包括咽炎、喉炎、會厭炎。

具有下列條件任一項者：

*㈠有發燒、咽部發紅、喉嚨痛、咳嗽、聲音沙啞、喉部有膿樣滲出液等臨床症狀任兩項，且有下列條件任一項者：

**1. 咽、喉、會厭等部位經培養分離出微生物者。

**2. 血液培養分離出微生物者。

***3. 血液或呼吸道分泌物測得有陽性之抗原反應者。

***4. 血清學檢查測得陽性 IgM 抗體或者四倍效價上升之 IgG 抗體。

*5. 醫生之診斷。

*㈡經由直接視檢，或手術中或以病理組織切片檢查發現膿瘍者。

*㈢一歲以下之嬰兒，具有發燒、體溫過低、呼吸中止、心跳徐緩、流鼻水、喉部有膿樣滲出液等臨床症狀任兩項，且有下列條件任一項者：

**1. 咽、喉、會厭等部位經培養分離出微生物者。

**2. 血液培養分離出微生物者。

***3. 血液或呼吸道分泌物測得有陽性抗原反應者。

***4. 血清學檢查測得陽性 IgM 抗體或者四倍效價上升之 IgG 抗體。

*5. 醫生之診斷。

流行病學暨生物統計專欄㈢

簡單直線迴歸分析

林明瀅

-----台北榮民總醫院感染管制委員會------

當我們翻到報紙的財經版時，常可看到類似的報導，預測我國今年的經濟成長率為百分之八，某某公司的股票會上漲百分之二十，諸如此類的報導，讀者是否注意到這些財經專家是以什麼依據來下此結論呢？其實他們就是將可能的變項套入統計學上的「迴歸分析」的分析模式，求出這些答案。迴歸分析是為連續性資料與連續性資料間的統計分析方法，而可依其變項的多寡予以區分為單一變項迴歸分析及多變項迴歸分析；例如在探討，身高與體重兩者間是否有關係時，我們不可用卡方檢定或 t 檢定來分析，因為這兩種方法的分析結果並不適當，唯有利用迴歸分析才可得到適當的結論。本期將介紹簡單直線迴歸分析，可探究二個等距／比率變項間的因果關係，且希望以一變項（自變項），來預測或解釋另一變項（依變項）的變化，即自變項每增加一個測量單位時，依變項應增加或減少若

干測量單位的數值。

在進行迴歸分析之前，一般需先將全部個體的二個變項值逐一的畫在 XY 座標上，看其分佈情形，此圖稱之為相關分散圖，為了解二個連續性變項間關係的最簡單的方法，請見本通訊的第四卷第一期。我們都知道，二變項呈直線關係可用一次方的方程式來表示如 $Y=a+bX$，若呈拋物線則是二次方程式，$Y=cX^2+bX+a$。故由分散圖的分佈情形，可以讓我們了解這二個變項間，是否成直線關係或成拋物線關係或是其他多次方的線性關係。簡單直線迴歸分析顧名思義就是探討自變項與依變項之間是否有直線關係的存在；一般臨床上的資料，其二變項間的關係不太可能完全成為一直線，直線迴歸分析就是要找出某一條線能夠代替所有的點，即是 $Y=a+bX+\varepsilon$（ε 為誤差值），數理學家發現當 ε 值最小，此方程式的代表性最好，要讓 ε 最小，以各點與此直線之平行於 Y 軸的距離平方和最小，故此線稱之為「最小平方值迴歸直線」。所以在進行簡單直線迴歸分析，事實上是決定斜率 b 與截距 a 的過程，使各點至此線之平行於 Y 軸之距離之平方和成為最小，以求出根據 X 變項來預測 Y 變項時之直線迴歸方程式，利用此方程式來預知 Y 變項的數值，由於計算的方式較以往的統計方法複雜，借助電腦的幫忙可減少很多時間及計算錯誤，因此本文著重在觀念的介紹。

簡單直線迴歸分析

(一)適用條件：二個變項都是連續性變項

(二)基本假設：直線迴歸方程式之斜率 (b) 是否為 0？

(三)檢定程序：1. 檢定直線迴歸方程式之斜率 (b) 是否為 0？

2. 檢查殘差圖？

(四)檢定統計量：

截距 $a=Y-bX$

$$斜率\ b=\frac{\Sigma(X-\bar{X})\Sigma(Y-\bar{Y})}{\Sigma(X-\bar{X})^2}$$

$$=\frac{\Sigma XY-[(\Sigma X)(\Sigma Y)/n]}{\Sigma X^2-(\Sigma X)^2/n}$$

$$=\frac{\Sigma XY-\bar{X}\bar{Y}\times n}{\Sigma X^2-\bar{X}^2\times n}$$

檢定值 t=

$$\frac{b-0}{\sqrt{\Sigma(Y-\bar{Y})^2-[b^2\Sigma(X-\bar{X})^2/n-2]}\times\sqrt{1/\Sigma(X-Y)^2}}$$

(五)範例：研究者收集三十位患者的血中細菌含量與白血球數目探討二者之間是否有直線相關，其結果如下：（見表一）

1. 求出迴歸方程式的斜率及截距

$n=30$, $\bar{X}=969333$, $\bar{Y}=19810$,

$\Sigma X=2.908\times10^8$, $\Sigma Y=5.943\times10^6$

$\Sigma XY=7.03658\times10^{11}$,

$\Sigma X^2=4.02784\times10^{15}$,

$\Sigma Y^2=1.344053\times10^{10}$

$$斜率\ b=\frac{\Sigma XY-(\Sigma X)(\Sigma Y)/n}{\Sigma X^2-(\Sigma X)^2/n}$$

$$=0.001055$$

截距 $a=Y-bX$，將上述 $\bar{X}$, $\bar{Y}$, b 三個數值代入，即可得知 $a=9580.98$。

表　一

患者編號	細菌量	白血球
1	3.4×10^6	1.25×10^4
2	1.3×10^6	1.36×10^4
3	7.8×10^6	1.78×10^4
4	1.45×10^7	3.22×10^4
5	8.7×10^6	1.96×10^4
6	4.6×10^6	1.48×10^4
7	4.4×10^6	1.41×10^4
8	2.13×10^7	2.60×10^4
9	1.82×10^7	2.59×10^4
10	7.8×10^6	1.88×10^4
11	2.9×10^6	1.38×10^4
12	2.24×10^7	2.95×10^4
13	1.17×10^7	2.41×10^4
14	5.4×10^6	1.32×10^4
15	7.8×10^6	1.57×10^4
16	4.3×10^6	1.26×10^4
17	9.6×10^6	2.05×10^4
18	1.44×10^7	2.58×10^4
19	6.4×10^6	1.71×10^4
20	4.3×10^6	1.28×10^4
21	5.9×10^6	1.13×10^4
22	7.4×10^6	1.93×10^4
23	1.7×10^6	1.27×10^4
24	2.14×10^7	4.22×10^4
25	4.8×10^6	1.26×10^4
26	9.3×10^6	1.70×10^4
27	1.16×10^7	2.56×10^4
28	1.54×10^7	2.73×10^4
29	2.38×10^7	3.12×10^4
30	8.3×10^6	1.47×10^4

註：假設資料，結果由統計軟體 SAS 計算

直線迴歸方程式血中白血球 =0.001055x 血中細菌量 +9580.98 。

2. 找出最大值及最小值

在此範例中，血中細菌量之最小值為 1700000 ，最大值為 22400000 ，我們可依直線迴歸方程式在最大值及最小值之間，任選二個數值代入求出的方程式中，算出白血球的數值，將此二點在座標上連成一直線，此線即為迴歸直線。（見圖一）

3. 迴歸係數的檢定

$H_0:\beta=0$ 白球數量與血中細菌量無關直線關係

檢定值 t=

$$\frac{b-0}{\sqrt{\Sigma(Y-\bar{Y})^2-[b^2\Sigma(X-\bar{X})^2/n-2]}\times\sqrt{1/\Sigma(X-Y)^2}}$$

$=10.84$

查表七 $t_{28(0.001)}=2.46$

4. 相關係數

$$r=\frac{n\Sigma XY-\Sigma X\Sigma Y}{\sqrt{[\Sigma X^2-(\Sigma X)^2][\Sigma Y^2-(\Sigma Y)^2]}}$$

$=0.8986$

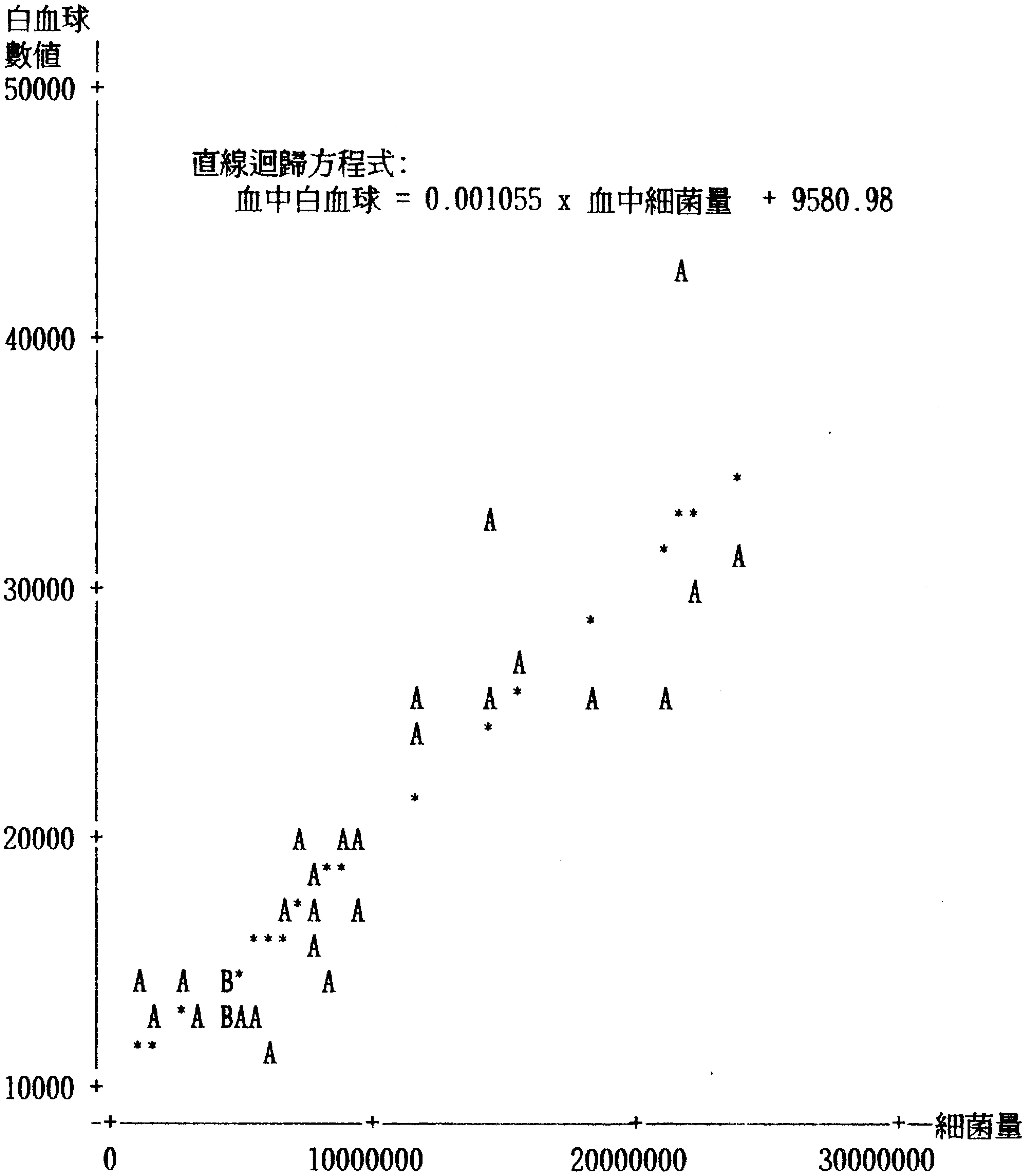
5. 判定係數 (r^2)

判定係數為相關係數的平方，判定係數愈大，表示迴歸直線愈好，預測能力愈高，其值介於 0 與 1 之間，是用來判定迴歸直線好壞的指標，所代表的是 Y 的總變異量 $\Sigma(Y-\bar{Y})^2$ 中有多少百分比可由 X 值的變化所解釋。

㈥決策原則

在範例中統計量 t 值 10.84>2.46 拒絕接受虛無假設，即是斜率 b 不等於 0 ，所以血中細菌濃度與白血球之間有直線關係存在，而 $r^2=0.807425$ ，可知此方程式能有 80.74% 正確性由血中細菌量來預估血中白血球的數量。由依變項的觀察值與預期值的差值稱為殘差值，各個殘差值在 XY 座標上亦可構成殘差圖，檢視殘差圖（圖二），亦可幫助我們判定是否有異常趨勢。若研究者對 80.74% 的預估能力不滿意時，當然也可以用二次方程式迴歸來預估，經 SAS 統計軟體的計算後，其判定係數只有 73.33% ，正確的

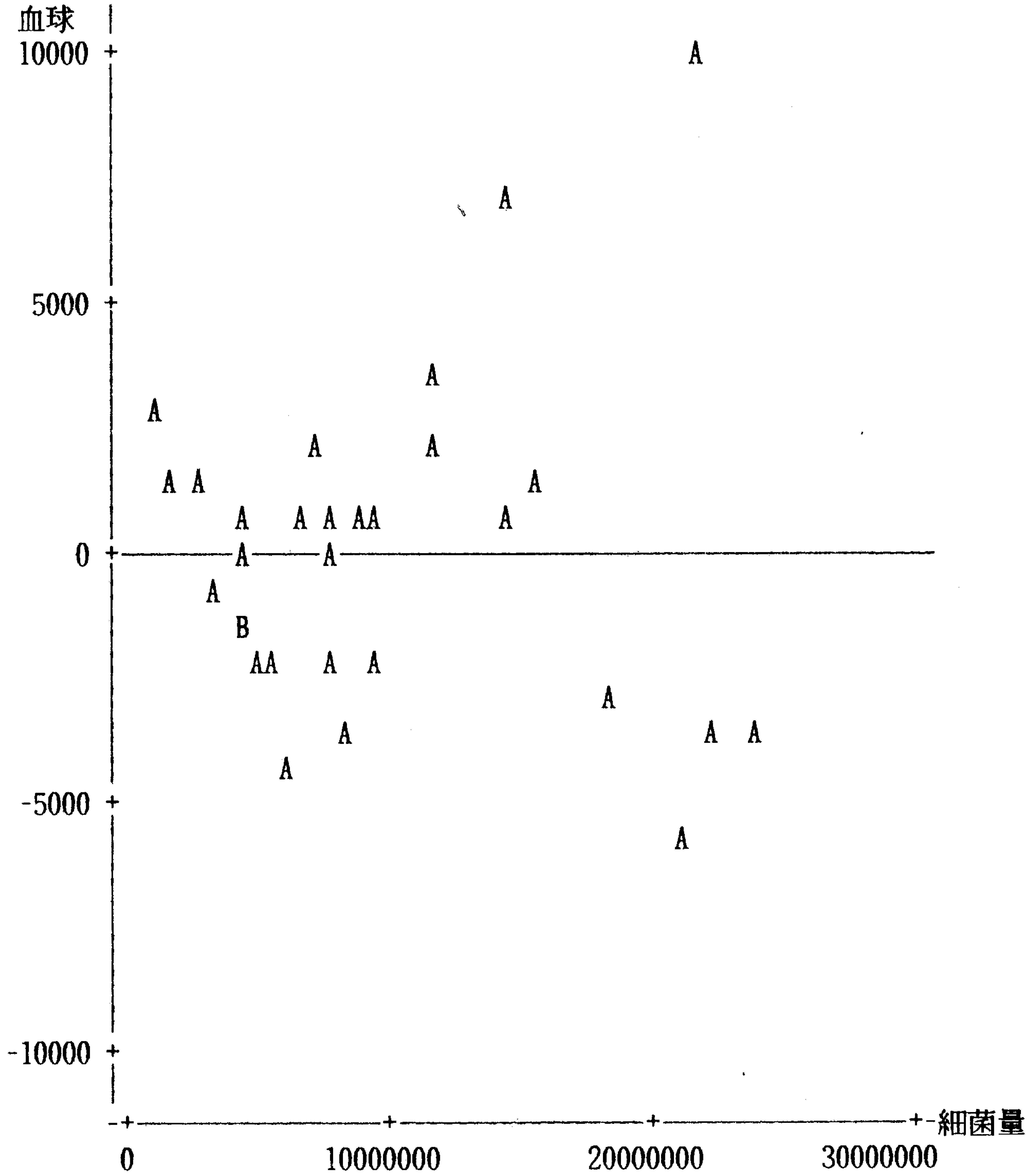
圖一　資料分佈圖及迴歸線



備註: A,B 所代表的爲相關分散圖, A = 1 obs, B = 2 obs.
* 所連結的線爲直線迴歸線

預估能力反而降低了，因此本範例不適宜用拋物線來代表。由上述的分析，可以知道有 19.26% 的其它因素會影響血中白血球的數值，這時我們便可應用多變數迴歸分析來找出最適當的代表方程式，如方程式 $Y=aX1+bX2+cX3+\cdots+\varepsilon$，其中 X2, X3 即是其它因素，例如性別、年齡、細菌種類、服用的藥物、病人的潛在疾病，臨床症狀的嚴重度等等，這些因素，研究者可依研究的目的及所收集的資料予以一一測試，來找出最具代表性的迴歸方程式，使我們的預測能力提高。

圖二 殘差圖



上期勘誤（第四卷第一期）：

腦神經外科加護病房 Serratia marcescens
院內感染群突發調查之作者

黃正華→ 黃政華
黃美貞→ 黃美真