國立台灣師範大學教育心理學系
教育心理學報,民70, 14期, 189—204頁

不完整資料對共變數矩陣估計正確性的探討：實徵性研究*

吳 鐵 雄

本研究以電腦模擬資料方式，探討變數取樣各種因素對共變數矩陣估計之影響。這些因素包括：變數數目，變數間平均相關係數、樣本數、變數取樣比率、設計型態等，全部探討的設計組合共有56種。並以概化變異數、軌跡、和最大根做爲探討估計正確性之數量。一般地說，估計偏差相當大，尤以概化變異數爲甚。在所有因素中，以變數間平均相關係數估計影響最大，變數數目則無甚作用，其他如取樣比率、設計型態則結果不甚一致。意外地，樣本數的大小對估計之正確性並無明顯影響。

行爲科學研究者在研究過程中，由於涉及多變數、長期追踪研究、或其他原因，常遭遇漏失資料(missing data)的問題，此種不完整資料的統計分析是一個極難處理的難題，統計學者一直在探究各種可能的解決途徑。本文即以變數取樣方法，利用電腦模擬 (computer simulation) 的方式，產生系統不完整資料，以探討各種因素對共變數矩陣特質估計的影響。

不完整資料處理的研究，一直是一個熱門的課題，尤以近一、二十年更受重視。 Wilks (1932a)可能是第一個從事不完整資料之研究的統計學者，他首先建議以組平均數代替不完整資料，然後以最小平方法(least square method) 來估計母數，只是他只研究雙變數的情形。Matthai (1951)延伸 Wilks 的研究到三變數的情形。Dear (1959)更進一步建議以所有完全資料的總平均 (grand mean) 來代入所有的漏失分數，在研究隨機漏失資料的情形，他主張先以完整資料計算因素分數 (factor scores) 和因素負荷量 (factor loading)，再以它們利用主成份分析 (principal component analysis)獲得"原始分數"。其主要觀念是將資料矩陣分解成已知和未知兩部分，並就已知部分求出主成份解答，來估計未知部分。

Buck (1960) 首先採用一種第一階迴歸方法 (first order regression method) 的或然性評估(probabilistic evaluation)。他發展出一組估計的歷程，此歷程是由所有完整資料的受試者來計算一系列的迴歸函數 (regression function)。每一個變數被用爲依變項，剩下的變數的可能組合做爲自變項，然後再根據適當的迴歸函數估計未知的值。Kosobud (1963) 延伸 Buck 的方法，提出以一對變項的所有完整資料來計算相關係數，而不採取在有一變項資料漏失時即捨棄該受試的方法，他認爲如此可使剩餘變數 (residual variance) 變小。上述兩者，Buck 方法是採取凡是一個受試者在所有變項中有一變項資料漏失，即捨棄不用，而 Kosobud 的方法則採用變項兩兩處理的方法，這兩種方式目前仍廣被採用。

有些統計學者則採用可能性函數 (likelihood function) 方法估計不完整資料的母數 (Parameters)。Wilks (1932a) 以常態雙變項不完整資料樣本，計算出估計的母平均數 (μ) 和母共變數矩陣(Σ) 的最大可能性方程式 (maximum likelihood equation)。當不完整資料是樣本裏的每一個個體沒有接受所有的變項時，Lard (1955a) 考慮一個特例以獲得有效的估計值而估計常態多變數母羣

* 本研究電腦作業承美國紐約州立大學教育心理系 S. David Farr 博士協助

。他自不完整資料導出一個估計常態三變數母數的明確解答（explicit solution）的最大可能性方程式，同時他也導出計算最大可能性估計值的樣本，變數和共變數的方程式。

Trawinski 和 Bargmann (1964) 研究一個比較複雜的漏失資料問題，他們利用系統安排某些受試者不接受某些變數，然後推導計算平均數向量 μ 和共變數矩陣的最大化可能性方程式。他們並寫成以反覆法（iteration）解方程式的電腦程式，並能做假設的統計考驗。Hocking 和 Smith (1968) 則研究一般漏失資料的情形，而不局限於系統的不完整資料，他們利用啓發性歷程 (heuristic procedure) 發展具有最大可能性特徵（properties），但他們只能推出巢狀（nested）情形的估計值，未能普及一般情形。

有些統計學者嘗試着比較利用不同方法代入漏失資料及不同估計母數方法的結果，到目前尚未有定論。Afifi 和 Elashoff (1966, 1967) 評述了不完整資料的研究。他們利用簡單迴歸方法探討四種不同指派數值給漏失資料的方法：(1)以傳統的方式；(2)用平均數；(3)用簡單迴歸；(4)用他們的加權法。他們演算了這四種的分配理論（distribution theory），並評估它們的有效性，他們的結論是沒有一種方法在所有情形下都比其他方法爲佳，一般來說，第二種方法對非常低相關變數的估計最好，第四種方法對低相關較佳，第一種方法則對中等相關變數最好，第三種則有利於高相關變數的估計。Haitovsky (1968) 利用電腦模擬資料研究兩種處理不完整資料方法在迴歸分析的結果，其中一種方法是捨棄所有不完整資料，然後利用最小平方技術分析完整資料。另一種方法是利用所有完整的一對分數計算其共變數，並用這些共變數建立常態方程式系統（system of normal equation)。他製造八組迴歸資料，每一組產生漏失資料的方式都不同，他發現幾乎在所有的情形，用最小平方法分析只包括完整資料（第一種方式）的結果最佳。但是，當不完整資料的比例相當高或當資料之漏失爲非隨機時，則我們必須利用某一種方法去塡補漏失資料。

Timm (1970) 利用模擬資料系統地研究四種不同估計方法對共變數矩陣和相關係數矩陣的估計，他研究的四種方法是 Wilks (1932a), Dear (1959), Buck (1960), 和傳統只用完整資料將有漏失資料的個案捨棄的方法。爲了避免普遍存在於多變項研究中的混淆現象，Timm有系統的改變樣本數 (sample size)、變數數目（number of variable）、漏失資料百分比（percent of missing data）、和變數的平均相關係數（average intercorrelations of variables），而探討上述四種方法的效果。他發現要選擇一個最好的方法估計共變數矩陣比估計相關係數矩陣較少問題。一般來說，每一種方法都有其優點，完全視各種情境的配合，Buck 和 Dear 的方法則稍佳。但是，在估計相關係數矩陣時，Wilk 和只利用完整資料的方法有降低變數平均相關的趨勢，Dear 的方法則趨向增高相關，而 Buck 的方法只在原矩陣爲低相關時會降低相關，對其他矩陣則有增高相關的現象。

上述研究已有統計學者採用變數取樣的方法，造成系統的不完整資料，用之以探討各種估計母數的方法。變數取樣似乎源自於 William Turnball 項目取樣（item sampling 或 matrix sampling）的觀念。此種觀念 Lord (1955b) 將之應用於測驗以解決測驗編製上取樣及常模問題。他在 1962 年有系統地研究項目取樣的模式對測驗總分及其他統計數的估計，並發展出若干基本公式。其研究架構亦成爲以後項目取樣研究的一個典型。他嘗試瞭解一個包含有70個選擇題的測驗的常模分配是否可以用項目取樣可靠地估計。他將70個題目分成10個分測驗，每個分測驗7個題目，這10個分測驗分別施測於10組受試，每組 100 人。並同時用受試取樣的方法，選取 100 個人，每人做所有的70題。他認爲項目取樣的結果，不論對測驗分數的平均數，變異數或分配都有較佳的估計。幾年後他與 Novick (1968) 系統地討論了項目取樣時母平均數（μ）和母變異數（σ^2）的估計問題，他們有幾項結論：第一，如果增加受試人數，對平均數和標準差估計的標準誤（standard error）會逐漸降低。第二、如受試人數一定，則增加分測驗的數目會比增加每一分測驗的題數，或增加每一分測驗的受測人數得到更佳的估計結果。第三、對某一固定長度的測驗具有相同數目的受試者的各種取樣，在估計變異數時，一般都會有相同的標準誤，但估計平均數的標準誤則不相同。

此後研究者針對 Lord (1962) 提出的若干問題繼續研究，Plumlee (1964) 和 Barcikowski

探討不重置取樣 (sampling without replacement) 的問題，均認爲項目取樣有較佳的結果。Cook 和 stufflebeam (1967) 以重置取樣 (sampling with replacement) 的方法研究各種樣本數的影響，亦發現項目取樣結果較佳 。 Owens 和 Stufflebeam (1968) 比較項目取樣和受試取樣在估計一些常模統計數 (norm statistics) 的正確性，發現不同比例的試題數在估計上，對不同能力的受試有不同的結果。Jacobs 和 Wildemann (1969) 比較不同比率的試題取樣對敎育測驗的影響，結果顯示不同的取樣比率並沒有產生多大差別，但發現項目取樣偶而會產生負值變異數，此種結果亦經其他研究證實 (Husek 和 Sirotnik, 1968).。Shoemaker (1970a, 1970b) 的兩篇研究很有系統地改變幾個變數：分測驗的數目，每個分測驗的題數，及受試者的人數，他認爲在利用項目取樣以估計常模分配時，要達到最大效果的關鍵在於實際獲得「觀察」數目，亦卽受試者對他們的題目所做的全部反應數目，而不是項目取樣本身（如分測驗數目，每個分測驗的題數，每一個分測驗施測人數），甚至與總分分配形狀也沒有關係，此結果受到 Hambleton, Rovinelli 和 Gorth (1971) 的支持。

Lord (1965)曾建議在從事項目取樣研究時，可以利用統計學上的平衡不完全區組設計(balanced incomplete block design, BIBD)的概念組織題目。 BIBD 一詞首由 Yates (1936) 提出，並與 Fisher (1938) 建立了一些 BIBD 的設計表，不過它們只考慮了區組數少於10的情形，Bose (1939) 更進一步建立較完整的設計表。Ramanujacharyulu (1966) 利用數學上 Galois fields的概念討論平衡不完全區組設計。他導出了 BIBD 設計中一些因素之間的關係。

平衡不完全區組設計雖有其統計分析上令人滿意的特質 (Clatworthy, 1956)，但在實驗研究情形下，殊難做平衡的設計。因此統計學家研究出部分平衡不完全區組設計 (partially balanced incomplete block design, PBIBD) (Bose, 1951; Bose 和 Clatworthy, 1954; Bose 和 Nair, 1939)。Clatworthy (1973) 編製了相當完整的 PBIBD 表，分別適用於各種不同的區組設計。

Knapp (1968a) 應用 BIBD 設計研究項目取樣，他發現在估計平均數，變異數和信度係數時，此種設計有相當令人滿意的結果。在另一篇文章， Knapp (1968b) 比較平衡不完全區組設計及部分平衡不完全區組設計在項目取樣的效果，他發現前者的估計結果非常接近母數，但後者的結果則沒有前者那麼好。

筆者在1979年曾以上述兩種區組設計的方法，設計變數取樣的架構，研究各種不同情形下的不完整資料對於母共變數矩陣的估計。在該研究中，考慮了不同的變數數目、變數取樣比率、樣本大小、取樣設計型態 、 變數間的平均相關 、 變數間的因素結構等因素 ， 探討對共變數矩陣幾種歸結數量 (summarizing number) 的估計，這些數量包括概化變異數 (generalized variance)、軌跡(trace)、最大根 (largest root)、和估計誤差量等等。結果發現變數取樣造成的不完整資料對各種共變數矩陣數量的估計有不同的效果，大概可歸納如下：

1. 在幾種因素中，樣本數對估計的影響最大。
2. 兩種不同的區組設計對共變數矩陣的估計，效果差不多，且估計誤差不比傳統的受試取樣大，甚至爲小。
3. 一般地說，對概化變異數的估計誤差較大，而且當變數間的相關高時，常常會產生負值的概化變異數，但對於軌跡的估計則相當理想。

筆者在該研究中曾指出，爲了牽就部分平衡不完全區組設計的有限設計，及顧及共變數的平均人數相等，結果每一種設計的總取樣人數及變異數之人數都不相等，因此在比較時產生很大的困難，本研究針對此問題，兼顧變異數的人數，共變數的平均人數及總取樣人數的儘可能接近，進一步探討共變數矩陣的估計問題，希望能提供更多資料，以便對不完整資料的使用有更進一步的瞭解。

方 法 與 設 計

本研究仍採用筆者 1979 年研究的架構，以蒙地卡羅 (Monte Carlo) 方法，系統地研究各種可

能因素對估計共變數矩陣的效果。這些因素可歸納爲兩大類，第一類爲界定共變數矩陣的變數，第二類爲涉及取樣設計的變數。本研究首先利用第一類變數，經由因素分析的方法，建立母共變數矩陣，再由此矩陣，系統地變換第二類變數，利用電腦模擬資料而計算樣本共變數矩陣，此種電腦資料的模擬，每種取樣設計都重複 200 次，以建立其分配而瞭解估計的情形。

一、研究變項

本研究所採用的設計共包含下列幾種因素：

1. 變數數目 (number of variables)：本研究共用兩種，即 6 個變數和10個變數。
2. 變數間的平均相關 (average correlation between variables)：也有二層次，平均相關在 .65 至 .75 之間，是爲高相關變數；平均相關在 .20 至 .30 之間即爲低相關變數。
3. 變數取樣比率 (ratio of variable sampling)：此爲每一個變數組包含的變數數目，有二層次，分別爲約包含變數的二分之一及三分之一。例如在 6 個變數的情形下，每一個變數組包含了 3 個及 2 個變數。
4. 設計型態(design type)：有三個層次，即 BIBD 爲一種，再加上兩種 PBIBD 設計。在 BIBD 取樣中，構成每一個共變數的人數都一樣，但共變數的人數與變異數的人數則不相等。至於 PBIBD 則不但共變數與變異數的人數不相等，而且其中一種情形是共變數間之人數不相等，其不相等情形有兩種，即 Clatworthy (1973) 所稱之兩種結合類別部分平衡設計 (two-associate-class partially balanced designs) 其中之一爲共變數間人數不相等，但都不等於零，一部分人數爲另一部分的二倍（即 Clatworthy 所用之 $\lambda_1=1, \lambda_2=2$ 之設計）；一爲有些共變數人數爲零（即 $\lambda_1=0, \lambda_2=1$ 設計）。
5. 樣本數 (sample size)：樣本數之決定係根據變異數之人數而定，分小樣本和大樣本兩種，前者爲 80人，後者爲 240人，但爲牽就 BIBD 和 PBIBD 之設計，部分設計之人數只約等於此二數目。本研究所用之設計及其人數如表 1 。

表 1　研究設計及取樣人數

變數	設計型態	每人接受的變數數目	小樣本 共變數平均人數	小樣本 變異數人數	小樣本 需要總人數	大樣本 共變數平均人數	大樣本 變異數人數	大樣本 需要總人數	比率*
6	完全取樣 (15)+	6	80	80	80	240	240	240	1.00
	BIBD (15)	3	32	80	160	96	240	480	.40
	PBIBD1 (12, 3)	3	32.4	81	162	97.2	243	486	.40
	PBIBD2 (12, 3)	3	32	80	160	96	240	480	.40
	BIBD (15)	2	16	80	240	48	240	720	.20
	PBIBD1 (12, 3)	2	16.2	81	243	48.6	243	729	.20
	PBIBD2 (12, 3)	2	16	80	240	48	240	720	.20
10	完全取樣 (45)	10	80	80	80	240	240	240	1.00
	BIBD (45)	5	36	81	162	108	243	486	.44
	PBIBD1 (15, 30)	5	36	81	162	108	243	486	.44
	PBIBC2 (40, 5)	5	35.6	80	160	106.8	240	480	.44
	BIBD (45)	3	18	81	270	54	243	810	.22
	PBIBD1 (15, 30)	3	17.3	78	260	52	234	780	.22
	PBIBD2 (30, 15)	3	17.3	78	260	52	234	780	.22

*：比率等於共變數平均人數除於變異數人數。

+：括號內數目是共變數之中，其人數有λ_1，和λ_2結合的數目。

上述研究變項中，前二項爲界定母共變數矩陣的特性因素，共有 4 種（2×2）結合，後三項爲界定取樣的特性因素，共有 12 種（2×3×2）結合，因此本研究所系統研究的取樣情形共有 48 種（4×12），再加上傳統的完全取樣設計 8 種（4種母共變數矩陣 ×2 種樣本數），總共研究了56種取樣的情形，有關詳細取樣的設計請參閱筆者1979年的論文。

二、母共變數矩陣的界定和樣本共變數矩陣的計算

雖然 Odell 和 Fieveson（1966）提出直接概化樣本共變數矩陣的方法，且經一些研究者使用（Montanelli, 1971; Tatsuoka, 1973; Dawson, 1977），但因限於變數取樣的複雜性，本研究仍採用 Kaiser 和 Dickman（1962）所提之傳統方法，先由共變數矩陣概化樣本多變數常態隨機變數，然後再計算樣本共變數矩陣。

本研究採用具有 p 個變數的實際資料，計算有 q 個因素的因素組型矩陣(factor pattern matrix)，然後每一變數的獨特性（uniqueness）再由下面公式計算：

$$U_i = (1 - \sum_{j=1}^{q} P_{ij}^2)^{.5},$$

式中 P_{ij}^2 是第 i 變數在第 j 因素上的因素負荷量（factor loading）的平方，然後整個因素矩陣則由 p×q 的共同因素（common factor）和 p×p 的誤差主軸矩陣（diagonal matrix of random error）組成。爲了校正計算誤差（runing error），因素負荷量（l_{ij}）和獨特性（u_i）都用下面公式加以校正並常態化（normalized）：

$$l_{ij} = l_{ij} / (\sum_{j=1}^{q} l_{ij}^2 + u_i^2)^{.5},$$

$$u_i = u_i / (\sum_{j=1}^{q} l_{ij}^2 + u_i^2)^{.5},$$

爲了簡化起見，所有變數的變異數都定爲 4，所有變數的因素負荷量和獨特性再以標準差加以加權，母共變數矩陣（Σ）因此可用下面公式計算出：

$$\Sigma = PP + U^2,$$

其中P是經由標準差加權後的因素組型矩陣，U^2 則是加權後的獨特性主軸矩陣。

根據母共變數矩陣，再利用隨機數目概化函數的電腦程式模擬因素分數向量（factor score vector），此隨機數目概化程式係根據 Marsaglia 和 Bray（1964）的方法所設計，其模擬出的數值平均數爲 0，標準差爲 1，因此概化出的因素分數矩陣之共變數矩陣爲一同一性矩陣（identity metrix），即 $\Sigma_F = I$。然後每一個人 p 變數的樣本觀察數再根據此 F 向量予以模擬：

$$X_{ij} = \sum_{k=1}^{q} f_{ik} p_{kj} + a_{ij} u_j,$$

式中 X_{ij} 是第 i 個人在第 j 變數的觀察數，f_{ik} 是第 i 個人在第 k 個因素的因素分數，p_{kj} 是第 j 個變數在第 k 個因素上的因素負荷量，a_{ij} 則是第 i 個受試者的第 j 個獨特性因素分數，而 u_j 則是第 j 個變數的獨特性，如用矩陣表示，則上述成爲 $X = FP' + AU$。最後，用此樣本觀察數計算樣本共變數矩陣。缺失的共變數再用 Spearman（1927）所提四分法（tetrad method）加以估計。

三、分　　析

在多變數統計分析中，目前對共變數矩陣所採用的歸納性統計數大概可分爲兩種，第一爲多變數變異數分析（MANOVA）所常用之概化變異數(generalized variance），此爲 Wilks（1932b）首提。Wilks 界定概化變異數爲共變數矩陣的行列式（determinant）。因此它在計算中同時考慮各變數的變異數及各變數間的共變數，唯其如此，當變數間爲高相關時，概化變異數便接近於零。此種特質增加了其應用時的困難。第二種爲因素分析（factor analysis），主成份分析（principal component analysis) 等所採用之軌跡 (trace) 觀念。此概念首由 Pearson (1901) 提出，再由 Hotelling (1933; 1936) 延伸應用到因素分析方法 (factor analytic approches) 的多變數統計。它卽爲共變數矩陣中主軸元素的簡單和 (simple sum)，亦卽各變數變異數之和，此種數量的一個特徵是它並不考慮變數之間的相關。第三，最大根 (largest root) 則與軌跡觀念及概化變異數有關，且經常見於 Lawley 之統計方法。最後，在使用項目取樣時，有些研究者發現有變異數爲負值之情形，而本研究中亦曾發生，因此，共變數矩陣具有行列式爲負值之數目亦加以統計。筆者（民69）曾對這些統計數量有過詳細的介紹與討論。本研究因此用這四者作爲探討共變數矩陣估計的統計數量。每種統計數量均利用前述過程重複 200 次，以所得之分配計算其平均數，標準差，偏態及峯度等予以討論。

四、電腦程式

本研究所用之電腦程式 SAMCOV 爲筆者於1979年所設計，其中用了 Farr (1975) 的 XRAN 函數副程式 (function subprogram) 概化隨機常態數目，及 HOW 副程式 (Cooley 和 Lohnes, 1971) 計算特徵值 (eigenvalue)。本程式需要 60k 來執行。資料則利用美國紐約州立大學之 CDC Cyber 173 型電子計算機所做，此型電腦有 60 bit words，當可提供非常準確之結果。

結　　果

對上述所提四種描述共變數矩陣之統計數量的估計，將分別討論其結果。針對本研究的設計，每一個估計數的結果都有四個表，每個表都包括設計的因素及估計結果的平均數及標準差，其中有一個設計在資料模擬時發生困難，無法得到結果，在表中以「一」表示。茲將表中的符號分別說明如下：

1. N_c＝共變數的人數。
2. $N\lambda_1$＝具有λ_1型聯結之共變數數目。
3. $N\lambda_2$＝具有λ_2型聯結之共變數數目。
4. $\overline{N}_c$＝整個矩陣中共變數的平均人數。
5. N_v＝每一個變異數的人數。
6. N_P＝某一設計的全部取樣人數。
7. v/p＝每個受試者所接受變數的數目。
8. r_L＝低相關係數矩陣的平均相關係數。
9. r_H＝高相關係數矩陣的平均相關係數。

一、軌跡的估計

軌跡依其定義是變異數的和，並不計算共變數在內，其估計將受取樣時變異數的取樣情形影響較大。表2、表3、表4、和表5分別爲樣本軌跡的平均數與標準差。在不影響估計問題之下，本研究假設每一變數的變異數都是4.00。因此，在六個變數共變數矩陣的軌跡母數爲 24.00，在十個變數的矩陣則爲 40.00。首先，變數間的平均相關似乎是一個重要因素，從四個表中很顯然地可以發現，對於軌跡的估計，無論是在六個變數或在十個變數的情形，低相關矩陣的結果比高相關矩陣的結果更接近母數，而且相差極爲懸殊。但是，無論是低相關矩陣或高相關矩陣，對於軌跡的估計都不算理想。再看兩者在估計時的變異情形，高相關矩陣似乎比低相關矩陣稍爲小些。

當比較表2、表3、和表4、表 5 時，意外地發現，當取樣人數增大三倍時，對軌跡的估計却沒有改變，亦卽大樣本和小樣本的結果幾乎完全相同，至於其他取樣設計的因素，諸如取樣比率、設計型態等，對於估計結果都沒有什麼影響。

二、概化變異數的估計

Wilks (1932b) 定義概化變異數爲共變數矩陣的行列式，因在計算時包含共變數在內，其數值大小受變數之間相關的影響甚大，原則上，高相關之矩陣其概化變異數會比較小，而低相關矩陣則較大。此種現象可以從表6、表7、表8、和表9、看出。這四個表呈現各種取樣設計對概化變異數之估計結果。另方面，

表 2　小樣本六個變數矩陣軌跡之估計

取樣設計							平均數		標準差	
每一分子的觀察				設計			低相關	高相關	低相關	高相關
$N_O(N\lambda_2)$	$N_C(N\lambda_1)$	$\overline{N}_C$	N_V	N_P	V/P	型態	($r_L=.22$)	($r_H=.72$)	($r_L=.22$)	($r_H=.72$)
						母數	24.00	24.00		
	80(15)	80	80	80	6	傳統取樣	22.20	10.49	1.58	.84
	32(15)	32	80	160	3	BIBD	22.04	10.64	1.61	1.06
36(12)	18(3)	32.4	81	162	3	PBIBD1	22.30	10.46	1.52	.85
40(12)	0(3)	32	80	160	3	PBIBD2	22.28	10.58	1.60	.95
	16(15)	16	80	240	2	BIBD	22.24	10.61	1.62	.97
18(12)	9(3)	16.2	81	243	2	PBIBD1	22.27	10.55	1.65	.96
20(12)	0(3)	16	80	240	2	PBIBD2	22.29	10.53	1.64	.96

表 3　大樣本六個變數矩陣軌跡之估計

取樣設計							平均數		標準差	
每一分子的觀察				設計			低相關	高相關	低相關	高相關
$N_O(N\lambda_2)$	$N_C(N\lambda_1)$	$\overline{N}_C$	N_V	N_P	V/P	型態	($r_L=.22$)	($r_H=.72$)	($r_L=.22$)	($r_H=.72$)
						母數	24.00	24.00		
	240(15)	240	240	240	6	傳統取樣	22.20	10.49	1.58	.84
	96(15)	96	240	480	3	BIBD	22.04	10.64	1.61	1.06
108(12)	54(3)	97.2	243	486	3	PBIBD1	22.30	10.40	1.52	.85
120(12)	0(3)	96	240	480	3	PBIBD2	22.28	10.58	1.60	.95
	48(15)	48	240	720	2	BIBD	22.24	10.61	1.62	.97
54(12)	27(3)	48.6	243	729	2	PBIBD1	22.27	10.55	1.65	.96
60(12)	0(3)	48	240	729	3	PBIBD2	22.29	10.55	1.64	.96

表 4　小樣本十個變數矩陣軌跡之估計

取樣設計							平均數		標準差	
每一分子的觀察				設計			低相關	高相關	低相關	高相關
$N_O(N\lambda_2)$	$N_C(N\lambda_1)$	$\overline{N}_C$	N_V	N_P	V/P	型態	($r_L=.22$)	($r_H=.68$)	($r_L=.22$)	($r_H=.68$)
						母數	40.00	40.00		
	80(45)	80	80	80	10	傳統取樣	34.90	16.65	1.90	1.17
	36(45)	36	81	162	5	BIBD	34.81	16.63	1.77	1.10
54(15)	27(30)	36	81	162	5	PBIBD1	34.78	16.52	1.89	1.05
40(40)	0(5)	36.5	80	160	5	PBIBD2	34.87	16.52	1.70	1.18
	18(45)	18	81	270	3	BIBD	35.02	16.64	1.85	1.14
261(5)	13(30)	17.3	78	260	3	PBIBD1	34.78	16.61	1072	.98
26(30)	0(15)	17.3	78	260	3	PBIBD2	34.71	16.60	1.93	1.01

表 5　大樣本十個變數矩陣軌跡之估計

取樣設計							平均數		標準差	
每一分子的觀察				設計			低相關	高相關	低相關	高相關
$N_C(N\lambda_2)$	$N_C(N\lambda_1)$	$\overline{N}_C$	N_V	N_P	V/P	型態	($r_L=.22$)	($r_H=.68$)	($r_L=.22$)	($r_H=.68$)
						母數	40.00	40.00		
	240(45)	240	240	240	10	傳統取樣	34.88	16.65	1.06	1.62
	108(45)	108	243	486	5	BIBD	34.81	16.61	1.14	.65
162(15)	81(30)	106	243	486	5	PBIBD1	34.78	16.61	1.13	.60
120(40)	0(5)	106	240	480	5	PBIBD2	34.68	16.64	1.10	.68
	54(45)	54	243	810	3	BIBD	34.67	16.52	.97	.61
78(15)	39(30)	52	234	780	3	PBIBD1	34.76	16.59	1.10	.66
78(30)	0(15)	52	234	780	3	PBIBD2	34.83	16.56	1.07	.61

從純理論的觀點，一個共變數矩陣只能爲正定 (positive definite) 或半正定 (positive semidefinite)。但是，使用變數取樣估計母共變數時，由於取樣誤差或其他理由 (Thompson, 1962; Sirotnik, 1972)，估計的共變數矩陣的行列式可能會產生負值，尤以母共變數矩陣的行列式接近零時爲然。本研究發生此種現象，因此上述四個表中亦列舉了具有負行列式的矩陣數目做爲另一種描述估計的數量。

從四個表中，我們可以發現傳統的取樣方法，無論是六個變數或十個變異的矩陣，矩陣的行列式都不會有負值，其他的取樣設計型態則多少會產生負值行列式的矩陣。在大樣本時，此種現象較少發生，但在小樣本則發生的次數比較多。其次，取樣比率似乎亦與產生負值行列式的數目有關，當取樣三分之一變數時，其產生負值行列式矩陣的數目比取樣二分之一變數時爲多。在三種設計型態中，PBIBD1 設計產生負值行列式矩陣的可能性似乎比 BIBD 和 PBIBD2 爲高。至於二種界定矩陣的因素，如同前述，高相關矩陣產生負值行列式矩陣的次數比低相關矩陣所產生的爲多。而十個變數樣本矩陣亦比六個變數的矩陣產生較多的負值行列式。

表 6　小樣六個變數矩陣概化變異數之估計

取樣設計							平均數		標準差	
每一分子的觀察				設計			低相關	高相關	低相關	高相關
$N_C(N\lambda_2)$	$N_C(N\lambda_1)$	$\overline{N}_C$	N_V	N_P	V/P	型態	($r_L=.22$)	($r_H=.72$)	($r_L=.22$)	($r_H=.72$)
						母數	2273.03	23.10		
	88(15)	80	80	80	6	傳統取樣	1296.92(0)*	3.17(0)	542.75	1.23
	32(15)	32	80	160	3	BIBD	808.08(0)	1.93(5)	408.76	1.20
36(12)	18(3)	32.4	81	162	3	PBIBD1	746.60(4)	1.80(11)	408.26	1.10
40(12)	0(3)	32	80	160	3	PBIBD2	1068.58(0)	2.46(0)	502.32	1.23
	16(15)	16	80	240	2	BIBD	451.26(71)	1.08(99)	354.13	1.07
18(12)	9(3)	16.2	81	243	2	PBIBD1	342.73(79)	.87(102)	266.42	.67
20(12)	0(3)	16	80	240	2	PBIBD2	656.59(23)	1.56(45)	422.26	1.05

* 括號內數目爲個矩陣中，出現負行列式矩陣之數目。

表 7　大樣本六個變數矩陣概化變異數之估計

取樣設計							平均數		標準差	
每一分子的觀察				設計			低相關	高相關	低相關	高相關
$N_C(N\lambda_2)$	$N_C(N\lambda_1)$	$\overline{N}_C$	N_V	N_P	V/P	型態	$(r_L=.22)$	$(r_H=.72)$	$(r_L=.22)$	$(r_H=.72)$
						母數	2273.03	23.10		
	240(15)	240	240	240	6	傳統取樣	1465.90(0)*	3.80(0)	348.22	.80
	96(15)	96	240	480	3	BIBD	1316.52(0)	3.13(0)	349.82	.79
108(12)	54(3)	97.2	243	486	3	PBIBD1	1236.86(0)	3.20(0)	281.25	.85
120(12)	0(3)	96	240	480	3	PBIBD2	1388.09(0)	3.50(0)	331.87	.74
	48(15)	48	240	720	2	BIBD	1056.16(0)	2.47(0)	340.88	.98
54(12)	27(3)	48.6	243	729	2	PBIBD1	1014.25(1)	2.32(1)	292.19	.92
60(12)	0(3)	48	240	729	2	PBIBD2	1210.30(0)	—	343.96	—

* 括號內數目為在 200 個矩陣中，出現負行列式矩陣之數目

表 8　小樣本十個變數矩陣概化變異數之估計

取樣設計							平均數		標準差	
每一分子的觀察				設計			低相關	高相關	低相關	高相關
$N_C(N\lambda_2)$	$N_C(N\lambda_1)$	$\overline{N}_C$	N_V	N_P	V/P	型態	$(r_L=.22)$	$(r_H=.68)$	$(r_L=.22)$	$(r_H=.68)$
						母數	301918.38	151.25		
	80(45)	80	80	80	10	傳統取樣	90620.73(0)*	9(47(0)	46631.75	4.86
	36(45)	36	81	162	5	BIBD	26709.89(12)	264(16)	19035.29	2.18
54(15)	27(30)	36	81	162	5	PBIBD1	21373.76(22)	2.4(27)	19920.81	2.18
40(40)	0(5)	36.5	80	160	5	PBIBD2	46064.85(2)	4.62(3)	28927.81	3.18
	18(45)	18	81	270	3	BIBD	9253.97(177)	60(179)	13370.92	.76
26(15)	13(30)	17.3	78	260	3	PBIBD1	4406.01(192)	27(193)	3283.72	.29
26(30)	0(15)	17.3	78	260	3	PBIBD2	27166.13(27)	3.25(29)	24897.00	2.81

* 括號內數目為在 200 個矩陣中，出現負行列式矩陣之數目

就概化變異數而論，估計的結果都不太理想，樣本數與母數相去甚為懸殊。一般的說，高相關矩陣的估計比低相關矩陣為差。而十個變數矩陣也比六個變數矩陣不理想。在幾個界定取樣設計的因素中，樣本數的大小是一個重要的因素，大樣本的結果比小樣本的結果為佳。樣本的變異在十個變數的矩陣也比小樣本為大，但在六個變數矩陣則反比小樣本為小，這倒是一個有趣的發現。再者，變數取樣比率對概化變異數的估計也有影響，當我們取樣二分之一變數時，估計的結果比取樣三分之一變數時為佳，但一般地說，前者樣本數的變異却比後者為大，這是可以理解的，因為後者產生負值行列式矩陣的數目為大，而這些矩陣則不在計算之內。至於設計型態，除了傳統取樣方法的結果稍佳，其他三種變數取樣方法的型態之間却沒有明顯的差異。

三、最大根的估計

最大根是共變數矩陣的第一個特徵值，就四個母共變數矩陣而言，其值就不會一樣。當變數數目

表 9　大樣本十個變數矩陣概化變異數之估計

取樣設計							平均數		標準差	
每一分子的觀察				設計			低相關	高相關	低相關	高相關
$N_c(N\lambda_2)$	$N_c(N\lambda_1)$	$\bar{N}_c$	N_v	N_P	V/P	型態	(r_L=.22)	(r_H=.68)	(r_L=.22)	(r_H=.68)
						母數	301918.38	151.25		
	240(45)	240	240	240	10	傳統取樣	133262.(0)*	4.90(0)	37252.72	4.47
	108(45)	108	243	486	5	BIBD	100246.18(0)	10.61(0)	33956.53	3.41
132(15)	81(30)	108	243	486	5	PBIBD1	94686.93(0)	10.01(0)	29406.54	3.30
120(40)	0(5)	106	240	480	5	PBIBD2	108417.16(0)	11.86(0)	35090.12	3.92
	51(45)	54	243	810	3	BIBD	51540.34(1)	5.66(0)	22828.94	2.87
78(15)	39(30)	52	234	780	3	PBIBD1	42855.88(3)	4.52(7)	21494.52	2.34
78(30)	0(15)	52	234	780	3	PBIBD2	98943.07(0)	9.93(0)	33877.86	3.64

* 括號內數目為在 200 個矩陣中，出現負行列式矩陣之數目

相等時，高相關矩陣的最大根將比低相關矩陣的最大根爲大。本研究假設每一變數的變異數都是4.00，因此，十個變數矩陣的軌跡就比六個變數矩陣的軌跡來得大，由於矩陣的特徵值 (θ) 與它們的變數有關，其關係爲：軌跡$=\Sigma\sigma_i{}^2=\Sigma\theta_i$ ，十個變數矩陣的最大根比六個變數的最大根爲大是可預期的，我們也就不能將兩個矩陣的最大根拿來比較。

從表10、表11、表12和表13，首先我們可以看出在全部情形下，都低估了母數，但在六個變數，低相關矩陣，變數取樣三分之一，小樣本的情形下，所有的三個設計型態却都高估了母數，是很意外的結果。大體說來，低相關矩陣所得之結果比高相關之結果偏差較小，而高相關矩陣所得到的估計值更小於其相對應的取樣設計在低相關所得之值，是個很意外的結果。至於取樣標準差，估計值在高相關矩陣的分散情形却也比在低相關矩陣的分散來得小。在所有界定取樣設計因素中，變數取樣比率似乎是比較重要，一般地說，取樣三分之一變數的結果比取樣二分之一的結果偏差較小，而前者估計值的分散也比後者爲大。其次，各種設計型態之間的估計結果雖有不同，其差異却沒有什麼意義，且不穩定，而三種取樣設計與傳統取樣所得到的結果也差不多。最後，非常意外地，對於最大根的估計，小樣本的

表10　小樣本六個變數矩陣最大之估計

取樣設計							平均數		標準差	
每一分子的觀察				設計			低相關	高相關	低相關	高相關
$N_c(N\lambda_2)$	$N_c(N\lambda_1)$	$\bar{N}_c$	N_v	N_P	V/P	型態	(r_L=.22)	(r_H=.72)	(r_L=.22)	(r_H=.72)
						母數	8.60	18.48		
	8.(15)	80	80	80	6	傳統別樣	7.78	5.69	1.15	.77
	32(15)	32	80	160	3	BIBD	8.06	5.93	1.32	1.04
36(12)	18(3)	32.4	81	162	3	PBIBD1	8.31	5.67	2.28	.86
40(12)	0(3)	32	80	160	3	PBIBD2	8.05	5.84	1.26	.97
	16(15)	16	80	240	2	BIBD	8.85	6.10	1.39	1.14
18(12)	9(3)	16.2	81	243	2	PBIBD1	8.90	6.02	1.61	1.07
20(12)	0(3)	16	80	280	2	PBIBD2	8.62	5.62	5.86	1.50

表11 大樣本六個變數矩陣最大根之估計

取樣設計							平均數		標準差	
每一分子的觀察				設計			低相關	高相關	低相關	高相關
$N_C(N\lambda_2)$	$N_C(N\lambda_1)$	$\overline{N}_C$	N_V	N_P	V/P	型態	$(r_L=.22)$	$(r_H=.72)$	$(r_L=.22)$	$(r_H=.72)$
						母數	8.60	18.48		
	240(15)	240	240	240	6	傳統取樣	7.62	5.72	.67	.50
	96(15)	96	240	480	3	BIBD	7.75	5.70	.71	.54
108(12)	54(3)	97.2	243	486	3	PBIBD1	7.79	5.71	.70	.54
120(12)	0(3)	96	240	480	3	PBIBD2	7.78	5.69	.77	.56
	48(15)	48	240	720	2	BIBD	7.85	5.85	.81	.55
54(12)	27(3)	48.6	243	729	2	PBIBD1	7.88	5.76	077	.57
60(12)	0(3)	48	240	729	2	PBIBD2	7.86		.76	

表12 小樣本十個變數矩陣最大根之估計

取樣設計							平均數		標準差	
每一分子的觀察				設計			低相關矩陣	高相關矩陣	低相關	高相關
$N_C(N\lambda_2)$	$N_C(N\lambda_1)$	$\overline{N}_C$	N_V	N_P	V/P	型態	$(r_L=.22)$	$(r_H=.68)$	$(r_L=.22)$	$(r_H=.68)$
						母數	12.22	28.59		
	80(45)	80	80	80	10	傳統取樣	8.07	6.16	1.14	1.00
	36(45)	36	81	162	5	BIBD	8.76	6.39	1.25	.98
54(15)	27(30)	36	81	162	5	PBIBD1	8.71	6.28	1.24	.91
40(40)	0(5)	36.5	80	160	5	PBIBD2	8.39	6.10	1.13	.96
	18(45)	18	81	270	3	BIBD	9.80	6.64	1.39	1.02
	13(30)	17.3	78	260	3	PBIBD1	9.74	6.78	1.30	1.02
	0(15)	17.3	78	260	3	PBIBD2	8.57	6.24	1.30	.87

表13 大樣本十個變數矩陣最大根之估計

取樣設計							平均數		標準差	
每一分子的觀察				設計			低相關	高相關	低相關	高相關
$N_C(N\lambda_2)$	$N_C(N\lambda_1)$	$\overline{N}_C$	N_V	N_P	V/P	型態	$(r_L=.22)$	$(r_H=.68)$	$(r_L=.22)$	$(r_H=.68)$
						母數	12.22	28.59		
	240(45)	240	240	240	10	傳統取樣	7.67	6.33	.96	1.30
	108(45)	108	243	486	3	BIBD	7.83	6.05	.70	.47
162(15)	81(30)	108	243	486	3	PBIBD1	7.88	6.07	.78	.52
120(40)	0(5)	106	240	480	5	PBIBD2	7.63	6.07	.66	.59
	54(45)	54	243	810	3	BIBD	8.19	6.10	.75	.59
78(15)	39(30)	52	234	780	3	PBIBD1	8.33	6.20	.80	.64
78(30)	0(15)	52	234	780	3	PBIBD2	7.75	6.03	.72	.52

結果比大樣本更接近母數，雖然這種差異並未達到實質的意義。

討論與結論

項目取樣方法曾被研究者多方探討，並證實對平均數及變異數的估計具有與傳統取樣方法相同的良好結果（Cook 和 Stufflebeam, 1967; Lord, 1962; Plumlee, 1964; Shoemaker, 1970a; 1970b)。Knapp (1968a) 更進一步利用平衡不完全區組設計建立取樣架構探討估計問題，本研究利用平衡不完全區組設計及部分平衡不完全區組設計建立變數取樣，探討多變數共變數矩陣之估計，為瞭解估計結果，乃利用共變數矩陣的幾個歸納數量，諸如概化變異數（generalized variance)，軌跡（trace)，和最大根（largest root）做為指標。

在變數取樣研究中，由於其設計的特質，首先遭到的一個問題是，變異數和共變數的樣本數不一樣，如果利用部分平衡不完全區組設計建立取樣架構的話，不但變異數與共變數人數不相等，而且共變數間人數也不相等，甚至有些共變數的人數會等於零，取樣人數之決定在變數取樣研究也就居於決定性地位。本研究在考慮取樣設計時，為便利估計結果之比較，求各設計間變異數之取樣人數相近為目標，亦即在小樣本中變異數之人數儘可能接近 80人，大樣本接近 240人。如此，則各設計間共變數之平均人數便有所差別，且與變異數之人數相去頗多，從各表之左半邊可以看到此現象，有些設計共變數之平均人數只有16個人，也許由此原因，本研究之估計結果偏差相當大，尤以軌跡之估計更是出乎意外。筆者在1979年曾從事相似之研究，唯在該研究中，取樣人數以各設計間共變數取樣人數之相近為考慮，其結果雖對概化變異數之估計不甚理想，對軌跡與最大根之估計則偏差不大，可見，在設計變數取樣時，似乎應以共變數之人數相近為主。

其次，部分平衡不完全區組設計有共變數漏失之取樣設計（即PBIBD2)，本研究中這種取樣設計之估計結果雖與其他設計之結果無甚差異，但它具有一個相當困難的問題，亦即在估計概化變異數與最大根時，因兩者都必求共變數矩陣之特徵值或行列式， 當利用 PBIBD2 設計，某些共變數便無法獲得，欲求其概化變異數便必先估計這些共變數，本研究使用 Spearman 所提之四分法(tatrad method)，求其所有可能之估計值之平均數來代入這些無法獲得之共變數，因此其概化變異數之估計的可靠性與正確性便有問題，這是值得考慮的。

最後，共變數矩陣產生負定（negative definite）結果，乃變數取樣之一個嚴重問題。筆者在1979年之研究中，曾有過負定矩陣的結果，此問題在本研究依然存在，可見這是個相當普遍的問題。Hůsek 與 Sirotnik (1968)，和 Jacobs 與 Wildeman (1969) 在研究項目取樣對變異數之估計時，亦報告了這些問題的存在，Sirotnik (1970) 認為當 Cronback 的概化係數 (generalizability coefficient 或稱∝係數）在樣本中小於零便會造成負值估計。筆者曾試過幾種方法 (Mosteller 和 Tukey, 1968; Bock 和 Vandenberg; 1967)，但都無法解決此問題，就此而言，以軌跡做為共變數矩陣的歸納數量可能比概化變異數理想。

本研究採用變數取樣方法系統地探討各種可能因素對共變數矩陣的估計。幾種因素中，似乎只有變數的平均相關對估計產生較大的影響，其他諸如變數的數目、樣本的大小、變數取樣比率、設計型態等因素都沒有什麼顯著的影響，此結果與筆者 (1979) 之另一研究出入頗多。根據 Lord 和 Novick (1968)的研究，增加受試人數及降低變數取樣比率都會減少估計誤差。 Shoemaker (1970a, 1970) 的研究結果則認為變數取樣比率，樣本數等都對估計沒有什麼效果，主要因素為受試者的總反應數，又似乎與本研究有所符合，到底事實如何，實有待進一步研究。

參考書目

吳鐵雄：多變數離散觀念的起源和發展。教育學刊，民69年，第2期，199～229。

Afifi, A. A. & Elashoff, R. M. Missing observations in multivariate statistics I. Review of the literature. *Journal of the American Statistical Association,* 1966, **61,** 595-604.

Afifi, A. A. & Elashoff, R. M. Missing observations in multivariate statistics II. Point estimation in simple linear regression. *Journal of the American Statistical Association,* 1967, **62,** 10-29.

Barcikowski, R. S. A Monte Carlo study of item sampling (versus traditional sampling) for norm construction, *Journal of Educational Measurement,* 1972, **9,** 209-214.

Bose, R. C. On the construction of balanced incomplete block designs. *Annals of Eugenics,* 1939, **9,** 353-399.

Bose R. C. Partially balanced incomplete designs with two associate classes involving only two replications. *Calcutta Statistical Association,* 1951, **3,** 120-125.

Bose, R. C., Clatworthy, W. H. & Shrikhande, S. S. *Tables of partially balanced designs with two associate classes.* North Carolina Agricultural, Experimental and Statistical Bulletin, No. 107, 1954.

Bose, R. C. & Nair, K. R. Partially balanced incomplete block designs, *Sankhya,* 1939, **4,** 337-372.

Buck, S. F. A method of estimation of missing values in multivariate data suitable for use with an electronic computer. *Journal of the Royal Statistical Society, Series B,* 1960, **22,** 302-307.

Clatworthy, W. H. *Contributions on partially balanced incomplete block designs with two associate classes.* National Bureau of Standards Applied Mathematics, Series No, 47, 1956.

Clatworthy, W. H. *Tables of two-associate-class partially balanced designs.* National Bureau of Standards Applied Mathematics, Series 63, 1973.

Cook, D. L. & Stufflebeam, D. L. Estimating test norms from variable size item and examinee samples. *Educational and Psychological Measurement,* 1967, **27,** 601-610.

Cooley, W. W. & Lohnes, P. R. *Multivariate data analysis.* New York: Wiley, 1971.

Dawson, E. K. *The samplling distribution of the canonical redundancy statistic.* Unpublished doctoral dissertation, University of Illinois at Urbana-Champaign, 1977.

Dear, R. E. *A principal-component missing data method for multiple regression models.* System Development Corporation, Technical Report SP-86, Santa Monica, California, 1959.

Fisher, R. A. & Yates, F. *Statistical tables for biological, agricultural and medical research.* London and Edinburgh: Oliver and Boyd, 1938.

Haitovsky, Y. Missing data in regression analysis. *Journal of the Royal Statistical Society, Series B,* 1968, **30,** 67-82.

Hambelton, R. K., Rovinelli, R., & Gorth, W. P. *Efficiency of various item-examinee sampling designs for estimating test parameters.* Paper presented at the Annual Meeting of the American Psycological Association, September, 1971.

Hocking, R. R. & Smith, W. B. Estimation of parameters in the multivariate normal distribution with missing observations. *Journal of the American Statistical Association,* 1968, **63,** 159-173.

Hotelling, H. Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology,* 1933, **24,** 417-441; 495-520.

Hotelling, H. Relations between two sets of variates. *Biometrika,* 1936, **28,** 321-377.

Husek, T. R. & Sirotnik, K. *Item sampling in educational research.* Occational report, No.

2, Center for the Study of Evaluation of Instructional Program, University of California at Los Angeles, December, 1967.

Jacobs, S. S. & Wildemann, C. *Matrix sampling with classroom tests: An empirical investigation.* Paper presented at the Annual Convocation of the Educational Research Association of New York State, November, 1969.

Kaiser, H. F. & Dickman, K. Sample and population score matrices and sample correlation matrices from an arbitrary population correlation matrix. *Psychometrika,* 1962, **27,** 179-182.

Knapp. T. R. An application of balanced incomplete block designs to the estimation of test norms. *Educational and Psychological Measurement,* 1968, **28,** 265-272, (a).

Knapp, T. R. *BIBD vs. PBIBD: An example of a priori item sampling.* Unpublished manuscript, University of Rochester, August, 1968, (b).

Kosobud, R. A note on a problem caused by assignment of missing data in sample surveys. *Econometrica,* 1963, **31**, 562-563.

Lord, F. M. Estimation of parameters from incomplete data. *Journal of the American Statistical Association,* 1955, **50,** 870-876, (a).

Lord, F. M. Sampling fluctuations resulting from the sampling of test items. *Psychometrika,* 1955, **20,** 1-22, (b).

Lord, F. M. Estimating norms by item sampling. *Educational and Psychological Measurement,* 1962, **22,** 259-267.

Lord, F. M. *Item sampling in test theory and in research design.* Princeton, N.J.: Educational Testing Service Research Bulletin, RB-65-22, 1965.

Lord, F. M. & Novick, M. R. *Statistical theories of mental test scores.* Reading, Mass: Addison-Wesley, 1968.

Marsaglia, G. & Bray, T. A. A convenient method for generating normal variables. *SIAM Review,* 1964, **6,** 260-264.

Matthai, A. Estimation of parameters from incomplete data with application to design of sample surveys. *Sankhya,* 1951, **11,** 145-152.

Montanelli, R. G., Jr. *An investigation of the goodness of fit of the maximum likelihood estimation procedure in factor analysis.* Unpublished doctoral dissertation, University of Illinois, Urbana, 1971.

Odell, P. I. & Fieveson, A. H. A numerical procedure to generate a sample covariance matrix. *Journal of the American Statistical Association,* 1966, **61,** 199-203.

Owens, T. R. & Stufflebeam, D. L. An experimental comparison of item sampling and examinee sampling for estimating test norms. *Journal of Educational Measurement,* 1969, **6,** 75-83.

Pearson, K. On lines and planes of closest fit to systems of points in space. *Philosophical Magzine, Series 6,* 1901, **2,** 569-572.

Plumlee, L. B. Estimating means and standard deviations from partial data—An empirical check on Lord's item sampling technique. *Educational and Psychological Measurement,* 1964, **24,** 623-630.

Ramanujacharyulu, C. A new general series of balanced incomplete block designs. *Proceedings* *of the American Mathematical Society*, 1966, **17**, 1064-1068.

Shoemaker, D. M. Allocation of items and examinees in estimating a norm distribution by item-sampling. *Journal of Educational Measurement*, 1970, **7**, 123-128. (a).

Shoemaker, D. M. Item-examinee sampling procedures and associated standard errors in estimating test parameters. *Journal of Educational Measurement*, 1970, **7**, 255-262, (b).

Sirotnik, K. An analysis of variance framework for matrix sampling. *Educational and Psychological Measurement*, 1970, **30**, 891-908.

Tatsuoka, M. M. *An examination of the statistical properties of a multivariate measure of strength of relationship.* Final Report, Project No. 2-E-020, HEW, 1973.

Thompson, W. A. The problem of negative estimates of variance components. *Annals of Mathematical Statistics*, 1962, **33**, 273-289.

Timm, N. H. The estimation of variance-covariance and correlation matrices from incomplete data. *Psychometrika*, 1970, **35**, 417-437.

Trawinski, I. M. & Bargmann, R. E. Maximum likelihood estimation with incomplete multivariate data. *Annals of Mathematical Statiscs*, 1964, **35**, 647-657.

Wilks, S. S. Moments and distributions of estimates of population parameters from fragmentary samples. *Annals of Mathematical Statistics*, 1932, **3**, 163-195, (a).

Wilks, S. S. Certain generalizations in the analysis of variances. *Biometrika*, 1932, **24**, 471-484, (b).

Wu, T. H. *The estimation of covariance matrices from systematically incomplete data.* Unpublished doctoral dissertation, State University of New York at Buffalo, 1979.

Yates, F. Incomplete randomized blocks. *Annals of Eugenics*, 1936, **7**, 121-140.

Bulletin of Educational Psychology, 1981, 14, 189-204.
Taiwan Normal University, Taipei, Taiwan, China.

THE ESTIMATION OF COVARIANCE MATRICES ON INCOMPLETE DATA: AN EMPIRICAL STUDY

TIEH-HSIUNG WU

This Monte Carlo study, utilizing variable sampling approach, is to systematically investigate the estimation of covariance matrices from incomplete data. The parameters studied were generalized variance, trace of a matrix, and the largest root. Five factors which related to sampling plan were included: number of variables, average correlation among variables, sample size, proportion of variable sampling, pattern of design. The sampling plans were organized by balanced incomplete block design (BIBD) and partially balanced incomplete block design (PBIBD). There were 56 conditions studied.

The results of estimates were quite biased, especially those of generalized variance. Some negative-definite matrices occured for high correlation matrices, small sample. Among the factors studied, average correlation among variables seems to be the most important one. There were no consistent pattern for the rest of the factors in terms of accuracy of estimation. Surprisely, the estimation from the large sample size was not substantially better than those from the small sample size.