生物統計學與公共衛生之關係

梁明第

一、前言

近數十年來，統計學不但在自然科學方面漸漸被人重視；就是在社會科學中，同樣被人所注意。由於統計學是一種科學的方法，而科學之研究不外藉着實驗，由實驗之結果再運用統計之方法予以整理與分析，如此就可獲得其間之通理與規則。因爲統計學具有如此之功能，故其運用範圍非常廣泛，其所研究之對象則包括自然科學與社會科學。此等經驗科學，需要以實驗之事實，證明其因果關係，並推測假設之是否成立，進而建立科學之定律，故在經濟學、社會學、生物學、心理學、一般醫學及公共衛生學等學科中，都需要以統計方法作爲研究之工具。

二、生物統計學之意義及其重要性

生物統計學，爲統計學之一部分，是一門用途相當廣泛之科學。在醫學或生物學之領域中，亦佔有極重要之地位。所謂生物統計學，乃係運用各種之統計方法，分別對有關醫學與生物學領域中，發現之各種生物現象，加以數量之研究，並推論該種事象之未來發展之科學。

生命統計學，是生物統計學中之主要部分，對公共衛生極爲重要，所以亦稱爲公共衛生統計。爲研究人類全部生活成長之現象，也就是以人口之生命爲對象，從事研究分析之應用統計學。所謂生命統計事項，舉凡出生、死亡、死產、結婚、離婚、遷徙、疾病及死因等都包括在研究範圍以內。

由於生命統計，是有關人口及生命之賬簿，藉着此種統計，顯示疾病率、出生率及死亡率之高低，以反映該地區國民之健康狀況及公共衛生之成效。促使衛生當局對疾病之流行與傳染，獲得正確之資料，進而採取有效之預防措施。如對某些死亡最多之疾病予以遏止或預防，就能使一般死亡率降低。

藉著生命統計之方法，除獲知人口自然增加率外，吾人更可計算國民平均期待之餘命，而平均餘命之延長，爲醫藥及公共衛生改善之重要指標，其主要原因是各年齡別死亡率之下降，尤其是嬰兒與新生兒死亡率（ infant and neonatal mortality rate ）之下降及環境衛生之改進，促使肺炎、胃腸炎及結核病等疾病死亡率之減少。

三、公共衛生與生命統計（Vital statistics)

公共衛生的對象是人，凡是人類日常生活，衣、食、住、行及由受孕起到出生至死亡止，以及在此段期間所患之疾病等等，無一不與公共衛生有關，而包括在公共衛生問題範圍之內。因此一國　之國民，不管是自己或一個家庭之健康，和整個社會、民族之健康，都朝夕發生密切之關係與重大之影響。

健康與人之生命有密切之關係，很多人認爲無病症之徵候便是健康，但醫學上認爲此種觀念，不很適合，因爲無病症並不指絕對無病，亦即不能說絕對健康。世界衛生組織給健康所下之定義爲：「健康是身體的、心理的與社會的一種完全安寧狀態，不僅無病或不虛弱而已……」。欲瞭解健康狀況之是否正常，必須依據住所之環境衛生與其身體組織器官的生理情形而定。由此可見，維護健康並不是一件簡

梁明第：台灣省政府衛生處主計室副主任

單而容易的事，特別是公共衛生，面對着群體大衆更非易事。如何揭示各種疾病侵害人類所隱藏之奧妙，這就需要疾病統計與生命統計發揮其功能了。

公共衛生，以預防全體社會民衆的疾病為主要目的之一，也就是預防重於治療，此爲公共衛生防止疾病傳染之目標，然而疾病之患者，以老、弱、及妊娠婦女期間爲最多，其死亡之機會亦大。因此公共衛生所需要接觸之對象，就是生命統計之事項，也就是說，公共衛生除直接預防傳染疾病之發生，以確保國民身體健康外，如何促使出生率之降低，減少疾病之死亡，延長人民之平均餘命與增加國民之健康與工作效率，再再都與公共衛生之加強有關。這也是何以世界衛生組織將生物統計，特別是生命統計與衛生統計列爲公共衛生重要部門之原因。

四 公共衛生何以需要生物統計 (*Biostatistics*)

宇宙間之生物現象，千變萬化，而各種事物之現象，又顯得零亂無章，毫無頭緒。吾人如欲研究其各種現象時，經常感覺無從下手。雖然生物之現象有如此大之差別，然而其中也有些共同之原則可尋，甚至事物與事物之間，更有相互之關係存在。欲尋求這些共同之原則與若干事物之間的關係，進而研究其間之差別現象，甚至推測所研究結果之可靠性程度等等都需要運用統計之研究方法，予以歸納或比較，由此所得之結果，就可以解釋大自然間之各種生物現象與其蘊藏之眞理。

茲僅舉兩例說明公共衛生需要生物統計之原因：在公共衛生及生物學或醫學上，經常會遭遇著二種比率之比較，如吸香煙與不吸香煙之肺癌罹患率等，所謂罹患率（ Attack rate or morbidity rate ）係某地某年患某種疾病之人數對該地年中人口總數之千分比。而疾病致死率（ fatality or case fatality ），則系指死於某種疾病之人數對患染該種疾病人數之百分比，因疾病致死率之高低，則說明某種疾病對人類嚴重威脅之程度，此種百分比最易於瞭解與引人注意。

例一 吾人隨機調查（即對被調查之各個個體，不作任何有目的之選擇，純用偶然之方法調查，使全體中每一個體，均有同等被調查之機會）某種傳染病之患者235人，未接受預防注射者93人中，已死去61人；已接受預防注射者142人中，死去53人，故未接受預防注射之致死率爲 $\hat{p}_1=61/93=65.6\%$，而接受預防注射之致死率 $\hat{p}_2=53/142=37.3\%$，則接受預防注射之致死率，遠較未接受者少28.3%。由此二樣本中，吾人能否推斷出未接受預防注射之全體患者與已接受預防注射之全體患者的致死率，是否亦有差別？

此處必須注意者，因二者之致死率相差很大，一般人一定認爲此種預防注射很有效，但究竟有效否，吾人尚不能就數值上之差，予以斷定，必須利用生物統計的客觀之方法，予以推定或檢定，才能作最後斷定此項注射是否有效。欲推斷上述接受或未接受預防注射全體患者之比率，是否有差別？必須以一種「虛無假設」（ Null hypothesis ）來測驗。所謂虛無假設即假設「兩個全體 p_1 與 p_2 患者的致死率，彼此並無差別，即 $p_1-p_2=0$」。在這樣之假設下，其兩個全體致死率 p_1 及 p_2 並不知道，必須以 $\hat{p}_1$ 及 $\hat{p}_2$ 來推測或估計 p_1 及 p_2 值。即取 $\hat{p}=\dfrac{x_1+x_2}{n_1+n_2}$，$x_1$ 與 x_2 各爲樣本死亡者61與53人，$\hat{p}$ 爲

全體致死率之估計值，最後必須求出Z值（爲標準常態之Z值）其公式如下：

$$Z=\frac{(\hat{p}_1-\hat{p}_2)-0}{\sqrt{\hat{p}\hat{q}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$$

式中 $\hat{q}=(1-\hat{p})$，n_1 = 未接受預防注射者，n_2 = 接受預防注射者，求Z時，須以數值較大之致死率爲 $\hat{p}_1$，較小者爲 $\hat{p}_2$。

當樣本甚大時，推測或估計兩比率相等之假設，按下列步驟進行：

第一步　假設（H_0）：　p_1 ＝ p_2

第二步　對立假設（H_1）；$p_1 > p_2$

第三步　$\alpha=0.025$（即採用顯著水準爲5％之兩尾測驗）

第四步　棄却區域：$Z>1.96$

第五步　計算：$\hat{p}_1=\frac{61}{93}=0.656$，

$$\hat{p}_2=\frac{53}{142}=0.373$$

$$\hat{p}=\frac{61+53}{93+142}=\frac{114}{235}$$

$$=0.4851$$

$$Z=\frac{(\hat{p}_1-\hat{p}_2)-0}{\sqrt{\hat{p}\hat{q}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$$

$$=\frac{0.656-0.373}{\sqrt{0.4851(0.5149)\left(\frac{1}{93}+\frac{1}{124}\right)}}$$

$$=\frac{0.283}{\sqrt{0.2498\,(0.010752+0.008064)}}$$

$$=\frac{0.283}{\sqrt{0.0047}}$$

$$=\frac{0.283}{0.0686}=4.1254$$

第六步　結論：所計算之Z值爲4.13 >1.96，在棄却區域以內，應摒棄虛無假設（即推翻假設），可斷定未接受預防注射之致死率，大於接受預防注射者之致死率。亦即是二者之差異非常顯著，並非由於抽樣之機遇而來。

例二　吾人隨機抽出某種疾病患者450人，任意分成二組，其中第一組150人，以A法治療，經過一段期間後，治癒者爲135人，死亡者15人。第二組爲300人，以B法治療，同樣經過相等之一段期間後，治癒者爲255人，死亡者爲45人。亦即前者（第一組）用A法治療之死亡率爲 $15/150\times100=10\%$，用B法治療之死亡率爲 $45/300\times100=15\%$，從百分數看，雖然前者死亡率較後者低5％，吾人是否認爲A法治療較B法有效？

吾人知道生物現象的變異很大，不能僅憑數字就下結論，吾人必須知道兩組之死亡率，到底是由於眞正相差？還是由於機遇而來？現在將上述治療法列於下表以便用卡平方 x^2（讀做 Chi-Square ）之方法分析之。

療法	死亡		生存		合計
	觀察值	期望值	觀察值	期望值	
A	15	20	135	130	150
B	45	40	255	260	300
合計	60	60	390	390	450

註：1.表中觀察值之合計爲60，其期望值合計與此相等亦應爲60。

2.生存觀察值之合計爲390，其期望值合計亦應爲390。

上表爲便於比較，吾人必須計算兩組合併之死亡率，在450人共死亡60人佔 $60/450\times100=13.3333\%$，按此百分數計算，則

A法在150人中，其期望數應有 $0.133333\times150=20$ 人死亡，其生存期望數 $=150-20=130$ 人。

B法在300人中，其期望數應有 $0.133333\times300=40$ 人死亡，其生存

期望數＝300-40＝260人。

於是觀察值與期望值各有四個，依照x^2之公式為 $X^2=\sum\left[\frac{(觀察數)-(期望數))^2}{期望數}\right]$

$$=\frac{(15-20)^2}{20}+\frac{(135-130)^2}{130}+\frac{(255-260)^2}{260}$$

$$=5^2\left[\frac{1}{20}+\frac{1}{130}+\frac{1}{40}+\frac{1}{260}\right]=2.16$$

在上列計算中，有一個虛無假設（見例一解釋），即假定一、二兩組都從平均死亡佔13.3333％的全體中隨機抽出，如果吾人從此全體中抽出許多個含有150人和300人的樣本，並且各分別求其x^2值，由於機遇的關係，吾人可以預料有的樣本靠近全體平均數，其x^2值自然小；有的樣本離全體平均數遠，其x^2值大。惟究竟大到如何程度才能表示彼此有顯著之差別，這點則可以查x^2值表就可以決定。表中v或df表示自由度，上端橫行標有α或p值者之下的數目為機率，表之本身為x^2值。例如當自由度為1時，相當於機率0.05的x^2值為3.841，亦即在100隨機抽樣中能夠得到x^2值大於或等於3.841的則有5次，如所計算的x^2值小於5％，則為不顯著。

上表又稱為四格表，所謂自由度即當四格表之橫行合計為150及300，與縱列合計60與390以及450等五數固定後，則四個方格內只有一數可以自由填寫，其餘三格內，就由已知的一數與合計相減而得，故其自由度為1，如表內，當死亡期望值20一值填入後，其餘期望值為60-20＝40，150-20＝130 以及300-40＝260等三數立即相減而得。

自由度1決定後，就可以推測估計了。由於自由度為1時，x^2的5％，為3.841。而吾人計算之x^2＝2.16（見第七頁之計算x^2值），較3.841為小，即計算之x^2值2.16＜3.841，在棄却區域以外並不顯著，亦即用A法治療與用B法治療之結果，毫無差別。

五、結論

人類為生物之一，故人口之出生率、死亡率、數量之自然增加以及素質的改進等，雖然受生物之進化與遺傳之影響，但這些都與公共衛生、家庭計劃有密切之關係，由於近半世紀來，公共衛生之推展，極為迅速，所為之範圍又不斷之擴大，由受孕、出生到死亡以及疾病等無一不與公共衛生有關，又無一不與生物統計尤其是生命統計發生關係。因為出生數與死亡數之增減，不僅可以測知一國人口數量增殖趨勢，壽命之長短，更可以測知衛生、醫療、保健等事業之是否發達與進步。

吾人研究治療某種疾病之方法，而將患者分成A、B兩組並儘可能使二組患者之環境完全一致，然後對A組給予特殊之治療法，而對B組施予普通之治療。經過一段時間後，再比較A、B兩法之治癒率，即可斷定治療之效果這就需要生物統計之知識為參考，由此可知生物統計對公共衛生與一般醫學之重要了。

最後用世界衛生組織衛生統計學家，Dr. Satya Swaroop 所著衛生統計學中一段話作為結束：「據有經驗的公共衛生行政專家描述，衛生統計學是公共衛生工作者之眼睛與耳朵，確是事實近於眞理，如果缺少了統計的資料，對他的職責，無所適從，便是既盲且聾了」。