha}{\alpha+\beta}$

b已出現a出現之機率為 $P_c = \frac{\alpha}{\beta+\gamma}$

b不出現a出現之機率為 $P_3 = \frac{\beta}{\beta+\gamma}$

a不出現b出現之機率為 $P_5 = \frac{\gamma}{\gamma+\delta}$

由是得 $P_1+P_2 = P_3+P_4$

或 $P_3 = P_1+P_2-P_4$

即 a,b最少有一出現之機率為a及b二者出現之機率之和,去其同時出現之機率,即得總和機率定理

又因 $P_4 = \frac{\alpha}{\alpha+\beta+\gamma+\delta} = \frac{\alpha+\beta}{\alpha+\beta+\gamma+\delta} \times \frac{\alpha}{\alpha+\beta}$

$= P_5 P_3$ 得複合機率定理

第四節 定理

(一) 總和機率定理

設依N事故可能發生之總和其中凡有α個為適於a,β個為其發生之事實,乃個為適於a而非b之事實,γ個為適於

b 而非 a 之事實 δ 個為適於非 a 非 b 之事實 $\alpha+\beta+\gamma+\delta=N$，凡此事實發生機會相等

$$\frac{ab}{n}=\frac{\alpha}{N} \qquad \frac{a}{n}=\frac{\alpha+\beta}{N} \qquad \frac{b}{n}=\frac{\alpha+\gamma}{N}$$

$$\frac{(a+b)}{n}=\frac{\alpha+\beta+\gamma}{\alpha+\beta+\gamma+\delta}=\frac{\alpha+\beta+\gamma}{N}=\frac{\alpha+\beta}{N}+\frac{\alpha+\gamma}{N}-\frac{\alpha}{N}$$

$$=\frac{a}{n}+\frac{b}{n}-\frac{ab}{n}$$

此公式即總和機率定理若以文詞表之可云：設有一偶然事件 E 適於於兩事件者乃為 E_1 或 E_2 之發生則該偶然事件之機率稱為 E_1 E_2 之總和機率當等於各偶然事件 E_1 E_2 之機率之和減去 E_1 E_2 同時發生之機率

例：

於人壽險中有夫婦兩人合保一養老金，依其合同之規定凡於 n 年後若有一人尚能生存則保險公司即當付其養老金 S，試問其數學保險費幾何？

欲求此數學保險費先當求此數學機率然欲於 n 年後夫婦兩人中至少有一人生存其不同之情形凡三 (a) 夫婦俱存 (b) 夫存而妻亡 (c) 妻存而夫亡，乃可依上述總和機率定理得有下列公式為

夫妻二人中至少有一人能再活n年之机率

＝(夫能再活n年之机率)+(妻能再活n年之机率)−(夫妻二人均能再活n年之机率)

若欲以数学公式表示之则可以

$P^n_{(xy)}$ ＝夫婦二人中至少有一人能再活n年之机率

P^n_x ＝夫能再活n年之机率

P^n_y ＝妻能再活n年之机率

P^n_{xy} ＝夫妻均能再活n年之机率

则可得下列公式

$$P^n_{\overline{xy}} = P^n_x + P^n_y - P^n_{xy}$$

既知其机率，乃甚易求其数学保險費矣

設若 π＝所求之数学保險費

i＝年息.

则可有下列公式求之

$$P^n_{\overline{xy}}\, S = \pi\,(1+i)^n$$

$$\Pi = \sum P_{xy}^{n}(1+i)^n = \sum(P_x^n + P_y^n - P_{xy}^n)(1+i)^n$$

上例公式中，所以減去 $\frac{ab}{n}$ 之理由以 $\frac{a}{n}$、$\frac{b}{n}$ 中各有一項故須減去，以免重複

上述完全机率定理，当於 a b 二事為互相排斥時不能適用，設以 a b 二者既相互排斥而不能同時出現，則 $ab/n = 0$

$$\therefore (a+b)/n = \frac{a}{n} + \frac{b}{n}$$

例如有紙牌52張中抽得一張問得鷄心一点或二点之机率若何？

令：a = 抽得抽鸡心一点　b = 抽得鸡心二点

$$\frac{a+b}{n} = \frac{a}{n} + \frac{b}{n} = \frac{1}{52} + \frac{1}{52} = \frac{1}{26}$$

(一) 積和机率定理

$E_1 E_2$ 之積和机率乃為 $E_1 E_2$ 一致發生之偶然現象之机率其数值可直於依據机率之理義公式求之

$$\frac{ab}{n} = \frac{\text{既然於a又属於b之事果}}{\text{可能事件之總数}} = \frac{x}{n}$$

然則吾人來考查求其發生E_1之機率，而剛纔知其E_1已發生矣，而再求其發生E_2之機率。今E_1之機率等於 [illegible]，但既知E_1已發生矣，則E_2之機率即不能按其原有之機率 [illegible]。設以 [illegible] 表示其E_1實現後再發生E_2之機率，則上述之命題乃可釋如下式

$$[illegible]$$

蓋因已知E_1之發實現，則E_2之可能事件數即不復為N，而既為$N-1$個矣，且其中適於E_2者乃僅有 [illegible] 個，故此情形下E_2之機率當改為 [illegible]

因此E_1E_2之積合機率可以寫如下式

$$[illegible]$$

積上兩式即得 [illegible]

上公式即積合機率與單獨機率之關係……設有一偶然現象E由E_1,E_2……而組成……則其機率……E_1,E_2之積合機率當等於其所因現象之一（如E_1）之機率乘其既知前一現象……

已發生再發生其他現象(如 E_2)之機率

例. 求E等於在一副52張撲克牌中甲乙兩人各抽得A之機率

設 E_1 = 甲在一副52張撲克牌中任抽一張A

E_2 = 乙在一副52張撲克牌中任抽一張A

a = 甲抽得一張A

b = 乙抽得一張A

ab = 甲乙抽得一張牌適為一張A

$\frac{a}{n}$ = 甲抽得一張A之機率

$\frac{b}{n}$ = 乙抽得一張A之機率

$\frac{ab}{n}$ = 甲乙均抽得一張A之機率.

$\frac{b}{an}$ = 既知某甲所抽得一張為A後而使乙所抽得者亦為一A之機率

$\frac{a}{bn}$ = 既知某乙所抽得者乃為一A後而將使甲所抽得者亦為A之機率

今欲求E之機率可先由機率定義公式計算之

甲乙二人於52張撲克牌中任抽一張其可能遇合數N等於 52^2

此即52個物件兩個一組之組合

$$C_{52}^2 = \frac{52\times 51}{1\times 2} = 1326$$

其有利事實數 C_4^2

$$C_4^2 = \frac{4\times 3}{1\times 2} = 6$$

故其機率 $\frac{ab}{n} = \frac{1}{221}$

$$\frac{ab}{n} = \frac{\text{屬於 } ab \text{ 之事實數}}{\text{可能事實之總數}} = \frac{x}{n} = \frac{6}{1326} = \frac{1}{221}$$

以上所論僅就個別應用總和機率定理與積合機率定理而言，惟於實際計算機率時往往同時應用上述定理。例如擲二骰於地，試計算獲得五點之機率為若干？

如何由二骰獲得五點之可能性有下述四種

2點	3點	$ab/n = a/n \cdot b/n = \frac{1}{6}\cdot\frac{1}{6} = \frac{1}{36}$
3點	2點	$ab/n = a/n \cdot b/n = \frac{1}{6}\cdot\frac{1}{6} = \frac{1}{36}$
4點	1點	$ab/n = a/n \cdot b/n = \frac{1}{6}\cdot\frac{1}{6} = \frac{1}{36}$
1點	4點	$ab/n = a/n \cdot b/n = \frac{1}{6}\cdot\frac{1}{6} = \frac{1}{36}$

依積合機率定理擲二骰於地獲得二點與三點之機率，等於

獲得二点之机率與獲得三点之机率相乘之積数，故為 $\frac{1}{6}\times\frac{1}{6}=\frac{1}{36}$ 其他依此類推。惟擲二骰於地獲得五点之情形，除上述第一骰子二点與第二骰子三点外，更有其他不同情形求得之。故擲二骰於地所得之五点可由2,3，或3,2，或由1,4，或由4,1求得之。由此由總和机率定理可得五点之机率如下。

命 P_5 = 獲得五点之机率，則

$$P_5=\frac{1}{36}+\frac{1}{36}+\frac{1}{36}+\frac{1}{36}=\frac{4}{36}=\frac{1}{9}$$

註：推普通一般分為机率之乘法與加法而言，則可分別如下：

(甲) 加法定理—共成立不共立均可用之 $\frac{(a+b)}{n}=\frac{a}{n}+\frac{b}{n}-\frac{ab}{n}$

又若多一 c 則推廣 $\frac{(a+b+c)}{n}=\frac{a+b}{n}+\frac{c}{n}-\frac{(a+b)c}{n}$

$$=\frac{a}{n}+\frac{b}{n}+\frac{c}{n}-\frac{ab}{n}-\frac{ac}{n}-\frac{bc}{n}+\frac{abc}{n}$$

因 $\frac{(a+b)c}{n}=ac+bc/n=\frac{ac}{n}+\frac{bc}{n}-\frac{abc}{n}$ 可推廣至 n

(乙) 乘法定理：—共立不共立均可（不共立者 $\frac{ab}{n}=\frac{a}{n}\cdot\frac{b}{n}$）共立者 $\frac{ab}{n}=\frac{a}{n}\cdot\frac{b}{n}$ 推廣之 $abc/n=ab/n\cdot c/abn=c/n\cdot ab/cn$

$= a/k \cdot b/ak \cdot c/abk$ 四個時 再增至 n 個 $abc \ldots\ldots n/k$

$= a/k \cdot b/ak \cdots\cdots n/ak - (n-1)/k$

第三章　二項分配

第一節　二項分配之來源

次數表可用以估計樣本所由抽出集團之次數分配，但此種方法，並不能滿意，蓋曲線之地位大部隨個人判斷而移動。又往往由於取樣机誤，可呈不整齊，此不對於之觀來而集團能無此情形。為此又如作一適合樣本之曲線，以作理論次數，再將此理論次數認為傳說然後決定此樣本資料是否與理論相符合。

凡計算每一理論次數，均可直接應用機率之基本理論推出之。由一是對數之樣本，進而討論任何對數之問題，即可推得二項分配。

如袋內N球之分配如何？本可隨意決定，並無一定標準，唯所研究者，則為N球之一種特殊情形分配，而於統計上有極大之功用者，即有袋內抽出之球僅有二種不同之可能性。例如袋內之N球係以二種不同之顏色而將N球分為二大部。

其一部為N′個白球，其另一部則為N″黑球，至黑球白球相加之和，當等於袋內總球數（N=N′+N″），故每次由袋內抽出一球，不為白球，或為黑球。至由袋內抽得之n球（抽出之球應放入袋內從再行抽取）其可能之分配數當為N^n，然則在N^n可

有其餘…

能分配数中含有某种颜色一定球数之分配数，例如含有 r 黑球之可能分配数共有若干乎？

吾人为计算此项可能分配数（含有 r 黑球之可能分配数）起见，以 $x_1, x_2, x_3 \cdots\cdots x_{n'}$ 代表 n' 個白球，$y_1, y_2, y_3 \cdots\cdots y_{n''}$ 代表 n'' 個黑球。n' 個白球與 n'' 個黑球相加之和，則為袋中之 n 總球数。如以上述方法表示之，此袋中之 n 球即可構成下列之 n 数列

$$x_1,\ x_2,\ x_3 \cdots\cdots x_{n'} \qquad y_1,\ y_2,\ y_3 \cdots\cdots y_{n''}$$

依此重複記載 n 次則得相同之 n 数列

據上所言，以袋内抽取 n 次獲得 n 球之試驗，實無異由上列 n 数列由一数列中各得抽取一項，由每数列抽取一項，則由 n 数列當可抽得 n 項。且 n 数列中，任何一数列之各項，即為袋内之 n 球，每一数列被取一次之结果，亦有與其他数列各項抽取之機會。故由上列 n 数列，每一数列中抽取一項，實與由袋内抽取 n 球之試驗無異。

欲求含有 r 個黑球（取 r 項 y）之可能分配数，吾人可假定 r 項 y 完全由 r 数列中所取得，至其他之 $(n-r)$ 数列，則為供給所餘 $(n-r)$ 個白球之数列（即 $(n-r)$ 項 x）。換言之，r 項 y 既完全由

r 數列中取得，則吾人應假定，r 數列之各項全部皆為 N''（即黑球）亦即 r 數列中之任意一數列均含有 N'' 黑球。同樣 $(n-r)$ 項亦既完全由 $(n-r)$ 數列中取得，則此 $(n-r)$ 數列中之任何一數列亦皆含有 N' 白球。含有 N'' 黑球之數列共有 r 個，是則含有 r 個黑球之可能分配數當為 N''^{r}，含有 N' 白球之數列共 $(n-r)$ 個，其可能之分配數當為 $N'^{(n-r)}$。而前者之任何一分配皆能與後者任何一分配相配合，故含有 r 個之黑球之可能分配數應為 $N''^{r}N'^{(n-r)}$

然此含有 r 個黑球之 r 數列，其在 n 數列中之分配情形，依上述吾人之假定，其實不過其中分配之一耳。此 r 數列在 n 數列中之可能分配情形，當為數學上 n 項中取 r 之組合 $_nC_r$

$$_nC_r = \frac{n(n-1)\cdots\cdots(n-r+1)}{r!} = \frac{n!}{r!(n-r)!}$$

是則含有 r 個黑球之一切可能分配數應為：

$$P_r^n = \frac{n(n-1)\cdots\cdots(n-r+1)}{r!} N''^{r} N'^{(n-r)}$$

設以 p 代表其中總球數 N 與黑球數 N'' 之比例，$p=\frac{N''}{N}$，以 q 代表 N 與白球數 N' 之比例，$q=\frac{N'}{N}$，則 $p+q=1$ 而含有 r 黑球之可能分配數與 n 球可能分配數之比例將為：

$$P_r = \frac{P_r^n}{N^n} = \frac{n(n-1)\cdots(n-r+1)}{r!}\left(\frac{N'}{N}\right)^{(n-r)}\left(\frac{N''}{N}\right)^r$$

$$P_r = \frac{n(n-1)\cdots\cdots(n-r+1)}{r!}p^r q^{n-r}$$

是則由袋內抽取n球含有r黑球之机率，即為上列公式之P_r。至r之數值，則以由袋內抽取之n球中所得之黑球數，或為0，或為1，或為2，……或為(n−1)或為n，變動于0與n之間，計共有(n+1)項，即可取下列0，1，2，……(n−1)，n，(n+1)項中之任何一項。r之數值既變動于0與n之間，則含有r黑球之可能分配數，當亦隨隨之變動，以r之數值代入P_r^n，是於該數值時之可能分配數。

例如　$r=0$　$P_r^n = N'^n$

此種為當然之結果。試以r既等于零，則僅由白球之N'數列中繼續抽取n項之結果，其可能分配數當為N'^n

$r=1$　$P_1^n = n \cdot N'^{(n-1)} N''$

$r=2$　$P_2^n = \frac{n(n-1)}{2} N'^{(n-2)} N''^2$

- - - - - - - - - - - - - - - -

$r=n$　$P_r^n = N''^n$

但以比例表示之得上述之 P_r 以 r 之數值代入上列 P_r 之公式即可獲得 r 等於該數值時之机率 由袋內抽取 n 球含有 r 黑球之可能比例計其為下列之 $(n+1)$ 項

$$0,\ \frac{1}{n},\ \frac{3}{n} \cdots\cdots \frac{r}{n} \cdots\cdots 1$$

其相對之机率為: $q^n + {}_nC_1 q^{n-1}p + \cdots\cdots + {}_nC_r q^{n-r}p^r + \cdots + p^n$

由此可見比數亦相對之各項机率與二項列式 $(q+p)^n$ 展開後所得之各項無异.

$$(q+p)^n = q^n + {}_nC_1 q^{n-1}p + {}_nC_2 q^{n-2}p^2 + \cdots + {}_nC_r q^{n-r}p^r + \cdots + p^n$$

故由上列比數亦所組成之數列,稱為二項分配

因此以後二項分配與於擲錢於地,或得錢背或得錢面之試驗相似,試以袋內之球,分為二部,其一部為 N' 白球,其另為 N'' 黑球,故由袋內抽得之一球,不出白色黑色二種可能,於是吾人自可將 N' 白球上均刻一字,於 N'' 黑球上均刻零字,則由袋內抽得一球或為一字或為零字,此為二項分配原則相符

袋內之 N'' 球均等於零, N' 球均等於1,則整个 N 球之平均數

$$\frac{N''\cdot 0 + N'\cdot 1}{N} = \frac{N'}{N} = p$$

由袋內抽得n球之平均數

$$\frac{[(n-r)\times 0]+[r\times 1]}{n}=\frac{r}{n}$$

此n球之平均數，依前節所得之結果言，如抽取之次數n極大，則各n球之平均數與p相接近

然是之故，吾人如依Bernoulli定理所得之結果論斷，可得下列之定理

「黑球在抽得n球中所佔之比例與實際之p之差，如以對之比例"$\frac{r}{n}$"，兩者間之差誤甚小於某種程度，只要加大，不論規定之差誤如何，祇須將抽得之次數n增大，可使於之"

第二節　二項展開式

假定二數a, b之和共有n個乘積，如$(a+b)(a+b)\cdots$至n即$(a+b)^n$從各式中每取一字相乘，即得此連乘積之一項（任何一項均如此）再將各乘相加，即得此連乘積之一切項亦即二數n方之展開式。

先自n因子中各取a，此只一法，故a^n為此連乘積之一項。二次從n因子中取一b，其餘$(n-1)$因子中各取a。取b只一法，取a有n法

亦即從n物中取一之組合 C_1^n 故 $C_1^n a^{n-1}b$ 為其第二項。又次取二b 從(n-2)因子中取各a 有 C_2^n 個 故 $C_2^n a^{n-2}b^2$ 為其第三項……因此 從n因子中取r個b ($r<n$ 且為正整數) 其法數即等於 [illegible] 中取r之組合法 亦即 $C_r^n a^{n-r}b^r$ 為其第r+1項 至其最後一項為 $C_n^n a^{n-n}b^n$ 亦即為 b^n 因此 可得二數n次乘積(乘方)的展開式

$$(a+b)(a+b)(a+b)\cdots\cdots = (a+b)^n$$

$$(a+b)^n = a^n + C_1^n a^{n-1}b + C_2^n a^{n-2}b^2 + \cdots\cdots + C_r^n a^{n-r}b^r + \cdots\cdots + C_n^n a^{n-n}b^n$$

$$= a^n + n a^{n-1}b + C_2^n a^{n-2}b^2 + \cdots + C_r^n a^{n-r}b^r + \cdots + b^n$$

例如、$(a+b)^3 = a^3 + C_1^3 a^{3-1}b + C_2^3 a^{3-2}b^2 + C_3^3 a^{3-3}b^3$

$$= a^3 + 3a^2b + 3ab^2 + b^3$$

$$(a+b)^4 = a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4$$

二項式中其各項之係數、然 其第一項 [illegible] 其係數為1 第二項 其第二項為n，第三項 其第三項為 $\frac{n(n-1)}{2!}$ (因為 $C_2^n = \frac{n(n-1)}{2!}$ 故) 由此我們可以知道 二項式的展開式之兩端 以其中一二 二項為中點而對稱 (就像 [illegible])

至於 $(a-b)^n$ 亦適用此展式，但為正負相間的展式，[illegible] 以 $-b = d$

因此：　$(a-b) = [a+(-b)] = (a+d)$

$\therefore (a+d)^n = a^n + na^{n-1}d + C_2^n a^{n-2}d^2 + \cdots + d^n$

代入 $(-b)$　$\therefore (a-b)^n = a^n + na^{n-1}(-1)b + C_2^n a^{n-2}(-1)^2 b^2 + \cdots (-1)^n b^n$

$= a^n - na^{n-1}b + C_2^n a^{n-2}b^2 - C_3^n a^{n-3}b^3 + \cdots + (-1)^n b^n$

至其末項視 n 之為奇為偶而定。

例：　$(a-b)^3 = a^3 - 3a^2b + 3ab^2 - b^3$　（$n=3$ 為奇末項為負）

$(a-b)^4 = a^4 - 4a^3b + 6a^2b^2 - 4ab^3 + b^4$

$n=4$ 為偶末項為正

第三節　二項分配之求微數

(1)　$m_1' = np$

証　$\because \dfrac{\Sigma fx}{N} = 0 + nq^{n-1}p + (n-1)q^{n-2}p^2 + \dfrac{n(n-1)(n-2)}{2!}q^{n-3}p^3$

$+ \cdots\cdots np^n$

$= np[q^{n-1} + (n-1)q^{n-2}p + \dfrac{(n-1)(n-2)}{2!}q^{n-3}p^2 \cdots + p^{n-1}]$

$= np[q+p]^{n-1} = np$

$\therefore m_1' = np$

(2)　$m_2 = npq$　$\sigma = \sqrt{npq}$

證 $m_3' = $ [illegible]

[illegible]

$= np(1-p) + n^2p^2$ [illegible]

$= np - np^2 = np(1-p) = npq$

$\therefore \sigma = \sqrt{npq}$

(3) $\alpha_3 = \frac{m_3}{\sigma^3} = \frac{q-p}{\sqrt{npq}}$

高等統計學

茲先証明。

$$m_3' = \frac{\Sigma f x^3}{\Sigma f} = np[1+p(n-1)\{3+(n-2)p\}]$$

$$\because \Sigma f x^3 = 0 + nq^{n-1}p + 4n(n-1)q^{n-2}p^2 + \frac{9}{2!}n(n-1)(n-2)q^{n-3}p^3 + \cdots + n^3p^n$$

$$= np[q^{n-1} + 4(n-1)q^{n-2}p + \frac{9}{2!}(n-1)(n-2)q^{n-3}p^2 + \cdots + n^2p^{n-1}]$$

$$= np[\{q^{n-1} + (n-1)q^{n-2}p + \frac{1}{2!}(n-1)(n-2)q^{n-3}p^2 + \cdots + p^{n-1}\} + 3(n-1)q^{n-2}p + 4(n-1)(n-2)q^{n-3}p^2 + \cdots + (n^2-1)p^{n-1}]$$

$$= np[(q+p)^{n-1} + (n-1)p\{3q^{n-2} + 4(n-2)q^{n-3}p + \cdots (n+1)p^{n-2}\}]$$

$$= np[1+(n-1)p\{3q^{n-2} + 3(n-2)q^{n-3}p + \cdots + 3p^{n-2} + (n-2)q^{n-3}p + \cdots (n-2)p^{n-2}\}]$$

$$= np[1+(n-1)p\{3(q^{n-2} + (n-2)q^{n-3}p + \cdots p^{n-2}) + (n-2)p(q^{n-3} + \cdots p^{n-3})\}]$$

$$= np[1+(n-1)p\{3(q+p)^{n-2} + (n-2)p(q+p)^{n-3}\}]$$

$$= np[1+(n-1)p\{3+(n-2)p\}]$$

$\therefore m_3 = m_3' - 3m_2'm_1' + 2m_1'^3$

$= np[1+(n-1)p\{3+(n-2)p\}] - 3np[np\{1+(n-1)p\}] + 2n^3p^3$

$= np[1+(np-p)\{3+np-2p\} - 3np\{1+np-p\} + 2n^2p^2]$

$= np[1+3np-3p+n^2p^2-2np^2-np^2+2p^2-3np-3n^2p^2+3np^2+2n^2p^2]$

$= np[1-3p+2p^2]$

$= np[1-2p+p^2-p+p^2]$

$= np[(1-p)^2-(1-p)p]$

$= np(q^2-qp) = npq(q-p)$

$\therefore \alpha_3 = \frac{m_3}{\sigma^3} = \frac{npq(q-p)}{(\sqrt{npq})^3} = \frac{npq(q-p)}{(npq)^{3/2}} = \frac{npq(q-p)}{npq(npq)^{\frac{1}{2}}} = \frac{q-p}{\sqrt{npq}}$

(4) $\alpha_4 = \frac{m_4}{\sigma^4} = 3 + \frac{1-6pq}{npq}$

証. $m_4 = m_4' - 4m_3'm_1' + 6m_2'm_1'^2 - 3m_1'^4$

而 $m_1' = \frac{\Sigma f x^4}{N} = [0 + np^{n-1}p + 8n(n-1)q^{n-2}p^2 + \frac{27}{2!}n(n-1)(n-2)q^{n-3}p^3$
$+ \cdots\cdots n^4p^n]$

$= np[q^{n-1} + 8(n-1)q^{n-2}p + \frac{27}{2!}(n-1)(n-2)q^{n-3}p^2 + \cdots n^3p^{n-1}]$

$$= np\{q^{n-1} + (n-1)q^{n-2}p + \tfrac{1}{2!}(n-1)(n-2)q^{n-3}p^2 + \cdots\ p^{n-1}$$

$$+ 7(n-1)q^{n-2}p + 13(n-1)(n-2)q^{n-3}p^2 + \cdots (n^3 \cdots)p^{n-1}]$$

$$= np\{(q+p)^{n-1} + (n-1)p\{7q^{n-2} + 13(n-2)q^{n-3}p + \cdots (n^2-n+1)p^{n-2}$$

$$= np\{4(n-1)p\{7q^{n-2} + 7(n-2)q^{n-3}p + \cdots 7p^{n-2} + 6(n-2)q^{n-3}p$$

$$+ \cdots (n^2+n-1)p^{n-2}\}]$$

$$= np[1 + (n-1)p\{7(q+p)^{n-2} + (n-2)p\{6q^{n-3} + \cdots (n-3)p^{n-2}\}]$$

$$= np[1 + (n-1)p\{7 + (n-2)p\{6q^{n-3} + \cdots 6p^{n-3} + p(n-3)\}\}]$$

$$= np[1 + (n-1)p\{7 + (n-2)p\{6(q+p)^{n-3} + (n-3)p\}\}]$$

$$= np[1 + (n-1)p\{7 + (n-2)p\{6 + (n-3)p\}\}]$$

$$\therefore m_4 = m_4' - 4m_3' m_1' + 6m_2' m_1'^2 - 3m_1'^4$$

$$= np[1 + (n-1)p\{7 + (n-2)p\{6 + (n-3)p\}\}] - 4np \cdot np[1$$

$$+ (n-1)p\{3 + (n-2)p\}] + 6n^2p^2 \cdot np[1 + (n-1)p] - 3n^4p^4$$

$$= np[[1 + np - p\{7 + (np-2p)(6 + (np-3p))\}] - 4np[1 + (np-p)$$

$$(3 + np - 2p) + 6n^2p^2(1 + (np-p)) - 3n^3p^3]]$$

$$= np[[1 + (np-p)\{7 + 6np - 12p + n^2p^2 - 2np^2 - 3np^2 + 6p^2\}$$

$$-4np[1+3np-3p+n^2p^2-np^2-4np^2+2p^2+6n^2p^2+6n^3p^2-6n^2p^3$$
$$-3n^3p^4]]$$

$$= np[1+7np-7p+6n^2p^2-6np^2-2np^2+12p^2+n^3p^3-n^2p^3-5n^2p^3$$
$$+5np^3+6np^3-6p^3 \text{[illegible]} +12np \text{[illegible]} +4n^2p^3$$
$$+8n^2p^3-8np^3+6n^2p^3+6n^3p^3 \text{[illegible]} -3n^3p^3]$$

$$= np[1-7p+12p^2-6p^3+3np-6np^2+3np^3]$$

$$= np[(1-p)(1-6p+6p^2)+3np(1-2p+p^2)]$$

$$= np[\{(1-p)(1-6p)(1-p)\}+3np(1-p)^2]$$

$$= np[q(1-6pq)+3npq^2]$$

$$= npq(1-6pq+3n^2p^2q^2$$

$$= npq[1-6qp+3npq]$$

$$\alpha_4 = \frac{\mu_4}{\sigma^4} = \frac{npq[(1-6pq)+3npq]}{(\sqrt{npq})^4} = \frac{npq(1-6pq+3npq)}{(npq)^2}$$

$$= \frac{3npq}{npq} + \frac{1-6pq}{npq} = 3 + \frac{1-6pq}{npq}$$

第四節　二項分配之應用

(I) 用以估計理論次數

例：今抽選800粒種子分成80列種下，若發芽之種子為零粒不發芽為6次，求其每列種子發芽之實際次數和理論次數之比較。如表

每列發芽之個數	0	1	2	3	4	5	6	7	共計
次數	6	20	28	12	8	6	-	-	80

據上表求得每列平均發芽數為 $\frac{174}{80}=2.175$ 粒

解：發芽之机率為 $P=\frac{2.175}{10}=0.2175$

不發芽之机率為 $q=\frac{10-2.175}{10}=0.7825$

依二項展開式 $N(p+q)^n$ 以求其理論次數

$80(0.7825+0.2175)^{10}$

$=80\times(.7825)^{10}+80\times10\times(.7825)^{9}\times0.2175+\cdots\cdots$

$=6.9+19.1\cdots\cdots$

(II) 測定机率。

例：估計擲銅元10,000次求擲出面之机率

解：設擲出面之机率為P，則 $P=\frac{1}{2}$

擲出背之机率為q，則 $q=\frac{1}{2}$

則據本例數字 $\bar{x}=np$ 其擲出面之數學希望數為：

10,000×0.5=5000次，此為最佳之估計值。

又據 $\sigma=\sqrt{npq}=\sqrt{10000\times\frac{1}{2}\times\frac{1}{2}}=\sqrt{2500}=50$

則 $\bar{x}\pm2\sigma=5000\pm100$ 即 5000−100 與 5000+100 = 4900 到 5100，而此現象之概率為 約95% 在 4900 到 5100 之範圍內

又 $\bar{x}\pm3\sigma=5000\pm3\times50=5000\pm150=4850$ 到 5150

即此現象之概率有 99%在 4850 到 5150 之範圍內

(四) 在保險上之應用

例 某保險公司，保有 10,000，將其保額為 10,000,000，平均每家保費為 1000元，每年該公司的管理費為 5000元，試問此公司須向保戶每人收費若干為適當，按照統計發生火災概率為 $\frac{1}{1000}$

解 該公司 10,000家中被火災的為 $10,000\times\frac{1}{1000}=10$ 家

則 $\bar{x}=np=10$ 家 $n=10,000$ $p=\frac{1}{1000}$ $q=1-\frac{1}{1000}=\frac{999}{1000}$

$\sigma=\sqrt{10000\times\frac{1}{1000}\times\frac{999}{1000}}=\sqrt{10}=3.3$

99.7%被火災之家數應為 $\bar{x}\pm3\sigma=10\pm3\times3.3=10\pm9.9$

即為 10−9.9 到 10+9.9

即在 0—20 家範圍內

設20家為被火災，則應賠償之保費為 20×10,000=200,000元

高等統計學 P.30

其保費20,000元＋管理費5000 總計25000元 則每家1000元應納保費為 25000 ÷ 1000 = 25元

IV 測定死亡率

例：設25人都是30歲，問此25人在一年內死亡一人之機率為若何？

解：30歲的人之死亡率為 $\frac{8}{1000}=0.008=p$ $\quad 1-0.008=0.992=q \quad n=25$

依二項式：$(0.008+0.992)^{25}=(0.008)^{25}+25(0.008)^{24}(0.992)+25(0.008)^{23}(0.992)^{2}$

$+\cdots\cdots+25(0.008)(0.992)^{24}+(0.992)^{25}$

其中 $(0.992)^{25}$ 代表25人均長壽，今要求死亡一人之機率為：

$25(0.992)^{24}(0.008)+(0.992)^{25}$ 其0.992之計算出來結果近於一，而0.008近於零，故知其屬於偏態分配，則偏斜非常之利害

又如此例每團500人中最大可能死若干人 $(0.992+0.008)^{500}$

其計算方法有二：

(1) 如將原式展開最大係數項即為所求之數

(2) 即為平均數法 $\bar{x}=np$

V 測定統計數字之可靠性

例一 如某國測定各地所報嬰孩出生數是否正確等

如所報之數合於吾人所希望者（由計算得出之機率）為正確，不合於吾人之希望者（與機率不合）為不正確，其特例如下

如測定男女分配

由麥在1860年—1879年嬰孩出生比率統計如下表

年份	男	女
1860—4	130089	138389
1856—9	138324	142828
1870—4	139733	148360
1875—9	154214	162823

解：設以P為出生男孩

$$P_1=\frac{138289}{268378}=0.5156 \text{（經驗數）}$$

$$P_2=\frac{142828}{278152}=0.5135$$

$$P_3=\frac{148360}{288093}=0.5149$$

$$P_4=\frac{162823}{317037}=0.5136$$

$P=0.5144$ 平均數作為機急的理想的P

$\bar{x}=np=268,378\times0.5144=138,072$

$\sigma=\sqrt{npq}=\sqrt{268378\times.5144\times0.4856}$

故99.7%之男性之出生率在 ±3σ 之間時

即在138289−137530之间而與实際数字差不多的，故知

1860—4年之男孩出生率是相当正確的 徐劼雅

第四章 常態分配

第一節 常態曲線之意義、演進及其發展

1. 意義：

曲線之幾何圖形，在橫軸之上方呈對稱鐘形者，即曲線由最低點漸升，經過一轉曲點，而繼續上升至頂峯，然後開始下降，再經過另一轉曲點，而漸次下降至最低點，如斯情形之兩邊相等，則此所得之曲線，即稱為常態曲線（Normal curve）。如將此種曲線應用於二項定理在機率問題上者，則稱之為或率曲線（probability curve）

法國數學大師拉普拉斯（Laplace）曾由研究最小二乘法着手，而發現常態曲線公式，因此後人稱之為（Laplace Curve）拉普拉斯曲線。同時在德國亦有一位數學家高斯（Gauss）因研究誤差平均數之分配，而發現常態曲線之公式，故德人又喜稱之為高斯曲線（Gauss Curve）

2. 演進

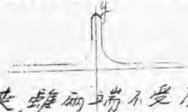
遠在十七世紀以前，由於數學本身之發展尚未完備，故所得者多係合於若干後降之公式者，並不是比較複雜的常態公式，乃是半圓形曲線公式：

$$y = +\sqrt{r^2 - x^2}$$

如圖所示，此種曲線雖具有常態曲線之輪廓，但其然用不能因其兩端有限制，頗不合實際資料之原理

先後有近似常態曲線公式由半圓形曲線演變而成，後人稱之為古代常態曲線（Cauchy's curve）其公式及圖形如下：

$$y = \frac{b}{\pi(b^2 + x^2)}$$



如圖所示，此種曲線因峰度太狹，雖兩端不受限制，但因次數過於集中，頗不合常態分配原則

由於伯努力 Bernouilli 首創二項定理在機率方面的應用所提示，故在拉普拉斯及高斯兩位數學大師，繼起研究，拉氏用研究最小二乘法，而發現常態曲線公式：

$$y = \frac{h}{\sqrt{\pi}} e^{-h^2 x^2}$$

高氏於研究算術平均數之性質，而同時發現常態曲線公式：

$$\phi(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2}$$

兩氏各自先後獨立研究，雖出發點不同，而所得結果完全相同，故常態曲線發展史中，兩氏實建立了特殊功勛。

最後英國生物統計學家皮爾生(Pearson)教授繼起研究此種曲線之各種偏態與峯度，而將常態曲線視為其所創公式之一型，可用微分方程式表之如下：

$$\frac{dy}{dx}=-\frac{1}{c_0}xy$$

3. 發展

過去一般統計學者，凡研究次數理論分配者，莫不以常態分配為研究之對象。蓋以常態分配合於統計推理法則及大量錯誤法則，同時又由於常態分配能夠說明社會多數現象。因之以前一般統計學者，以為一切社會現象均合乎常態分配，只要觀察現象之次數無限增加。

自從1904年荷蘭Kapstein發表其偏態分配，研究關於蘋果其徑和重量分配之結果時，方開始證明社會上有些現象不合常態分配。凱氏發現觀察相同數量之蘋果，其直徑最合乎常態，而其重量之分配則為偏態；而此種偏態現象，並不因蘋果數量之增加而有所改變。

雖然如是之後，研究偏態分配之統計學，風起雲湧，至皮爾生教授，始集各家之大成，創立皮爾生微分方程式：

$$\frac{dy}{dx}=\frac{(x-a)y}{C_0+C_1x+C_2x^2}$$

如鐘形曲線之常態，及各式偏態，均包含於此範式之中

近代統計學者，研究次數分配論者，除鐘形外，人口統計學家常視各國人口成長曲線為S型，社會統計學者現薪。一資之一分配，合於J型；天文統計學家，研究太空陰雲之分佈為U型

雖然統計資料之次數分配，已由常態發展至偏態，更要化各型，但其中任何一型分配平均數之分型，統計學者恒能證明之為常態分配，故研究常態分配，實為研究其他任何分配理論之基礎。

在許多適合常態分配之社會現象中，尤以人之身長最為適合常態分配。

第二節　常態公式證明

在二項分配當 $S\longrightarrow\infty$ 時可自斷的二項數列，成為連續的常態機率曲線

$$(p+q)^S=f(x)=\frac{S!}{x!(S-x)!}p^xq^{S-x}$$

$f(x) = y \quad mx = sp \quad \bar{x} = x - sp \quad x = \bar{x} + sp$

$$y = f(x) = f(\bar{x}) + sp$$

$$= \frac{s!}{(\bar{x}+sp)!(sq-\bar{x})!} p^{sp+\bar{x}} q^{(sq-\bar{x})}$$

$$= \frac{\sqrt{2\pi}\, s^{s+\frac{1}{2}} e^{-s} p^{sp+\bar{x}} q^{(sq-\bar{x})}}{\left[\sqrt{2\pi}(\bar{x}+sp)^{sp+\bar{x}+\frac{1}{2}} e^{-sp-\bar{x}}\right]\left[\sqrt{2\pi}(sq-\bar{x})^{sq-\bar{x}+\frac{1}{2}} e^{-sq+\bar{x}}\right]}$$

此利用 Stirling formula $\left(e^{-sp-\bar{x}} e^{-sq+\bar{x}} = e^{-s(p+q)} = e^{-s}\right)$

$$\therefore f(x) = \frac{s^{s+\frac{1}{2}} p^{sp+\bar{x}} q^{sq-\bar{x}}}{\sqrt{2\pi pq}\,(sp+\bar{x})^{sp+\bar{x}+\frac{1}{2}}(sq-\bar{x})^{sq-\bar{x}+\frac{1}{2}}}$$

$$= \frac{s^{s+\frac{1}{2}}}{\sqrt{2\pi pq}\,\frac{(sp+\bar{x})^{sp+\bar{x}+\frac{1}{2}}}{p^{sp+\bar{x}+\frac{1}{2}}} \cdot \frac{(sq-\bar{x})^{sq-\bar{x}+\frac{1}{2}}}{q^{sq-\bar{x}+\frac{1}{2}}}}$$

$$= \frac{s^{s+\frac{1}{2}}}{\sqrt{2\pi pq}\left(\frac{sp}{p}+\frac{\bar{x}}{p}\right)^{sp+\bar{x}+\frac{1}{2}}\left(\frac{sq}{q}-\frac{\bar{x}}{q}\right)^{sq-\bar{x}+\frac{1}{2}}}$$

$$= \frac{S^{s+\frac{1}{2}}}{\sqrt{2\pi pq}\,\left(S+\frac{S\bar{x}}{Sp}\right)^{Sp+\bar{x}+\frac{1}{2}}\left(S-\frac{S\bar{x}}{Sq}\right)^{Sq-\bar{x}+\frac{1}{2}}}$$

$$= \frac{S^{s+\frac{1}{2}}}{\sqrt{2\pi pq}\,S^{1+s}\left(1+\frac{\bar{x}}{Sp}\right)^{Sp+\bar{x}+\frac{1}{2}}\left(1-\frac{\bar{x}}{Sp}\right)^{Sp-\bar{x}+\frac{1}{2}}}$$

$$= \frac{1}{\sqrt{2\pi pqS}\left(1+\frac{x}{Sp}\right)^{Sp+\bar{x}+\frac{1}{2}}\left(1+\frac{x}{Sq}\right)^{Sq-\bar{x}+\frac{1}{2}}}$$

$$= \frac{1}{\sqrt{2\pi pqS}}\left(1+\frac{x}{Sp}\right)^{-Sp+\bar{x}+\frac{1}{2}}\left(1-\frac{\bar{x}}{Sp}\right)^{-(Sq-\bar{x}+\frac{1}{2}}$$

求: $\log\left(1+\frac{\bar{x}}{Sp}\right)^{-(Sp+\bar{x}+\frac{1}{2})}$ $\quad$ $\log\left(1-\frac{\bar{x}}{Sf}\right)^{-(Sp-\bar{x}+\frac{1}{2})}$

$-\left(Sp+\bar{x}+\frac{1}{2}\right)\log\left(1+\frac{\bar{x}}{Sp}\right)$ 展開

$$= -\left(Sp+\bar{x}+\frac{1}{2}\right)\left(\frac{\bar{x}}{Sp}-\frac{\bar{x}^2}{2S^2p^2}+\frac{\bar{x}^3}{3S^3p^3}+\cdots\cdots\right. \quad (1)$$

$-(sp-\bar{x}+\frac{1}{2})\log(1-\frac{x}{sp})$ 展开

$$=+(sq-\bar{x}+\frac{1}{2})\frac{\bar{x}}{sq}+\frac{\bar{x}^2}{2s^2q^2}+\frac{\bar{x}^3}{3q^3s^3}\text{-----------}(2)$$

(1)式乘出结果 $=-\bar{x}+\frac{\bar{x}^2}{2sp}-\frac{\bar{x}^3}{3s^2p^2}-\frac{\bar{x}^2}{sp}+\frac{\bar{x}^3}{2s^2p^2}+\frac{x^4}{3s^3p^3}-\frac{\bar{x}}{2sp}+\text{---}$

若 $s\to\infty$ 则 $sp\to$ 為确定數值

$\because \sigma=\sqrt{spq}=f(s^{\frac{1}{2}})$

$\bar{x}=x-sp$ $\quad M_x \pm N\sigma_x = sp \pm N\sqrt{spq}$

$\therefore x = f(s^{\frac{1}{2}})$

則 $\frac{\bar{x}^3}{3s^3p^3}=\frac{\{f(s^{\frac{1}{2}})\}^3}{3s^2p^2}=\frac{s^{\frac{1}{2}\cdot 3}}{3s^2p^2}=\frac{s^{\frac{3}{2}}}{3s^2p^2}=0$

則 $\cdots\to$ 其餘各項亦趨於零僅剩

$$-x-\frac{\bar{x}^2}{2sp}-\frac{2\bar{x}^2}{2sp}=-\bar{x}-\frac{\bar{x}^2}{2sp}\text{------------}(1)$$

同様得。$x-\frac{\bar{x}^2}{2sp}+\frac{2\bar{x}^2}{2sp}=+\bar{x}-\frac{\bar{x}^2}{2sp}\text{------------}(2)$

$$(1)+(2) = -\frac{x^2}{2sp} - \frac{x^2}{2sq} = -\frac{x^2(q+p)}{2spq} = -\frac{x^2}{2spq}$$

$$f(x) = \frac{1}{\sqrt{2\pi spq}}\, e^{-\frac{x^2}{2sp} - \frac{x^2}{2sq}} = \frac{1}{\sqrt{2\pi spq}}\, e^{-\frac{x^2}{2spq}}$$

$$= \frac{1}{\sqrt{2\pi}\,\sigma_x}\, e^{-\frac{(x-m_x)^2}{2\sigma_x^2}}$$

$\because spq = \sigma_x^2 \qquad x = X - sp = X - m_x \qquad t = \frac{X - m_x}{\sigma_x}$

則 $f(t) = [illegible] \frac{1}{\sqrt{2\pi}}\, e^{-\frac{t^2}{2}}$

第二節 常態曲線之統計性質

(a) 在常態次數分配中 [illegible] 由於 [illegible] 平均數、中數及眾數三值相等，

即

$\bar{x} = M_d = M_o$

(b) 在常態次數分配中 ~~[illegible]~~ [illegible]

證：

[illegible] $\int_{-\infty}^{x} [illegible] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} [illegible]\, dx$

$$=\frac{1}{\sqrt{2\pi}}\left\{\int_{-\infty}^{\infty} t^{2n+1}e^{-\frac{t^2}{2}}dt+\int_{-\infty}^{0} t^{2n+1}e^{-\frac{t^2}{2}}dt\right\}$$

令 $t=-s$ 代入上式之末項得

$$\mu_{2n+1}=\frac{1}{\sqrt{2\pi}}\left\{\int_0^{\infty} t^{2n+1}e^{-\frac{t^2}{2}}dt+\int_{\infty}^{0} s^{2n+1}e^{-\frac{s^2}{2}}ds\right\}$$

$$=\frac{1}{\sqrt{2\pi}}\left\{\int_0^{\infty} t^{2n+1}e^{-\frac{t^2}{2}}dt-\int_0^{\infty} s^{2n+1}e^{-\frac{s^2}{2}}ds\right\}$$

再令末項(上式)中之 $s=t$ 得

$$\mu_{2n+1}=\frac{1}{\sqrt{2\pi}}\left\{\int_0^{\infty} t^{2n+1}e^{-\frac{t^2}{2}}dt-\int_0^{\infty} t^{2n+1}e^{-\frac{t^2}{2}}dt\right\}=0$$

(C) 在常態次數分配中各偶次動差，可以下式表之：

$$\mu_{2n}=\frac{(2n)!}{2^n n!}$$

證： $\mu_{2n}=\int_{-\infty}^{\infty} t^{2n}\phi(t)dt=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} t^{2n}e^{-\frac{t^2}{2}}dt$

$$=\frac{1}{\sqrt{2\pi}}t^{2n-1}e^{-\frac{t^2}{2}}\Big|_{-\infty}^{\infty}+\frac{(2n-1)}{\sqrt{2\pi}}\int_{-\infty}^{\infty} t^{2n-2}e^{-\frac{t^2}{2}}dt$$

$$= -\frac{\mu_2}{D}(4\mu_2\mu_4 - 3\mu_3^2) = \frac{-\mu_2(4\mu_2\mu_4 - 3\mu_3^2)}{10\mu_2\mu_4 - 12\mu_3^2 - 18\mu_2^3}$$

$$= \frac{-\mu_2\left(\frac{4\mu_4}{\mu_2^2} - 3\frac{\mu_3^2}{\mu_2^3}\right)}{10\frac{\mu_4}{\mu_2^2} - 12\frac{\mu_3^2}{\mu_2^3} - 18} = \frac{-\mu_2(4\beta_2 - 3\beta_1)}{10\beta_2 - 12\beta_1 - 18}$$

$$= \frac{-\mu_2(4\beta_2 - 3\beta_1)}{K}$$

$$C_1 = \frac{1}{D}\begin{vmatrix} 1 & -\mu_2 & 2\mu_2 \\ 0 & -\mu_3 & 4\mu_3 \\ 3\mu_2 & -\mu_4 & 5\mu_4 \end{vmatrix} = \frac{1}{D}(-5\mu_3\mu_4 - 12\mu_2^2\mu_3 + 9\mu_2^2\mu_3 + 4\mu_3\mu_4$$

$$= \frac{1}{D}(-5\mu_3\mu_4 - 3\mu_2^2\mu_3 + 4\mu_3\mu_4) = \frac{1}{D}(-\mu_3\mu_4 - 3\mu_2\mu_3)$$

$$= \frac{-\mu_3}{D}(\mu_4 + 3\mu_2^2) = \frac{-\mu_3(\mu_4 + 3\mu_2^2)}{10\mu_2\mu_4 - 12\mu_3^2 - 18\mu_2^3} = \frac{-\mu_3\left(\frac{\mu_4}{\mu_2^2} + 3\right)}{\frac{10\mu_4}{\mu_2^2} - 12\frac{\mu_3^2}{\mu_2^2} - 18}$$

$$= \frac{-\mu_3(\beta_2 + 3)}{\mu_2 K} = \frac{(\beta_2 + 3)\frac{\mu_3}{\mu_2}}{K}$$

(1) 常態曲線之全面積(亦即相對次數之和)等於一
證:

$$A=\int_{-\infty}^{\infty}\phi(x)\,dx=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-\frac{x^2}{2}}\,dx=\frac{2}{\sqrt{2\pi}}\int_{0}^{\infty}e^{-\frac{x^2}{2}}\,dx$$

令 $\sqrt{\theta}=\frac{x}{\sqrt{2}}$　$dx=\frac{d\theta}{\sqrt{2\theta}}$　代入上式則

$$A=\frac{1}{\sqrt{\pi}}\int_{0}^{\infty}e^{-\theta}\theta^{-\frac{1}{2}}\,d\theta=\frac{1}{\sqrt{\pi}}\Gamma\left(\tfrac{1}{2}\right)=\frac{1}{\sqrt{\pi}}\cdot\sqrt{\pi}=1$$

(2) 在常態次數分配中, 自算術平均數(M)與向左右取等距離之條其間面積恰好為全面積之半, 則此兩點為 $\pm P.E$ 稱為机誤差

由. $A=\int_{-P.E}^{+P.E}\phi(x)\,dx=2\int_{0}^{P.E}\phi(x)=\frac{1}{2}$

因根據皮爾生常態曲線面積查得 $P.E=0.6745$

第四節　常態曲線表格之作法及用法

常態曲應用之表格有兩種:
a, 常態曲線下面積表 Areas of the normal curve
b, 常態曲線下縱坐標表 ordinate of the normal curve
1, 今分別敘述兩種表格之作法如下:

由方程式

$$\frac{1}{y}\frac{dy}{dx} = \frac{x-a}{c_0+c_1x+c_2x^2}$$

$$\int_u^v (c_0+c_1x+c_2x^2)\,dy = \int_u^v (x-a)\,y\,dx \qquad (11)$$

上式兩端乘以 x^n 得:

$$\int_u^v (c_0+c_1x+c_2x^2)x^n\,dy = \int_u^v (x-a)x^n\,y\,dx$$

用部分積分法使

$$(c_0+c_1x+c_2x^2)x^n = u \qquad du = dy$$

$$du = nc_0x^{n-1}+(n+1)c_1x^n+(n+2)c_2x^{n+1}\,dx \qquad v = y$$

(11)之左端可變為:

$$\left|(c_0+c_1x+c_2x^2)x^n y\right|_u^v - \int_u^v \{y(nc_0x^{n-1}+(n+1)c_1x^n+(n+2)c_2x^{n+1})\}\,dx$$

即 $-\int_u^v[nc_0x^{n-1}+(n+1)c_1x^n+(n+2)c_2x^{n+1}]\,y\,dx = \int_u^v (x-a)yx^n\,dx$

$$-nc_0\mu_{n-1}-(n+1)c_1\mu_n-(n+2)c_2\mu_{n+1} = \mu_{n+1}-a\mu_n$$

$$nc_0\mu_{n-1}+[(n+1)c_1-a]\mu_n+[(n+2)c_2+1]\mu_{n+1} = 0$$

令 $n = 0, 1, 2, 3$, 代入上式得

$$c_1 = a$$

依照計算各項和數方法，吾人可以求出至任何位小數已確定之答案：

例求 $t=0$；$t=.05$ 間之面積：

$$\frac{1}{\sqrt{2\pi}}\int_0^{.05} e^{-\frac{t^2}{2}}dt=\frac{1}{\sqrt{2\pi}}\left[.05-\frac{(.05)^3}{2\cdot3}+\frac{(.05)^5}{5\cdot2^2 2!}-\frac{(.05)^7}{7\cdot2^3 3!}+\cdots\cdots\right]=.39894(.04998)=.01993$$

b. t 為較大數值計算面積之方法

若吾人欲求常態曲線下自 $t=0$ 至 $t=t_1$ 之面積而 t_1 為較大數值則可以下列式：

$$\frac{1}{\sqrt{2\pi}}\int_0^{t_1} e^{-\frac{t^2}{2}}dt=\frac{1}{2}-\frac{1}{\sqrt{2\pi}}\int_{t_1}^{\infty} e^{-\frac{t^2}{2}}dt \quad 而$$

$$\int_{t_1}^{\infty} e^{-\frac{t^2}{2}}dt\cdot\frac{t}{t}=\frac{-e^{-\frac{t^2}{2}}}{t}\Big]_{t_1}^{\infty}-\int_{t_1}^{\infty}\frac{1}{t^2}e^{-\frac{t^2}{2}}dt$$

但

$$\int_{t_1}^{\infty}\frac{t}{t^3}e^{-\frac{t^2}{2}}dt=\frac{-e^{-\frac{t^2}{2}}}{t^3}\Big]_{t_1}^{\infty}-\int_{t_1}^{\infty}\frac{3}{t^4}e^{-\frac{t^2}{2}}dt$$

$$=\frac{e^{-\frac{t_1^2}{2}}}{t_1}+\frac{e^{-\frac{t^2}{2}}}{t^3}\Big]_{t_1}^{\infty}+3\int_{t_1}^{\infty}\frac{e^{-\frac{t^2}{2}}}{t^4}dt=\frac{e^{-\frac{t_1^2}{2}}}{t_1}-\frac{e^{-\frac{t_1^2}{2}}}{t_1^3}+\frac{3e^{-\frac{t_1^2}{2}}}{t_1^5}-3\cdot5\int_{t_1}^{\infty}\frac{e^{-\frac{t^2}{2}}}{t^6}dt$$

因 $t=\infty$，$\frac{e^{-\frac{t^2}{2}}}{t}=\frac{1}{te^{\frac{t^2}{2}}}=\frac{1}{\infty}=0$

依此方法繼續進行，吾人得：

$$\int_{t_1}^{\infty}\frac{t}{t}e^{-\frac{t^2}{2}}dt=\frac{e^{-\frac{t_1^2}{2}}}{t_1}\left[1-\frac{1}{t_1^2}+\frac{1}{t_1^4}-\frac{3\cdot5}{t_1^6}+\frac{3\cdot5\cdot7}{t_1^8}+\cdots\frac{(-1)^k3\cdot5\cdots(2k-1)}{t_1^{2k}}\cdots\right]$$

此為常態曲線之形式

設 $p \neq q$ 時,吾人可得

$$\frac{1}{y}\frac{dy}{dx}=\frac{2(sp-(x-\frac{1}{2})-q)}{sp+q+(q-p)(x-\frac{1}{2})}=\frac{2(sp-q-x+\frac{1}{2})}{sp+q+(q-p)x-\frac{1}{2}q+\frac{1}{2}p}$$

$$=\frac{2(sp-q-x+\frac{1}{2}p+\frac{1}{2}q)}{sp+\frac{1}{2}+(q-p)x}=\frac{2[(sp-x)+\frac{1}{2}(p-q)]}{sp+\frac{1}{2}+(q-p)(x-sp)+sp(q-p)}$$

$$=\frac{2[(sp-x)+\frac{1}{2}(p-q)]}{2spq+\frac{1}{2}+(q-p)(x-sp)}=\frac{(sp-x)+\frac{1}{2}(p-q)}{spq+\frac{1}{4}+\frac{1}{2}(q-p)(x-sp)}$$

令 $-\frac{p-q}{\sqrt{spq}}=\sqrt{\beta_1}$, $\frac{x-sp}{\sqrt{spq}}=\frac{x-\bar{x}}{\sigma}=t$ 代入上式可得

$$\frac{1}{y}\frac{dy}{dx}=\frac{-\sqrt{spq}\,t-\frac{1}{2}\sqrt{spq}\sqrt{\beta_1}}{spq+\frac{1}{4}+\frac{1}{2}\sqrt{spq}\sqrt{\beta_1}\sqrt{spq}\,t}=\frac{-t-\frac{1}{2}\sqrt{\beta_1}}{\sqrt{spq}+\frac{1}{4\sqrt{spq}}+\frac{1}{2}\sqrt{spq}\sqrt{\beta_1}\,t}$$

即 $\frac{1}{y}\frac{dy}{dx}=\frac{x-a}{c_0+c_1x}$

此方程式適為皮尔生曲線中系中之第三型(Type III),上式中若使 $\beta_1\rightarrow 0$ 則变成常態曲線也

前已述及皮尔生氏所选定之微分方程式為:

第五章　皮尔生曲線系

第一節　緒　言

常態曲綫不足以代表一切同類事体的次數分配，在皮尔生之前早已有人提及。但是大部份統計學者，均認為次數分配皆為常態，因而視常態曲綫為配合曲綫次數之萬应符。其間有不合常態者，实由于材料不充足的緣故，而非事項之本身不為常態也。此說雖為当時多数學者所主張，但按諸事实，未免太牽強附会，削足適履，頗貽人以攻擊之處。蓋宇宙之間，其本身不為常態分配甚多，然自皮尔生之次数理論問世後，始得到明確之証明。

皮尔生從許多生物上的例子，說明次数分配，未必皆為常態，皮氏在1895年，發表他的次数曲綫理論，根据經驗得到几种次數分配極其普遍的性質，於是大家才漸漸的明白了德謨尔(De More)拉不拉斯(Laplace)以及高斯(Gauss)諸氏的常態分配理論，不能包括宇宙間所有的次数分配。

皮尔生之次数曲綫系，初非机率之假定而得，故又称為經驗曲綫系。皮氏之經驗假定為：

1. 倘观察之材料為同類者，次数曲綫之起点之縱坐標為零，其後漸增至極大，再漸次降為零

于零小于一,其三為 K_2 之值大于一,此三种曲綫,通常称為皮氏之主要曲線,其餘十型(常態在内)皆為前款之特殊情形,故常称為皮氏之変型曲綫。

第二節　皮尔生微分方程式

前面已經説过皮尔生氏之次数曲綫,系初非由机率之假定而得,実為經驗曲線,然吾人仍可由机率之立場説明之。

設一袋内有形狀大小輕重相等之球 n 个,其中 np 个為紅球,nq 个為白球,混合在袋内任意摸取 S 个,求 S 球中得 x 紅球之机率。

S 為一固定数,若 n 趨近於無窮大時,則 S/n 之值趨近於零,而此問題之解答,全視其是否趨近于零而定,茲先設

$n \longrightarrow \infty$,　$S/n \longrightarrow 0$

由于伯 意(Bernoulli)定理,吾人知道 S 次試驗中,取得紅球為 0, 1, 2, 3, …… S 个之机率為二項式 $(q+p)^S$ 展開之各項,今設其 $(x+1)$ 項之值用縱坐標 y_x 代表之得

$$y_x = {}_SC_x\, p^x q^{S-x}$$

$$y_{x+1} = {}_SC_{x+1}\, p^{x+1} q^{S-(x+1)}$$

而

$$y_{x+1} \cdot y_x = \frac{S!}{(S-x-1)!\,x!}\, p^x q^{S-x-1} \left(\frac{Sp-q-x}{S-x+Sx-x^2}\right)$$

于零小于一，其三為K_2之值大于一，此三種曲綫通常称為皮氏之主要曲線，其餘十型(常態在内)皆為前款之特殊情形，故常称為皮氏之变型曲綫。

第二節　皮尔生微分方程式

前面已經說过皮尔生氏之次数曲綫系初非由机率之假定而得，实為經驗曲線，然吾人仍可由机率之立場説明之。

設一袋内有形狀大小輕重相等之球n个，其中np个為紅球，nq个為白球，混合在袋内任意摸取S个，求S球中得x紅球之机率。

S為一固定数，若n趨近於無窮大時則S/n之值趋近於零，而此问題之解答，全視其是否趋近于零而定，兹先設

$n\longrightarrow\infty$，　$S/n\longrightarrow 0$

由于伯努意(Bernoulli)定理，吾人知道S次試驗中取得紅球為$0, 1, 2, 3, \cdots, S$个之机率為二項式$(q+p)^S$展開之各項，今設其$(x+1)$項之值用縱坐標y_x代表之得

$$y_x = {}_SC_x\, p^x q^{S-x}$$

$$y_{x+1} = {}_SC_{x+1}\, p^{x+1} q^{S-(x+1)}$$

而

$$y_{x+1}\cdot y_x = \frac{S!}{(S-x-1)!\,x!}\, p^x q^{S-x-1}\left(\frac{Sp-q-x}{S-x+Sx-x^2}\right)$$

第五章　皮尔生曲線系

第一節　緒　言

常態曲綫不足以代表一切同類事体的次數分配，在皮尔生之前早已有人提及。但是大部份統計學者，均認為次數分配皆為常態，因而視常態曲綫為配合曲綫次數之萬应符。其間有不合常態者，实由于材料不充足的緣故，而非事項之本身不為常態也。此説雖為当時多数學者所主張，但按諸事实，未免太牽強附会，削足适履，頗貽人以攻击之處。蓋宇宙之間，其本身不為常態分配甚多，然自皮尔生之次数理論問世後，始得到明確之証明。

皮尔生从許多生物上的例子，説明次數分配，未必皆為常態。皮氏在1895年，發表他的次数曲綫理論，根据経驗得到几種次数分配極其普遍的性質，於是大家才漸漸的明白了德謨尔(De More)拉不拉斯(Laplace)以及高斯(Gauss)諸氏的常態分配理論，不能包括宇宙間所有的次数分配。

皮尔生之次数曲綫系，初非机率之假定而得，故又称為経驗曲綫系。皮氏之経驗假定為：

1. 倘观察之材料為同類者，次数曲綫之起点之縦坐標為零，其後漸增至極大，再漸次降為零

此為常態曲綫之形式

設 $p \neq q$ 時，吾人可得

$$\frac{1}{y}\frac{dy}{dx}=\frac{2(sp-(x-\frac{1}{2})-q)}{sp+q+(q-p)(x-\frac{1}{2})}=\frac{2(sp-q-x+\frac{1}{2})}{sp+q+(q-p)x-\frac{1}{2}q+\frac{1}{2}p}$$

$$=\frac{2(sp-q-x+\frac{1}{2}p+\frac{1}{2}q)}{sp+\frac{1}{2}+(q-p)x}=\frac{2[(sp-x)+\frac{1}{2}(p-q)]}{sp+\frac{1}{2}+(q-p)(x-sp)+sp(q-p)}$$

$$=\frac{2[(sp-x)+\frac{1}{2}(p-q)]}{2spq+\frac{1}{2}+(q-p)(x-sp)}=\frac{(sp-x)+\frac{1}{2}(p-q)}{spq+\frac{1}{4}+\frac{1}{2}(q-p)(x-sp)}$$

令 $-\frac{p-q}{\sqrt{spq}}=\sqrt{\beta_1}$，$\frac{x-sp}{\sqrt{spq}}=\frac{x-\bar{x}}{\sigma}=t$ 代入上式可得

$$\frac{1}{y}\frac{dy}{dx}=\frac{-\sqrt{spq}\,t-\frac{1}{2}\sqrt{spq}\sqrt{\beta_1}}{spq+\frac{1}{4}+\frac{1}{2}\sqrt{spq}\sqrt{\beta_1}\sqrt{spq}\,t}=\frac{-t-\frac{1}{2}\sqrt{\beta_1}}{\sqrt{spq}+\frac{1}{4\sqrt{spq}}+\frac{1}{2}\sqrt{spq}\sqrt{\beta_1}\,t}$$

即 $\frac{1}{y}\frac{dy}{dx}=\frac{x-a}{c_0+c_1x}$

此方程式適為皮尔生曲綫体系中之第三型(Type III)，上式中若使 $\beta_1\to 0$ 則變成常態曲綫也

前已述及皮尔生氏所選定之微分方程式為：

依照計算各項和數方法，吾人可以求出至任何位小數正確性之答案：

例求 $t=0$；$t=.5$ 間之面積：

$$\frac{1}{\sqrt{2\pi}}\int_0^{.05} e^{-\frac{t^2}{2}}dt = \frac{1}{\sqrt{2\pi}}\left[.05 - \frac{(.05)^3}{2\cdot 3} + \frac{(.05)^5}{5\cdot 2^2\, 2!} - \frac{(.05)^7}{7\cdot 2^3\, 3!} + \cdots\cdots\right] = .39894(.04998) = .019939$$

b. t 為較大數值計算面積之方法

設吾人欲求常態曲線下自 $t=0$ 至 $t=t_1$ 之面積而 t 為較大數值則可以下列計：

$$\frac{1}{\sqrt{2\pi}}\int_0^{t} e^{-\frac{t^2}{2}}dt = \frac{1}{2}\cdot\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{t^2}{2}}dt \text{ 而}$$

$$\int_{t_1}^{\infty} e^{-\frac{t^2}{2}}dt\cdot\frac{t}{t} = \frac{-e^{-\frac{t^2}{2}}}{t}\Bigg]_{t_1}^{\infty} - \int_{t_1}^{\infty}\frac{1}{t^2}e^{-\frac{t^2}{2}}dt$$

故
$$\int_{t_1}^{\infty}\frac{t}{t}e^{-\frac{t^2}{2}}dt = \frac{-e^{-\frac{t^2}{2}}}{t}\Bigg|_{t_1}^{\infty} - \int_{t_1}^{\infty}\frac{t}{t^3}e^{-\frac{t^2}{2}}dt$$

$$= \frac{e^{-\frac{t_1^2}{2}}}{t_1} + \frac{e^{-\frac{t^2}{2}}}{t^3}\Bigg|_{t_1}^{\infty} + 3\int_{t_1}^{\infty}\frac{e^{-\frac{t^2}{2}}}{t^4}dt = \frac{e^{-\frac{t_1^2}{2}}}{t_1} - \frac{e^{-\frac{t_1^2}{2}}}{t_1^3} + \frac{3e^{-\frac{t_1^2}{2}}}{t_1^5} - 3\cdot 5\int_{t_1}^{\infty}\frac{e^{-\frac{t^2}{2}}}{t^6}dt$$

因 $t=\infty$，$\frac{e^{-\frac{t^2}{2}}}{t} = \frac{1}{t e^{\frac{t^2}{2}}} = \frac{1}{\infty} = 0$

依此方法繼續進行吾人得：

$$\int_{t_1}^{\infty}\frac{t}{t}e^{-\frac{t^2}{2}}dt = \frac{e^{-\frac{t_1^2}{2}}}{t_1}\left[1 - \frac{1}{t_1^2} + \frac{1}{t_1^4} - \frac{3\cdot 5}{t_1^6} + \frac{3\cdot 5\cdot 7}{t_1^8} + \cdots \frac{(-1)^k 3\cdot 5\cdots(2k-1)}{t_1^{2k}}\cdots\right]$$

由方程式

$$\frac{1}{y}\frac{dy}{dx} = \frac{x-a}{c_0+c_1x+c_2x^2}$$

$$\int_u^v (c_0+c_1x+c_2x^2)\,dy = \int_u^v (x-a)\,y\,dx \qquad (11)$$

上式兩端乘以 x^n 得：

$$\int_u^v (c_0+c_1x+c_2x^2)x^n\,dy = \int_u^v (x-a)x^n y\,dx$$

用部分積分法使

$$(c_0+c_1x+c_2x^2)x^n = u \qquad du = dy$$

$$du = nc_0x^{n-1}+(n+1)c_1x^n+(n+2)c_2x^{n+1}\,dx \qquad v = y$$

(11)之左端可變為：

$$\left|(c_0+c_1x+c_2x^2)x^n y\right|_u^v - \int_u^v \{y(nc_0x^{n-1}+(n+1)c_1x^n+(n+2)c_2x^{n+1})\}\,dx$$

即 $-\int_u^v [nc_0x^{n-1}+(n+1)c_1x^n+(n+2)c_2x^{n+1}]\,y\,dx = \int_u^v (x-a)yx^n\,dx$

$$-nc_0\mu_{n-1}-(n+1)c_1\mu_n-(n+2)c_2\mu_{n+1} = \mu_{n+1}-a\mu_n$$

$$nc_0\mu_{n-1}+[(n+1)c_1-a]\mu_n+[(n+2)c_2+1]\mu_{n+1} = 0$$

令 $n = 0, 1, 2, 3,$ 代入上式得

$c_1 = a$

(1) 常態曲線之全面積(亦即和率次數之和)等於一

證

$$A=\int_{-\infty}^{\infty}\varphi(x)\,dx=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-\frac{x^2}{2}}\,dx=\frac{2}{\sqrt{2\pi}}\int_{0}^{\infty}e^{-\frac{x^2}{2}}\,dx$$

令 $\sqrt{\theta}=\frac{x}{\sqrt{2}}$　$dx=\frac{d\theta}{\sqrt{2\theta}}$　代入上式則

$$A=\frac{1}{\sqrt{\pi}}\int_{0}^{\infty}e^{-\theta}\theta^{-\frac{1}{2}}\,d\theta=\frac{1}{\sqrt{\pi}}\Gamma\left(\tfrac{1}{2}\right)=\frac{1}{\sqrt{\pi}}\sqrt{\pi}=1$$

(2) 在常態次數分配中，以算術平均數($\bar{x}$)與向左右取等距離之依其間面積恰好為全面積之半，則此兩點為 $\pm P.E$ 稱為机誤矣

由。$A=\int_{-P.E}^{+P.E}\varphi(x)\,dx=2\int_{0}^{P.E}\varphi(x)=\frac{1}{2}$

因根據度當之常態曲線面積查得 $P.E=0.6745$

第四節　常態曲線表格之作法及用法

常態曲應用之表格有兩種。

(a) 常態曲線下面積表 Areas of the normal curve

(b) 常態曲線下縱坐標表 ordinate of the normal curve

1. 今分別斜述兩種表格之作法如下：

$$= \frac{-\mu_2}{D}\left(4\mu_2\mu_4 - 3\mu_3^2\right) = \frac{-\mu_2\left(4\mu_2\mu_4 - 3\mu_3^2\right)}{10\mu_2\mu_4 - 12\mu_3^2 - 18\mu_2^3}$$

$$= \frac{-\mu_2\left(\frac{4\mu_4}{\mu_2^2} - 3\frac{\mu_3^2}{\mu_2^3}\right)}{10\frac{\mu_4}{\mu_2^2} - 12\frac{\mu_3^2}{\mu_2^3} - 18} = \frac{-\mu_2(4\beta_2 - 3\beta_1)}{10\beta_2 - 12\beta_1 - 18}$$

$$= \frac{-\mu_2(4\beta_2 - 3\beta_1)}{K}$$

$$C_1 = \frac{1}{D}\begin{vmatrix} 1 & -\mu_2 & 2\mu_2 \\ 0 & -\mu_3 & 4\mu_3 \\ 3\mu_2 & -\mu_4 & 5\mu_4 \end{vmatrix} = \frac{1}{D}\left(-5\mu_3\mu_4 - 12\mu_2^2\mu_3 + 9\mu_2^2\mu_3 + 4\mu_3\mu_4\right)$$

$$= \frac{1}{D}\left(-5\mu_3\mu_4 - 3\mu_2^2\mu_3 + 4\mu_3\mu_4\right) = \frac{1}{D}\left(-\mu_3\mu_4 - 3\mu_2\mu_3\right)$$

$$= \frac{-\mu_3}{D}\left(\mu_4 + 3\mu_2^2\right) = \frac{-\mu_3\left(\mu_4 + 3\mu_2^2\right)}{10\mu_2\mu_4 - 12\mu_3^2 - 18\mu_2^3} = \frac{-\mu_3\left(\frac{\mu_4}{\mu_2^2} + 3\right)}{\frac{10\mu_4}{\mu_2^2} - 12\frac{\mu_3^2}{\mu_2^2} - 18}$$

$$= \frac{-\mu_3(\beta_2 + 3)}{\mu_2 K} = \frac{(\beta_2 + 3)\frac{\mu_3}{\mu_2}}{K}$$

$$=\frac{1}{\sqrt{2\pi}}\left\{\int_{-\infty}^{\infty} t^{2n+1} e^{-\frac{t^2}{2}}\,dt+\int_{-\infty}^{0} t^{2n+1} e^{-\frac{t^2}{2}}\,dt\right]$$

令 $t=-s$ 代入上式之末項得

$$\mu_{2n+1}=\frac{1}{\sqrt{2\pi}}\left\{\int_{0}^{\infty} t^{2n+1} e^{-\frac{t^2}{2}}\,dt+\int_{\infty}^{0} s^{2n+1} e^{-\frac{s^2}{2}}\,ds\right]$$

$$=\frac{1}{\sqrt{2\pi}}\left\{\int_{0}^{\infty} t^{2n+1} e^{-\frac{t^2}{2}}\,dt-\int_{0}^{\infty} s^{2n+1} e^{-\frac{s^2}{2}}\,ds\right]$$

再令末項(上式)中之 $s=t$ 得

$$\mu_{2n+1}=\frac{1}{\sqrt{2\pi}}\left\{\int_{0}^{\infty} t^{2n+1} e^{-\frac{t^2}{2}}\,dt-\int_{0}^{\infty} t^{2n+1} e^{-\frac{t^2}{2}}\,dt\right\}=0$$

(C) 在常態次數分配中各偶次動差, 可以下式表之:

$$\mu_{2n}=\frac{(2n)!}{2^n n!}$$

証: $$\mu_{2n}=\int_{-\infty}^{\infty} t^{2n}\phi(t)\,dt=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} t^{2n} e^{-\frac{t^2}{2}}\,dt$$

$$=\frac{1}{\sqrt{2\pi}}\, t^{2n-1} e^{-\frac{t^2}{2}}\Big]_{-\infty}^{\infty}+\frac{(2n-1)}{\sqrt{2\pi}}\int_{-\infty}^{\infty} t^{2n-2} e^{-\frac{t^2}{2}}\,dt$$

類數 主要類	方程式	原点	準
I	$y=y_0(1+\frac{x}{a_1})^{m_1}(1-\frac{x}{a_2})^{m_2}$	在衆数	K_2
IV	$y=y_0(1+\frac{x^2}{a^2})^{-m}e^{-\nu\tan^{-1}x/a}$ 或	平均数+($\nu a/r$ × 組距) 即在平均数後 $\nu a/(2n-2)$	K_2
VI	$y=y_0(x-a)^{q_2}x^{-q_1}$	平均数$-\mu_1$ (即在平均数後)	K_2
變化類 常態曲綫	$y=\frac{N}{\sigma\sqrt{2\pi}}e^{-x^2/2\sigma^2}$	在众数(=平均数)	K_2 β_1 β_2
II	$y=y_0(1-\frac{x^2}{a^2})^m$	在众数(=平均数)	$K_2=$ $\beta_1=$ $\beta_2=$
III	$y=y_0(1+\frac{x}{a})^{\gamma a}e^{-\gamma x}$	在众数	$2\beta_2=$ 或
V	$y=y_0x^{-p}e^{-\nu/x}$	在曲綫始端	$K_2=$
VII	$y=y_0(1+\frac{x^2}{a^2})^{-m}$	在众数=(平均数)	K_2 β_1 β_2

則	以平均數為原点之方程式	特点
負	$y=y_0(1+\frac{x}{A_1})^{m_1}(1-\frac{x}{A_2})^{m_2}$ 其中$(m_1+1):A_1=(m_2+1):A_2$	屬於此型之次数分配,係不对称,平常為鈴形,但可為U,J,反J手形,并且数值全距有限度
$\geqq 0$ 及 $\leqq 1$	$y=y_0[1+(\frac{x}{a}-\frac{\nu}{r})^2]^{-m}e^{-\nu\tan^{-1}(x/a-\nu/r)}$ 其中$r=2m-2$	偏態鈴形且全距無限
>1	$y=y_0(1+\frac{x}{A_1})^{-q_1}(1+\frac{x}{A_2})^{q_2}$ 其中$(q_1-1):A_1=(q_2+1):A_2$	次数分配一端之数值無限,偏測鈴形,亦可為J形比屬之
$=0$ $=0$ $=3$	$y=y_0e^{-\frac{x^2}{2\sigma^2}}$	对称,鈴形,两端無限.
$=0$ $=0$ $=3$	$y=y_0(1-\frac{x^2}{a^2})^m$	两边数值無限对称而常為鈴形,但有時為U形,当$\beta_2<1.8$時
$6+3\beta_1$	$y=y_0(1+\frac{x}{s})^{s_1}e^{s_2x}$ $y=y_0(1+\frac{x}{A})^{p}e^{-rx}$	一端数字無限,通常為鈴形,但亦可為J形
$=1$	$y=y_0x^{t_0}e^{-\frac{t_1}{x}}$	一端為数值無限并且為鈴形
$=0$ $=0$ >3	$y=y_0(1+\frac{x^2}{a^2})^{-m}$	無限对称鈴形

P. 42.

高級統計學

類數	方程式	原點
VIII	$y=y_0\left(1+\frac{x}{a}\right)^{-m}$	在曲線終點
IX	$y=y_0\left(1+\frac{x}{a}\right)^{m}$	在曲線終點
X	$y=y_0e^{-x/\sigma}$	在曲線始端
XI	$y=y_0x^{-m}$	在曲端始端前b
XII	$y=y_0\left[\frac{\sigma\{\sqrt{3+\beta_1}+\sqrt{\beta_1}\}+x}{\sigma\{\sqrt{3+\beta_1}-\sqrt{\beta_1}\}-x}\right]^{\sqrt{\beta_1/3+\beta_1}}$	在平均數

$$\lambda=\frac{\{(4\beta_2-3\beta_1)(10\beta_2-12\beta_1-18)-\beta_1(\beta_2+3)^2(8\beta\ldots}{\{\beta_1(\beta_2+3)^2+4(\beta_2-3\beta_1)(3\beta_1-2\beta_2+6)\}}$$

$\beta_1=\mu_3^2/\mu_2^3 \qquad \beta_2=\mu_4/\mu_2^2$

$$K_1=2\beta_2-3\beta_1-6 \qquad K=\frac{\beta_1(\beta_2+3)}{4(4\beta_2-3\beta_1)(2\beta_2-3\beta_1\ldots}$$

別	以平均數爲原点之方程式	特　点
$=0$ … $-6\beta_1-9>0$ 負	$y=y_0(1-x/A)^{-m}$	全距由∞縱坐標至有定縱坐標或由$-a(1-m)/(2-m)$至$a/(2-m)$以平均數爲原点.
0 負 0 $-6\beta_1-9>0$ 正 $-3\beta_1-6<0$ 負	$y=y_0(1+x/A)^{m}$	全距由$x=-a$至$x=0$即$y=0$至$y=y_0$或$-a(m+1)/(m+2)$至$a/(m+2)$以平均數爲原点.由有定縱坐至無限縱坐標J形.
$=4$ $=9$	$y=y_0e^{-x/\sigma}$	
$>0, \lambda=0$ $-3\beta_1-6>0$ 正	$y=y_0(1+\frac{x}{A})^{-m}$	J形始端$x=b$或$(s+b/m-2)$以平均數爲原点.
$-6\beta_1-9=0$	$y=y_0(1+x/A_1)^{m_1}(1-\frac{x}{A_2})^{m_2}$	扭J字形,型I之特別情形.
$-9\beta_1-12)(3\beta_1-2\beta_2+6)$		
$-6)$		

第六章　皮尔生曲綫系第三秋

第一節　導　言

由前述知皮尔生曲綫系有十二秋曲線型而細究研之僅有四个基本的曲綫式及一个特例式為同類个体抽樣問題之解答即：

(1) $K=\infty$　第三類

(2) $K=1$　第五類

(3) $K>1$　第六類

(4) $0\leqq K<1$　第四類

(5) 特例　第七類　（常態 $K=0$）

是。其餘七類皆不認為解答，而此兩秋曲綫其性質有基本的區别，盖前五者與常態曲線一樣，有先驗的机率的根拠，故不妨称為机率曲綫，後比七類與梅開姆(Makeham)公式等同類并非由机率的根据而得，故称修匀曲綫。事实上後者七類修匀曲線之用处甚少，一則因為兩端有明確限制之分佈甚少，再則因為J字形分佈本来甚少，且社經过考慮之後，并非J字形，然則并非謂用修匀曲綫配合次数分佈為非法者，不过不能用机率曲綫配合之。其次数分佈形式之源則決非同類個体抽樣問題所能解答者，則可断言也。

由此以观皮氏曲綫系分為主型與变化型兩類，而主型I并無例1

率的根據,故常受人批評,但主型Ⅲ屬於變化型而為偏態不甚理論次數分配最重要之一種,因此公式如其偏斜度($K=0$)為零,則為常態分配,故知偏態為一般情形,而常態為特例也,由是在研究次數分配之偏斜度時,由於偏態之利害,適中,或不十分偏,仍有各種之理論次數分配,而在偏態不甚理論次數分配皮氏曲線之型Ⅲ為最重要之一種,在應用上功用亦極大,蓋吾人於曲線配合中,欲測定該曲線配合之良窳,往往計算 χ^2 之值以測定之,而 χ^2 本身之分配,適合皮爾生氏之曲線型Ⅲ故也.

第二節 公式之導來

Type Ⅲ 適用於連續性的不甚偏態之公式之導來,仍不離乎二項分配理論,蓋設 $p=q$ 則对于不連續之二項次數分配為对称的,而在 Type Ⅲ $p \neq q$ 則对於不連續性之二項次數分配為偏斜的.

展開式 $(p+q)^s = p^s + sC_1 p^{s-1} q + \cdots\ sC_r p^r q^{s-r} + \cdots\cdots q^r$

取其普通項 $sC_r p^r q^{s-r} = \dfrac{s!}{s!(s-r)!} p^r q^{s-r}$

則其前一項為:

$$y_{r+1} = \frac{s!}{(r+1)!(s-r-1)!} p^{r+1} q^{s-r-1}$$

$$= \frac{s!\,(s-r)}{r!\,(r+1)\,[(s-r)(s-r-1)]!}\;\frac{p^r p\cdot q^{s-r}}{q}$$

$$= \left[\frac{s!}{r!\,(s-r)!}\,p^r q^{s-r}\right]\frac{s-r}{(r+1)}\,\frac{p}{q}$$

$$\therefore\ y_{r+1} = \frac{s-r}{r+1}\,\frac{p}{q}\,y_r$$

$$\Delta y_r = y_{r+1} - y_r = \frac{(s-r)p}{(r+1)q}\,y_r - y_r$$

$$= y_r\left[\frac{s-r}{r+1}\,\frac{p}{q} - 1\right] = y_r\left[\frac{sp-rp-rq-q}{(r+1)q}\right]$$

故中央項 $y_{r+\frac{1}{2}} = \dfrac{y_{r+1}+y_r}{2} = \dfrac{y_r}{2}\left[\dfrac{sp-rp+rq+q}{(r+1)q}\right]$

如以 $x = r+\frac{1}{2}$ $\Delta x = 1$ 則 $r = x - \frac{1}{2}$

$$\frac{\Delta y_r}{y_{r+\frac{1}{2}}} = \frac{1}{y}\,\frac{\Delta y_r}{\Delta x} = y_r\left[\frac{(s-r)p-(r+1)q}{(r+1)q} \times \frac{2(r+1)q}{y_r(sp-rp+rq+q)}\right]$$

$$= 2\left[\frac{sp-rp-rq-q}{sp-rp+rq+q}\right] = 2\left[\frac{sp-q-r(q+p)}{sp+q-(q-p)r}\right]$$

$$= 2\left[\frac{sp-q-r}{sp+q-(q-p)r}\right]$$

则 $\frac{1}{y}\frac{dy}{dx} = \frac{2sp-q-(x-\frac{1}{2})}{sp+q-(q-p)(x-\frac{1}{2})}$

令 $2sp-q+\frac{1}{2} = a$

$sp+(q-p)\frac{1}{2} = c_0$

$-(q-p) = c_1$

$$\therefore \frac{1}{y}\frac{dy}{dx} = \frac{a-x}{c_0+c_1x} = \frac{1}{c_1}\left[\frac{x-a}{x+\frac{c_0}{c_1}}\right] = \frac{1}{c_1}\left[\frac{x+\frac{c_0}{c_1}}{x+\frac{c_0}{c_1}}\right] - \frac{a+\frac{c_0}{c_1}}{c_1(x+\frac{c_0}{c_1})}$$

$$= \frac{1}{c_1} - \frac{a+\frac{c_0}{c_1}}{c_1}\frac{1}{(x+\frac{c_0}{c_1})}$$

積分之：

$$\int\frac{dy}{y} = \frac{1}{c_1}\int dx - \frac{a+\frac{c_0}{c_1}}{c_1}\int\frac{dx}{a+\frac{c_0}{c_1}}$$

$$\log y = \frac{x}{c_1} - \left(\frac{a+\frac{c_0}{c_1}}{c_1}\right)\log\left(x+\frac{c_0}{c_1}\right) + c$$

以 $x'+a$ 代 x

則 $\log y = \frac{1}{c_1}(x'+a) - \frac{a+\frac{c_0}{c_1}}{c_1}\log(x'+a+\frac{c_0}{c_1}) + c$

令 $s_0 = \frac{c_0}{c_1} + a$

$s_1 = \frac{s_0}{c_1}$

$s_2 = \frac{s_1}{s_0}$

即 $y = y_0(1+\frac{x}{s_0})^{s_1} e^{-s_2 x}$

此即著名之偏態的型 III 公式也

定 y_0 之值

$$N = \int_{-s_0}^{\infty} y\,dx = \int_{-s_0}^{\infty} y_0(1+\frac{x}{s_0})^{s_1} e^{-s_2 x}\,dx$$

$$= y_0\int_{-s_0}^{\infty}(1+\frac{x}{s_0})^{s_1} e^{-s_2 x}\,dx$$

$$y_0 = N \div \int_{-s_0}^{\infty}(1+\frac{x}{s_0})^{s_1} e^{-s_2 x}\,dx$$

令 $z = s_2(s_0+x)$，則 $x = \frac{z - s_0 s_2}{s_2}$， $dx = \frac{1}{s_2}dz$ 得

$$y_0 = N \div \int_0^{\infty} \left(1 + \frac{z - S_0 S_2}{S_0 S_2}\right)^{S_1} e^{-S_2 \frac{z - S_0 S_2}{S_2}} \frac{dz}{S_2}$$

$$= N \div \frac{e^{S_0 S_2}}{S_2^{S_1+1} S_0^{S_1}} \int_0^{\infty} z^{S_1} e^{-z} dz$$

$$= N \div \left[\frac{e^{S_0 S_2}}{S_2^{S_1+1} S_0^{S_1}} \Gamma(S_1+1)\right] = N S_2^{S_1+1} S_0^{S_1} \Big/ e^{S_0 S_2} \Gamma(S_1+1)$$

$$= \frac{N S_1^{S_1} S_2}{e^{S_1} \Gamma(S_1+1)} = \frac{N S_1^{S_1+1}}{S_0 e^{S_1} \Gamma(S_1+1)}$$

$\because S_0 S_2 = S_1$ 即 $y_0 = \dfrac{N S_1^{S_1+1}}{S_0 e^{S_1} \Gamma(S_1+1)}$

而通常應用之公式可寫為

$$y = y_0 (A + t)^{A^2-1} e^{-At}$$

證：

假定 $a = -\frac{1}{A}$ $c_1 = \frac{t}{A}$ 而 $c_0 = 1$ 吾人得

$$\frac{1}{y}\frac{dy}{dx} = \frac{-\frac{1}{A} + t}{1 + \frac{t}{A}}$$

積分之 $\int \frac{dy}{dx} = \int - \frac{\frac{1}{A}+t}{1+\frac{t}{A}}\,dt = -\int \frac{1+At}{A+t}\,dt$

$$= \int -A + \frac{A^2-1}{A+t}\,dt$$

$$\int \frac{dy}{dx} = -A\int dt + A^2-1\int \frac{dt}{A+t}$$

令 $u = A+t \quad du = dt$

取对数 $\log y = -At + (A^2-1)\log(A+t) - c$

即 $(A+t)^{A^2-1} e^{-At}$

或 $y = \frac{A^{A^2} e^{-A^2}}{\Gamma(A)}(A+t)^{A^2-1} e^{-At}$

又設 $y_0 = \frac{A^{A^2} e^{-A^2}}{\Gamma(A)}$

在决定 y_0 之数值前,先須研究曲線之性質

令 $t = -A \qquad y = 0$

$t = +\infty$

$$y = \frac{y_0(A+\infty)^{A^2-1}}{e^{\infty}} = \frac{\infty}{\infty} = \frac{0}{\infty} = 0$$

而 $t = \dfrac{1}{-A}$

故此曲綫係為 α_3 所決定（$\because A = \dfrac{2}{\alpha_3}$）如 α_3 為負

A 亦為負 曲綫形像相反

使曲綫下之面積為一單位

$$\int_{-A}^{\infty} y_0 (A+t)^{A^2-1} e^{-At} dt = 1$$

$$= y\int_{A}^{\infty} (A+t)^{A^2-1} e^{-At} dt = 1$$

$$y_0 = \frac{1}{\int_{-A}^{\infty} (A+t)^{A^2-1} e^{-At} dt}$$

將 $\int_{-A}^{\infty} (A+t)^{A^2-1} e^{At} dt$ 用代替方法使 $A(A+t) = z$　$A\,dt = dz$

則 $\int \left(\dfrac{z}{A}\right)^{A^2-1} e^{-(z-A^2)} \dfrac{dz}{A}$ ……

因 $t = -A$

$$\int_0^{\infty}\left(\frac{z}{A}\right)^{A^2-1} e^{-(z-A^2)}\frac{dz}{A} = \frac{e^{A^2}}{A^{A^2-1}A}\int_0^{\infty} z^{A^2-1}e^{-z}\,dz = \frac{e^{A^2}}{A^{A^2}}\Gamma(A^2)$$

$\because z = A(A+(-A)) = 0$

以合於 $\Gamma(n)$ 形式，惟改 $\Gamma(A^2)$ 而已，故等。

即 $y_0 = \dfrac{1}{\dfrac{e^{A^2}\Gamma(A^2)}{A^{A^2}}}$

$$y_0 = \frac{A^{A^2}}{e^{A^2}\Gamma(A^2)} = \frac{A^{A^2}e^{-A^2}}{\Gamma(A^2)}$$

故型 III 全式

$$y = \frac{A^{A^2}e^{-A^2}}{\Gamma(A^2)}(A+t)^{A^2-1}e^{-At}$$

第三節 Type III 之性質

1. 屬於此型之次數分配曲線其一端無限故其起点的
2. 通常為鐘形，但可為J字形
3. 型 III 公式中設 $A^2-1>0$ 若僅使 $A^2-1>0$ 一條件不然 [illegible]

為正或負偏態，故不能不再定其第二條件

再設 $\alpha_3 > 0$ 從而 $A > 0$ 故當 $t = -\frac{1}{A} = -\frac{\alpha_3}{2}$ 為最高點

(4) 設 $A^2-1 > 0$, $\alpha_3 < 0$ 即 $A < 0$ 則為負偏態

$$y = \frac{y_0}{(A+t)^{A^2-1}} e^{-At}$$

(5) 設 $A^2-1 < 0$, α_3 為正或負曲綫形式已變無大关係

$$y = y_0 \frac{e^{-At}}{(A+t)^{1-A^2}}$$

則為最偏斜的了，称 J-Chaped Curucs

(6) 設 $A = 1$ 則 $\alpha_3 = 0$ 則公式

$$y = \frac{e^{-1}}{\Gamma(1)}(1+t)^0 e^{-t} = e^{-t} = \frac{e^{-t}}{e} = e^{-(t+1)}$$

$t = -1 \quad y = 1$

$t = +\infty \quad y = 0$ 称為前偏斜

此又為 Type X 之形式

(7) $A^2-1 < 0 \quad \alpha_3 < 0 \quad y = y_0(A+t)^{A^2-1} e^{-At}$ 稍偏

(8) $A = \infty \quad \alpha_3 = 0$ 其公式為 $y = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$

8. Type Ⅲ 分配為 χ^2 測驗及小樣本原理之根據前証型 Ⅲ 公式導成 χ^2 條可由 χ^2 証明之

$$y = y_0\left(1+\frac{t}{S_0}\right)^{S_1} e^{-S_0 t}$$

$$y_0 = \frac{N S_1^{S_1+1}}{S_0 e^{S_1} \Gamma(S_1+1)}$$

$S_2 = \frac{2}{\mu_{3t}}$ （設 $m = S_2$） $S_1 = m^2-1$ 且 $S_0 = m - \frac{1}{m}$

因此可將 (A) 式寫成

$$y = y_0 S_0^{-S_1}\left(m-\frac{1}{m}+t\right)^{m^2-1} e^{-mt}$$

使 $m^2 = \frac{n}{2}$ 及 $m\left(m-\frac{1}{m}+t\right) = \frac{\chi^2}{2}$ 則上式変成

$$y = y_0 S_0^{-S_1}\left(\frac{\chi^2}{2m}\right)^{\frac{n}{2}-1} e^{-\frac{\chi^2}{2}+m^2-1}$$
$$= K\left(\frac{\chi^2}{2}\right)^{\frac{n}{2}-1} e^{-\frac{\chi^2}{2}}$$

式中 $K = y_0 S_0^{-S_1} m^{\frac{n}{2}-1} e^{m^2-1} = \frac{N}{\Gamma(\frac{n}{2})}$

上式即為χ^2之机率分配

有了上式吾人可用下式以示P之值

$$P=\frac{\int_{\chi^2}^{\infty}K\left(\frac{\chi^2}{2}\right)^{\frac{n}{2}-1}e^{-\frac{\chi^2}{2}}d\chi^2}{\int_0^{\infty}K\left(\frac{\chi^2}{2}\right)^{\frac{n}{2}-1}e^{-\frac{\chi^2}{2}}d\chi^2}$$

$$=\frac{\int_x^{\infty}x^{n-1}e^{-x}dx}{\int_0^{\infty}x^{n-1}e^{-x}dx}$$

第四節 Type Ⅲ之应用

Type Ⅲ应用之便利更加一層,有表可查,又表格分為縱坐標和與面積表兩种,作此表者為Solvasa在1931—1935间成於美國某大學,作四年始完成,該表只涉及$\bar{x}$, σ, α_3三种常数,表中偏態数自之二0⟶1.1,如$\alpha_3=0$, $\alpha_3=1$……$\alpha_3=1.1$即為最大数,致其他所以不能用再大者,因此公式只求其偏態,不甚分配故也。

至於应用今設例以明之。

例一:皮尔生偏態曲線解决二項分配所不能解决之问題

如求60歲在一年内死亡率為.03,生存率為.97,人数為100,

若以二項分配法，計展开極為麻煩，又因此為偏態，而不能用常態去做，故祇能用 Type Ⅲ 作之可得

$\bar{X} = m = SP = 100 \times .03 = 3$

$\sigma = 1.7$

$\alpha_3 = \frac{q-p}{\sqrt{N \cdot qp}} = .55$

如求在此一百人中死亡四人到十人之机率，若用展開式則非常麻煩，若用 type Ⅲ 之面積法，則甚簡即求 $t_1 \to t_2$ 之面積

$t_1 = \frac{4-3}{1.7} = \frac{1}{1.7} = .6 \qquad t_2 = \frac{10-3}{1.7} = 4$

$.6 \to 4$ 之面積即為所求之机率

	$\alpha_3 = .5$	$\alpha_3 = .2$	$\alpha_3 = .55$
$t_1 = .6$	.74472	.74073	.74553
$t_2 = 4$	.99909	.99934	.99921

如每一百個六十歲在一年内可死之机率為：

高級統計學 P. 52

$.77921 - .74553 = .25308 = 25\%$

例二：設有 20,000,000 人每人每年平均收入

$\bar{x} = 1193$ 元　$\sigma = 190.5$　$\alpha_3 = 1.05$　此由抽樣得來。如抽所得稅，最少生活費用定每年 2000 元，問收入在 2000 元以上人數為多少

$$t = \frac{2000 - 1193}{190.5} = 4.23$$

查 $t = 4.2$　則 $t = \infty$ 之面積表即為所求之机率

表中 $t = 4.2$　$\alpha_3 = 1.0$ 之面積為 .99832 ⎫ 相加除2/3

$t = 4.2$　$\alpha_3 = 1.1$　　　　　　　　.99785 ⎭

$$2 \overline{)1.99617}$$

$$.99808$$

$t = -4.2$ 時，$\alpha_3 = 1.05$ 之面積

即 $1 - .99815 = .00185$

$.00185 \times 20{,}000{,}000 = 37{,}000$ 人

抽所得稅每人抽 200 元可抽得 $37000 \times 200 = 7{,}400{,}000$ 元

例三：設一組 1000 工人之年齡統計常数如次

$\bar{x} = 30$ 歲　$\sigma = 5$ 歲　$\alpha_3 = .8$

試估計(a)小於23歲之工人数目
(b)大於50歲之工人数目

a)化成標準单位

$$t = \frac{x-\bar{x}}{\sigma} = \frac{23-30}{5} = -1.4$$

由 $t = 1.4$ 及 $\alpha_3 = .8$ 查表得 0.04797

故知一千工人中共小於23岁者為 $.04797 \times 1000 = 48$ 人

b)代入標準单位

$$t = \frac{x-\bar{x}}{\sigma} = \frac{50-30}{5} = 4$$

由 $t = 4$ 及 $\alpha_3 = .8$ 查表得 .99846

故知1000工人中大於50歲者有 $(1-.99846) \times 1000 = 1.54$ 人

例四：全市工商業共有萬　　為徵收过多所得稅起见作抽樣調查得一萬家每年贏利之統計常数如次

$m = 60,000$ 元　$\sigma = 13,000$ 元　$\alpha_3 = .9$

如欲求全市工商業中贏利在100,000元以上之家数若干

化成標準单位

$$t = \frac{x-m}{\sigma} = \frac{100,000-60,000}{13,000} = 3.1$$

查表得 .99198

故知全市工商業中贏利在 100,000 元以上之家數為

$(1-0.99198)\times 10,000 = 80$ 家

曲綫配合之例

設有一次數分配資料為下表并求動差

i	f	x	xf	x^2f	x^3f	x^4f
0	87	-3	-261	783	-2347	7047
2	152	-2	-384	768	-1536	3072
4	128	-1	-128	128	-128	128
6	71	0	0	0	0	0
8	12	1	12	12	12	12
10	7	2	14	28	56	112
12	3	3	9	27	81	243
	500		-738	1,746	-3864	10,614

于是求得 $\bar{x} = 4.048$, $\mu_3 = 1.30342$ $\sigma = 1.15$ [illegible]

則:

(一) 面積表法

x下限	$x-\bar{x}=x$	x/σ	$K\int_A^t (A+t)^{A^2-1} e^{A} dt$	在組距中面積	y
0	-4.048	-1.09	0048620	.181600	90.8
2	-2.048	-.89	1864600	.3585296	182.5025
4	-.048	-.0208	5449896	2729550	136.4750
6	1.952	.848	8179440	1256144	62.8014
8	2.952	1.718	9437588	0366178	18.3391
10	5.952	2.587	9804175	0148981	7.4491
12	7.952	3.456	9953156	0036558	1.8219
14	9.952	4.327	9989714		
					500.1872

(二)縱坐標法

中點	$x'=x-\bar{x}$	x'/σ	y_0	$y_n-\frac{N}{\sigma}y_0$
1	3.048	-1.325	20946	90.069
3	1.048	.455	43671	189.872
5	.952	.413	30659	133.298
7	2.952	1.283	13650	59.347
9	4.952	2.153	04751	20.355
11	6.952	3.022	01444	6.278

13	8.952	3.892	00343	$\frac{1.419}{500716}$

$$\frac{N}{\sigma} = 434.78 = \frac{500}{1.15} \qquad \alpha_3 = .9$$

(三)用公式 $y = y_0 (1 + \frac{x}{S_0})^{S_1} e^{-S_2 x}$ 直接計算法

計算時所用常数如下

$$S_2 = \frac{2u_2}{u_3}$$

$$S_1 = (\frac{4}{\beta_1}) - 1$$

$S_0 = \frac{S_1}{S_2}$ 由众数至两限之距離

$$y_0 = \frac{10}{S_0} \left[\frac{S_1^{S_1+1}}{e^{S_1} p(S_1+1)} \right]$$

大始头係在众数，众数離平均数之距離為 $\frac{1}{S_2}$

組距	m	x	S_0+x	[illegible]	[illegible]	$x-a$	[illegible]	第五項之对数相加 $10/35$	計算数 y_c	観察数 y
2-	1	-1.0238	1.1233	.0504858	2189245	-2.0715	.8896411	[illegible]	29.81	87
3-	[illegible]	-.0238	2.1233	.3270114	1.4177579	-.0560	.0213205	[illegible]	179.30	192
4-	[illegible]	[illegible]	3.1233	.4946137	2.0443977	1.3595	$\bar{1}$.1489999	2128555	134.45	128
6-	7	[illegible]	4.1233	.6152445	2.6173843	3.7750	$\bar{2}$.2731754	1.7762296	59.74	71
8-	[illegible]	[illegible]	5.1233	.7095648	3.0762532	5.9505	$\bar{3}$.3182388	[illegible]	20.41	[illegible]
10-	[illegible]	3.9722	6.1233	.7869855	3.4119756	8.0060	$\bar{4}$.5330382	[illegible]	5.51	7
12-	[illegible]	[illegible]	7.1233	.8526812	3.6967943	10.0415	$\bar{5}$.6473177	[illegible]	1.51	[illegible]

[illegible]其常数如下 $\mu_2 = 1.3134$ $\beta_1 = .7497$

$\mu_3 = 1.3033$ $\beta_2 = 4.0663$

$\mu_4 = 7.0092$

由上之常数求得 r P a 及对数 y_0 之值如下

$$S_{[illegible]} = \frac{2.6268}{1.3033} = 2.0155$$

$$r = \frac{4}{[illegible]} - 1 = 4.3355$$

$$S_0=\frac{4.3855}{2.0155}=2.1511$$

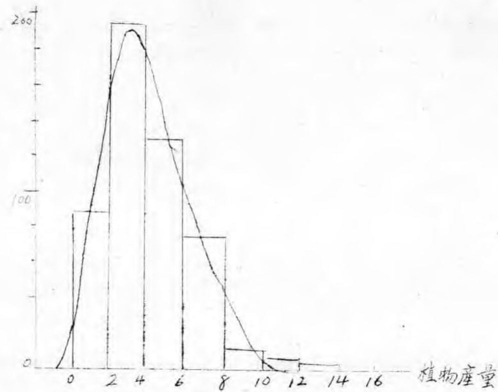
$$y_0=\frac{500}{2.1511}\frac{4.3355^{5.3355}}{2.7182818^{4.3355}\Gamma(4.3355+1)}$$

若以 y_0 自对数计算则得

对数 $y_0=2.2774059$ $\quad \bar{X}=4.0480$

$$z=4.0480-\left(\frac{1}{2.0155}\times c\right)=3.0556$$

註：附圖詳下頁



圖表諭拟矩形及其配合型 III 曲線圖

第七章 χ^2分配論

第一節 χ^2分配之来源与特性

一 来源 最初發明χ^2分配的是K. pearson 在1912年他決定χ^2(Chi Square)分配之形式，不過在他決定χ^2分配才有一点錯[illegible]，[illegible]自由度，後經过R.A. Fisher之修正自由度以後，大概相当完[illegible]。[illegible]已經発達，可利用表格，目表格有二种即：

(1) Elderton表格　(2) R.A. Fisher表格

Karl pearson之貢献：他在於一种這樣情形下決定χ^2[illegible]，此种[illegible]可決定理想与現象[illegible]之一种[illegible]，[illegible]是實際所得觀察分配与理想分配差异之研究。

在此差異中有不[illegible]之分量大部，[illegible]討論而非絶対決定，故應[illegible]驗次数以m_x表示理想次数。

$$\chi^2 = \sum \frac{(n_x - m_x)^2}{m_x}$$

吾人知道之[illegible]一种分配有兩种方法 (1)理論方法 (2)[illegible]方法 [illegible]經驗方法而是実際實驗而得，為一個[illegible] 義者，皆若干差異可解釋χ^2之分配，吾人知χ^2之分配是[illegible] Type III

之分配亦可屬於 Type III 之函数，故 χ^2 之函数公式為：

χ^2 之函数公式 $= Ke^{-\chi/2}\chi^{n-1}d\chi$

以 Fisher 之記号，形式為 $Ke^{-\chi^2/2}\chi^{n-1}d\chi$

據 Fisher 氏之應用 χ^2 來測驗 χ 分流全数 Population 是否服從常態，如此作至實上含有之偏斜太小時，仍可用 χ^2 測驗方法以驗之

III)特性：由實驗吾人已知 χ^2 分配是合於 Type III 之函数，並且知道全数是常態時或不十分偏斜時，由其實際次数與理論次数之差，是否此項之 χ^2 分配是合於 Type III 之形式，即當全数是常態或不太偏斜時，可用 χ^2 測驗

χ^2 分配合於 Type III 形式，Type III 之原点是 $\chi=0$ 而 χ^2 只有正的意義，故原点是 Type 為正偏態時之起点 $\chi^2=0$，$P=1.00$ 也即是百分之一百，即表示一百次中有一百次居於出自同一全数，亦即兩者絕對沒有顯著差異，也即是理想與實際完全符合

$\chi^2=$ 大 $P=0.00$ 因所有面積都包括在內，其值之面積甚少，也即百次中無一次同出自一個全数的，或說百次中有一百次是出自二個全数，也即絕對沒有顯著差，這二种形式，完全是理想的，不過我們可以知道 P 愈小，χ^2 愈大，愈不符合而有顯著差，χ^2 之偏斜情形以自由度決定之，自由度愈小，偏斜度愈厲害，自由度愈

大偏斜度發入，自由度超过了 n 時何候 χ^2 所呈現之偏斜[illegible]漸趨常態。

故此偏態曲線漸趨於常態曲線時，[illegible] χ^2 [illegible]之機率[illegible] $\sqrt{2\chi^2}-\sqrt{2n-1}$ 此点也為 Type Ⅲ之要点[illegible]

[illegible]如在不合情形中投擲 24 個骰子，以觀察其結果[illegible]結果[illegible]之次數（即 1、2、3、4、5、6 点）記錄之，投擲 24 個骰[illegible]所希望得 4 個 1 点，四個 2 点，四個 3 点，等，之結果如下

点數	1	2	3	4	5	6
觀察次數	2	5	6	4	4	3
希望次數	4	4	4	4	4	4

此种結果之 χ^2 值如下

$$\chi^2=\frac{(2-4)^2}{4}+\frac{(5-4)^2}{4}+\frac{(6-4)^2}{4}+\frac{(4-4)^2}{4}+\frac{(4-4)^2}{4}+\frac{(3-4)^2}{4}=[illegible]$$

此共[illegible]有六組成分子，但祇五組成分子[illegible]之值，因在事實上總数是固定的 24 個，故在試驗中計算 χ^2 [illegible]自由度[illegible]如 24 個骰子投擲一千次，則有千個 χ^2 值，[illegible]吾人[illegible]討論过 χ^2 是合于 Type Ⅲ之曲表的。

[illegible]用於各种問題，根據于合理之基礎[illegible]

高等[illegible]

[illegible]

數求觀察均數比較，此種測驗法，方極適合，蓋直而經濟，均數間之差異是否有意義，至於其之應用下節再述

第二節 自由度之研究

如有一組數值，吾人指定其有三個數字，若已知其總數是12，於是三個數字之數值只有兩個可以隨意列，可以注意其三之中，有自由選擇之餘地，因為當我們已知總數是12，而我們已決定二個數字是1及8，第三個數字之值，當然不能隨意決定，而必須是已經選擇之方能如：12－1－8＝3 第三個數字必是3，故當三個數字已知其總數時，只有兩個自由度

至過去一般對自由度之決定，大概除去總數，如有n個數值，除去總數，得自由度為 $n-1$ 後經Fisher之修正，除去已知之總數外，又當加一平均數，則自由度又減少一度，共計減少二個自由度，故Fisher以為知一統計常數，則減少一個自由度，如在一分數分配表中，有一組數值，如我們已知其總數平均數，標準差，偏斜度，峯度，則我們只有 10－5＝5 五個自由度

如： X f 則 $4\times1+2\times12+3\times9=16\times2=52$

1 4 $2x+3(16-4-x)=52-4=28$

2 $x=8$ $2x+48-12-3x=28$ $x=48-12=8$

3 $y=4$ $y=16-4-8=4$

$T=16$

$m=2$

由此可知一個平均數，即少一個自由度。如上例，已知總數及平均數，我們只能任意決定一個次數；其餘兩個就失自由選擇之可能，故上例只有一個自由度。我們可知每一個統計常數，即少一個自由度，上例亦可證明矣。

第三節 χ^2分配之應用

χ^2之分配為比較經驗次數與理論次數之差數之分配，χ^2之驗法可用於各種問題。根據合理之基礎求得若干之理論次數而比較其次數之大小，此法驗定極適合決定兩組數字間之差異，吾人常由此得手求得理論次數分配與之觀察次數。

依理論次數之差數為$\chi^2=0$，即無限之觀察次數，以不確實符合於所有之理想次數，χ^2誰不測定準確，也即觀察次數

合於理想分数，吾人研究之目的，也即当χ^2之值以什么程度时，就不能适据检决，χ^2表格為一种標準决定。以决定观察某种差异之和平，是否纯為机会所造成者，或此种差过大，可以χ^2之应用可分三种，兹申述之：

(1)適合性之测驗(Test of goodness fit)，一種理論分数，不論為常态或其他形式，配合於观察次数分配，须依理論与观察之比較，而以用观察之方法测驗二者间之配合適度，即適合性也，此种测驗不适当，如应用χ^2测驗法，可以正确无误，故χ^2之测驗也可说测驗用一种公式之估计数目最適合实際情形最正確。

何	0	1	2	3	4	5	6	7
理論数	12	84	250	400	430	252	84	12
经驗数	12	78	270	486	386	252	69	13

吾人已知總数同為1536，故自由度＝8−1＝7 查χ^2表格，由表中求例得 $\chi^2=9.803$时 $P=.20$ $\chi^2=12.017$时 $P=.10$ 求例由插補法得 $\chi^2=10.315$时 $P=.18$ 即表明百次中十八次是否於同出自同一 Population 如：

何度：	61.5以下	61.5	62.5	63.5	64.5	65.5	66.5	67.5	68.5	69.5	70.5	71.5	72.5	统计	
经驗数：	11.5	19.0	33.5	61.5	95.5	142.0	137.5	154	141.5	116.0	78	49	28.5	9.5	1078

[illegible]數 11.8 17.9 35.5 62.8 16.7 130.1 153.0 157.1 141.0 110.5 75.7 [illegible] [illegible] [illegible]

同 χ^2 測驗法討論經驗次數是否合於理論次數，求〔illegible〕在上例

[illegible] 平均數，標準差（所謂已知即是用經驗之數）[illegible]，

[illegible]

故一例自由度 $= 14 - 3 = 11$，用 χ^2 之公式求得：$\chi^2 = 8.57$ [illegible] [illegible] 即表示 1 百次中 66 次是同一 population 中出來，可知結果 [illegible] 合理論次數

不合理論次數，而要經驗次數之原因甚多，此其 [illegible] [illegible] 大量測驗，吾人檢查其精密程度，事實上不可能一一 [illegible]，只能用此種方法推求全體之實際情形，χ^2 測驗法 [illegible] [illegible] 一 [illegible] 配也

[illegible] χ^2 測驗不一定得決定關於指定 χ^2 值之 [illegible]，[illegible] [illegible]，[illegible] 在 [illegible] 觀察之結果，是否合於 [illegible] [illegible] [illegible]，[illegible] 在此種目的只須知相關於一指定 χ^2 之 P 值 [illegible] 於 [illegible]

Fisher 指出 χ^2 之低空值為極低，使 P 之值為 .99 或 [illegible] [illegible]，此種低空值指明假定與觀察極為吻合，[illegible] [illegible] [illegible] 有力之可能性，認為假定之錯誤，或者 [illegible] 方面 [illegible]

錯誤之弊病　理論與理想值之吻合表示一种情形，即如以
色格有行常具之方程式，配合一曲線於各點，以人為之力量得
迎合原點與理論值之吻合．若欲使χ²之測驗之應用不能治當
則用χ²之測驗法，不僅以用習慣之標準(99%或95%机率)以
指二及對於其結論之是否較為方便，不難推求，若引用极端之
標準應用絕對之意義上，則可能引用對於此种經驗之其他智
識以能解此种測驗所產生之結果；並代及測驗之假定之理想之
基礎．
此外理論次數與經驗次數比較時，有將組值合併(如例二中615
其附以下．所以要合併之原因，是因經驗次數很小，得福之數
，不僅防實在比較，如理論次數过小，經驗與理論間輕微之
差異在實際上在χ²之值上有合併之意義，即在避免此种危現
，因選擇所要合併之各組係根于研究者個人之判斷，此仁之
需常備有一种因素，舉之依照Kendall and Yule之意見，歸併之
標準以次數小於十者歸併使次數大於十組數，宜由十組之一
，普通以十五組左右為佳，至於尾端組個加以合併，要試之
測驗之結果，通常不太受影響．
引用χ²測驗法，以測驗理論次數與曲線配合度，更有一种更

大之限制。其計算χ^2并不代為觀察及理論次數差並之分配，但此種差異之分配狀況實質上可以影响吾人對於某種配合之是否適度之判斷。因此吾們引用χ^2測驗，並以差異之為任分配觀察配合情形之解釋χ^2不能純以技術之應用χ^2測驗法測驗之。至於相合性之測驗與有同性質之測驗法，各有優處。特特性之所在，由此可見，卽知χ^2對於統計上之重要性質

第八章 嘉氏(Charlier)及卜瓦松(Poisson)次數分配論

第一節 導言

嘉氏次數分配是能適合於偏斜較利害之次數分配其來源大概是1901年以後英人皮尔生作偏態分配以後約~~同時作出……次數分配時~~時在Scandinavia半島有Charlier,以普通常態函數作基礎作出偏態次數分配,于1903年成功此种偏態曲線名之Type A, Type B ---- 是以与法人Poisson之Type I, Type II ----- 等區别,Type A是常態函數之展開式,連續性之偏態函數,又名之為Laplace-Gauss-Charlier Type A.

Type B是Poisson函數之展開式,為不連續性之偏態函數,亦名之Gram-Charlier-Poisson Type B.

嘉氏曲線之原本登載于"Ueber das Fehlergesetz and Die Zweite Form des Fehlergesetzes" Charlier 此称在假定的原始誤差……任何差誤率不出下之兩公式中之一

Type A $F(x) = A_0 \psi(x) + A_3 \overset{(3)}{\psi}(x) + A_4 \overset{(4)}{\psi}(x) + \cdots$

式中 $\psi(x) = \dfrac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\bar{x})^2}{2\sigma^2}}$

Type B $F(x) = B_0 \phi(x) + B_1 \Delta\phi(x) + B_2 \Delta^2 \phi(x) \cdots$

式中 $\phi(x) = \frac{e^{-\lambda}\sin\pi x}{\pi}\left[\frac{1}{x} - \frac{\lambda}{(x-1)1!} + \frac{\lambda^2}{(x-2)2!} - \frac{\lambda^3}{(x-3)3!}\right.$

如 x 為整數(正)或為零則其於 x 數列可簡寫如下:

$$\phi(x) = \frac{e^{-\lambda}\lambda^x}{x!}$$

第二節 嘉氏 TYPE A 次數分配

Charler TYPE A 是常態函數之展開式是連續性之偏態曲線其公式之導來与一切曲線公式同樣以二項展開式為基礎 $(p+q)^s f x = \frac{s!}{x!(s-x)!} p^x q^{s-x}$ 得來,共有兩種寫法。

(a) $f(x) = \psi(x) + A_3 \frac{d^3\psi(x)}{dx^3} + A_4 \frac{d^4\psi(x)}{dx^4} + \text{-----}$

如改換一下符号亦可寫成

$$f(x) = \psi(t) + a_3\psi^{(3)}(t) + a_4\psi^{(4)}(t) + \text{------}$$

此為一收斂級數且其收斂之速度與分配之对称與否有关即在对称或近於对称之分配收斂甚速而在強度偏態分配,此級数收斂甚緩,在統計學上可不必去研究,故從畧而不計,僅取其前二項或三項即足。

而式中 $\psi(x) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2}$ 即常態曲綫標準公式 $\psi^{(3)}(t)$ 為第三次微分

餘以此類推.

(b) 計算便利上可寫為

$$f(x) = \frac{N}{\sigma}\left[\phi(t) - \frac{c_3}{3!}\phi^{(3)}_{(t)} + \frac{c_4}{4!}\psi^{(4)}_{(t)} - \frac{c_5}{5!}\psi^{(5)}_{(t)}\right.$$

式中 $c_3 = d_3$　$c_4 = d_4 - 3$　$c_5 = d_5 - 10d_3$　$c_6 = d_6 - 15d_4 + 30$ 等等

吾人已知理想公式中常数不应太多,因高次動差可使其正確性減低,故在上式看來,本是一个無窮級数,是一个理想的公式。因要使理想切合实际,而合乎所設之條件,即常数不宜太多,所以取前三項

$$\phi(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} \qquad d_n = \frac{\mu_n}{\sigma^n}$$

然在 type A 中只有四个統計常数 $\bar{X}, \sigma, d_3, d_4$,而有的峯度加入無大等級,故只取前二項 Bowley 称之為 Generalized Curve of error 其公式為

$$f(x) = \psi(t) + a_3\psi^{(3)}_{(t)}$$

Bowley 普通公式為:

$$y = \frac{N}{\sqrt{2\pi}}\left[1 - \frac{\mu_3}{2\sigma^2}\left(\frac{x}{\sigma} - \frac{1}{3}\frac{x^3}{\sigma^3}\right)\right] e^{-\frac{x^2}{2\sigma^2}} \qquad \text{式中 } x = X - \bar{X}$$

如再加上峯度則更甚確实,故取兩項只包括 $\bar{X}, \sigma, d_3$ 三常数而丁

ype Ⅲ同，到不如用 type Ⅲ 為簡便，但偏態較利害時仍以用 Type A 為宜。

第三節　Type A 曲線之形式與表格

在 Type Ⅲ 中其曲線之形式視 $A^2-1 \gtreqless 0$ 而定，当 $A^2>0$，又視 $\alpha_3>0$ $A>0$ 或 $\alpha_3<0$，$A<0$，而定其曲線形式，則当視 α_3 及 α_4（即偏斜度及峯度）而修正之偏斜之位置，曲线之形式当視 α_3 及 α_4 為轉移。

在 Type Ⅲ 中之偏態先有起点，由 $-A$ 到 $+A$ 而在 Type A 中之起点而為無限大（∞），由 $-\infty$ 到 $+\infty$。其 α_4 乃由 $A_4\psi'''(t)$ 糾正之数 Type A 是与常態曲線分配之理論差不多，式中 α_3 愈大，則偏斜愈大，α_3 為正則為正偏態，α_3 為負，則為負偏態，如是 Type Ⅲ 的問題在 Type A 中可以解決，但 Type A 之問題不能在 Type Ⅲ 中求得解決，峯度高則所作出之圖亦高，低則亦低。

因 Type A 来源是根据常態曲線函数，故 Type A 公式之計算可利用常態曲線表格，如 Type A 公式中之第二項可利用常態曲線表中第三次微分之縱坐標表或二次微分之面積表，因求面積時積分與微分可互相約去一次。（附表）

Bowley 教授藉用動差相等之測驗，從 n 私抽樣之結果将其総

和或平均数组成之次数曲线,在已知條件之下如取其第一近似值 略去前所得 $/\sigma\sqrt{2\pi}\, e^{-x^2/2\sigma^2}$

如取其第二近似值,即得偏態曲線

$$y = \frac{N}{\sigma\sqrt{2\pi}}\left[1-\frac{\mu_3}{2\sigma^3}\left(\frac{x}{\sigma}-\frac{1}{3}\frac{x^3}{\sigma^3}\right)\right]e^{-\frac{x^2}{2\sigma^2}}$$

Bowley 據此造成兩種表

(1) $F_3(z) = \frac{1}{3!\sqrt{2\pi}}\left(\frac{d}{dz}\right)^2 e^{-\frac{1}{2}z^2}\Big|_0^z$

(2) $F_4(z) = \frac{1}{4!\sqrt{2\pi}}\left\{\left(\frac{d}{dz}\right)^3 e^{-\frac{1}{2}z^2}\right\}_0^z$

但 Bowley 著 El. of St. 僅有 $F_3(z)$ 表一種

即 $F_3(z) = \frac{1}{\sqrt{2\pi}}\left\{1-(1-z)^2 e^{-\frac{1}{2}z^2}\right\}$ 之数值表 表附 $\left(z=\frac{x}{\sigma}\right)$

第四節 Type A 之应用

在此吾人所当注意者為 type A 与 Type Ⅲ 之不同,除公式中所含統計常数不同外,其曲線基本有不同,即 type Ⅲ 是有制限,因而 type A 是無制限的,至於 Type A 之应用甚廣,今举例以明之,分面積法、縱坐標法及 F_3 表配合法。

(一)面積法：

例：10,000人中其統計常数為

$m = 119.23$ 磅　m為平均数

$\sigma = 13.35$

$\alpha_3 = 0.65$

$\alpha_4 = 3.31$

求重於150磅的估計人数，用 normal Form 與 type III Type A 三法比較之

令 $t = \dfrac{X - m}{\sigma} = \dfrac{150 - 119.23}{13.35} = \dfrac{30.77}{13.35} = 2.3$

用 normal Type 為 $.5 - .489276 = .010724$

$.010724 \times 10{,}000 = 107$人　即所求之数用此数列為偏態，故這種做法(正確求之)之不正確

用 Type III 做之　取 $\alpha_3 = .6$ 與 $\alpha_3 = .7$ 之平均数 $\alpha_3 = .65$

則為 $1 - .97706 = .02294$

$.02294 \times 10{,}000 = 229$人　即為所求之数（比較正確因合偏態性質）

用 Type A 做之一種取兩項捨其 α_4 則為

$$\int_{2.3}^{\infty} f(t) = \int_{2.3}^{\infty} \phi(t)\,dt + a_3 \int_{2.3}^{\infty} \phi^{3}(t)\,dt$$

由上求之得 $.010724$

$$a_3 = \frac{-\alpha_3}{3!} = \frac{-.65}{6} = -1.1$$

$$\int_{2.3}^{\infty} \phi(t) = \int_{2.3}^{\infty} \phi(t)\,dt + a_3\left[\phi^{2}(\infty) - \phi^{(2)}(2.3)\right]$$

$$\because \phi^{2}(2.3) = -.12152$$

上式 $= .010724 + (-1.1)\times(-.12152) = .010724 + .0133672$

$= .024091$

$.024091 \times .0001 = 240$ 即為所求之数(与 type Ⅲ 相近)

用 Type A 取到第三項時

$$a = \frac{\alpha_4 - 3}{4!} = \frac{3.33-3}{4!} = .026$$

例 $\int_{2.3}^{\infty} f(t) = \int_{2.3}^{\infty} \phi(t)\,dt + a_3 \int_{2.3}^{\infty} \phi(t)\,dt + a_3\left[\underset{0}{\phi^{(3)}(\infty)} - \underset{.14920}{\phi^{(3)}(2.3)}\right]$

$= .010724 + .0133672 + .026 \times .14920 = .00278$

P. 65

高級統計學

$.00278 \times 10{,}060 = 278$ 人 最為正確 (因為有 α_4 故也)

(二) 縱坐標法

$$y = \frac{1}{\sqrt{2\pi}} e^{-\frac{t}{2}}$$

$$y' = \frac{-t}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$$

$$y'' = \frac{1}{\sqrt{2\pi}} (t^2-1) e^{-\frac{t^2}{2}}$$

$$y''' = \frac{1}{\sqrt{2\pi}} \left[(2t) e^{-\frac{t^2}{2}} - (t^2-1) t e^{-\frac{t^2}{2}}\right]$$

$$= \frac{1}{\sqrt{2\pi}} \left[-t^3+3t\right] e^{-\frac{t^2}{2}} = \frac{1}{\sqrt{2\pi}} \left[(t^2-3)t\right] e^{-\frac{t^2}{2}}$$

查 $t = .5$ 時為 $.484090$

$t = -.5$ 時為 $-.484090$

故 t 為負數時則其面積变一符号亦可用.

故縱坐標法 $f(x) = \phi(t) + a_3 \phi^{(3)}(t) + a_4 \phi^{(4)}(t)$

如 $t = 1$ 代入則

$$f(1) = \phi(1) + a_3 \phi^{(3)}(1) + a_4 \phi^{(4)}(1)$$

如 $\alpha_3 = 0$ 則可用常態曲線求理論次数

但 α_3 在（−1.1 到 +1.1）之間者用 Type Ⅲ 求理論次数在起过此数者用 type A 以求理論次数分配，如上例分三種方法比較已可証实.

(三) 应用 Bowley $F_3(z)$ 表配合法

例：(Bowley El. of SS. p. 309)

St Louis 公立学校报告中曾将第六级学龄不同之学生人数調查如下.

x 10 11 12 13 14 15 16 17 18 19

y 26 201 673 1001 739 390 80 13 1 3044

求得 $m_1 = 13.665$　$m_2 = 1.698$　$m_3 = 0.356$

$\sigma = 1.190$　$K = .2059$　$s^2 = 1.498 - 0.083 = 1.415$

$s = 1.190$　$1/s = 0.840$　$a_3 = m_3/s^3 = 0.211$

分解計算列成如下表

x	$z = \frac{x-(1)}{s}$	$\frac{1}{2}f(z)$	$F_3(z)$	$a_3F_3(z)$	(3)+(5)	Δ	$n\Delta$		normal
10	$-\infty$	−0.5000	−0.0665	−0.0140	−.5140				
11	−2.24	−.48750	−.0882	−.0186	−.5061	0.0079	24	25	38
12	−1.40	−.4192	−.0904	−.0191	−.4383	.0678	206	201	208

13	-0.56	-.2123	-.0275	-.0058	-.2181	.2202	676	673	630
14	0.28	.1103	-.0076	-.0016	-.1087	.3268	995	1001	982
15	1.12	.3686	-.1755	-.0159	-.3527	.2840	743	759	786
16	1.96	.4750	-.0942	-.0199	-.4551	.1024	312	310	521
17	2.80	.4974	.0755	-.0159	-.4815	.0264	81	80	[illegible]
18	3.64	.4999	.0672	-.0142	-.4857	.0042	13	13	
19	-∞	.5000	.0565	-.0140	-.4860	.0003	1	1	[illegible]
							3044	3044	3044

在第一近似值在平均数左右2б以内尚能配合,第二近似值则与观察值密切配合

第五節 (Charlier Type B) 嘉氏B型曲線

当变数為整数時,Type A 頗不适用,困此又利用 Pisson 二项极限分配函数而得 Type B,此亦如 Type A 之以常態曲線公式為基礎修改而成,故在不連續曲線中 Paisson 曲線是一种重要基本之曲线,当專章論之,今先述 Type B.

(一) 公式

$$f(x) = \phi(x) + B_0 \Delta^2 \phi(x) + B_3 \Delta^3 \phi(x) + \cdots$$

$M = 3.0715$ $\mu_2 = 14.8121$ $B_2 = \frac{\mu_2 - \lambda^2}{2!} = -.0864$

$\Delta^2\phi(x) = \phi(x) - 2\phi(x-1) + \phi(x-2) = .020827 - 0 + 0 = .020827$

$\phi(x) = \ldots \phi(x-1) + \ldots \phi(x-2) = .080632$

其 b. Sepia granules Rescharged from [illegible] Rutherford and Geiger

$N = 2608$ $m = 3.55$ $B_2 = -.0125$

(1)	(2)	(3)	(4)	(5)	(6)
x	ϕx	$\Delta^2\phi(x)$	$N\times(2)$	$N\times(3)\times B_2$	(4)+(5)
0	.020668	.020668	53.9	− 6.7	47
1	.050156	+.038820	209.1	−12.1	196
2	.153035	+.013611	406.4	− 5.2	[illegible]
3	.201015	−.020937	524.2	9.7	[illegible]
4	.190962	−.051125	508.5	16.8	525
5	.151166	−.039151	358.5	12.3	447
6	.097660	−.007918	254.9	3.2	258
7	.050109	−.003814	141.2	−3.2	138
8	.026316	.015668	68.7	−5.1	64

9	.011351	.008021	29.6	−4.2	25
10	.004407	.004092	11.5	−2.6	9
11	.001555	.001500	4.1	−1.2	3
12	.000503	.000699	1.3	−0.6	1
13	.000150 .000012	.000245	0.4	−0.2	0
14	.000010	.000076	0.1	−0.1	0
15	.000003	.000025	0	−0	0
16	.000001	.000005	—	—	0

算法 (C)

(1) x	(2) $n\psi$	(3) $n\Delta\psi$	(4) $n^2\Delta\psi$	(5) $\frac{1}{2}nb_3\Delta^4\psi$	(6) (2)+(5)	(7) Obs.	(8) Poisson
0	54.40	54.40	50.40	−5.25	49	57	54
1	210.52	156.12	101.72	−9.82	201	203	21[illegible]
2	407.37	196.85	40.73	−3.93	403	383	[illegible]
3	525.49	118.12	−78.73	7.60	533	525	526
4	508.43	−17.06	−135.18	13.05	522	532	509
5	393.52	−114.91	−97.85	9.44	403	408	394

6	253.81	-139.71	-24.80	2.39	256	273	254
7	140.34	-113.47	26.20	-2.55	138	137	140
8	67.89	-72.45	41.02	-3.86	64	65	68
9	29.18	-38.71	33.74	-3.26	26	27	29
10	11.29	-17.89	20.80	-2.01	9	10	11
11	3.96	-7.33	10.56	-1.02	3	4	4
12	1.28	-2.68	4.15	-0.45	1	2	1
13	0.39	-0.89	1.79	-0.17	0	0	0
					2605	2608	2608

第六節 Poisson 次数之分配論

1. Poisson Type 公式之由来

Poisson Type 是小数机率函数是不連續性的曲線其有常態曲線及偏態曲線之存在,是因宇宙間發生的事情有呈常態或偏態,故有常態曲線及偏態曲線之必要,然某種統計發生之数目很多,可說是大量現象,假如社会上还有其他不常發生的事情,如社会統計中之自殺統計,疾病統計,災害統計,工廠机械停工統計,汽車撞死人之統計,馬跌人統計統計等,這或許是不常發生之事件,但也必需有統

計既要統計不得不設法想出合乎這統計理論之公式以應付一切困為這粒既不适合事態曲线又不适合 Type Ⅲ 及 Type A 故 Poisson Type A 及 Type B 因此而發生,至於曲线公式之証明,亦以二項分配定理為出發矣

該曲线公式 $y = f(x) = e^{-\lambda}\dfrac{\lambda^{x}}{x!}$ 　發生之平均數 $\lambda = m$,　發生次數 $x = $ 正洮

今吾人取二項展開式之任意一項,并取其極限值

并先假定 p 相当小,取 P_m 之極限(当 $S \longrightarrow \infty$) p 雖然小,但很可能使 Sp 趋於極限 λ 故有

$$P_m = \frac{S!}{m!\,(S-m)!}p^m q^{S-m}$$

$$= \frac{S(S-1)(S-2)\cdots(S-m+1)}{m!}p^m(1-p)^{S-m}$$

$$= \frac{S(S-1)(S-2)\cdots(S-m+1)}{S^m\,m!}S^m p^m(1-p)^{S-m}$$

$$= \left(1-\frac{1}{S}\right)\left(1-\frac{2}{S}\right)\cdots\left(1-\frac{m-1}{S}\right)\frac{(Sp)^m}{m!}(1-p)^{S-m}$$

当 $S \longrightarrow \infty$. $Sp \longrightarrow \lambda$ 　$p = \dfrac{\lambda}{S}$

$$\therefore P_m = \left(1-\frac{1}{S}\right)\left(1-\frac{2}{S}\right)\left(1-\frac{3}{S}\right)\cdots\left(1-\frac{m-1}{S}\right)\frac{\lambda^m}{m!}\left(1-\frac{\lambda}{S}\right)^{S-m}$$

此公式以 $\phi(\lambda,x) = \frac{e^{-\lambda}\lambda^x}{x!}$ 為母函数，將分配函数 ??? 展為（此時 x 為整数即不連續分配）$\phi(\lambda,x)$，$\Delta\phi(\lambda,x-1)$，$\Delta^2\phi(\lambda,x-2)$ $\Delta^3\phi(\lambda,x-3)$ ------- $\Delta^n\phi(\lambda,x-n)$ 之級数即

$$F(x) = B_0\phi(\lambda,x) + B_1\Delta\phi(\lambda,x-1) + B_2\Delta^2\phi(\lambda,x-2) + B_3\Delta^3\phi(\lambda,x-3) +$$
$$+ B_n\Delta^n\phi(\lambda,x-n) +$$

式中 $\Delta^n\phi(\lambda,x-n)$ 為 $\phi(\lambda,x-n)$ 之 n 級差分

$$\Delta\phi(\lambda,x-1) = \phi(\lambda,x) - \phi(\lambda,x-1)$$
$$\Delta^2\phi(\lambda,x-2) = \phi(\lambda,x) - 2\phi(\lambda,x-1) + \phi(\lambda,x-2)$$

$$\Delta^n\phi(\lambda,x-n) = \Sigma_0^n r(-1)^r\left(\frac{n}{r}\right)\phi(\lambda,x-r)$$

式中 $\phi(x) = e^{-\lambda}\frac{\lambda^x}{x!}$ 使 $x = 0,1,2$

$\phi_0(x) = 1$ $\Delta\phi(x) = \phi(x) - \phi(x-1)$

$$\Delta^2\phi(x) = \Delta\phi(x) - \Delta\phi(x-1) = \phi(x) - 2\phi(x-1) + \phi(x-2)$$

$B_1 = 0$ $B_0 = 1$ $B_2 = \frac{1}{2!}(\mu_2 - \lambda)$

$$B_3 = \frac{\mu_3 - 3\mu_2 + 2\lambda}{3!}$$

（二）計算方法

如不計其校正数与 Poisson type 相同 正如 TYPE Ⅲ 不加入校

高級統計學 p.69.

正項之與 Type A 相同一樣。

算法 (a)

x	f_x	Poisson $\phi(x)$	$\Delta^2\phi(x)$ (4)	Poisson理論次數 $n\phi(x)$ (5)	校正數 (6)	(7)
0	57	.020827	.020827	58.3	-4.8	49.5 (50)
1	203	.080632	.031978	210.5	-.9	201.3 (201)
2	283	.156033	.075576	407.1	-3.6	403.5 (403)
3	525	.201425	.030105	525.3	6.9	532.5 (532)
4	532	.194954	.051813	508.4	11.9	520.3 (52[illegible])
5	408	.150952	.037551	393.7	8.6	402.3 (402)
6	273	.097402	.069548	254.3	2.2	256.5 (256)
7	139	.053869	.010017	140.5	-2.3	138.2 (13[illegible])
8	45	.026069	.015733	68.0	-3.6	64.4 (6[illegible])
9	27	.011214	.012945	29.0	-3.0	26.1 (26)
10	10	.004341	.007982	14.3	-1.8	9.5 (9)
11	4	.001528	.000060	4.0	-.9	3.1 (3)
12	0	.000413	.111778	1.7	-.4	.9 (1)
13	1	.000147	.000689	.4	-.3	.2 (0)
14	1	.00041	.000401	.1	-.1	0 (0)
	2607					

而 $(1-\frac{1}{s})(1-\frac{2}{s})\cdots\cdots(1-\frac{m-1}{s})\longrightarrow 1$, 当 $s\longrightarrow\infty$

$(1-\frac{\lambda}{s})^{s-m}=(1-\frac{\lambda}{s})^{s}(1-\frac{\lambda}{s})^{-m}=e^{-\lambda}$

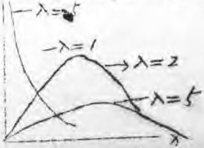
因 $\lim_{s\to\infty}(1-\frac{\lambda}{s})^{s}=e^{-\lambda}$; $\lim_{s\to\infty}(1-\frac{\lambda}{s})^{-m}\longrightarrow 1$

故 $P_m=\frac{\lambda^m}{m!}e^{-\lambda}$

此即 Paisson Type 公式

2, Paisson Type: 曲線形式

Paisson Type 為不連續性的曲线,且普通函数等不能統計的小数函数而此函数可計小数之統計,其曲線之形式当以λ之值而定,λ之值越小,曲線之偏斜亦越利害. 如圖:



吾人今將其式推廣之即 x 之值不僅為一正整数,此時公式可改書如下,比較合理,即当 x 不限定為正整数時可寫成

$$y=f(x)=e^{-\lambda}\frac{\lambda^{x}}{\Gamma(x')}$$

(3) 應用:

paisson Type 之應用以便利計算發生很稀少事件之統计,故常態曲線者又称 Law of Small numbers 今举例概畧述之.

(1) 利用 Paisson Type 估計理論次數之配合即配合曲線

例如根據德人 Bostkewitack 統計,在陸軍中每 50,000 人共染疾病死亡之人數之次數分配如次,因在 50,000 人中死亡之次數很小

x 死亡數	0	1	2	3	4	5	共計
f 次數	109	65	22	3	1	0	200
理論次數	108.7	66.3	20.2	4.1	.6	.1	200.00

如用常態曲線 Type A, Type Ⅲ 等配合結果不確,故以 paisson Type 來配合曲線結果才合理

如 $\lambda = \frac{\Sigma xf}{\Sigma f} = \frac{122}{200} = .61$

代入公式 $y = e^{-\lambda}\frac{\lambda^x}{x!} = e^{-.61}\frac{(.61)^x}{x}$

$x = 0 \quad y = e^{-.61}\frac{(.61)^0}{0!} = e^{-.61} = .5435 \quad \because 0! = 1$

理論次數 $.5435 \times 200 = 108.7$

$x = 1 \quad Ny = e^{-1.61} \times 200 \times 1.61 = 108.7(1.61) = 66.3$

$x = 2 \quad Ny = e^{-1.61} \frac{1.61^2}{2} \times 200 = 66.3 \left(\frac{1.61}{2}\right) = 20.2$

$x = 3 \quad Ny = 20.2 \times 1.61/3 = 4.1$

$x = 4 \quad Ny = .6$

$x = 5 \quad Ny = .1$

又例如白喉之死亡統計如次(以每年10,000人為次數單位)求理論次數

死亡人數	0	1	2	3	4	5	6	共計
次數	4	6	5	3	4	0	1	25
理論次數	3.54	6.93	6.79	4.02	1.77	1.07	[illegible]	24.7

解. $\lambda = \frac{\Sigma x f}{\Sigma f} = \frac{49}{25} = 1.96$

$y = e^{-1.96} \left(\frac{1.96^x}{x!}\right)$

$x = 0. \quad 0! = 1 \quad y = e^{-1.96}$ 以對數求之

$\log y = -1.96 \log e = -1.96 \times .4342945 = -1.96 \times .4342945$

$= -.8491\,6592 = \bar{1}.1508\,2794$

$\therefore y = \text{Anti}\log \bar{1}.150834 = .141472 = .1415$

例. 設台灣之死亡統計 以每(10,000人為單位)

死亡人數之次數

x	f_x	xf_x	理論次數
0	4	0	3.5
1	8	8	6.9
2	5	10	6.8
3	3	9	4.4
4	4	16	2.2
5	0	0	.8
6	1	6	.3
	25	49	

$$\lambda = \frac{\Sigma x f}{n} = \frac{49}{25} = 1.96$$

$$f_0' = \frac{e^{-\lambda}\lambda^x}{x!} = \frac{e^{-1.96}(1.96)^0}{(0)!} \times 25 = \frac{25\times e^{-1.96} x!}{1} = 25\times e^{-1.96}$$

$$= 25\times .140858 = 3.5$$

$$f_1' = \frac{25\times e^{-1.96}(1.96)^1}{1!} = 3.5214\times 1.96 = 6.9019$$

$$f_2' = \frac{25 \times e^{-1.96}(1.96)^2}{2!} = \frac{6.9019 \times 1.96}{2} = 6.7635$$

$$f_3' = \frac{25 \times e^{-1.96}(1.96)^3}{3!} = \frac{6.7638 \times 1.96}{3} = 4.4167$$

$$f_4' = \frac{25 \times e^{-1.96}(1.96)^4}{4!} = \frac{4.4167 \times 1.96}{4} = 2.1641$$

$$f_5' = \frac{25 \times e^{-1.96}(1.96)^5}{5!} = \frac{6.1641 \times 1.96}{5} = .8483$$

$$f_6' = \frac{25 \times e^{-1.96}(1.96)^6}{6!} = \frac{.8485 \times 1.96}{6} = .2773$$

總結上述若用 Paisson 小数机率法其结果与二項展開法之結果相同

$(\ldots)^{100}$

$$p = \frac{e^{-\lambda}\lambda^x}{x!} \qquad \lambda = m = np = 100 \times (\ldots)^{100}$$

Poisson TYPE 之应用

1. 可以估计理論次数

2. 可以求出二項分配之各种估計數字

高级統計学

P. 72

附錄

A型曲線為B型曲線之証明

I. 各式之研究。

如果伊在城中取S個為一樣本，r個成功之机率當以

$$B(r) = \frac{s!}{r!(s-r)!}p^r q^{s-r}$$

表示。此為非連續机率之形式，如為連續机率則取其極限值。歐洲大陸學者常取一函數為各$B(r)$的机率之協補公式。

即 (1) $$B_0(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\theta(w)e^{-xwi}dw$$

式中 $i^2 = -1$。

(2) $$\theta(w) = (pe^{wi}+q) = \sum_{r=0}^{\infty}B(r)e^{rwi}$$

茲先証明公式(1)

查里愛(Charlier)去導演其公式時擬定一函數$\theta(w)$為下列之級數

(3) $$\theta(w) = \alpha\left[B(0)+B(\alpha)e^{\alpha wi}+B(2\alpha)e^{2\alpha wi}+B(3\alpha)e^{3\alpha wi}+\cdots\right.$$
$$\left.+B(-\alpha)e^{-\alpha wi}+B(-2\alpha)e^{-2\alpha wi}+B(-3\alpha)e^{-3\alpha wi}+\cdots\right]$$

或 $$\theta(w) = \alpha\sum^{\infty} B(\alpha m)e^{\alpha mwi}$$

$m=-\infty$

依柯西定理(Theorem of Cauchy)下列積分(參看: Goursat: Mathematical analysis. P.364)

$$(4)\quad I(w) = \int_{-\infty}^{+\infty} B(x) e^{xwi} dx$$

爲有限且收斂。若使 $m\alpha = x$，並使 0 爲 α 之極限值，則 α 變爲與 dx 相等，則 $B(m\alpha) = B(x)$，其結果可寫爲

$$\lim_{\alpha=0} \theta(w) = I(w)$$

將(3)式乘以 $e^{-r\alpha wi} dw$，並於 $-\frac{\pi}{2}$ 與 $\frac{\pi}{2}$ 間積分，則其左邊之形式爲 $\int_{-\pi/2}^{\pi/2} \theta(w) e^{-r\alpha wi} dw$，右邊爲若干積分之總和，且除含 $B(r\alpha)$ 是因子之一項外，其他各項均爲零(參看: Goursat: Mathematical Analysis. P.418—420)。

此不爲零之一項即爲

$$\alpha \int_{-\pi/2}^{\pi/2} B(\alpha r) dw \quad 或 \quad 2\pi B(r\alpha)$$

因此可得 $$B(r\alpha) = \frac{1}{2\pi} \int_{-\pi/2}^{\pi/2} \theta(w) e^{-r\alpha wi} dw \qquad (4a)$$

若使 $\alpha=1$, $r\alpha=x$, 則得

$$B(x)=\frac{1}{2\pi}\int_{-\pi}^{\pi}\theta(w)e^{-xwi}dw \qquad (14b)$$

即 $B_0(x)=\frac{1}{2\pi}\int_{-\pi}^{\pi}\theta(w)e^{-xwi}dw$

若使 α 漸趨於零為極限 $r\alpha=x$ 則有

$$B(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\theta(w)e^{-xwi}dw \qquad (14c)$$

即 $B_0(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\theta(w)e^{-xwi}dw$

亦即為所欲証者

$B_0(x)$之來源已知,進而即可演繹$B_0(x)$,將(2)代入(1)並使x等於正整數m,則有

$$B_0(x)=\sum_{r=0}^{s}B(r)\frac{\sin(r-x)\pi}{(r-x)\pi}$$

$$=B(0)\frac{\sin(-x\pi)}{-x\pi}+B(1)\frac{\sin(1-x)\pi}{(1-x)\pi}+\cdots\cdots+B(s)\frac{\sin(s-x)\pi}{(s-x)\pi}$$

以x等於正整數m時,除$B(m)$之一項外,其餘各項均為零,即

$$B_0(m) = B(m)$$

因此若 $x=m$（正整数），$B_0(m)$ 即為 $(p+q)^s$ 展開式中之一項
由此結果，亦可証明(1)式

上列証明均了解後，即可推究A型與B型級数最关推究方法先將(2)式展開再取 w 與 p 之冪級数，即可得A型級数與B型級数。

(a) A型曲線之推究

由(2)

$$\frac{d\log\theta(w)}{dw} = \frac{ispe^{wi}}{q+pe^{wi}} \qquad (5)$$

展開(5)式右边（取 w 之冪級数）則為

$$\frac{d\log\theta(w)}{dw} = ispe^{wi}(q+pe^{wi})^{-1}$$

$$= isp\left(1+wi+\frac{1}{2!}(wi)^2+\frac{1}{3!}(wi)^3+\cdots\cdots\right)$$
$$\left[q-(pe^{wi})+(pe^{wi})^2-\cdots\cdots\right]$$

$$= isp\left(1+wi+\frac{1}{2!}(wi)^2+\frac{1}{3!}(wi)^3+\cdots\cdots\right)$$
$$\left[q-p\left\{1+wi+\frac{1}{2!}(wi)^2+\frac{1}{3!}(wi)^3+\cdots\cdots\right\}\right.$$
$$\left.+p^2\left\{1+wi+\frac{1}{2!}(wi)^2+\frac{1}{3!}(wi)^3+\cdots\right\}+\cdots\right]$$

$$= isp\left[1+qwi-\frac{1}{2}q(p-q)(wi)^2+\cdots\right]$$

$$\therefore \log \Theta(w) = s\left[pwi + \frac{1}{2!}pq(wi)^2 + \frac{1}{3!}pq(p-q)(wi)^3 + \cdots\right]$$

$$[\Theta(0) = 1]$$

或寫為 (6) $\Theta(w) = e^{b_1 wi + \frac{b_2}{2!}(wi)^2 + \frac{b_3}{3!}(wi)^3 + \cdots\cdots}$

及 (7) $b_1 = sp,\quad b_2 = spq,\quad b_3 = spq(p-q), \cdots\cdots$

若使 $A_3 = \frac{1}{6}spq(p-q),\quad A_4 = \frac{1}{24}spq(1-6pq), \cdots\cdots$

則(6)式可寫為

(8) $\Theta(w) = e^{b_1 wi - b_2 w^2/2}\left[1 - A_3(wi)^3 + A_4(wi)^4 - \cdots\right]$

(因(8)式為w之完全函數，右邊為一致收斂級數，故可如此也)

以(8)代入(1)，則有

(9) $B_0(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} dw\, e^{-(x-b_1)wi - b_2 w^2/2}\left[1 - A_3(wi)^3 + A_4(wi)^4 - \cdots\right]$

若使 (10) $\Phi(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} dw\, e^{-(x-b_1)wi - b_2 w^2/2}$

則由(9)得 (11) $B_0(x) = \Phi + A_3 \frac{d^3\Phi(x)}{dx^3} + A_4 \frac{d^4\Phi(x)}{dx^4} + \cdots$

若 b_2 不甚小，变换其積分限後（即由 $\pm\pi$ 变至 $\pm\infty$），可以 $\varphi(x)$ 代替 $\Phi(x)$，即 $\varphi(x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} dw\, e^{-(x-b_1)wi - b_2 w^2/2}$

但 $\varphi(x)$ 可寫為

$$(12) \quad \varphi(x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-b_2 w^2/2} \cos[w(x-b_1)]\, dw$$

將此式對 x 微分，則有

$$\frac{d\varphi(x)}{dx} = -\frac{1}{\pi}\int_{0}^{\infty} e^{-b_2 w^2/2} w \sin[w(x-b_1)]\, dw$$

其式右边以分部積分法積分，則得

$$\frac{d\varphi(x)}{dx} = -\frac{(x-b_1)}{b_2 \pi}\int_{0}^{\infty} e^{-b_2 w^2/2} \cos[w(x-b_1)]\, dw$$

$$= -\frac{(x-b_1)}{b_2}\varphi(x)$$

以分離变数法求此方程式之解，得

$$(13) \quad \varphi(x) = A e^{-(x-b_1)^2/2b_2}$$

式中A為常数，於(12)與(13)中，使 $x = b_1$，則有

$$A = \frac{1}{\pi}\int_0^{\infty} e^{-b_2\omega^2/2}\,d\omega = \frac{1}{(2\pi b_2)^{\frac{1}{2}}}$$

故 (14) $$\varphi(x) = \frac{1}{(2\pi b_2)^{\frac{1}{2}}}\,e^{-(x-b_1)^2/2b_2}$$

上式可將(11)式之A型級數寫為

(15) $$B_0(x) = \varphi(x) + A_3\,\frac{d^3\varphi(x)}{dx^3} + A_4\,\frac{d^4\varphi(x)}{dx^4} + \cdots\cdots$$

式中若 $b_2 = \sigma^2$, 則

$$\varphi(x) = \frac{1}{\sigma(2\pi)^{1/2}}\,e^{-(x-b_1)^2/2\sigma^2}$$

但所須研究者積分限由 $\pm\pi$ 變至 $\pm\infty$ 時接近程度如何,相差多少,欲論此點可書

$$\varphi(x) - \Phi(x) = \frac{1}{\pi}\int_{\pi}^{\infty} d\omega\, e^{-\sigma^2\omega^2/2}\cos[(x-b_1)]$$

因之 $$[\varphi(x) - \Phi(x)] < \frac{1}{\pi}\int_{\pi}^{\infty} d\omega\, e^{-\sigma^2\omega^2/2}$$

$$= \frac{1}{\sigma\pi}\int_{\sigma\pi}^{\infty} e^{-\lambda^2/2}\,d\lambda \quad 若\ \lambda^2 = \sigma^2\omega^2$$

由此式可知 σ 愈大, $\varphi(x)$ 和 $\Phi(x)$ 之差數愈小 接近程度愈甚.

(b) B 型級数之推究

首先展開(5)式取 p 之冪級数，則得 $[p+q=1]$

$$(16)\quad \frac{d\log\Theta(w)}{dw} = \frac{ispe^{wi}}{1-p(1-e^{wi})}$$

$$= ispe^{wi}[1+p(1-e^{wi})+p^2(1-e^{wi})^2+\cdots\cdots]$$

此式為一收斂級数，因 $|p(1-e^{wi})|<1$，而 $\Theta(0)=1$，故得

$$\log\Theta(w) = -sp[1-e^{wi}+\frac{p}{2}(1-e^{wi})^2+\frac{p^2}{3}(1-e^{wi})^3+\cdots]$$

故 (17) $\Theta(w) = e^{-sp(1-e^{wi})}[1+B_2(1-e^{wi})^2+B_3(1-e^{wi})^3+\cdots\cdots]$

式中 $B_2 = \frac{-sp^2}{2}$，$B_3 = -\frac{sp^3}{3}$，$B_4 = -\frac{sp^4}{4}+\frac{s^2p^4}{8}\cdots\cdots$

由(1)及(17)得

$$(18)\quad B_0(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} dw\, e^{-xwi-sp(1-e^{wi})}[1+B_2(1-e^{wi})^2 + B_3(1-e^{wi})^3+\cdots\cdots\cdots]$$

使 $\varphi(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} dw\, e^{-xwi-sp(1-e^{wi})}$

$$= \int_{-\pi}^{\pi} Q(w,x)\,dw.$$

所後 $\Delta\varphi(x) = \varphi(x) - \varphi(x-1) = \frac{1}{2\pi}\int_{-\pi}^{\pi} dw\, e^{-xwi - sp(1-e^{wi})}$

$$- \frac{1}{2\pi}\int_{-\pi}^{\pi} dw\, e^{-(x-1)wi - sp(1-e^{wi})}$$

$$= \int_{-\pi}^{\pi} (1-e^{wi})\,\Phi(w, x)\, dw$$

同様 $\Delta^2\varphi(x) = \int_{-\pi}^{\pi} (1-e^{wi})^2\,\Phi(w, x)\, dw$

$$\Delta^3\varphi(x) = \int_{-\pi}^{\pi} (1-e^{wi})^3\,\Phi(w, x)\, dw$$

- - - - - - - - - - - - - - - - - -

因之 (19) $B_0(x) = \varphi(x) + B_2\Delta^2\varphi(x) + B_3\Delta^3\varphi(x) + \cdots\cdots$

將 $\varphi(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-xwi - sp(1-e^{wi})}\, dw$

寫為第一形式

$$\varphi(x) = \frac{e^{-sp}}{2\pi}\int_{-\pi}^{\pi} e^{-xwi + spe^{wi}}\, dw$$

$$= \frac{e^{-sp}}{2\pi}\int_{-\pi}^{\pi} e^{-xwi}\left[1 + spe^{wi} + \frac{s^2p^2}{2!}e^{2wi} + \cdots\cdots\right] dw$$

$$= \frac{e^{-sp}}{2\pi}\left[\int_{-\pi}^{\pi} e^{-xwi}\, dw + sp\int_{-\pi}^{\pi} e^{-wi(x-1)}\, \Phi(w) + \frac{s^2p^2}{2!}\int_{-\pi}^{\pi} e^{-wi(x-2)} \cdots + \cdots\right]$$

$$(20)\quad = \frac{e^{-sp}}{\pi}\left\{\frac{\sin x\pi}{x} + sp\frac{\sin(x-1)\pi}{x-1} + \cdots + \frac{(sp)^r}{r!}\frac{\sin(x-r)\pi}{x-r} + \cdots\right]$$

$$= \frac{e^{-sp}}{\pi}\sin\pi x\left\{\frac{1}{x} - \frac{sp}{x-1} + \frac{s^2p^2}{2!(x-2)} - \cdots + (-1)^r\frac{(sp)^r}{r!}\frac{1}{x-r} + \cdots\right]$$

$$(21)\quad = e^{-\lambda}\frac{\sin\pi x}{\pi}\left\{\frac{1}{x} - \frac{\lambda}{x-1} + \frac{\lambda^2}{(x-2)2!} - \cdots + \frac{(-1)^r\lambda^r}{(x-r)r!} + \cdots\right]$$

式中 $\lambda = sp$

上列方法乃用一致收斂級数之性質(properties of uniformly convergent series)分析者. 当 x 接近於整数 r 時, (20)中除

$$\frac{e^{-sp}}{\pi}\frac{(sp)^r}{r!}\frac{\sin(x-r)\pi}{x-r}$$ 外,

其餘各項均为零. 而此一項之極限值即為 Poisson 函数

$$\frac{e^{-sp}(sp)^r}{r!} = \frac{e^{-\lambda}\lambda^r}{r!}$$

公式(21)可作为 x 为非整数之 poisson 函数

*　　　*　　　*

II. A型級数和B型級数中常数之確定.

A型級数中常数之数可利用正交性質(Biorthogonal property)確定之.

公式(15)中若以平均数為原点，标准差為単位，則可以下式代替(15)式，即

(22) $$F(x) = \varphi(x) + a_3\varphi^{(3)}(x) + a_4\varphi^{(4)}(x) + \cdots + a_n\varphi^{(n)}(x) + \cdots$$

式中 $$\varphi(x) = \frac{1}{(2\pi)^{\frac{1}{2}}} e^{-\frac{x^2}{2}}$$

$\varphi^{(n)}(x)$ 為 $\varphi(x)$ 對 x 之 n 級導数

A式中係数 a_n $(n=0,3,4,\cdots)$ 可以動差表示，$\varphi^{(n)}(x)$ 與 Hermite 多項式 $H_n(x)$ 有正交関係：

$$\varphi^{(n)}(x) = (-1)^n H_n(x)\varphi(x)$$

因此(23) $$\int_{-\infty}^{\infty} \varphi^{(n)}(x) H_m(x)\, dx = 0 \qquad (m \neq n)$$

與 (24) $$\int_{-\infty}^{\infty} \varphi^{(n)}(x) H_m(x)\, dx = (-1)^n \frac{n!}{\sigma} \qquad (m = n)$$

因正交性質，可確定A型級数中之係数

茲先証明(23)與(24)式，因(23)可寫為

(25) $$\left\{\begin{aligned} \int_{-\infty}^{\infty} \varphi^{(n)}(x) H_m(x)\, dx &= (-1)^n \int_{-\infty}^{\infty} \varphi(x) H_n(x) H_m(x)\, dx \\ &= (-1)^{m+n} \int_{-\infty}^{\infty} \varphi^{(m)}(x) H_n(x)\, dx \end{aligned}\right.$$

以分部積分法積分

$$\int_{-\infty}^{\infty}\varphi^{(n)}(x)H_m(x)dx=\left[\varphi^{(n-1)}(x)H_m(x)\right]_{-\infty}^{+\infty}-\int_{-\infty}^{\infty}\varphi^{(n-1)}(x)H_m'(x)dx$$

$$=-\int_{-\infty}^{\infty}\varphi^{(n-1)}(x)H_m'(x)dx$$

假定 $n>m$, 連續 $m+1$ 次積分則得

$$\int_{-\infty}^{\infty}\varphi^{(n)}(x)H_m(x)dx=(-1)^{m+1}\int_{-\infty}^{\infty}\varphi^{(n-m-1)}(x)H_m^{(m+1)}(x)dx.$$

式中 $H_m^{(m+1)}(x)$ 為 $H_m(x)$ 之 $(m+1)$ 次導數, 但 $H_m(x)$ 為 x 之 m 次多項式, 第 $(m+1)$ 次導數為零, 故

$$(26)\quad \int_{-\infty}^{\infty}\varphi^{(n)}(x)H_m(x)dx=0 \qquad (n>m)$$

由(25) 設 $m>n$, 亦可証明得同樣結果.

若 $m=n$, 則由(25)

$$\int_{-\infty}^{\infty}\varphi^{(n)}(x)H_n(x)dx=(-1)^n\int_{-\infty}^{\infty}\varphi^{(n-n)}(x)H_n^{(n)}(x)dx$$

$$=(-1)^n\int_{-\infty}^{\infty}\varphi(x)H_n^{(n)}(x)dx$$

但多項式 $H_n(x)$ 之第 n 級導數 $H_n^{(n)}(x)$ 為 $n!$ 故

$$(27)\quad \int_{-\infty}^{\infty}\varphi^{(n)}(x)H_n(x)\,dx = (-1)^n n!\int_{-\infty}^{\infty}\varphi(x)\,dx$$

$$= \frac{(-1)^n n!}{\sigma(2\pi)^{1/2}}\int_{-\infty}^{\infty}e^{-x^2/2}\,dx = \frac{(-1)^n n!}{\sigma}$$

(22)式兩邊同
乘以 $H_n(x)$ 後積分，則得

$$\int_{-\infty}^{\infty}F(x)H_n(x)\,dx = a_n\int_{-\infty}^{\infty}\varphi^{(n)}(x)H_n(x)\,dx$$

$$= (-1)^n\frac{a_n n!}{\sigma} \qquad \text{(因其各項均為零)}$$

故得(28) $$a_n = \frac{\sigma(-1)^n\int_{-\infty}^{\infty}F(x)H_n(x)\,dx}{n!}$$

至相當次數之配中，a_n 之決定，則以相當次數函數 $f(x)$ 代替(28)中之 $F(x)$。

(28)中所決定之常數，乃以標準差(σ)為單位，為適合數學應用起見，仍回至原來單位，即將(28)中之 x 以 $\frac{x}{\sigma}$ 代替，故得

$$(29)\quad \begin{cases} a_n = \sigma\dfrac{(-1)^n}{n!}\displaystyle\int_{-\infty}^{\infty}f(x)H_n\left(\frac{x}{\sigma}\right)d\frac{x}{\sigma} \\ \quad = \dfrac{(-1)^n}{n!}\displaystyle\int_{-\infty}^{\infty}f(x)H_n\left(\frac{x}{\sigma}\right)dx \end{cases}$$

* * *

B 型函數之係數

決定B型級數中之係數 $c_0, c_1, c_2 \cdots\cdots$ 時，在此只限制於相等距離之縱坐標之均為正整數，且為方便起見，只以前三項代表此式，即

$$F(x) = c_0\varphi(x) + c_1\Delta\varphi(x) + c_2\Delta^2\varphi(x)$$

式中 $$\varphi(x) = \frac{e^{-\lambda}\lambda^x}{x!}$$

$$(x = 0, 1, 2 \cdots\cdots)$$

設 $f(x)$ 為觀察分配相對次數之縱坐標，故 $\Sigma f(x) = 1$，

$$\Sigma x f(x) = m_1', \quad \Sigma x^2 f(x) = m_2';$$

由此三動差，即可由下列方程式決定各係數之近似值

$$(30)\quad \begin{cases} \Sigma\left[c_0\varphi(x) + c_1\Delta\varphi(x) + c_2\Delta^2\varphi(x)\right] = \Sigma f(x) = 1 \\ \Sigma x\left[c_0\varphi(x) + c_1\Delta\varphi(x) + c_2\Delta^2\varphi(x)\right] = \Sigma x f(x) = m_1' \\ \Sigma x^2\left[c_0\varphi(x) + c_1\Delta\varphi(x) + c_2\Delta^2\varphi(x)\right] = \Sigma x^2 f(x) = m_2' \end{cases}$$

在決定各係數 $c_0, c_1, c_2 \cdots\cdots$ 以前，先以大數律中之各數代替作為近似值，即 $\Sigma\varphi(x) = 1$

$$m_1' = \sum x\varphi(x) = \lambda$$

$$m_2' = \sum x^2\varphi(x) = \lambda + \lambda^2$$

上列各数知後，即可進而推究其他各近似常数

$$\sum \Delta\varphi(x) = \sum\left[\varphi(x) - \varphi(x-1)\right] = 1 - 1 = 0$$

$$\sum \Delta^2\varphi(x) = \sum\left[\varphi(x) - 2\varphi(x-1) + \varphi(x-2)\right] = 0$$

$$\begin{aligned}\sum x\Delta\varphi(x) &= \sum x\left[\varphi(x) - \varphi(x-1)\right] \\ &= \sum x\varphi(x) - \sum(x-1)\varphi(x-1) - \varphi(x-1) \\ &= \lambda - \lambda - 1 \\ &= -1\end{aligned}$$

同样可得

$$\sum x\Delta^2\varphi(x) = 0$$

$$\sum x^2\Delta\varphi(x) = -2\lambda - 1$$

和 $\sum x^2\Delta^2\varphi(x) = 2$

將求得各数代入(30)中得

$$c_0 = 1;\quad \lambda c_0 - c_1 = m_1';\quad (\lambda+\lambda^2)c_0 - (2\lambda+1)c_1 + 2c_2 = m_2'$$

若按 $\lambda = m_1'$ 採数 $c_1 = 0$，m_2' 对换至以平均数为原点，则有下列関係

$$m_1' = \sum x\varphi(x) = \lambda$$

$$m_2' = \sum x^2\varphi(x) = \lambda + \lambda^2$$

上列各数知后，即可進而推究其他各近似常数

$$\sum \Delta\varphi(x) = \sum[\varphi(x) - \varphi(x-1)] = 1 - 1 = 0$$

$$\sum \Delta^2\varphi(x) = \sum[\varphi(x) - 2\varphi(x-1) + \varphi(x-2)] = 0$$

$$\sum x\Delta\varphi(x) = \sum x[\varphi(x) - \varphi(x-1)]$$

$$= \sum x\varphi(x) - \sum(x-1)\varphi(x-1) - \varphi(x-1)$$

$$= \lambda - \lambda - 1$$

$$= -1$$

同樣可得

$$\sum x\Delta^2\varphi(x) = 0$$

$$\sum x^2\Delta\varphi(x) = -2\lambda - 1$$

和 $$\sum x^2\Delta^2\varphi(x) = 2$$

將求得各数代入(30)中得

$$c_0 = 1, \quad \lambda c_0 - c_1 = m_1'; \quad (\lambda + \lambda^2)c_0 - (2\lambda + 1)c_1 + 2c_2 = m_2'$$

若按 $\lambda = m_1'$ 係数 $c_1 = 0$，m_2' 對換至以平均数為原点，則有下列關係