

研究設計之樣本數決定

林正祥

在規劃研究設計時，研究人員往往忽視樣本多寡的問題，因而造成統計推論上的嚴重偏差。忽視樣本的結果，將使得：(1)由於樣本數不足，雖已達到臨床上的顯著差異，卻無法達到統計上的顯著差異；(2)因樣本過多，雖研究結果沒有臨床上的意義，卻達到統計上的顯著差異及(3)由於樣本有限，雖然研究結果皆達臨床及統計上的意義，但檢力過低無法肯定其對立假說是否成立。本文提出了在單一及變常態母體平均值檢定、二項分佈檢定、邏輯斯迴歸分析、變方分析世代死亡率之樣本計算公式，並援引數例說明之，俾提醒研究人員能藉此瞭解樣本數之重要性並妥為運用之。(中華衛誌 1993；12(3)：282-290)

Key words: Sample size determination

及第Ⅱ型的誤差，吾等可以下表表示之：

Ⅰ、前言：

在規劃一研究設計進行統計分析前，樣本數之決定對研究之品質關係至鉅，惟此一問題往往被忽視。有些研究在完成後才發覺樣本數不足造成檢力過低，使得研究結果毫無意義；此外有些實驗則採用了過量的樣本雖然大大提高了檢力但確發現在臨床上毫無意義的差距竟然具有統計上顯著的差異。由於在事先規劃研究時沒有處理好樣本的數目，其間所浪費的人力、物力實難以想像。本研究旨在探討樣本數的多寡對研究品質的影響，並提供一些研究決定樣本數的公式，俾能使研究人員藉此了解樣本數的重要性並合理的運用之。

決策 結果	真 實 情 況	
	H_0	H_a
H_0	$1 - \alpha$	β (第Ⅱ型誤差)
H_a	α (第Ⅰ型誤差)	$1 - \beta$

H_0 ：虛擬假說

H_a ：對立假說

Ⅱ、檢力的定義：

在進行假設檢定時，會產生所謂第Ⅰ型

東海大學應用統計研究所

聯絡人：林正祥

聯絡地址：台中市東海大學統計學系

所謂檢力(power)即是當虛擬假說為假時拒絕其之機率，即是：

檢力 $= 1 - \beta = \Pr(\text{拒絕 } H_0 \mid H_a)$

換句話說，檢力乃是肯定對立假說成立之機率

我們可以利用圖一求出在單樣本檢定時之檢力

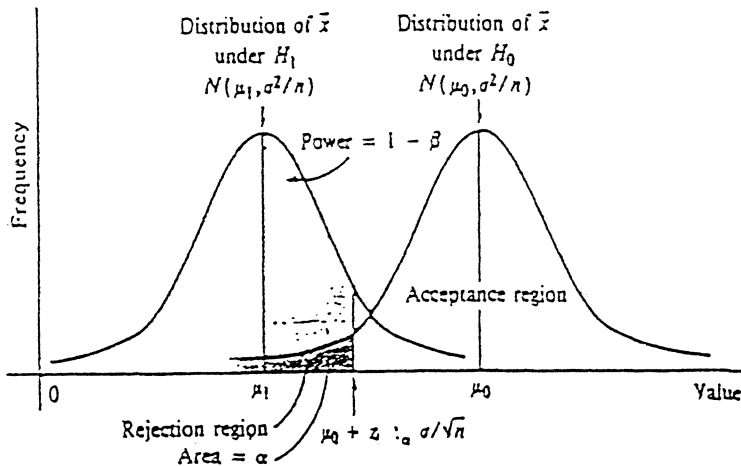


圖 1. 單樣本檢力之檢定(Rosner [1])

(1)單尾檢定

$$H_0: \mu = \mu_0, H_a: \mu = \mu_1 < \mu_0$$

$$\Pr \left[Z < Z_{\alpha/2} + \frac{(\mu_1 - \mu_0)\sqrt{n}}{\sigma} \right] \text{ 當 } \mu_1 > \mu_0$$

$$\text{檢力} = \Pr \left(x < \mu_0 + Z_{\alpha} \frac{\sigma}{n} \mid \mu = \mu_1 \right)$$

$$= \Pr \left[\frac{x - \mu_1}{\sigma / \sqrt{n}} < \frac{\mu_0 - \mu_1}{\sigma / \sqrt{n}} + Z_{\alpha} \right]$$

$$= \Pr \left[Z < Z_{\alpha} + \frac{(\mu_0 - \mu_1)\sqrt{n}}{\sigma} \right]$$

$$\text{若 } \mu_1 > \mu_0 \text{ 則檢力為 } \Pr \left[Z < Z_{\alpha} + \frac{(\mu_1 - \mu_0)\sqrt{n}}{\sigma} \right]$$

檢力會受到以下幾項因素所影響：

- (i) 若 $\alpha \downarrow$ 則 $Z_{\alpha} \downarrow$ 則檢力下降。
- (ii) 若 $|\mu_0 - \mu| \uparrow$ 則檢力上升。
- (iii) 若 $\sigma \uparrow$ 則檢力下降。
- (iv) 若樣本數 $n \uparrow$ 則檢力上升。

(2)雙尾檢定

$$H_0: \mu = \mu_0, H_a: \mu = \mu_1 \neq \mu_0$$

雙尾檢定和單尾檢定之檢力求法類似，惟稍複雜而已，其檢力為

$$\Pr \left[Z < Z_{\alpha/2} + \frac{(\mu_0 - \mu_1)\sqrt{n}}{\sigma} \right] \text{ 當 } \mu_1 < \mu_0$$

III、對假設檢定之誤解：

一般人在進行假設檢定時，往往只考慮到去拒絕虛擬假設，而未能考慮在拒絕虛擬假設的同時，他有多少信念肯定對立假設是正確的，因此拒絕虛擬假設並不意謂其結果就是合理的，因有可能檢力過低造成對立假設沒有意義，一般而言檢力最好是在90%以上，若由於某些特殊條件的限制(如樣本數之多寡，平均值間的差異)，至少也要有70%方可。

IV、統計上的顯著差異與臨床上的顯著差異：

樣本數的決定對研究計畫的成敗攸關至鉅，主要乃基於研究人員希望研究結果能顯示出統計及臨床上的顯著意義，許多研究由於(i)樣本數不夠，雖已達臨床上的顯著差異，卻無法達到統計上的顯著差異。(ii)樣本過多，雖然研究結果沒有臨床上的意義，卻能達到統計上的顯著差異及(iii)樣本有限，雖然研究結果皆達臨床及統計上的意

義，但檢力過低無法肯定其對立之假設是否成立。此三種情況的發生將造成推論上的偏差，然而這類問題卻不斷地重覆發生在許多的研究中，茲援引數例說明之：

例一：有一研究在探討口服避孕丸是否會造成血壓的上升，現將35-39歲的婦女中就(A)服用口服避孕丸者及(B)未服用者二組中分別隨機抽取8人及21人，其中(A)組之平均收縮壓為135.86 mmHg, $S_a=15.34$; (B)組之平均收縮壓為125.02 mmHg, $S_b=18.23$ 問(A)、(B)之間是否有顯著的差異？

$$H_0: \mu_a = \mu_b, H_a: \mu_a > \mu_b$$

假設(A)、(B)二組之變方相等，則吾等可用2-sample t檢定：

$$S_p^2 = \frac{7(15.34)^2 + 20(18.23)^2}{27} = 307.18$$

$$S_p = \sqrt{S_p^2} = 17.527$$

$$t = \frac{135.86 - 125.02}{17.527 \sqrt{1/8 + 1/21}} = 1.5 < t_{27,0.95} = 1.703$$

$$\text{令 } \alpha = 0.05$$

二者皆無顯著差異。

若臨床上專業人員認為血壓有8 mmHg以上之差距即是有顯著差異，則此一結果顯示(A)、(B)二組之有臨床上的顯著差異10.84 mmHg，但卻無統計上的顯著差異，在此我們可以將其檢力算出[2]：

$$\begin{aligned} \text{檢力} &= \Pr(t < t_{27,\alpha} + \frac{n_1(\bar{x}_1 - \bar{x}_2)}{\sqrt{S_1^2 + S_2^2/(n_2/n_1)}}) \\ &= \Pr(t < -1.703 + \frac{8(10.84)}{\sqrt{(15.34)^2 + (18.23)^2/(21/8)}}) \end{aligned}$$

$$= \Pr(t < -0.897) \quad (\text{d.f.} = 27)$$

$$= 0.14$$

$$(\alpha = 0.05)$$

由此可看出其第二型錯誤為0.86 (= 1 - 0.14)，換句話說我們接受 H_0 可能會有問題， H_0 為假被接受的機會高達86%。

此例之樣本數不足，故雖已達臨床上的顯著差異，卻無法達到統計上之顯著差異，其第二型錯誤之機率高達86%乃不足為奇。

例二：若例一的情況為

(A)平均收縮壓130.2 mmHg $S_a=15.34$ $n_a=2000$

(B)平均收縮壓129 mmHg $S_b=18.23$ $n_b=2000$

$$S_p^2 = \frac{(15.34)^2 + (18.23)^2}{2} = 283.82$$

$$\therefore S_p = 16.85$$

$$t = \frac{130.2 - 129}{16.85 \sqrt{1/2000 + 1/2000}} = 2.25 > t_{\infty,0.95} = 1.645$$

此例(A)、(B)之血壓僅差1.2 mmHg，臨床上並無顯著差異，惟統計上有顯著差異。其檢力幾近100%，此即意謂當樣本數過大，臨床上毫無意義的差距亦有可能成為統計上顯著的差異。

例三：若將例一中之 S_a 及 S_b 改變則可得到以下結果：

$$\begin{aligned} < I > S_a^2 = (15.34)^2 & < II > S_a^2 = S_b^2 = (20)^2 \\ S_b^2 = (18.23)^2 & x_1 - x_2 = 10.84 \\ x_1 - x_2 = 10.84 & \end{aligned}$$

樣本數 ($n_a=n_b$) (單尾)	t值	檢力	樣本數 ($n_a=n_b$) (單尾)	t值	檢力
40	2.76**	0.89	40	2.42**	0.78
30	2.49**	0.81	30	2.01**	0.67
25	2.18*	0.74	25	1.92*	0.61
20	1.96*	0.65			

$$* P < 0.05 \quad ** P < 0.01$$

由此不難看出< I >、< II >兩組同時具有臨床上及統計上的顯著差異，但其檢力在< II >組之變方為(20)²及樣本數30以下時降至0.7以下，將無法肯定對立假設是否正確。這已說明許多人認為在做雙樣本t檢定時只要各組樣本維持在15左右即可的說法是很有問題的，主要係其對如何決定樣本之觀念不清，造成以訛傳訛的後果。進行雙樣本t檢定時由於樣本受到二組平均值差異，變方及檢力的影響，有許多的條件限制，斷不可用所謂的一般概念或經驗來決定之。有關臨床上及統計上差異的觀念可參閱Mark [3]，及Freiman [4]等之著作或研究。

V、樣本數之決定：

當抽樣目的是對母體作估計時，則樣本數決定的原則需考慮母體的變異、抽樣誤差及信賴區間係數；然若抽樣目的是在檢定虛無假設時，樣本數決定之原則除需考慮以上的因素外，尚需考慮檢力。現將各種較為普遍統計方法樣本決定之公式說明如下，並舉例說明之：

- (a) 估計：假設研究人員希望有 $1-\alpha$ 之信念肯定其抽樣誤差在 $\pm C$ 間或 $\pm \theta$ 間，則其樣本數為：

$$Z_{1-\alpha/2} \frac{\sigma}{n} = c \rightarrow n = \frac{\sigma^2}{c^2} Z_{1-\alpha/2}^2 \quad (\text{變數為常態分佈})$$

$$Z_{1-\alpha/2} \frac{pq}{n} = \theta \rightarrow n = \frac{pq}{\theta^2} Z_{1-\alpha/2}^2$$

(變數為二項分佈， θ 為誤差百分比)

例四：做某心理測驗時，給予受測者一種刺激作用，結果可能產生反應A或B。有位心理學家想估計受測者產生A反應之比率P。他預期P在0.6左右，並假設他希望估計誤差小於0.04之機率

等於0.95，試問受測之人數應為若干？

$$\begin{aligned} n &= \frac{pq}{\theta^2} Z_{1-\alpha/2}^2 \\ &= \frac{(0.6)(0.4)}{(0.04)^2} (1.96)^2 \\ &= 576 \end{aligned}$$

故需要有576人參加受測

若不知道 $p=0.6$ 則必須用 $p=0.5$ 以求得 n 之極大值。

- (b) 單一常態母體平均值檢定：

$$n = \frac{\sigma^2 (Z_{1-\beta} + Z_{1-\alpha/2})^2}{(\mu_0 - \mu_1)^2} \quad (\text{雙尾})$$

$$n = \frac{\sigma^2 (Z_{1-\beta} + Z_{1-\alpha/2})^2}{(\mu_0 - \mu_1)^2} \quad (\text{單尾})$$

- (c) 兩常態母體平均值檢定：

$$n = \frac{(\sigma_1^2 + \sigma_2^2) (Z_{1-\alpha/2} + Z_{1-\beta})^2}{(\mu_2 - \mu_1)^2} \quad (\text{兩組樣本相等時})$$

$$\begin{cases} n_1 = \frac{(\sigma_1^2 + \sigma_2^2/k) (Z_{1-\alpha/2} + Z_{1-\beta})^2}{(\mu_2 - \mu_1)^2} \\ n_2 = \frac{(k\sigma_1^2 + \sigma_2^2) (Z_{1-\alpha/2} + Z_{1-\beta})^2}{(\mu_2 - \mu_1)^2} \end{cases}$$

(兩組樣本不等時)

$$k = n_2/n_1$$

例五：有人懷疑服用口服避孕丸者可能會罹患高血壓，現擬進行一實驗，探討是否口服避孕丸會引起高血壓，吾等針

對30-35歲之婦女為研究對象，假設服用口服避孕丸者和未服者血壓均為常態分佈。故其檢力為 $H_0: \mu_1 = \mu_2$, $H_a: \mu_1 \neq \mu_2$ ，問在 $\alpha = 0.05$ 之情況下需要多少樣本方足以使其檢力達到0.8？

假設臨床上之認定 $\Delta = \mu_1 - \mu_2 \geq 10$ ，即意味血壓有顯著之差異，且由於 σ_1^2 及 σ_2^2 無法得知，故乃進行一小型試驗得到

$$\bar{x}_1 = 130 \quad s_1 = 16 \quad \bar{x}_2 = 140 \quad s_2 = 19$$

$$\text{則 } n = \frac{(16^2 + 19^2) (1.96 + 0.84)^2}{10^2} = 49$$

每一組需要49人

(d) 二項分佈之單樣本及雙樣本檢定

單樣本：

$$n = \frac{p_0 q_0 (Z_{1-\alpha} + Z_{1-\beta} \sqrt{p_1 q_1 / p_0 q_0})^2}{(p_1 - p_0)^2}$$

雙樣本：

$$n_1 = \lceil \sqrt{\bar{p}\bar{q}} / (1 + 1/k) Z_{1-\alpha} + \sqrt{p_1 q_1 + (p_2 q_2 / k)} Z_{1-\beta} \rceil^2 / \Delta^2$$

$$n_2 = k n_1$$

p_1, p_2 = 二群體之發生率(or死亡率)

$$\Delta = p_1 - p_2$$

$$q_1, q_2 = 1 - p_1, 1 - p_2$$

$$\bar{p} = \frac{p_1 + k p_2}{1 + k}$$

$$\bar{q} = 1 - \bar{p}$$

雙尾：單、雙樣本 n_1 公式中之 $Z_{1-\alpha}$ 以 $Z_{1-\alpha/2}$ 取代之。

例六：欲研究維他命A是否有預防乳癌之效果，將45-50歲之婦女分二組，一為控制組只服用安慰劑，另一為實驗組

服用維他命A，且希望能減少乳癌罹患率20%，現假設某年45-50歲婦女之乳癌發生率為150/100,000，問需要多少樣本方足以達到 $\alpha = 0.05$ (單尾)及檢力為0.8之需求？(假設控制組與實驗組之樣本比為1:1)

$$p_1 = \frac{150}{100000} = 0.0015$$

$$q_1 = 1 - p_1 = 0.9985$$

$$p_2 = (0.0015) (1 - 0.2) = 0.0012$$

$$q_2 = 0.9988$$

$$\bar{p} = \frac{0.0015 + 0.0012}{2} = 0.00135$$

$$\bar{q} = 0.99865$$

$$n_1 = 185043$$

每組各需185043個婦女參與實驗。

(e) 迴歸分析(regression analysis)

目前並無關於迴歸分析之樣本之計算公式，根據Neter et al. [3]指出通常樣本數約為自變項總數的6到10倍。Kleinbaum et al. [4]亦同意此一看法，但其亦指出在迴歸分析中誤差項之自由度至少得大於10乃最低的基本需求。顯然二者之理念係基於實證經驗而來，此種理念並無客觀的理論來支持它，惟目前在對決定迴歸分析樣本所知有限的情況下，吾等不得不接受樣本數為自變數總和的6至10倍。

(f) 變方分析(ANOVA)

欲決定變方分析之樣本，首先得了解F檢定之檢力

$$\text{檢力} = \Pr \{ F_{1-\alpha, r-1, n-r} \mid \Phi \}$$

Φ 為離中性參數(noncentral parameter)，係用以衡量各因子之水準平均值 μ_i 之差異程度



$$\Phi = \frac{1}{\sigma} \sum n_i (\mu_i - \mu)^2 / r$$

$$\mu = \frac{\sum n_i \mu_i}{r_i}$$

當所有因子各水準之樣本相等時(n)，其參數即為

$$\Phi = \frac{1}{\sigma} n \sum (\mu_i - \mu)^2 / r$$

$$\mu = \frac{\sum \mu_i}{r}$$

欲計算出檢力之機率，得算出離中性F分佈，此為對立假設成立時之分佈。根據此並依以下條件吾等可以推算出變方分析中各因子之水準所需要之樣本數。

① $\Delta = \max(\mu_i) - \min(\mu_i)$

② 決定 α

③ 定出 Δ / σ 之比值

④ 決定 β

Neter et al. [7]之書中列有在不同 α ， Δ / σ 及 $1 - \beta$ 之情形下不同因子之水準所需之樣本數。

例七：假設有四種廠牌所需的車胎之性能，且吾等認定最好和最差之車胎平均壽命大約為3 (千哩)，且根據過去的經驗知道車胎的標準差約為2 (千哩)。假設 $\alpha = 0.05$ ，檢力為0.9時各種廠牌需要多少樣車來進行實驗方足以達到 $\alpha = 0.05$ 且檢力為0.9之效果？

$$\text{此題之 } \Delta = \max(\mu_i) - \min(\mu_i) = 3$$

$$\sigma = 2$$

$$\frac{\Delta}{\sigma} = 1.5$$

查表可以得到 $n=14$

即每一廠牌需要14個車胎來進行實驗

(g) 邏輯斯迴歸分析(logistic regression analysis)

Whittemore [8]曾估算出邏輯斯迴歸分析所需之樣本數：

令P為控制組之發生率或死亡率

而N為世代樣本數，則

(i) 若自變項為類別變項時

$$N = \{[Z_{1-\alpha} F(0)^{1/2} + Z_{1-\beta} F(b)^{1/2}]^2 / b^2\} / P$$

$$\text{其中 } F(\mu) = (1 - P')^{-1} + (P' e^{\mu})^{-1}$$

P' 為危險群體(實驗組)之發生率或死亡率

b 為邏輯斯迴歸係數(e^b 為實驗組對控制組之odds ratio)

(ii) 若自變項為常態變數時

$$N = \{[Z_{1-\alpha} F(0)^{1/2} + Z_{1-\beta} F(b)^{1/2}]^2 / b^2\} / P$$

$$\text{其中 } F(\mu) = \exp(-\mu^2/2)$$

其中b為標準化邏輯斯迴歸係數

(iii) 自變項為多變量常態分佈

令自變項 x_1 為對所欲探討影響因變項之因子， x_2, \dots, x_k 為共變項。假設 $\mathbf{B}_0' = (0, b_2, \dots, b_k)$ 及 $\mathbf{B}_1' = (b, b_2, \dots, b_k)$ 為標準化迴歸係數，且b為吾等所欲探討之變項 x_1 之係數，則

$$N = [Z_{1-\alpha} F(\mathbf{B}_0)^{1/2} + Z_{1-\beta} F(\mathbf{B}_1)^{1/2}]^2 / b^2 P$$

$$\text{其中 } F(\mathbf{u}) = [\exp(-\mathbf{u}'\mathbf{v}\mathbf{u}/2) (1 - R^2_{1,2,\dots,k})]^{-1}$$

其中 \mathbf{v} 為 x_1, \dots, x_k 之相關係數矩陣且 $R^2_{1,2,\dots,k}$ 為 x_1 可被 x_2, \dots, x_k 解釋之判定係數。

(i), (ii), (iii)所求出之 N 為近似值，需以 $N' = N(1+2P)$ 加以修正。 N' 即是所求世代樣本數。

Hsieh [9]根據Whittemore公式算出簡單邏輯斯迴歸分析(自變項僅有一個)所需之樣本數。設定在不同 α 、檢力及訂出(i)在自變項平均值時之致病率(ii)當自變項平均值增加一個標準差時，疾病odds ratio增加為原來之 r 倍的情況下。至於多元邏輯斯迴歸分析之樣本之決定和簡單迴歸分析類似，惟必須訂出(i)在其所有自變項平均值疾病之發生率為 P 及(ii)在其他自變項被控制的情況下特定自變項大於其平均值1個標準差時之疾病的odds ratio增加為原來的 r 倍時即可查表得到樣本值再將此值除以 $1-P$ 此即為所需之樣本值， P 為特定自變項與其他共變項之複相關係數。

(h) 世代死亡率(mortality of cohort study)

根據假設 $SMR (= O/E) \geq \theta$ 是研究人員所欲探討的結果，則

$$\sum_{k=0}^{\infty} [\exp(-E) E^k/k!] \leq \alpha \quad (1)$$

若 $O/E = \theta$ ，在對立假設時 $\theta > 1$ ，則

$$\sum_{k=0}^{O-1} [\exp(-O/\theta) (O/\theta)^k/k!] \geq 1-\alpha \quad (2)$$

令 R_i 為參考母體各年齡層之死亡率

$$\text{則 } \sum_{i=1}^I R_i n_i = E = O/\theta \quad (3)$$

$$\sum_{i=1}^I (\theta R_i) n_i = O \quad (4)$$

在此吾等係假設 SMR 在每一年齡層 i 皆相等

n_i ：第 i 個年齡層暴露的人年數

O ：觀察值 E ：期望值

(4)式之 n_i 取決於(2)式所得之 O

若 R_i 小，則 n_i 必得增大，意謂世代之追蹤時間得須加長。若(3)式之 E 大於5，則吾等可以根據卜瓦松分佈逼近常態分佈之理念，利用標準化常態分佈

$$\frac{(O-E)}{E} = Z_{1-\alpha} \quad \text{算出 } O$$

有關世代死亡率的樣本估算可參閱 Symons and Taulbee [10]。

VI、結語：

- (1) 由於樣本數對研究之良窳具舉足輕重之地位，研究人員在規劃研究時非得小心謹慎不可。惟研究設計之樣本決定係一相當複雜之過程，因樣本受到 α 、 β ，處理間的差異及變方大小之影響，因此在規劃研究時研究人員得就其專業理念對 α 、 β ，處理間之差異做合理的規定，統計人員的地位純是處於諮詢的性質，他們可根據研究人員的決定算出研究所需的樣本數，而非由專業人員一味地依靠統計學家做盲目的摸索。
- (2) 本文第(V)節中(a)–(d)之樣本公式需要有母體資料如 σ_1 、 σ_2 ，或 P_1 、 P_2 及 $\mu_1 - \mu_2$ 等，惟事實上有很多情況這些資料是得不到的，通常這些資料可從以往的研究中得到，或是抽取一組較小的樣本來估計。此外可用間接的方法來得到樣本數，如單、雙樣本 t 檢定，吾等可根據專業人員之經驗認定 $|\mu_0 - \mu_1|$ 為 σ 之 n 倍(或 $|\mu_1 - \mu_2|$ 為 $\sqrt{\sigma_1^2 + \sigma_2^2}$ 之 n 倍)時 μ_0 和 μ_1 (或 μ_1 和 μ_2)之差異會有臨床上之顯著差異，因而定出 $|\mu_0 - \mu_1| / \sigma$ 及 $|\mu_1 - \mu_2| / \sqrt{\sigma_1^2 + \sigma_2^2}$ 之比值來推算樣本數。
- (3) 第V節之公式係普遍常用統計方法之樣本決定公式，惟迴歸分析就目前所知仍無樣本之決定公式，而是以自變數之6–10倍為樣本數，另外邏輯迴歸分析之樣本公式是否合理仍有待驗證，因利用標

準化迴歸係數來推算樣本並不是很合理。再者，目前對邏輯斯迴歸分析之樣本推算僅及於自變項全為類別或全為連續資料，並未考慮自變項同時具有連續及類別資料之情形，因此Whittemore等之樣本公式推算的實用性可能要大打折扣了。至於世代死亡率之樣本公式牽涉範圍亦廣，故不擬在此作更深入的探討。

- (4) 由於檢力函數在某些分佈不易求得，因此在決定樣本數恐仍有困難，此一方面可能得要有更深一層的探討。
- (5) 儘管某些統計分析之樣本決定有其侷限性，但本文已提供了重要統計分析方法之樣本公式，願研究人員能妥為運用之。

參考文獻：

1. Rosner B. Fundamentals of Biostatistics. 3rd ed. Boston: PWS-KENT. 1990; 208.
2. Rosner B. Fundamentals of Biostatistics. 3rd ed. Boston: PWS-KENT. 1990; 275.
3. Marks RG. Designing a Research Project. Belmont: Wadsworth. 1982; 115-128.
4. Freiman JA. The importance of data, the type II error and sample size in the design and interpretation of the randomized control trail. N. Engl. J. Med. 1978; **299**: 690-94.
5. Neter J, Wasserman WJ and kunter MH. Applied Linear Regression Model. 2nd ed. Homewood: Richard D. Irwin, 1989; 433-35.
6. Kleinbaum DG, Kupper LL. Applied Regression Analysis and Other Multivariable Methods. 2nd ed. Boston: Duxbury, 1988.
7. Neter J, Wasserman WJ and kunter MH. Applied Linear Statistical Models. 3rd ed. Homewood: Richard D. Irwin, 1990; 633-42.
8. Whittemore A. Sample size for logistic regression with small response probability. J. Am. Stat. Assoc. 1981; **76**: 27-32.
9. Hsieh FY. Sample size table for logistic regression. Statistics in Medicine. 1989; **8**: 795-802.
10. Symons MJ and Taulbee JD. Statistical evaluation of cancer mortality among industrial populations. In: Cornell, R.G. ed. Statistical Methods for Cancer Studies. New York: Marcel Dekker, 1984; 25-90.

DETERMINATION OF SAMPLE SIZE FOR RESEARCH DESIGN

CHENG-HSIANG LIN

Researchers always ignore the problem of sample size when they are planning a research design. This might lead to the incorrect result of statistical inference. In this paper, we present three examples to describe the problems when sample size are small or too large in research design. Also, we provide the formulas of sample size and some examples for comparing the means of one and two normal distribution, binomial test, logistic regression,

ANOVA and mortality of cohort study to remind researchers paying attention to the importance of sample size for research design. Although there are limitations on the formula of sample size related to some statistical analysis, this paper do provide a guideline for handling the quality of research design by carefully considering the requirement of sample size. (*J Natl Public Health Assoc (ROC)*: 1993; 12(3): 282-290)

Key words: *Sample size determination*

Graduate Institute of Applied Statistics, Tunghai University, Taichung, Taiwan, R.O.C.