DOCUMENT RESUME

ED 073 163 | TM 002 412

AUTHOR	Gleser, Leon Jay; Olkin, Ingram
TITLE	Multivariate Statistical Inference Under Marginal Structure, I.
INSTITUTION	Educational Testing Service, Princeton, N.J.
SPONS AGENCY	Air Force Office of Scientific Research, Washington, D.C.; National Science Foundation, Washington, D.C.
REPORT NO	ETS-RB-72-40
PUB DATE	Aug 72
NOTE	81p.
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	*Hypothesis Testing; *Mathematical Applications; *Mathematical Models; *Psychological Tests; *Statistical Studies

ABSTRACT

Statistical inference concerning the parameters of k multivariate normal populations is considered. Several models in which the parameters have certain hierarchical relationships are discussed, in particular as related to testing the hypothesis that k psychological tests are parallel forms of the same test. The report contains the following sections: Introduction; The Underlying Structure; Maximum Likelihood Estimation; Tests of Hypotheses; and An Application. An appendix is titled "Derivation of the Maximum Likelihood Estimators." References are provided. (DB)

RB-72-40

RESEARCH BULLETIN

MULTIVARIATE STATISTICAL INFERENCE UNDER MARGINAL STRUCTURE, I.

Leon Jay Gleser
The Johns Hopkins University
and Educational Testing Service

and

Ingram Olkin
Stanford University
and Educational Testing Service

This Bulletin is a draft for interoffice circulation. Corrections and suggestions for revision are solicited. The Bulletin should not be cited as a reference without the specific permission of the authors. It is automatically superseded upon formal publication of the material.

Educational Testing Service
Princeton, New Jersey
August 1972

Multivariate Statistical Inference under Marginal Structure, I.[1]

by

Leon Jay Gleser

The Johns Hopkins University and Educational Testing Service

and

Ingram Olkin

Stanford University and Educational Testing Service

1. Introduction

In this paper we are concerned with statistical inference concerning the parameters of k multivariate normal populations. Several models are considered in which the parameters have certain hierarchical relationships. These models may arise in a variety of scientific contexts, but our concern with this problem originated in the context of testing the hypothesis that k chological tests are parallel forms of the same test.

Suppose that we are utilizing k different (collections of) psychological tests. These (collections of) tests have one subtest T_0 in common, and are designed to be statistically equivalent (parallel) to one another. The components of the g -th test can be represented as (T_0, T_g) , where T_0 is the subtest common to all k psychological tests, and T_g is the subtest peculiar to the g -th test, $g = 1,2,\ldots,k$.

In one possible experimental design, each of the k psychological tests is given to a different group of persons. The k groups of persons are randomly constructed (of possibly unequal sizes), and are considered to be (statistically) homogeneous with respect to the psychological traits being measured. The score of a single person from the g -th group on the

[1] Work supported in part by the Educational Testing Service, the Air Force Office of Scientific Research (contract F44620-70-C-0066) at Johns Hopkins University, and the National Science Foundation (Grant GP-32326X) at Stanford University.

g -th test (T_0, T_g) is denoted by $(x_0^{(g)}, x_1^{(g)})$, where $x_0^{(g)}$ is the score on subtest T_0 and $x_1^{(g)}$ is the score on the remainder, T_g, of the test. The scores $x_0^{(g)}, x_1^{(g)}$ may be scalars, or they may be (row) vectors, depending on whether the subtests T_0, T_g themselves are considered to consist of one, or of more than one, parts. However, the dimensions of $x_0^{(1)}, x_0^{(2)}, \ldots,$ and $x_0^{(k)}$ are the same (since they are scores on the common subtest T_0), and the dimensions of $x_1^{(1)}, x_1^{(2)}, \ldots,$ and $x_1^{(k)}$ are the same (since the subtests $T_1, T_2, \cdots, T_k$ are designed to be statistically equivalent to one another). To be specific, let us assume that T_0 consists of q parts, so that the common dimensions of $x_0^{(1)}, x_0^{(2)}, \ldots, x_0^{(k)}$ are $1 \times q$, and let us assume that $T_1, T_2, \cdots, T_k$ each consist of $p - q$ parts $(q < p)$, so that the common dimensions of $x_1^{(1)}, x_1^{(2)}, \ldots, x_1^{(k)}$ are $1 \times (p - q)$.

It is assumed that the score of an individual on any test has a multivariate normal distribution, and that scores of individuals are mutually statistically independent. Thus, in describing a statistical model for this problem, it remains only to specify the mean vectors and covariance matrices of the score vectors $(x_0^{(1)}, x_1^{(1)}), (x_0^{(2)}, x_1^{(2)}), \ldots,$ $(x_0^{(k)}, x_1^{(k)})$. These parameters are perhaps best described by a table:

Parameters	Group (Test)			
	1	2	...	k
Mean Vector	$(\mu_0^{(1)}, \mu_1^{(1)})$	$(\mu_0^{(2)}, \mu_1^{(2)})$	...	$(\mu_0^{(k)}, \mu_1^{(k)})$
Covariance Matrix	$\begin{pmatrix} \Sigma_{00}^{(1)} & \Sigma_{01}^{(1)} \\ \Sigma_{10}^{(1)} & \Sigma_{11}^{(1)} \end{pmatrix}$	$\begin{pmatrix} \Sigma_{00}^{(2)} & \Sigma_{01}^{(2)} \\ \Sigma_{10}^{(2)} & \Sigma_{11}^{(2)} \end{pmatrix}$	...	$\begin{pmatrix} \Sigma_{00}^{(k)} & \Sigma_{01}^{(k)} \\ \Sigma_{10}^{(k)} & \Sigma_{11}^{(k)} \end{pmatrix}$

For the g -th group (the individuals who take the g -th test), $\mu_0^{(g)}$ is the expected score (vector) on the parts of the subtest T_0, $\mu_1^{(g)}$ is the expected score (vector) on the parts of subtest T_g, $\Sigma_{00}^{(g)}$ is the covariance matrix among the scores on the parts of subtest T_0, $\Sigma_{11}^{(g)}$ is covariance matrix among the scores on the parts of subtest T_g, and $\Sigma_{01}^{(g)} = (\Sigma_{10}^{(g)})'$ is the matrix of covariances between scores on parts of T_0 and scores on parts of T_g. We write

$$\mu^{(g)} = (\mu_0^{(g)}, \mu_1^{(g)}) \quad , \quad \Sigma^{(g)} = \begin{pmatrix} \Sigma_{00}^{(g)} & \Sigma_{01}^{(g)} \\ \Sigma_{10}^{(g)} & \Sigma_{11}^{(g)} \end{pmatrix} \quad , \quad g = 1,2,\ldots,k \quad .$$

Thus, $\mu^{(g)}$ is the mean vector and $\Sigma^{(g)}$ is the covariance matrix of the distribution of the score of a single person from the g -th group on the g -th test (T_0, T_g).

If the k groups are truly (statistically) homogeneous with respect to the psychological attributes being measured, then, since all k groups take subtest T_0, we would expect

$$(1.1) \qquad H_{m'vc'}: \mu_0^{(1)} = \mu_0^{(2)} = \ldots \mu_0^{(k)} \quad , \quad \Sigma_{00}^{(1)} = \Sigma_{00}^{(2)} = \ldots = \Sigma_{00}^{(k)}$$

to be true, regardless of whether or not the k tests $(T_0, T_1), (T_0, T_2), \ldots, (T_0, T_k)$ are parallel forms (statistically equivalent). The hypothesis that all k tests are parallel forms is

$$(1.2)\quad H_{mvc}: \mu^{(1)} = \mu^{(2)} = \dots = \mu^{(k)} \quad , \quad \Sigma^{(1)} = \Sigma^{(2)} = \dots = \Sigma^{(k)} \quad .$$

To verify that the k tests $(T_0,T_1),(T_0,T_2),\dots,(T_0,T_k)$ are indeed parallel forms, given that the k groups are statistically homogeneous, we can test the hypothesis H_{mvc} versus the more general alternative hypothesis $H_{m'vc'}$.

In some instances we may believe that the noncommon parts $T_1,T_2,\dots,T_k$ of the k tests are not necessarily statistically equivalent, and we may have some doubts as to whether or not the k groups have the same mean performance on the common subtest T_0 (i.e., whether $\mu_0^{(1)} = \mu_0^{(2)} = \dots = \mu_0^{(k)}$). However, we may continue to believe that the parts of subtest T_0 have the same interrelationships (variances and covariances) in all k groups. In such a case, our given hypothesis is

$$(1.3)\quad H_{vc'}: \Sigma_{00}^{(1)} = \Sigma_{00}^{(2)} = \dots = \Sigma_{00}^{(k)} \quad ,$$

and we may want to test $H_{m'vc'}$ or H_{mvc} against this hypothesis. (Note that H_{mvc} implies $H_{m'vc'}$ which in turn implies $H_{vc'}$.) Acceptance of the hypothesis $H_{m'vc'}$ as against the hypothesis $H_{vc'}$ means that all k groups respond similarly to subtest T_0 --in other words, the k groups are marginally homogeneous in their response to subtest T_0 . Acceptance of the hypothesis H_{mvc} as against $H_{vc'}$ means that the k tests are parallel forms and that the k groups are homogeneous in their responses to the tests $(T_0,T_1),(T_0,T_2),\dots,(T_0,T_k)$.

Besides hypotheses H_{mvc} , $H_{m'vc'}$, $H_{vc'}$, various intermediate hypotheses may be of interest. For example, the hypotheses:

(1.4) $H_{vc}: \Sigma^{(1)} = \Sigma^{(2)} = \ldots \Sigma^{(k)}$

and

(1.5) $H_{m'vc}: \mu_0^{(1)} = \mu_0^{(2)} = \ldots = \mu_0^{(k)}$, $\Sigma^{(1)} = \Sigma^{(2)} = \ldots = \Sigma^{(k)}$

may be of concern. Figure 1 indicates the logical relationships among the hypotheses (models) H_{mvc} , $H_{m'vc'}$, $H_{vc'}$, H_{vc} , and $H_{m'vc}$.

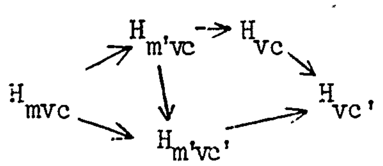


Figure 1. Logical relationships among the hypotheses. An arrow indicates implication. Thus, $H_{mvc} \rightarrow H_{m'vc}$ means that hypothesis H_{mvc} implies hypothesis $H_{m'vc}$.

The maximum likelihood estimators (MLE) of the parameters under the various models (the models defined by hypotheses H_{mvc} , $H_{m'vc'}$, $H_{vc'}$, $H_{m'vc}$, and H_{vc}) are listed in Section 3 and derived in an Appendix at the end of this paper. Using the results given in the Appendix, we may obtain likelihood ratio tests (LRT) between various pairs of hypotheses. These tests are given in Section 4. In general, the test statistics obtained from the likelihood ratio approach have distributions similar in form to the distributions for the LRT for the multivariate analysis of variance and to the distributions for Wilks's lambda test for the equality of covariance matrices [see Anderson (1958)]. The exact distributions are those of products of powers of independent beta variates, and are known to be very complicated in form. However, by using the Box (1949) approximation, we may obtain approximate levels of significance for these tests from a weighted sum of chi-square distributions (see Section 4). A numerical illustration of the computation and use of one of these likelihood ratio tests in a practical context appears in Section 5; this practical example is also used (and introduced) in Section 3 to illustrate the differences in value of the MLE under the various models discussed in this paper. Before entering into a discussion of the various estimators and tests of hypotheses, however, we have a few further comments to make concerning the underlying structure of the inferential problem.

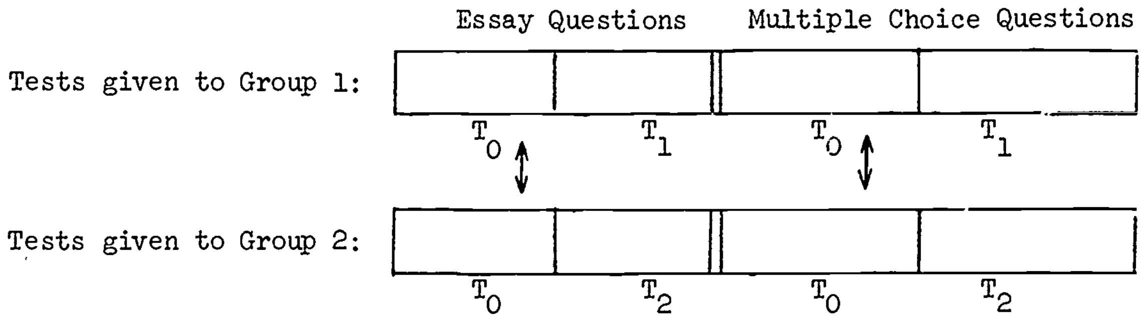
2. The Underlying Structure

Recall that we have k groups of individuals, and that each individual in the g-th group takes the test (T_0, T_g), $g = 1,2,\ldots,k$. The score for the i-th individual in group g is $(x_{0i}^{(g)}, x_{1i}^{(g)})$. If there are N_g individuals in group g, then we need only consider the sample means:

$$(2.1)\qquad \bar{x}_0^{(g)} = \frac{1}{N_g}\sum_{i=1}^{N_g} x_{0i}^{(g)}\ , \qquad \bar{x}_1^{(g)} = \frac{1}{N_g}\sum_{i=1}^{N_g} x_{1i}^{(g)}\ ,$$

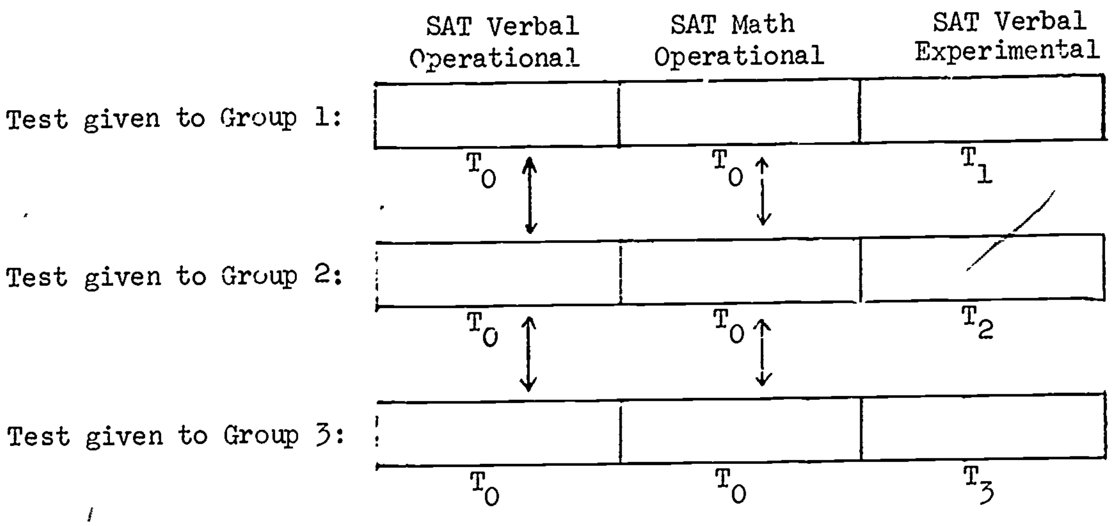
and can summarize these means by $\bar{x}^{(g)} = (\bar{x}_0^{(g)}, \bar{x}_1^{(g)})$.

If the subtests T_0 and T_g each have only one part (i.e., each of T_0 and T_g are summarized by one score), then $\bar{x}_0^{(g)}$ and $\bar{x}_1^{(g)}$ are scalars. However, we need not be restrictive about this, since the theory applies even when T_0 and T_g each have several parts (and are summarized by several scores), and $\bar{x}_0^{(g)}$ and $\bar{x}_1^{(g)}$ are vectors. For example, suppose there are two groups, and the tests given to the groups are made up as follows:



That is, some of the essay questions and some of the multiple choice questions are taken by both groups, while other essay questions and other multiple choice questions differ between the two groups. Suppose that the common essay questions are summarized by a single score, the distinct essay questions are summarized by a single score, the common multiple choice questions are summarized by a single score, and the distinct multiple choice questions are summarized by a single score. In this case, $\bar{x}_0^{(1)}$ and $\bar{x}_0^{(2)}$ are (row) vectors with 2 components (i.e., $\bar{x}_0^{(g)}$ is 1×2, $g = 1,2$), and $\bar{x}^{(1)}$ and $\bar{x}^{(2)}$ are (row) vectors with 2 components.

In the example provided in Sections 3 and 5, there are three groups, and we have the following format:



In this case, T_0 has 2 parts (and is summarized by 2 scores), and T_1, T_2, and T_3 have 1 part each. Hence $\bar{x}_0^{(1)}$, $\bar{x}_0^{(2)}$, $\bar{x}_0^{(3)}$ are all 1×2 vectors, while $\bar{x}_1^{(1)}$, $\bar{x}_1^{(2)}$, $\bar{x}_1^{(3)}$ are scalars.

Returning to our data, we can define sample cross-product matrices

$$(2.2) \qquad V^{(g)} = \begin{pmatrix} V_{00}^{(g)} & V_{01}^{(g)} \\ V_{10}^{(g)} & V_{11}^{(g)} \end{pmatrix} , \qquad g = 1,2,\ldots,k ,$$

where

$$(2.3) \qquad V_{00}^{(g)} = \sum_{i=1}^{N_g} (x_{0i}^{(g)} - \bar{x}_0^{(g)})'(x_{0i}^{(g)} - \bar{x}_0^{(g)})$$

is the sample cross-product matrix for the vector $x_0^{(g)}$ of scores of group g on subtest T_0, where

$$(2.4) \qquad V_{11}^{(g)} = \sum_{i=1}^{N_g} (x_{1i}^{(g)} - \bar{x}_1^{(g)})'(x_{1i}^{(g)} - \bar{x}_1^{(g)})$$

is the sample cross-product matrix for the vector $x_1^{(g)}$ of scores of group g on subtest T_g, and where

$$(2.5) \qquad V_{01}^{(g)} = \sum_{i=1}^{N_g} (x_{0i}^{(g)} - \bar{x}_0^{(g)})'(x_{1i}^{(g)} - \bar{x}_1^{(g)}) \quad , \quad V_{10}^{(g)} = (V_{01}^{(g)})' \quad ,$$

is the sample matrix of cross-products for group g between the scores on subtest T_0 and the scores on subtest T_g'. In the context of our first example, $V^{(1)}$ and $V^{(2)}$ are both 4 x 4 matrices, and $V_{00}^{(1)}$, $V_{00}^{(2)}$, $V_{01}^{(1)}$, $V_{01}^{(2)}$, $V_{11}^{(1)}$, $V_{11}^{(2)}$ are all 2 x 2 matrices. In the context of our second example, $V^{(1)}$, $V^{(2)}$, and $V^{(3)}$ are 3 x 3 matrices, $V_{00}^{(1)}$, $V_{00}^{(2)}$, $V_{00}^{(3)}$ are 2 x 2 matrices, $V_{01}^{(1)}$, $V_{01}^{(2)}$, and $V_{01}^{(3)}$ are 2 x 1 matrices, and $V_{11}^{(1)}$, $V_{11}^{(2)}$, $V_{11}^{(3)}$ are 1 x 1 matrices (i.e., scalars).

When all of the score vectors $(x_0^{(g)}, x_1^{(g)})$ have multivariate normal distributions, and the performances of individuals on the tests are mutually statistically independent, then it is well known that the mean vectors $\bar{x}^{(1)}, \bar{x}^{(2)}, \ldots, \bar{x}^{(k)}$ and sample cross-product matrices $V^{(1)}, V^{(2)}, \ldots, V^{(k)}$ together are jointly sufficient for inferences concerning the parameters $\mu^{(1)}, \mu^{(2)}, \ldots, \mu^{(k)}$, $\Sigma^{(1)}, \Sigma^{(2)}, \ldots, \Sigma^{(k)}$ of the model (see Anderson (1958)). Let us assume that we have already reduced our test score data to a summarization in terms of the quantities $\bar{x}^{(1)}, \bar{x}^{(2)}, \ldots, \bar{x}^{(k)}$ and $V^{(1)}, V^{(2)}, \ldots, V^{(k)}$.

It is known [Anderson (1958)] that under the assumptions described above, $\bar{x}^{(g)}$ and $V^{(g)}$ are statistically independent of one another for $g = 1,2,\ldots,k$. Also $(\bar{x}^{(1)},V^{(1)}), (\bar{x}^{(2)},V^{(2)}),\ldots, (\bar{x}^{(k)},V^{(k)})$ are mutually statistically independent. The distribution of $\bar{x}^{(g)}$ is multivariate normal with density function

$$(2.6) \quad p(\bar{x}^{(g)}) = c^{(g)}|\Sigma^{(g)}|^{-\frac{1}{2}} \exp -\frac{N_g}{2}[(\bar{x}^{(g)} - \mu^{(g)})(\Sigma^{(g)})^{-1}(\bar{x}^{(g)} - \mu^{(g)})'] \quad ,$$

where $c^{(g)}$ is a certain constant depending upon p and N_g. The distribution of $V^{(g)}$ is Wishart with parameters $n_g = N_g - 1$ and Σ_g. The density function of $V^{(g)}$ is given by

$$(2.7) \quad p(V^{(g)}) = d^{(g)}|\Sigma^{(g)}|^{-n_g/2}|V^{(g)}|^{(n_g-p-1)/2} \exp -\tfrac{1}{2}[\operatorname{tr} V^{(g)}(\Sigma^{(g)})^{-1}] \quad ,$$

where $d^{(g)}$ is a constant depending upon n_g and p.

Let

$$(2.8) \quad \begin{aligned} \bar{x} &= (\bar{x}^{(1)},\bar{x}^{(2)},\ldots,\bar{x}^{(k)}) \quad , & V &= (V^{(1)},V^{(2)},\ldots,V^{(k)}) \quad , \\ \underset{\sim}{\mu} &= (\mu^{(1)},\mu^{(2)},\ldots,\mu^{(k)}) \quad , & \underset{\sim}{\Sigma} &= (\Sigma^{(1)},\Sigma^{(2)},\ldots,\Sigma^{(k)}) \quad . \end{aligned}$$

To obtain maximum likelihood estimators (MLE) of $\underset{\sim}{\mu}$ and $\underset{\sim}{\Sigma}$ under the various models described in Section 1, we need to maximize the likelihood

$$(2.9) \quad p(\bar{x},V) = \prod_{g=1}^{k} [p(\bar{x}^{(g)})p(V^{(g)})] \quad ,$$

with respect to $\underset{\sim}{\mu}$ and $\underset{\sim}{\Sigma}$ under the restrictions upon these parameters imposed by the hypotheses H_{mvc}, $H_{m'vc}$, $H_{vc'}$, $H_{m'vc}$, and H_{vc}. For simplicity of exposition, we summarize the MLE's for $\underset{\sim}{\mu}$ and $\underset{\sim}{\Sigma}$ under the various hypotheses in Section 3; proofs of the results are deferred to the Appendix.

3. Maximum Likelihood Estimation

In this section we summarize the maximum likelihood estimators (MLE) of the parameters $(\underset{\sim}{\mu},\underset{\sim}{\Sigma})$ for each of the five models ($H_{vc'}$, $H_{m'vc'}$, H_{vc}, $H_{m'vc}$, and H_{mvc}) described in Section 1. However, it is helpful to first consider a reparameterization which simplifies the analysis and helps to clarify our understanding of the results.

In three of the five models described above (namely, in $H_{vc'}$, $H_{m'vc'}$, and $H_{m'vc}$), the restrictions on the parameters that are imposed by the model concern the parameters of the marginal distributions of the scores $x_{0i}^{(g)}$ made by individuals on subtest T_0. Thus, in the models defined by $H_{vc'}$ and $H_{m'vc'}$, the marginal covariance matrices $\Sigma_{00}^{(1)},\Sigma_{00}^{(2)},\ldots,\Sigma_{00}^{(k)}$ are constrained to be equal, while in the models defined by $H_{m'vc}$ and $H_{m'vc'}$, the marginal expected score vectors $\mu_0^{(1)},\mu_0^{(2)},\ldots,\mu_0^{(k)}$ are assumed to be equal. To isolate the marginal parameters $\mu_0^{(g)}$ and $\Sigma_{00}^{(g)}$, $g = 1,2,\ldots,k$, we are led [Lord (1955), Anderson (1957), Bhargava (1962)] to consider breaking the likelihood (2.9) into two factors: (i) the density function of the marginal quantities $\bar{x}_0^{(1)},\bar{x}_0^{(2)},\ldots,\bar{x}_0^{(k)}$, $V_{00}^{(1)},V_{00}^{(2)},\ldots,V_{00}^{(k)}$, and (ii) the conditional density function of the sufficient statistic $(\bar{x},V)$ given these marginal quantities. If we do this [see Equations (A.1) and (A.2) of the Appendix], we find that

$$\mu_0^{(g)} \; , \quad \alpha^{(g)} = \mu_1^{(g)} - \mu_0^{(g)}(\Sigma_{00}^{(g)})^{-1}\Sigma_{01}^{(g)} \; , \quad \Sigma_{00}^{(g)} \; ,$$

(3.1)

$$\beta^{(g)} = (\Sigma_{00}^{(g)})^{-1}\Sigma_{01}^{(g)} \; , \quad \text{and} \quad \Sigma_{11\cdot0}^{(g)} = \Sigma_{11}^{(g)} - \Sigma_{10}^{(g)}(\Sigma_{00}^{(g)})^{-1}\Sigma_{01}^{(g)} \; ,$$

$g = 1,2,\ldots,k$, appear as natural parameters in this representation. Note that $\beta^{(g)}$ is the $q \times p - q$ matrix of regression coefficients (slopes) and $\alpha^{(g)}$ is the $1 \times p - q$ vector of intercepts for the regressions of the elements of $x_1^{(g)}$ on $x_0^{(g)}$ (that is, $E[x_1^{(g)} | x_0^{(g)}] = \alpha^{(g)} + x_0^{(g)}\beta^{(g)}$) . Further, $\Sigma_{11.0}^{(g)}$ is the residual covariance matrix of $x_1^{(g)}$ after the dependence of $x_1^{(g)}$ on $x_0^{(g)}$ has been removed by regression. Thus, the parameters in (3.1) are not only of interest in connection with finding the MLE of $(\underset{\sim}{\mu},\underset{\sim}{\Sigma})$, but are also of interest in their own right.

It is not difficult to show that $(\underset{\sim}{\mu},\underset{\sim}{\Sigma})$ and (3.1) are equivalent parameterizations. Equation (3.1) represents $\mu_0^{(g)}$, $\alpha^{(g)}$, $\Sigma_{00}^{(g)}$, $\beta^{(g)}$, and $\Sigma_{11\cdot0}^{(g)}$, $g = 1,2,\ldots,k$, as functions of the parameters $(\underset{\sim}{\mu},\underset{\sim}{\Sigma})$. On the other hand,

$$\mu^{(g)} = (\mu_0^{(g)},\mu_1^{(g)}) = (\mu_0^{(g)},\alpha^{(g)} + \mu_0^{(g)}\beta^{(g)}) \; ,$$

(3.2)

$$\Sigma^{(g)} = \begin{pmatrix} \Sigma_{00}^{(g)} & \Sigma_{01}^{(g)} \\ \Sigma_{10}^{(g)} & \Sigma_{11}^{(g)} \end{pmatrix} = \begin{pmatrix} \Sigma_{00}^{(g)} & \Sigma_{00}^{(g)}(\beta^{(g)}) \\ (\beta^{(g)})'\Sigma_{00}^{(g)} & \Sigma_{11\cdot0}^{(g)} + (\beta^{(g)})'\Sigma_{00}^{(g)}(\beta^{(g)}) \end{pmatrix} ,$$

$g = 1,2,\ldots,k$, represents $(\underset{\sim}{\mu},\underset{\sim}{\Sigma})$ as functions of $\mu_0^{(g)}$, $\alpha^{(g)}$, $\Sigma_{00}^{(g)}$, $\beta^{(g)}$, and $\Sigma_{11\cdot 0}^{(g)}$, $g = 1,2,\ldots,k$.

Corresponding to the parameters $\alpha^{(g)}$, $\beta^{(g)}$, and $\Sigma_{11\cdot 0}^{(g)}$, we may define the sample quantities:

$$a^{(g)} = \bar{x}_1^{(g)} - \bar{x}_0^{(g)}(V_{00}^{(g)})^{-1}V_{01}^{(g)} \quad , \quad B^{(g)} = (V_{00}^{(g)})^{-1}V_{01}^{(g)} \; ,$$

$$V_{11\cdot 0}^{(g)} = V_{11}^{(g)} - V_{10}^{(g)}(V_{00}^{(g)})^{-1}V_{01}^{(g)} \; , \tag{3.3}$$

for $g = 1,2,\ldots,k$. Note [see the Appendix] that $\bar{x}_0^{(g)}$, $a^{(g)}$, $(N_g)^{-1}V_{00}^{(g)}$, $B^{(g)}$, and $(N_g)^{-1}V_{11\cdot 0}$ are the respective MLE of $\mu_0^{(g)}$, $\alpha^{(g)}$, $\Sigma_{00}^{(g)}$, $\beta^{(g)}$, and $\Sigma_{11\cdot 0}^{(g)}$, $g = 1,2,\ldots,k$, when the parameters $(\mu^{(g)},\Sigma^{(g)})$ of the distribution of the scores for the individuals in any one group are functionally unrelated to the parameters $(\mu^{(k)},\Sigma^{(k)})$ of the distribution of the scores of the individuals in any other group, $g \neq h$. In what follows, we refer to these maximum likelihood estimators as the "usual" (unrestricted) estimators of the corresponding parameters. For example, we refer to $B^{(g)}$ as the "usual" estimator of $\beta^{(g)}$, $g = 1,2,\ldots,k$.

We are now in position to give explicitly the MLE of the parameters $(\underset{\sim}{\mu},\underset{\sim}{\Sigma})$ under each of the models $H_{vc'}$, $H_{m'vc'}$, H_{vc} , $H_{m'vc}$ and H_{mvc} described in Section 1. Calculation of these MLE will be illustrated by the following practical example.

3.0 An Illustrative Example

The Scholastic Aptitude Test (SAT) of the College Entrance Examination Board contains items designed to measure verbal ability and items designed to measure mathematical ability. The test is given to a number of individuals at a time; different individuals who take the test may receive different forms of the test. In each such form, certain verbal items and certain mathematical items are common to all forms of the test. Other items, however, differ from form to form. The common items are used for the operational (measurement) purposes of the SAT; the differing items are included for certain experimental purposes. Suppose k such forms of the SAT exist. Then at a given administration of the test, each form is given to the same number of individuals, and forms are assigned to individuals by a process similar to the technique of (randomized) systematic sampling used in sample survey designs.

The score on the g-th form of the SAT can be summarized by three numbers (scores): (i) the total score on those verbal items common to all forms (SAT Verbal Operational Score), (ii) the total score on those mathematical items common to all forms (SAT Mathematical Operational Score), and (iii) the total score on those items peculiar to the g-th form (SAT Experimental Score). The common parts (SAT Verbal Operational, SAT Mathematical Operational) of each form constitute subtest T_0 in the terminology of Sections 1 and 2. Thus, the score vector $x_{0i}^{(g)}$ of the i-th individual in the group of individuals taking the g-th form is a 1×2 dimensional vector. The unique part (SAT Experimental) on the g-th form constitutes subtest T_g, $g = 1,2,\ldots,k$. The score vector

$x_{1i}^{(g)}$ of the i -th individual who takes the g -th form of the SAT is thus a scalar (a 1 x 1 dimensional vector).

In April, 1971, several thousand individuals took the SAT at testing centers across the country. A sample of 100 individuals was chosen from among all those individuals who took a given form of the SAT, for each of 3 different forms (T_0,T_1), (T_0,T_2), and (T_0,T_3) for which the experimental items were comparable (in the present case, all experimental items were verbal items). Thus, $q = 2$, $p = 3$, $k = 3$, and $N_1 = N_2 = N_3 = 100$. The test data have been summarized in terms of the sample mean vectors $\bar{x}^{(g)}$ and sample cross-product matrices $V^{(g)}$, separately for each form (group), $g = 1,2,3$. These summarizations appear in Table 1. Table 2 gives the "usual" estimators of the parameters $(\underset{\sim}{\mu},\underset{\sim}{\Sigma})$ and of the parameters defined in (3.1). The values of these "usual" estimators serve to provide comparisons to the values of the MLE of the parameters under each of the models $H_{vc'}$, $H_{m'vc'}$, H_{vc}, $H_{m'vc}$, H_{mvc}, which we discuss below.

3.1 Maximum Likelihood Estimators under $H_{vc'}$

We begin by considering the most general of the five models described in Section 1. In this model, which is defined by the hypothesis $H_{vc'}$, the score vectors $x_{0i}^{(g)}$ of individuals in all k groups have a common covariance matrix Σ_{00}. That is, under this model, the parameters $(\underset{\sim}{\mu},\underset{\sim}{\Sigma})$ of the distribution of the scores are restricted to belong to the parametric subspace $\omega_{vc'}$, where

Table 1. Summarization of Test Data for the Illustrative Example

$N_1 = N_2 = N_3 = 100$, $q = 2$, $p = 3$, $k = 3$,

$$\bar{x}^{(1)} = (33.86,\ 22.52,\ 14.77) \quad ,$$

$$\bar{x}^{(2)} = (33.62,\ 25.43,\ 14.55) \quad ,$$

$$\bar{x}^{(3)} = (36.05,\ 24.40,\ 16.21) \quad ,$$

$$V^{(1)} = \begin{pmatrix} 140,164 & 91,014 & 59,581 \\ 91,014 & 65,654 & 39,014 \\ 59,581 & 39,014 & 27,325 \end{pmatrix} \quad ,$$

$$V^{(2)} = \begin{pmatrix} 134,980 & 96,396 & 58,141 \\ 96,396 & 77,919 & 42,411 \\ 58,141 & 42,411 & 26,671 \end{pmatrix} \quad ,$$

$$V^{(3)} = \begin{pmatrix} 160,751 & 104,106 & 71,115 \\ 104,106 & 73,206 & 46,765 \\ 71,115 & 46,765 & 32,737 \end{pmatrix} \quad .$$

Table 2. The "Usual" Estimators of the Parameters $(\underset{\sim}{\mu},\underset{\sim}{\Sigma})$ for the Illustrative Example

$$\hat{\mu}^{(1)} = \bar{x}^{(1)} = (33.86,\ 22.52,\ 14.77) \quad ,$$

$$\hat{\mu}^{(2)} = \bar{x}^{(2)} = (33.62,\ 25.43,\ 14.55) \quad ,$$

$$\hat{\mu}^{(3)} = \bar{x}^{(3)} = (36.05,\ 24.40,\ 16.21) \quad ,$$

$$\hat{\Sigma}^{(1)} = \frac{1}{100} V^{(1)} = \begin{pmatrix} 1401.64 & 910.14 & 595.81 \\ 910.14 & 656.54 & 390.14 \\ 595.81 & 390.14 & 273.25 \end{pmatrix} \quad ,$$

$$\hat{\Sigma}^{(2)} = \frac{1}{100} V^{(2)} = \begin{pmatrix} 1349.80 & 963.96 & 581.41 \\ 963.96 & 779.19 & 424.11 \\ 581.41 & 424.11 & 266.71 \end{pmatrix} \quad ,$$

$$\hat{\Sigma}^{(3)} = \frac{1}{100} V^{(3)} = \begin{pmatrix} 1607.51 & 1041.06 & 711.15 \\ 1041.06 & 732.06 & 467.65 \\ 711.15 & 467.65 & 327.37 \end{pmatrix} \quad ,$$

$$\hat{\mu}_0^{(1)} = (33.86,\ 22.52) \quad , \qquad \hat{\alpha}^{(1)} = 0.3502 \quad ,$$

$$\hat{\mu}_0^{(2)} = (33.62,\ 25.43) \quad , \qquad \hat{\alpha}^{(2)} = -0.0705 \quad ,$$

$$\hat{\mu}_0^{(3)} = (36.05,\ 24.40) \quad , \qquad \hat{\alpha}^{(3)} = 0.1328 \quad ,$$

$$\hat{\Sigma}_{00}^{(1)} = \begin{pmatrix} 1401.64 & 910.14 \\ 910.14 & 656.54 \end{pmatrix} , \quad \hat{\beta}^{(1)} = \begin{pmatrix} 0.3928 \\ 0.0497 \end{pmatrix} , \quad \hat{\Sigma}_{11\cdot 0}^{(1)} = 19.82 \quad ,$$

$$\hat{\Sigma}_{00}^{(2)} = \begin{pmatrix} 1349.80 & 963.96 \\ 963.96 & 779.19 \end{pmatrix} , \quad \hat{\beta}^{(2)} = \begin{pmatrix} 0.3608 \\ 0.0980 \end{pmatrix} , \quad \hat{\Sigma}_{11\cdot 0}^{(2)} = 15.40 \quad ,$$

$$\hat{\Sigma}_{00}^{(3)} = \begin{pmatrix} 1607.51 & 1041.06 \\ 1041.06 & 732.06 \end{pmatrix} , \quad \hat{\beta}^{(3)} = \begin{pmatrix} 0.3630 \\ 0.1226 \end{pmatrix} , \quad \hat{\Sigma}_{11\cdot 0}^{(3)} = 11.90 \quad .$$

$$\omega_{vc'} = \{(\underset{\sim}{\mu},\underset{\sim}{\Sigma}): \ \Sigma_{00}^{(1)} = \Sigma_{00}^{(2)} = \cdots = \Sigma_{00}^{(k)} = \Sigma_{00} \ , \ \Sigma_{00} \text{ an arbitrary } q \times q \text{ covariance matrix}\} \ .$$

Because the marginal covariance matrices $\Sigma_{00}^{(1)},\Sigma_{00}^{(2)},\ldots,\Sigma_{00}^{(k)}$ have a common value Σ_{00}, the pooled estimator

$$(3.4) \qquad \frac{1}{N} W_{00} \equiv \frac{1}{N} \sum_{g=1}^{k} V_{00}^{(g)}$$

for Σ_{00}, where $N = \sum_{g=1}^{k} N_g$, has intuitive appeal. Note that although the hypothesis $H_{vc'}$ puts no explicit restrictions relating the parameters $\Sigma_{01}^{(g)}$ and $\Sigma_{11}^{(g)}$ to $\Sigma_{01}^{(h)}$ and $\Sigma_{11}^{(h)}$, respectively, $g \neq h$, implicit restrictions upon the relationships between these parameters are imposed by $H_{vc'}$ since

$$\Sigma^{(g)} = \begin{pmatrix} \Sigma_{00} & \Sigma_{01}^{(g)} \\ \Sigma_{10}^{(g)} & \Sigma_{11}^{(g)} \end{pmatrix}$$

must be a positive semi-definite matrix for all $g = 1,2,\ldots,k$. On the other hand, $H_{vc'}$ imposes no restrictions (explicit or implicit) relating the parameters $\beta^{(g)}$ and $\Sigma_{11\cdot0}^{(g)}$ to $\beta^{(h)}$ and $\Sigma_{11\cdot0}^{(h)}$, $g \neq h$. This fact suggests that $\beta^{(g)}$ and $\Sigma_{11\cdot0}^{(g)}$ be estimated by their "usual" estimators $B^{(g)}$ and $(N_g)^{-1}V_{11\cdot0}^{(g)}$, respectively, $g = 1,2,\ldots,k$. Similarly, since $H_{vc'}$ imposes no restrictions relating the parameters $\mu_0^{(g)}$ and $\alpha^{(g)}$ for one group to the corresponding parameters $\mu_0^{(h)}$ and $\alpha^{(h)}$ in any other group, $g \neq h$, we think of estimating these parameters by their "usual" estimators ($\mu_0^{(g)}$ by $\bar{x}_0^{(g)}$, and $\alpha^{(g)}$ by $a^{(g)}$, $g = 1,2,\ldots,k$).

It is shown in the Appendix that these estimators, namely

$$\hat{\mu}_0^{(g)}(vc') = \bar{x}_0^{(g)} \quad , \quad \hat{\alpha}^{(g)}(vc') = a^{(g)} \quad ,$$

$$(3.5) \qquad \hat{\Sigma}_{00}(vc') = \frac{1}{N}\sum_{g=1}^{k} V_{00}^{(g)} \quad , \quad \hat{\beta}^{(g)}(vc') = B^{(g)} \quad ,$$

$$\hat{\Sigma}_{11\cdot 0}^{(g)}(vc') = \frac{1}{N_g} V_{11\cdot 0}^{(g)} \quad ,$$

are indeed the MLE of $\mu_0^{(g)}$, $\alpha^{(g)}$, Σ_{00}, $\beta^{(g)}$, and $\Sigma_{11\cdot 0}^{(g)}$ respectively, $g = 1,2,\ldots,k$, when the model defined by $H_{vc'}$ holds. The MLE of $(\underset{\sim}{\mu},\underset{\sim}{\Sigma})$ under $H_{vc'}$ can now be found by substituting (3.5) into (3.2). This substitution results in the following MLE:

$$\hat{\mu}^{(g)}(vc') = (\bar{x}_0^{(g)}, \bar{x}_1^{(g)}) = \bar{x}^{(g)} \quad ,$$

$$(3.6) \qquad \hat{\Sigma}^{(g)}(vc') = \frac{1}{N}\begin{pmatrix} W_{00} & W_{00}(V_{00}^{(g)})^{-1}V_{01}^{(g)} \\ V_{10}^{(g)}(V_{00}^{(g)})^{-1}W_{00} & \frac{N}{N_g}V_{11\cdot 0}^{(g)} + V_{10}^{(g)}(V_{00}^{(g)})^{-1}W_{00}(V_{00}^{(g)})^{-1}V_{01}^{(g)} \end{pmatrix} \quad ,$$

for $g = 1,2,\ldots,k$. The actual values of the MLE (3.5) and (3.6) for the example described in Subsection 3.0 are summarized in Table 3.

Note that $\hat{\mu}^{(g)}(vc')$ is equal to the "usual" estimator $\bar{x}^{(g)}$ for $\mu^{(g)}$, but that $\hat{\Sigma}^{(g)}(vc')$ differs from the "usual" estimator $(N_g)^{-1}V^{(g)}$ for $\Sigma^{(g)}$. Indeed, the difference between $\hat{\Sigma}^{(g)}(vc')$ and $(N_g)^{-1}V^{(g)}$ equals

Table 3. Maximum Likelihood Estimators under $H_{vc'}$ for the Illustrative Example

$$\hat{\mu}^{(1)}(vc') = (33.86,\ 22.52,\ 14.77) \quad ,$$

$$\hat{\mu}^{(2)}(vc') = (33.62,\ 25.43,\ 14.55) \quad ,$$

$$\hat{\mu}^{(3)}(vc') = (36.05,\ 24.40,\ 16.21) \quad ,$$

$$\hat{\Sigma}^{(1)}(vc') = \begin{pmatrix} 1452.98 & 971.72 & 619.04 \\ 971.72 & 722.60 & 417.61 \\ 619.04 & 417.61 & 283.74 \end{pmatrix} \quad ,$$

$$\hat{\Sigma}^{(2)}(vc') = \begin{pmatrix} 1452.98 & 971.72 & 619.39 \\ 971.72 & 722.60 & 421.36 \\ 619.39 & 421.36 & 280.14 \end{pmatrix} \quad ,$$

$$\hat{\Sigma}^{(3)}(vc') = \begin{pmatrix} 1452.98 & 971.72 & 646.55 \\ 971.72 & 722.60 & 441.32 \\ 646.55 & 441.32 & 300.69 \end{pmatrix} \quad ,$$

$$\hat{\mu}_0^{(1)}(vc') = (33.86,\ 22.52) \quad , \qquad \hat{\alpha}^{(1)}(vc') = 0.3502 \quad ,$$

$$\hat{\mu}_0^{(2)}(vc') = (33.62,\ 25.43) \quad , \qquad \hat{\alpha}^{(2)}(vc') = -0.0705 \quad ,$$

$$\hat{\mu}_0^{(3)}(vc') = (36.05,\ 24.40) \quad , \qquad \hat{\alpha}^{(3)}(vc') = 0.1328 \quad ,$$

$$\hat{\Sigma}_{00}(vc') = \begin{pmatrix} 1401.64 & 910.14 \\ 910.14 & 656.54 \end{pmatrix} \quad ,$$

$$\hat{\beta}^{(1)}(vc') = \begin{pmatrix} 0.3928 \\ 0.0497 \end{pmatrix} \ , \quad \hat{\beta}^{(2)}(vc') = \begin{pmatrix} 0.3608 \\ 0.0980 \end{pmatrix} \ , \quad \hat{\beta}^{(3)}(vc') = \begin{pmatrix} 0.3630 \\ 0.1226 \end{pmatrix} \ ,$$

$$\hat{\Sigma}_{11\cdot 0}^{(1)} = 19.82 \ , \quad \hat{\Sigma}_{11\cdot 0}^{(2)}(vc') = 15.40 \ , \quad \hat{\Sigma}_{11\cdot 0}^{(3)}(vc') = 11.90 \quad .$$

$$\hat{\Sigma}^{(g)}(vc') - \frac{1}{N_g} V^{(g)} = \begin{pmatrix} I_q \\ V_{10}^{(g)}(V_{00}^{(g)})^{-1} \end{pmatrix} \left(\frac{1}{N} \sum_{s=1}^{k} V_{00}^{(s)} - \frac{1}{N_g} V_{00}^{(g)} \right) \begin{pmatrix} I_q \\ V_{10}^{(g)}(V_{00}^{(g)})^{-1} \end{pmatrix}' ,$$

where I_q is the $q \times q$ identity matrix.

3.2 Maximum Likelihood Estimators under $H_{m'vc'}$

In the model defined by the hypothesis $H_{m'vc'}$, the scores $x_{0i}^{(g)}$ of individuals on subtest T_0 are assumed to be identically distributed according to a q-variate normal distribution with mean vector μ_0 and covariance matrix Σ_{00}. That is, under the hypothesis $H_{m'vc'}$, the parameters $(\underset{\sim}{\mu}, \underset{\sim}{\Sigma})$ are restricted to belong to the parametric subspace $\omega_{m'vc'}$, where

$$\omega_{m'vc'} = \{\underset{\sim}{\mu}, \underset{\sim}{\Sigma}): \mu_0^{(1)} = \dots = \mu_0^{(k)} \equiv \mu_0 \text{ is an arbitrary } 1 \times q \text{ vector,}$$
$$\Sigma_{00}^{(1)} = \dots = \Sigma_{00}^{(k)} \equiv \Sigma_{00} \text{ is an arbitrary } q \times q \text{ covariance matrix}\} .$$

As is the case under $H_{vc'}$, the hypothesis $H_{m'vc'}$ explicitly requires $\Sigma_{00}^{(1)}, \Sigma_{00}^{(2)}, \dots, \Sigma_{00}^{(k)}$ to have the common value Σ_{00}, and implicitly relates the parameters $\Sigma_{01}^{(g)}$ and $\Sigma_{11}^{(g)}$ of the distribution of scores for one group to the corresponding parameters $\Sigma_{01}^{(k)}$ and $\Sigma_{11}^{(k)}$ of the distribution of scores for any other group, $g \neq h$. However, as before, $H_{m'vc'}$ places no restrictions relating the parameters $\beta^{(g)}$ and $\Sigma_{11\cdot 0}^{(g)}$ of one group to the corresponding parameters $\beta^{(h)}$ and $\Sigma_{11\cdot 0}^{(h)}$

of any other group; this suggests that we estimate $\beta^{(g)}$ and $\Sigma_{11\cdot 0}^{(g)}$ by their "usual" estimators $B^{(g)}$ and $(N_g)^{-1}V_{11\cdot 0}^{(g)}$, respectively, $g = 1,2,\ldots,k$.

Because $H_{m'vc'}$ requires that the marginal expected score vectors $\mu_0^{(1)},\mu_0^{(2)},\ldots,\mu_0^{(k)}$ must have a common value μ_0, we can estimate μ_0 by the pooled estimator:

$$(3.7) \qquad \bar{\bar{x}}_0 = \frac{1}{N}\sum_{g=1}^{k} N_g \bar{x}_0^{(g)} \quad .$$

The residual cross-product matrix

$$(3.8) \qquad A_{00} = \sum_{g=1}^{k} N_g(\bar{x}_0^{(g)} - \bar{\bar{x}}_0)'(\bar{x}_0^{(g)} - \bar{\bar{x}}_0)$$

can then be combined with $W_{00} = \sum_{g=1}^{k} V_{00}^{(g)}$ to provide an estimator

$$\frac{1}{N}(W_{00} + A_{00}) = \frac{1}{N}\sum_{g=1}^{k} [V_{00}^{(g)} + N_g(\bar{x}_0^{(g)} - \bar{\bar{x}}_0)'(\bar{x}_0^{(g)} - \bar{\bar{x}}_0)]$$

for Σ_{00}. Finally, since $H_{m'vc'}$ does not restrict the relationships among $\alpha^{(1)},\alpha^{(2)},\ldots,\alpha^{(k)}$, we are led to estimation of these parameters by their "usual" estimators $a^{(1)},a^{(2)},\ldots,a^{(k)}$, respectively.

In the Appendix we verify that the estimators

$$\hat{\mu}_0^{(g)}(m'vc') = \bar{\bar{x}}_0 \quad , \quad \hat{\alpha}^{(g)}(m'vc') = a^{(g)} \quad ,$$

$$(3.9) \qquad \hat{\Sigma}_{00}(m'vc') = \frac{1}{N}(W_{00} + A_{00}) \quad , \quad \hat{\beta}^{(g)}(m'vc') = B^{(g)} \quad ,$$

$$\hat{\Sigma}_{11\cdot 0}^{(g)}(m'vc') = \frac{1}{N_g} V_{11\cdot 0}^{(g)} \quad ,$$

are respectively the MLE of μ_0, $\alpha^{(g)}$, Σ_{00}, $\beta^{(g)}$, and $\Sigma_{11\cdot 0}^{(g)}$, $g = 1,2,\ldots,k$, under $H_{m'vc'}$. The MLE of the parameters $(\underset{\sim}{\mu},\underset{\sim}{\Sigma})$ can now be obtained by substituting (3.9) in (3.2). This substitution results in the following MLE:

$$\hat{\mu}^{(g)}(m'vc') = (\hat{\mu}_0^{(g)}(m'vc'),\hat{\mu}_1^{(g)}(m'vc')) = (\bar{\bar{x}}_0,\ \bar{x}_1^{(g)} - (\bar{x}_0^{(g)} - \bar{\bar{x}}_0)(V_{00}^{(g)})^{-1}V_{01}^{(g)}) ,$$

$$\hat{\Sigma}^{(g)}_{(m'vc')} = \frac{1}{N}\begin{pmatrix} (W_{00} + A_{00}) & (W_{00} + A_{00})(V_{00}^{(g)})^{-1}V_{01}^{(g)} \\ V_{10}^{(g)}(V_{00}^{(g)})^{-1}(W_{00} + A_{00}) & \frac{N}{N_g}V_{11\cdot 0}^{(g)} + V_{10}^{(g)}(V_{00}^{(g)})^{-1}(W_{00} + A_{00})(V_{00}^{(g)})^{-1}V_{01}^{(g)} \end{pmatrix} , \quad (3.10)$$

for $g = 1,2,\ldots,k$. The values of the MLE (3.9) and (3.10) for the example described in Subsection 3.0 are summarized in Table 4.

3.3 Maximum Likelihood Estimators under H_{vc}

In the model defined by the hypothesis H_{vc}, scores by all individuals (in any group) are assumed to have a common covariance matrix Σ, but not necessarily equal expected score vectors. That is, under this model, the parameters $(\underset{\sim}{\mu},\underset{\sim}{\Sigma})$ of the distributions of the scores are restricted to belong to the parametric subspace ω_{vc}, where

$$\omega_{vc} = \{(\underset{\sim}{\mu},\underset{\sim}{\Sigma}):\ \Sigma^{(1)} = \Sigma^{(2)} = \ldots = \Sigma^{(k)} = \Sigma \quad , \quad \Sigma \text{ an arbitrary } p \times p \text{ covariance matrix}\} .$$

The MLE of the parameters $(\underset{\sim}{\mu},\underset{\sim}{\Sigma})$ under this model are well known [Anderson (1958), p. 248]:

Table 4. Maximum Likelihood Estimators under $H_{m'vc'}$ for the Illustrative Example

$$\hat{\mu}^{(1)}(m'vc') = (34.51,\ 24.12,\ 15.10) \quad ,$$

$$\hat{\mu}^{(2)}(m'vc') = (34.51,\ 24.12,\ 14.94) \quad ,$$

$$\hat{\mu}^{(3)}(m'vc') = (34.51,\ 24.12,\ 15.61) \quad ,$$

$$\hat{\Sigma}^{(1)}(m'vc') = \begin{pmatrix} 1452.98 & 971.72 & 619.05 \\ 971.72 & 722.60 & 417.62 \\ 619.05 & 417.62 & 283.74 \end{pmatrix} \quad ,$$

$$\hat{\Sigma}^{(2)}(m'vc') = \begin{pmatrix} 1452.98 & 971.72 & 619.40 \\ 971.72 & 722.60 & 421.36 \\ 619.40 & 421.36 & 280.14 \end{pmatrix} \quad ,$$

$$\hat{\Sigma}^{(3)}(m'vc') = \begin{pmatrix} 1452.98 & 971.72 & 646.55 \\ 971.72 & 722.60 & 441.32 \\ 646.55 & 441.32 & 300.69 \end{pmatrix} \quad ,$$

$$\hat{\mu}_0(m'vc') = (34.51,\ 24.12) \quad ,$$

$$\hat{\alpha}^{(1)}(m'vc') = 0.3502\ ,\quad \hat{\alpha}^{(2)}(m'vc') = -0.0705\ ,\quad \hat{\alpha}^{(3)}(m'vc') = 0.1328 \quad ,$$

$$\hat{\beta}^{(1)}(m'vc') = \begin{pmatrix} 0.3928 \\ 0.0497 \end{pmatrix} ,\quad \hat{\beta}^{(2)}(m'vc') = \begin{pmatrix} 0.3608 \\ 0.0980 \end{pmatrix} ,\quad \hat{\beta}^{(3)}(m'vc') = \begin{pmatrix} 0.3630 \\ 0.1226 \end{pmatrix} ,$$

$$\hat{\Sigma}^{(1)}_{11\cdot 0}(m'vc') = 19.82\ ,\quad \hat{\Sigma}^{(2)}_{11\cdot 0}(m'vc') = 15.40\ ,\quad \hat{\Sigma}^{(3)}_{11\cdot 0}(m'vc') = 11.90 \quad .$$

$$(3.11)\quad \begin{aligned} &\hat{\mu}^{(g)}(vc) = \bar{x}^{(g)} \quad , \quad g = 1,2,\ldots,k \quad , \\ &\hat{\Sigma}(vc) = \frac{1}{N} \sum_{g=1}^{k} V^{(g)} = \frac{1}{N} W \quad , \end{aligned}$$

where

$$W = \begin{pmatrix} W_{00} & W_{01} \\ W_{10} & W_{11} \end{pmatrix} = \sum_{g=1}^{k} V^{(g)} \quad .$$

Here, we did not need to use the equivalent parameterization (3.1) in order to obtain MLE for $(\underset{\sim}{\mu},\underset{\sim}{\Sigma})$, since the results are directly and easily obtainable. However, for the sake of comparison to the results given in previous subsections, we can obtain the MLE of $\mu_0^{(g)}$, $\alpha^{(g)}$, $\Sigma_{00}^{(g)} \equiv \Sigma_{00}$, $\beta^{(g)} \equiv \beta$, $\Sigma_{11\cdot 0}^{(g)} \equiv \Sigma_{11\cdot 0}$, where

$$\Sigma = \begin{pmatrix} \Sigma_{00} & \Sigma_{00}\beta \\ \beta'\Sigma_{00} & \Sigma_{11\cdot 0} + \beta'\Sigma_{00}\beta \end{pmatrix} \quad ,$$

by substituting (3.11) into (3.1). The result of this substitution is the following list of MLE:

$$(3.12)\quad \begin{aligned} &\hat{\mu}_0^{(g)}(vc) = \bar{x}_0^{(g)} \quad , \quad \hat{\alpha}^{(g)}(vc) = \bar{x}_1^{(g)} - \bar{x}_0^{(g)} W_{00}^{-1} W_{01} \quad , \\ &\hat{\Sigma}_{00}(vc) = \frac{1}{N} W_{00} \quad , \quad \hat{\beta}(vc) = W_{00}^{-1} W_{01} \quad , \\ &\hat{\Sigma}_{11\cdot 0}(vc) = \frac{1}{N} (W_{11} - W_{10} W_{00}^{-1} W_{01}) \quad . \end{aligned}$$

The values of the MLE (3.11) and (3.12) for the example described in Subsection 3.0 are summarized in Table 5.

Table 5. Maximum Likelihood Estimators under H_{vc} for the Illustrative Example

$$\hat{\mu}^{(1)}(vc) = (33.86,\ 22.52,\ 14.77) \quad ,$$

$$\hat{\mu}^{(2)}(vc) = (33.62,\ 25.43,\ 14.55) \quad ,$$

$$\hat{\mu}^{(3)}(vc) = (36.05,\ 24.40,\ 16.21) \quad ,$$

$$\hat{\Sigma}(vc) = \begin{pmatrix} 1452.98 & 971.72 & 629.46 \\ 971.72 & 722.60 & 427.30 \\ 629.46 & 427.30 & 289.11 \end{pmatrix} \quad .$$

$$\hat{\mu}_0^{(1)}(vc) = (33.86,\ 22.52) \quad , \quad \hat{\alpha}^{(1)}(vc) = 0.1122 \quad ,$$

$$\hat{\mu}_0^{(2)}(vc) = (33.62,\ 25.43) \quad , \quad \hat{\alpha}^{(2)}(vc) = -0.2712 \quad ,$$

$$\hat{\mu}_0^{(3)}(vc) = (36.05,\ 24.40) \quad , \quad \hat{\alpha}^{(3)}(vc) = 0.5673 \quad ,$$

$$\hat{\Sigma}_{00}(vc) = \begin{pmatrix} 1452.98 & 971.72 \\ 971.72 & 722.60 \end{pmatrix} \quad , \quad \hat{\beta}(vc) = \begin{pmatrix} 0.3750 \\ 0.0871 \end{pmatrix} \quad , \quad \hat{\Sigma}_{11\cdot 0}(vc) = 15.87 \quad .$$

3.4 Maximum Likelihood Estimators under $H_{m'vc}$

In the model defined by hypothesis $H_{m'vc}$, the scores of all individuals on subtest T_0 have the same marginal distribution (with common mean vector μ_0 and common covariance matrix Σ_{00}). Further, the scores of all individuals on the remainder of the test have a common covariance matrix Σ_{11}, but not necessarily a common mean vector (that is, $\mu_1^{(1)},\mu_1^{(2)},\ldots,\mu_1^{(k)}$ are not necessarily equal). Finally, under $H_{m'vc}$, the score $x_{0i}^{(g)}$ of any individual on the common subtest T_0 serves as an equally good predictor of the score $x_{1i}^{(g)}$ of that individual on the remainder T_g of the test, regardless of the group to which the individual belongs. (That is, the correlations between elements of $x_{0i}^{(g)}$ and elements of $x_{1i}^{(g)}$ are the same for all individuals i in all groups g.) Thus, under $H_{m'vc}$ the parameters $(\underset{\sim}{\mu},\underset{\sim}{\Sigma})$ of the distributions of the scores of individuals on the various tests are restricted to belong to the parametric subspace $\omega_{m'vc}$, where

$$\omega_{m'vc} = \{(\underset{\sim}{\mu},\underset{\sim}{\Sigma}):\ \mu_0^{(1)} = \mu_0^{(2)} = \cdots = \mu_0^{(k)} \equiv \mu_0 \text{ is an arbitrary } 1 \times q \text{ vector},$$

$$\Sigma^{(1)} = \Sigma^{(2)} = \cdots = \Sigma^{(k)} \equiv \Sigma \text{ is an arbitrary } p \times p \text{ covariance matrix}\} .$$

The MLE of the parameters (μ,Σ) can be obtained as special cases of results obtained in a previous paper [Gleser and Olkin (1966)], or by direct analysis, as in the Appendix to the present paper. Note that the hypothesis $H_{m'vc}$ implies the following relationships among the parameters defined in (3.1):

$$\mu_0^{(1)} = \mu_0^{(2)} = \dots = \mu_0^{(k)} = \mu_0 \quad ,$$

$$\Sigma_{00}^{(1)} = \Sigma_{00}^{(2)} = \dots = \Sigma_{00}^{(k)} = \Sigma_{00} \quad , \tag{3.13}$$

$$\beta^{(1)} = \beta^{(2)} = \dots = \beta^{(k)} = \beta \quad ,$$

$$\Sigma_{11\cdot 0}^{(1)} = \Sigma_{11\cdot 0}^{(2)} = \dots = \Sigma_{11\cdot 0}^{(k)} = \Sigma_{11\cdot 0} \quad .$$

The hypothesis $H_{m'vc}$ imposes no restrictions concerning relationships between $\alpha^{(1)},\alpha^{(2)},\dots,\alpha^{(k)}$, and indeed differs from the hypothesis H_{vc} only in imposing the relationship $\mu_0^{(1)} = \mu_0^{(2)} = \dots = \mu_0^{(k)} = \mu_0$ on the marginal expected scores of individuals on subtest T_0. This additional restriction suggests estimating μ_0 by the pooled estimator $\bar{\bar{x}}_0$ defined in (3.7), thus freeing the residual cross-product matrix A_{00} to provide additional information about Σ_{00}. Since under H_{vc} we estimate Σ_{00}, β, and $\Sigma_{11\cdot 0}$ by $(N)^{-1}W_{00}$, $W_{00}^{-1}W_{01}$, and $(N)^{-1}W_{11\cdot 0} = [W_{11} - W_{10}W_{00}^{-1}W_{01}]$, respectively, the arguments used in Subsection 3.2 of this paper lead us to think of estimating β and $\Sigma_{11\cdot 0}$ by $W_{00}^{-1}W_{01}$ and $(N)^{-1}W_{11\cdot 0}$, respectively, and to think of estimating Σ_{00} by using the pooled estimator $(N)^{-1}[W_{00} + A_{00}]$. Finally, since $H_{m'vc}$ imposes no relationships among $\alpha^{(1)},\alpha^{(2)},\dots,\alpha^{(k)}$ beyond those imposed by H_{vc}, we can estimate $\alpha^{(1)},\alpha^{(2)},\dots,\alpha^{(k)}$ by the estimators used to estimate these parameters under H_{vc} ; namely, $\bar{x}_1^{(1)} - \bar{x}_0^{(1)}W_{00}^{-1}W_{01}, \bar{x}_1^{(2)} - \bar{x}_0^{(2)}W_{00}^{-1}W_{01}$, ..., and $\bar{x}_1^{(k)} - \bar{x}_0^{(k)}W_{00}^{-1}W_{01}$. In the Appendix, it is shown that these estimators:

$$\hat{\mu}_0(m'vc) = \bar{\bar{x}}_0 \quad , \quad \hat{\alpha}^{(g)}(m'vc) = \bar{x}_1^{(g)} - \bar{x}_0^{(g)}W_{00}^{-1}W_{01} \quad , \quad g = 1,2,\ldots,k \quad ,$$

$$(3.14) \qquad \hat{\Sigma}_{00}(m'vc) = \frac{1}{N}(W_{00} + A_{00}) \quad , \quad \hat{\beta}(m'vc) = W_{00}^{-1}W_{01} \quad ,$$

$$\hat{\Sigma}_{11\cdot 0}(m'vc) = \frac{1}{N}(W_{11} - W_{10}W_{00}^{-1}W_{01}) \quad ,$$

are indeed MLE of the corresponding parameters. To obtain MLE of $(\underset{\sim}{\mu},\underset{\sim}{\Sigma})$, we can substitute (3.14) into (3.2), taking account of the equalities (3.13), and obtain

$$\hat{\mu}^{(g)}(m'vc) = (\hat{\mu}_0(m'vc),\hat{\mu}_1^{(g)}(m'vc))$$

$$= (\bar{\bar{x}}_0, \bar{x}_1^{(g)} - (\bar{x}_0^{(g)} - \bar{\bar{x}}_0)W_{00}^{-1}W_{01}) \quad , \quad g = 1,2,\ldots,k \quad ,$$

(3.15)

$$\hat{\Sigma}(m'vc) = \frac{1}{N}\begin{pmatrix} W_{00} + A_{00} & (W_{00} + A_{00})W_{00}^{-1}W_{01} \\ W_{10}W_{00}^{-1}(W_{00} + A_{00}) & W_{11\cdot 0} + W_{10}W_{00}^{-1}(W_{00} + A_{00})W_{00}^{-1}W_{01} \end{pmatrix} \quad ,$$

as the MLE of $\mu^{(1)},\mu^{(2)},\ldots,\mu^{(k)}$, and Σ respectively. Values of the MLE (3.14) and (3.15) for the example described in Subsection 3.0 are summarized in Table 6.

Table 6. Maximum Likelihood Estimators under $H_{m'vc}$ for the Illustrative Example

$$\hat{\mu}^{(1)}(m'vc) = (34.51,\ 24.12,\ 15.15) \quad ,$$

$$\hat{\mu}^{(2)}(m'vc) = (34.51,\ 24.12,\ 14.77) \quad ,$$

$$\hat{\mu}^{(3)}(m'vc) = (34.51,\ 24.12,\ 15.61) \quad ,$$

$$\hat{\Sigma}(m'vc) = \begin{pmatrix} 1452.98 & 971.72 & 629.46 \\ 971.72 & 722.60 & 427.30 \\ 629.46 & 427.30 & 289.11 \end{pmatrix} \quad ,$$

$$\hat{\mu}_0(m'vc) = (34.51,\ 24.12) \quad ,$$

$$\hat{\alpha}^{(1)}(m'vc) = 0.1122 \ , \quad \hat{\alpha}^{(2)}(m'vc) = -0.2712 \ , \quad \hat{\alpha}^{(3)}(m'vc) = 0.5673 \quad ,$$

$$\hat{\Sigma}_{00}(m'vc) = \begin{pmatrix} 1452.98 & 971.72 \\ 971.72 & 722.60 \end{pmatrix} , \quad \hat{\beta}(m'vc) = \begin{pmatrix} 0.3750 \\ 0.0871 \end{pmatrix} , \quad \hat{\Sigma}_{11\cdot 0}(m'vc) = 15.87 \ .$$

3.5 Maximum Likelihood Estimators under H_{mvc}

Under the model defined by H_{mvc} the tests $(T_0,T_1),(T_0,T_2),\ldots,$ (T_0,T_k) when applied to the k randomly chosen groups produce statistically equivalent scores. That is, the scores $(x_{0i}^{(g)},x_{1i}^{(g)})$ for all individuals in all groups are identically distributed with identical expected score vectors $\mu = (\mu_0,\mu_1)$ and identical covariance matrices

$$\Sigma = \begin{pmatrix} \Sigma_{00} & \Sigma_{01} \\ \Sigma_{10} & \Sigma_{11} \end{pmatrix} .$$

The model defined by H_{mvc} thus requires the parameters of the distributions of scores to belong to the parametric subset ω_{mvc} , where

$$\omega_{mvc} = \{(\underset{\sim}{\mu},\underset{\sim}{\Sigma}):\ \mu^{(1)} = \mu^{(2)} = \ldots \mu^{(k)} \equiv \mu \text{ is an arbitrary } 1 \times p \text{ vector,}$$
$$\Sigma^{(1)} = \Sigma^{(2)} = \ldots = \Sigma^{(k)} \equiv \Sigma \text{ is an arbitrary } p \times p \text{ covariance matrix}\} .$$

The MLE of μ and Σ under H_{mvc} are well known [see Anderson (1958)]. These estimators are

$$\hat{\mu}(mvc) = \frac{1}{N} \sum_{g=1}^{k} N_g \bar{x}^{(g)} = \bar{\bar{x}} \quad ,$$

$$(3.16) \qquad \hat{\Sigma}(mvc) = \frac{1}{N} [\sum_{g=1}^{k} V^{(g)} + \sum_{g=1}^{k} N_g (\bar{x}^{(g)} - \bar{\bar{x}})' (\bar{x}^{(g)} - \bar{\bar{x}})]$$

$$= \frac{1}{N} [W + A] \quad .$$

The MLE of the parameters

$$\mu_0 \quad , \quad \alpha = \mu_1 - \mu_0 \Sigma_{00}^{-1} \Sigma_{01} \quad ,$$

$$(3.17)$$

$$\Sigma_{00} \quad , \quad \beta = \Sigma_{00}^{-1} \Sigma_{01} \quad , \quad \Sigma_{11 \cdot 0} = \Sigma_{11} - \Sigma_{10} \Sigma_{00}^{-1} \Sigma_{01} \quad ,$$

are obtained from (3.16) through an obvious substitution. The values of the MLE (3.16) and the MLE (3.17) for the example described in Subsection 3.0 are summarized in Table 7.

3.6 Some Comments

Going back over the lists of MLE under the various hypotheses, certain general rules can be observed to be at work. The assumption of the equality of the marginal covariance matrices $\Sigma_{00}^{(1)}, \Sigma_{00}^{(2)}, \ldots, \Sigma_{00}^{(k)}$ does not affect estimation of the mean vectors $\mu^{(1)}, \mu^{(2)}, \ldots, \mu^{(k)}$ (when compared to MLE for the mean vectors when equality of the $\Sigma_{00}^{(g)}$ is not assumed), but allows adjustment of our estimates of $\Sigma_{01}^{(g)}$ and $\Sigma_{11}^{(g)}$ through a

Table 7. Maximum Likelihood Estimators under H_{mvc} for the Illustrative Example

$$\hat{\mu}(mvc) = (34.51,\ 24.12,\ 15.18) \quad ,$$

$$\hat{\Sigma}(mvc) = \begin{pmatrix} 1452.98 & 971.72 & 629.46 \\ 971.72 & 722.60 & 427.30 \\ 629.46 & 427.30 & 289.11 \end{pmatrix} \quad ,$$

$$\hat{\mu}_0(mvc) = (34.51,\ 24.12) \quad ; \quad \hat{\alpha}(mvc) = 0.1390 \quad ,$$

$$\hat{\Sigma}_{00}(mvc) = \begin{pmatrix} 1452.98 & 971.72 \\ 971.72 & 722.60 \end{pmatrix} \quad , \quad \hat{\beta}(mvc) = \begin{pmatrix} 0.3750 \\ 0.0870 \end{pmatrix} \quad , \quad \hat{\Sigma}_{11\cdot 0}(mvc) = 15.87 \quad .$$

regression of the usual estimators $(N_g)^{-1}V_{01}^{(g)}$ and $(N_g)^{-1}V_{11}^{(g)}$ on the residual $(N_g)^{-1}V_{00}^{(g)} - (N)^{-1}\sum_{g=1}^{k} V_{00}^{(g)}$ around the pooled estimator $(N)^{-1}\sum_{g=1}^{k} V_{00}^{(g)}$ of the common value Σ_{00} of $\Sigma_{00}^{(1)},\Sigma_{00}^{(2)},\ldots,\Sigma_{00}^{(k)}$, $g = 1,2,\ldots,k$ (see Subsection 3.1).

The assumption of the equality of the marginal expected score vectors $\mu_0^{(1)},\mu_0^{(2)},\ldots,\mu_0^{(k)}$, similarly permits us to adjust estimation of the usual estimators $\bar{x}_1^{(1)},\bar{x}_1^{(2)},\ldots,$ and $\bar{x}_1^{(k)}$ of $\mu_1^{(1)},\mu_1^{(2)},\ldots,$ and $\mu_1^{(k)}$ by regressing these estimators on the residuals $\bar{x}_0^{(1)} - \bar{\bar{x}}_0, \bar{x}_0^{(2)} - \bar{\bar{x}}_0,\ldots,$ and $\bar{x}_0^{(k)} - \bar{\bar{x}}_0$ around the pooled estimator $\bar{\bar{x}}_0$ of the common value μ_0 of $\mu_0^{(1)},\mu_0^{(2)},\ldots,\mu_0^{(k)}$. This assumption also frees the residual cross-product matrix A_{00} to help provide additional information for estimating Σ_{00} when it is known that $\Sigma_{00}^{(1)} = \Sigma_{00}^{(2)} = \cdots = \Sigma_{00}^{(k)} = \Sigma_{00}$. The effects of such adjustments on the resulting estimators are illustrated in Tables 2 through 7.

Although the adjusted estimators may provide superior accuracy in comparison to the unadjusted estimators, the distributions of the adjusted estimators are usually more complicated than the distributions of the unadjusted estimators, and do not promise to be directly amenable for the purpose of forming confidence regions for the various parameters. In such cases, the indirect route of obtaining confidence regions for the parameters (3.1) is often more promising, since the MLE of these parameters in many cases have tractable distributions. Since the basic distribution theory for those estimators which do have convenient distributions is known [see Gleser and Olkin (1966), (1969), (1972a), Anderson (1958)], and since our remaining distributional results appear to be too cumbersome for practical use, distribution theory for the MLE is not given in the present paper.

4. Tests of Hypotheses

In previous sections we have described 5 separate hypotheses $H_{vc'}$, $H_{m'vc'}$, H_{vc}, $H_{m'vc}$, and H_{mvc}. These hypotheses specify relations among the parameters of the distributions of test scores on k psychological tests $(T_0,T_1),(T_0,T_2),\ldots,(T_0,T_k)$. In Section 3 we summarized the maximum likelihood estimators (MLE) of the parameters under each of these hypotheses. We also indicated what form the MLE of the parameters took under the general, all-inclusive hypothesis, H_t, in which the parameters ($\mu^{(g)},\Sigma^{(g)}$) of the g -th test score distribution are not necessarily functionally related to the parameters ($\mu^{(h)},\Sigma^{(h)}$) of any other test score distribution, $h \neq g$. In the present section, we describe statistical tests of hypotheses which, upon the basis of the given test score data, allow us to decide which hypothesis of any pair of these hypotheses best describes the parameters of the test score distributions.

4.1 Likelihood Ratio Test Statistic

Let H_a and H_b be any 2 of the 6 hypotheses: H_t, $H_{vc'}$, $H_{m'vc'}$, H_{vc}, $H_{m'vc}$, and H_{mvc}. For example, H_a may be the hypothesis H_{mvc} and H_b may be the hypothesis $H_{m'vc'}$. Assume that hypothesis H_a logically implies H_b. In this case, classical likelihood ratio test theory suggests comparing H_a to H_b by means of the likelihood ratio test statistic

$$(4.1) \qquad \lambda_{a,b} = \frac{\max\limits_{(\underset{\sim}{\mu},\underset{\sim}{\Sigma})\epsilon\omega_a} p(\bar{x},V)}{\max\limits_{(\underset{\sim}{\mu},\underset{\sim}{\Sigma})\epsilon\omega_b} p(\bar{x},V)} ,$$

where ω_a is the subspace of the total parameter space ω_t which corresponds to hypothesis H_a, and ω_b is the subspace of the parameter space which corresponds to hypothesis H_b. Since H_a implies H_b, ω_a is included in ω_b, and thus $\max\limits_{\omega_a} p(\bar{x},V) \leq \max\limits_{\omega_b} p(\bar{x},V)$. Since clearly $\lambda_{a,b} \geq 0$, it follows that $0 \leq \lambda_{a,b} \leq 1$. Values of $\lambda_{a,b}$ close to 1 favor hypothesis H_a, while values of $\lambda_{a,b}$ close to 0 favor hypothesis H_b. If we adopt the approach of Neyman and Pearson to hypothesis testing, we call H_a the null hypothesis, and reject H_a (not necessarily in favor of H_b) if

$$(4.2) \qquad \lambda_{a,b} < \lambda^* ,$$

where λ^* is a certain critical constant obtained from the null distribution of $\lambda_{a,b}$ (that is, the distribution of $\lambda_{a,b}$ when hypothesis H_a describes the parameters ($\underset{\sim}{\mu},\underset{\sim}{\Sigma}$)). If we wish to test H_a versus H_b at a level of significance of γ, $0 < \gamma < 1$, then we choose λ^* to satisfy

$$(4.3) \qquad P\{\lambda_{a,b} < \lambda^*\} \leq \gamma \quad , \text{ all } (\underset{\sim}{\mu},\underset{\sim}{\Sigma}) \epsilon\, \omega_a .$$

In very large samples (i.e., when $N_1, N_2, \ldots, N_k$ are all large, and of the same order of magnitude), it can be shown that the distribution of $-2\log \lambda_{a,b}$ when H_a is true is approximately a chi-square (χ^2)

distribution with $f_{a,b}$ degrees of freedom. Here $f_{a,b}$ is a certain integer which depends upon the hypotheses H_a and H_b, and upon q, p, and k. Let $\chi^2(f,\gamma)$ be that constant which is exceeded with probability γ by a random variable having a chi-square distribution with f degrees of freedom. Then, in large samples, it follows from the above discussion that the critical constant λ^* defined in Equations (4.2) and (4.3) is approximately equal to $\exp[-\frac{1}{2}\chi^2(f_{a,b},\gamma)]$. Hence, in large samples, a likelihood ratio test of H_a versus H_b, at a level of significance of approximately γ, rejects H_a if

$$(4.4) \qquad \lambda_{a,b} \leq \exp[-\tfrac{1}{2}\chi^2(f_{a,b},\gamma)] \quad .$$

Since H_a and H_b can be any 2 of the 6 hypotheses H_t, $H_{vc'}$, $H_{m'vc'}$, H_{vc}, $H_{m'vc}$, and H_{mvc}, a total of $\binom{6}{2} = 15$ pairs of hypotheses can be compared by means of a statistical test of hypothesis. In 14 of these pairs of hypotheses, one of the hypotheses to be compared logically implies the other, so that the classical likelihood ratio test theory described above can be applied to construct a test of these hypotheses. These 14 pairs of hypotheses are listed in the first two columns of Table 8.

In one of the 15 possible pairs of hypotheses, however, neither of the two hypotheses logically implies the other. This pair of hypotheses, $H_{m'vc'}$ and H_{vc}, cannot be compared using the classical likelihood ratio test theory sketched above. We can, of course, construct a likelihood ratio test statistic $\lambda_{a,b}$ of the form (4.1), but choice of a hypothesis to serve as the null hypothesis is arbitrary, $\lambda_{a,b}$ is not necessarily bounded above by 1, and the asymptotic distribution of this test statistic under either $H_{m'vc'}$ or H_{vc} is not necessarily the chi-square distribution. For these reasons, comparison of $H_{m'vc'}$ versus H_{vc} by means of a statistical test of hypothesis would require an entirely separate analysis and discussion. Since it is unlikely that a comparison of $H_{m'vc'}$ with H_{vc} would arise as an important problem in psychological testing contexts, we omit discussion of a test of significance for these two hypotheses.

For each of the 14 pairs of hypotheses for which the likelihood ratio test theory is applicable, we can construct the likelihood ratio test statistic $\lambda_{a,b}$ by making use of the various maxima of the likelihood $p(\bar{x},V)$ described in the Appendix. For example, suppose that we wish to make a statistical test of H_{vc} versus $H_{vc'}$. Note that H_{vc} logically implies $H_{vc'}$, so that $H_a = H_{vc}$ and $H_b = H_{vc'}$ in this comparison. From Equation (A.38) of the Appendix (remembering that $(\underset{\sim}{\mu},\underset{\sim}{\Sigma})$ and the parameterization in terms of the quantities defined in Equation (3.1) are equivalent parameterizations), we find that

$$(4.5) \qquad \max_{(\underset{\sim}{\mu},\underset{\sim}{\Sigma})\epsilon\omega_{vc}} p(\bar{x},V) = H(V)\left|\tfrac{1}{N}W_{00}\right|^{-\frac{1}{2}N}\left|\tfrac{1}{N}W_{11\cdot 0}\right|^{-\frac{1}{2}N} ,$$

where $H(V)$ is defined by Equation (A.21) in the Appendix. Similarly, from Equation (A.25) of the Appendix,

$$(4.6) \qquad \max_{(\underset{\sim}{\mu},\underset{\sim}{\Sigma})\epsilon\omega_{vc'}} p(\bar{x},V) = H(V)\left|\tfrac{1}{N}W_{00}\right|^{-\frac{1}{2}N}\prod_{g=1}^{k}\left|\tfrac{1}{N_g}V_{11\cdot 0}^{(g)}\right|^{-\frac{1}{2}N_g} .$$

Dividing (4.5) by (4.6), we obtain the likelihood ratio test statistic:

$$\lambda_{vc,vc'} = \frac{\max_{(\underset{\sim}{\mu},\underset{\sim}{\Sigma})\in\omega_{vc}} p(\bar{x},V)}{\max_{(\underset{\sim}{\mu},\underset{\sim}{\Sigma})\in\omega_{vc'}} p(\bar{x},V)}$$

(4.7)
$$= \prod_{g=1}^{k} \left(\frac{\left| \frac{1}{N_g} V_{11\cdot 0}^{(g)} \right|}{\left| \frac{1}{N} W_{11\cdot 0} \right|} \right)^{\frac{1}{2}N_g} .$$

Once we have calculated the likelihood ratio test statistic $\lambda_{a,b}$, then if the sample sizes $N_1, N_2, \cdots, N_k$ are large and of the same order of magnitude, we can test H_a versus H_b at an approximate level of significance γ by means of the test which rejects H_a if (4.4) holds. Use of (4.4) requires knowledge of the constant $f_{a,b}$, plus access to tables of the chi-square distribution. The constant $f_{a,b}$ can be obtained from the well-known asymptotic theory of likelihood ratio tests. Values of $f_{a,b}$ for each of the 14 possible tests of hypotheses are listed in the fourth column of Table 8. Thus, the constant $f_{vc,vc'}$ needed to apply the likelihood ratio test of H_{vc} versus $H_{vc'}$ in large samples is given by (see Table 8):

$$f_{vc,vc'} = \frac{(p - q)(k - 1)(p + q + 1)}{2} .$$

Suppose that $p = 3$, $q = 2$, $k = 3$. Then $f_{vc,vc'} = 6$. If we want to test H_{vc} versus $H_{vc'}$ at level of significance $\gamma = 0.05$, then in large samples we would reject hypothesis H_{vc} if

$$\lambda_{vc,vc'} < \exp[-\tfrac{1}{2}\chi^2(6,.05)] = .0018 \quad .$$

In Table 8, we give the likelihood ratio test statistics $\lambda_{a,b}$ for 4 of the 14 possible comparisons of hypotheses (namely, H_{mvc} versus $H_{m'vc}$, H_{mvc} versus H_{vc}, $H_{m'vc}$ versus H_{vc}, and $H_{m'vc'}$ versus $H_{vc'}$). For the remaining 10 comparisons, we recommend a modification of the likelihood ratio test statistic along the lines first suggested by Bartlett (1937). From the modified statistic $L_{a,b}$ given in Table 8, however, the likelihood ratio test statistic $\lambda_{a,b}$ may easily be obtained by merely substituting N_g for m_g or n_g, $g = 1,2,\ldots,k$, and N for m or n in the formula for $L_{a,b}$. For example, in Table 8, we suggest the statistic

$$(4.8) \qquad L_{vc,vc'} = \prod_{g=1}^{k} \left(\frac{\left| \frac{1}{m_g} V_{11\cdot0}^{(g)} \right|}{\left| \frac{1}{m} W_{11\cdot0} \right|} \right)^{\frac{1}{2}m_g}$$

for testing H_{vc} versus $H_{vc'}$ (here, $m_g = N_g - q - 1$, $g = 1,2,\ldots,k$, and $m = \sum_{g=1}^{k} m_g$). To obtain $\lambda_{vc,vc'}$, we substitute N_g for m_g and N for m everywhere in (4.8); the result is the formula for $\lambda_{vc,vc'}$ already obtained in (4.7).

4.2 Bartlett Modifications of the Likelihood Ratio Test

Consider the likelihood ratio test statistic $\lambda_{vc,t}$ for testing H_{vc} against general alternatives H_t. When $k = 2$ and $N_1 \neq N_2$, it is known that the test of hypothesis which rejects H_{vc} when $\lambda_{vc,t} < \lambda^*$ is a biased test [Das Gupta (1969)]. In the univariate case ($p = 1$), Bartlett (1937) suggests modifying the likelihood ratio test statistic for testing the equality of variances among k populations by replacing the sample sizes N_g by the degrees of freedom n_g of the estimators of the variances of the g -th population, $g = 1,2,\ldots,k$, everywhere these quantities (the N_g's) appear in the formula for the likelihood ratio test statistic. Anderson (1958; p. 249) proposes a similar modification of the likelihood ratio test statistic $\lambda_{vc,t}$ for testing the equality of the covariance matrices among k populations. When $N_1 = N_2 = \cdots = N_k$, the likelihood ratio test statistic $\lambda_{vc,t}$ and the Bartlett-type modification $L_{vc,t}$ (see Table 8) of this test statistic are monotonic functions of one another, so that in this case $\lambda_{vc,t}$ and $L_{vc,t}$ yield equivalent tests. That is, if we construct a test of H_{vc} versus H_t of level of significance γ which is based on $\lambda_{vc,t}$, and a test of H_{vc} versus H_t of level of significance γ which is based on $L_{vc,t}$ (and which rejects H_{vc} for small values of $L_{vc,t}$), then the test based on $\lambda_{vc,t}$ rejects H_{vc} if and only if the test based on $L_{vc,t}$ rejects H_{vc} .

When at least two N_g's are unequal, however, the tests of H_{vc} versus H_t based on $\lambda_{vc,t}$ and $L_{vc,t}$, respectively, are not the same. In particular, the test based on $\lambda_{vc,t}$ is biased [Das Gupta (1969), Sugiura and Nagao (1968)]. The difference between the tests is most pronounced for small and moderate sample sizes. For large samples $\lambda_{vc,t}$ and $L_{vc,t}$ are approximately equal to one another, and a test which rejects H_{vc} when

$$L_{vc,t} < \exp[-\tfrac{1}{2}\chi^2(f_{vc,t},\gamma)]$$

has level of significance approximately equal to γ .

Next consider the test of H_{mvc} versus H_t based on the likelihood ratio test statistic $\lambda_{mvc,t}$. It can be shown [see Anderson (1958; Chapter 10)] that

$$\lambda_{mvc,t} = \lambda_{mvc,vc}\lambda_{vc,t} \quad ,$$

where $\lambda_{mvc,vc}$ is the likelihood ratio test statistic for testing H_{mvc} versus H_{vc} . Anderson suggests that since a Bartlett modification of $\lambda_{vc,t}$ improves the properties of the test of H_{vc} versus H_t , the identical Bartlett modification of $\lambda_{mvc,t}$ will improve the properties of the test of H_{mvc} versus H_t . As far as the property of unbiasedness of a test is concerned, Anderson's conjecture is correct. That is, whereas the test of H_{mvc} versus H_t which rejects H_{mvc} for small values of $\lambda_{mvc,t}$ is a biased test when the sample sizes $N_1,N_2,\cdots,N_k$ are not all equal, the Bartlett modification $L_{mvc,t}$ of $\lambda_{mvc,t}$ always yields an unbiased test. The test statistics $\lambda_{mvc,t}$ and $L_{mvc,t}$ yield equivalent tests when $N_1 = N_2 = \cdots = N_k$, and are nearly equal for large sample sizes, regardless of whether the sample sizes $N_1,N_2,\cdots,N_k$ are equal or not. The Bartlett modification of the likelihood ratio test statistic $\lambda_{mvc,t}$ thus has greatest effect upon the properties of the resulting test of H_{mvc} versus H_t for small or moderate sample sizes $N_1,N_2,\cdots,N_k$, which are not all equal to one another.

Bartlett modifications for the test statistics $\lambda_{vc',t}$ and $\lambda_{m'vc',t}$ are justified by the same arguments used above to justify Bartlett modifications of $\lambda_{vc,t}$ and $\lambda_{mvc,t}$. In all of these cases the Bartlett modification $L_{a,b}$ of $\lambda_{a,b}$ is formed by replacing the sample sizes N_g by the degrees of freedom $n_g = N_g - 1$ of the appropriate estimator of the covariance matrix { $(N_g)^{-1}V^{(g)}$ as an estimator of $\Sigma^{(g)}$ in the case of $\lambda_{vc,t}$ and $\lambda_{mvc,t}$, and $(N_g)^{-1}V_{00}^{(g)}$ as an estimator of $\Sigma_{00}^{(g)}$ in the case of $\lambda_{vc',t}$ and $\lambda_{m'vc',t}$ } in the formula for $\lambda_{a,b}$. Note that in computing and analyzing $\lambda_{vc',t}$ (or $L_{vc',t}$) and $\lambda_{m'vc',t}$ (or $L_{m'vc',t}$), we can act as if only observations of scores $x_{0i}^{(g)}$ of individuals on subtest T_0 have been obtained.

Let us next turn to $\lambda_{vc,vc'}$. Note from (4.7) that

$$\lambda_{vc,vc'} = \prod_{g=1}^{k} \left(\frac{\left| \frac{1}{N_g} V_{11\cdot 0} \right|}{\left| \frac{1}{N} W_{11\cdot 0} \right|} \right)^{\frac{1}{2}N_g}$$

(4.9)

$$= \left[\prod_{g=1}^{k} \left(\frac{\left| \frac{1}{N_g} V_{11\cdot 0}^{(g)} \right|}{\left| \frac{1}{N} \sum_{g=1}^{k} V_{11\cdot 0}^{(g)} \right|} \right)^{\frac{1}{2}N_g} \right] \left(\frac{\left| \frac{1}{N} \sum_{g=1}^{k} V_{11\cdot 0}^{(g)} \right|}{\left| \frac{1}{N} W_{11\cdot 0} \right|} \right)^{\frac{1}{2}N} .$$

It is known that $V_{11\cdot 0}^{(g)}$ has a Wishart distribution with $m_g = N_q - q - 1$ degrees of freedom and expected value $E(V_{11\cdot 0}^{(g)}) = m_g \Sigma_{11\cdot 0}^{(g)}$, $g = 1,2,\ldots,k$.

The quantity

$$(4.10)\qquad U_1 = \prod_{g=1}^{k} \left(\frac{\left| \frac{1}{N_g} V_{11\cdot 0}^{(g)} \right|}{\left| \frac{1}{N} \sum_{g=1}^{k} V_{11\cdot 0}^{(g)} \right|} \right)^{\frac{1}{2}N_g}$$

has the form of a likelihood ratio test statistic for testing the equality of the residual covariance matrices $\Sigma_{11\cdot 0}^{(g)}$, $g = 1,2,\ldots,k$ against general alternatives. [Indeed, it can be shown that U_1 is the likelihood ratio test statistic for testing the hypothesis $\Sigma_{11\cdot 0}^{(1)} = \Sigma_{11\cdot 0}^{(2)} = \cdots = \Sigma_{11\cdot 0}^{(k)}$ against the hypothesis H_t .] Since

$$W_{11\cdot 0} = W_{11} - W_{10}W_{00}^{-1}W_{01}$$

$$(4.11)\qquad = \sum_{g=1}^{k} V_{11\cdot 0}^{(g)} + \sum_{g=1}^{k} (B^{(g)} - W_{00}^{-1}W_{01})' V_{00}^{(g)} (B^{(g)} - W_{00}^{-1}W_{01}) \quad ,$$

it can be seen that

$$(4.12)\qquad U_2 = \left(\frac{\left| \frac{1}{N} \sum_{g=1}^{k} V_{11\cdot 0}^{(g)} \right|}{\left| \frac{1}{N} W_{11\cdot 0} \right|} \right)^{\frac{1}{2}N}$$

somewhat resembles the likelihood ratio test statistic of MANOVA. [Actually, U_2 is the likelihood ratio test statistic for testing H_{vc} against the hypothesis that $\Sigma_{11\cdot 0}^{(1)} = \Sigma_{11\cdot 0}^{(2)} = \cdots = \Sigma_{11\cdot 0}^{(k)}$, and $\Sigma_{00}^{(1)} = \Sigma_{00}^{(2)} = \cdots = \Sigma_{00}^{(k)}$.] Using the arguments presented earlier which justified modifying $\lambda_{vc,t}$, it would appear that the performance of U_1 as the basis of a test of equality of covariances would be improved if in the formula (4.8) for U_1 we everywhere replaced N_g by the degrees of freedom m_g of $V_{11\cdot 0}^{(g)}$, $g = 1,2,\ldots,k$, and replaced $N = \sum_{g=1}^{k} N_g$ by $m = \sum_{g=1}^{k} m_g$. Then, using the arguments used earlier to justify modifying $\lambda_{mvc,t}$, it seems appropriate to make a similar modification in the formula (4.9) for $\lambda_{vc,vc'}$. That is, we modify $\lambda_{vc,vc'}$ by replacing N_g by m_g, $g = 1,2,\ldots,k$, and N by m everywhere in Equation (4.9). We call the resulting statistic,

$$L_{vc,vc'} = \prod_{g=1}^{k} \left(\frac{\left| \frac{1}{m_g} V_{11\cdot 0}^{(g)} \right|}{\left| \frac{1}{m} W_{11\cdot 0} \right|} \right)^{\frac{1}{2}m_g} ,$$

the Bartlett modification of $\lambda_{vc,vc'}$. Since $\lambda_{m'vc,m'vc'} = \lambda_{vc,vc'}$, $\lambda_{mvc,m'vc'} = \lambda_{vc,vc'}\lambda_{mvc,m'vc}$, $\lambda_{mvc,vc'} = \lambda_{vc,vc'}\lambda_{mvc,vc}$, and $\lambda_{m'vc,vc'} = \lambda_{vc,vc'}\lambda_{m'vc,vc}$, it seems appropriate to make modifications of the likelihood ratio test statistics $\lambda_{m'vc,m'vc'}$, $\lambda_{mvc,m'vc'}$, $\lambda_{mvc,vc'}$, and $\lambda_{m'vc,vc'}$ similar to the modification which we have just made to $\lambda_{vc,vc'}$. These Bartlett modifications are exhibited in Table 8.

For every Bartlett modification $L_{a,b}$ of a likelihood ratio statistic $\lambda_{a,b}$ shown in Table 8, the following comments apply:

(i) A test of H_a versus H_b based on $L_{a,b}$ rejects H_a when

$$L_{a,b} < L^* ,$$

where L^* is determined from the distribution of $L_{a,b}$ when H_a is true.

(ii) When the sample sizes are equal $L_{a,b}$ and $\lambda_{a,b}$ are monotone functions of one another and hence lead to equivalent tests of H_a versus H_b.

(iii) When the sample sizes are unequal, and are either small or moderate in size, the test of H_a versus H_b of level of significance γ which rejects H_a when $\lambda_{a,b} < \lambda^*$ is not the same test as the test of level of significance γ which rejects H_a when $L_{a,b} < L^*$. In certain cases, it is known that the former test is a biased test, while the latter test is unbiased. It is conjectured that the test based on $\lambda_{a,b}$ is always a biased test, while the test based on $L_{a,b}$ is always an unbiased test.

(iv) When the sample sizes are large, $\lambda_{a,b}$ and $L_{a,b}$ are approximately equal; further, the test which rejects H_a when

$$\lambda_{a,b} < \exp[-\tfrac{1}{2}\chi^2(f_{a,b},\gamma)] \quad ,$$

and the test which rejects H_a when

$$L_{a,b} < \exp[-\tfrac{1}{2}\chi^2(f_{a,b},\gamma)] \quad ,$$

both have level of significance approximately equal to γ.

For each of the 14 pairs of hypotheses which have been covered by our discussion in this section, Table 8 lists the test statistic ($\lambda_{a,b}$ or $L_{a,b}$) which is recommended for testing this pair of hypotheses, and the degrees of freedom $f_{a,b}$ of its asymptotic chi-square distribution under the null hypothesis (H_a). Table 8 also provides cross-references to other articles [or to Anderson's (1958) textbook] in which some of these hypothesis testing problems are considered.

4.3 Asymptotic Expansion for the Null Distribution

Each of the likelihood ratio test statistics listed in Table 8 is a ratio of products of powers of determinants of certain random Wishart-distributed matrices. The exact null distribution of each such test statistic can be shown to be the same as the distribution of a product of powers of certain independent beta variates. Thus [see Box (1949), Anderson (1958; pp. 203-209)], when $N_1, N_2, \ldots, N_k$ are all moderately large (say, $N_g \geq 3(p^2 + k^2)$, $g = 1,2,\ldots,k$), the null cumulative distribution function of $-2 \log T_{a,b}$, where $T_{a,b} = \lambda_{a,b}$ or $T_{a,b} = L_{a,b}$ depending on the hypotheses H_a and H_b to be compared, may be approximated as follows:

$$P\{-2\log T_{a,b} \leq \tau\} \doteq (1 - \phi_{a,b})P\{X^2(f_{a,b}) \leq \rho_{a,b}\tau\} + \phi_{a,b}P\{X^2(f_{a,b} + 4) \leq \rho_{a,b}\tau\} + O(N^{-3}) \quad , \tag{4.13}$$

where $\chi^2(f)$ represents a random variable having a chi-square distribution with f degrees of freedom, $f_{a,b}$ is the degrees of freedom of the asymptotic null distribution of the likelihood ratio test statistic (given in column 4 of Table 8), and $\phi_{a,b}$ and $\rho_{a,b}$ are constants depending on $N_1, N_2, \cdots, N_k$, p, q, k, and the hypotheses (H_a and H_b) being compared. Given a desired level of significance γ, we may use (4.13) to obtain the critical constant T^* for the test of H_a versus H_b which rejects H_a when

(4.14) $$T_{a,b} < T^* \quad .$$

To do this, we first find a number $t_{a,b}(\gamma)$ which satisfies

(4.15) $$(1 - \phi_{a,b})P\{\chi^2(f_{a,b}) \leq t_{a,b}(\gamma)\} + \phi_{a,b}P\{\chi^2(f_{a,b} + 4) \leq t_{a,b}(\gamma)\} = \gamma \quad .$$

Then

(4.16) $$T^* = \exp -\tfrac{1}{2} \left[\frac{t_{a,b}(\gamma)}{\rho_{a,b}} \right] \quad .$$

Thus, an approximate test of significance of level γ for H_a versus H_b rejects H_a when (4.14) holds, where T^* is given by (4.16), and $T_{a,b} = \lambda_{a,b}$ or $T_{a,b} = L_{a,b}$ depending upon the hypotheses H_a and H_b which are to be compared.

Table 9 gives formulas for obtaining values of $\rho_{a,b}$ and $\phi_{a,b}$. With few exceptions, explicit formulas for $\rho_{a,b}$ and $\phi_{a,b}$ in terms of the basic dimensions q , p , k , $N_1, N_2, \ldots, N_k$ are very long and complex (this is particularly the case for $\phi_{a,b}$). Thus, for the sake of compactness, we have found it expedient to give explicit formulas for ρ and ϕ in only a few cases; the other ρ's and ϕ's in Table 9 are then expressed as certain functions of these explicitly defined ρ's and ϕ's. The functions needed to achieve the above-mentioned compactification in Table 9 are defined as follows.

$$(4.17) \qquad \Gamma(a,b,c;\kappa) = \frac{\kappa f_{a,b}\rho_{a,b} + f_{b,c}\rho_{b,c}}{f_{a,b} + f_{b,c}} \quad ,$$

and

$$(4.18) \qquad \phi(a,b,c;\kappa) = \frac{1}{\Gamma^2(a,b,c;\kappa)} [\kappa^2\rho^2_{a,b}\phi_{a,b} + \rho^2_{b,c}\phi_{b,c} + \frac{f_{a,b}f_{b,c}}{4(f_{a,b} + f_{b,c})} (\kappa\rho_{a,b} - \rho_{b,c})^2] \quad .$$

Motivation for use of the functions defined by (4.17) and (4.18) can be found in Gleser and Olkin (1972b). Here, we illustrate how to use these two functions, and Table 9, to obtain $\rho_{mvc,m'vc'}$, and $\phi_{mvc,m'vc'}$ when (as in the illustrative example of Subsection 3.0) $q = 2$, $p = 3$, $k = 3$, and $N_1 = N_2 = N_3 = 100$.

Looking at Table 9 and Equations (4.17) and (4.18), we see that to determine $\rho_{mvc,m'vc'}$ and $\phi_{mvc,m'vc'}$, we need to first find the values of $f_{mvc,m'vc}$, $\rho_{mvc,m'vc}$, $\phi_{mvc,m'vc}$, $f_{m'vc,m'vc'}$, $\rho_{m'vc,m'vc'}$, and $\phi_{m'vc,m'vc'}$. Explicit formulas for these quantities (in terms of p , q , k , $N_1, N_2, \ldots, N_k$) are given in Tables 8 and 9. From Table 8,

$$f_{mvc,m'vc} = (p - q)(k - 1) = 2 \quad ,$$

$$f_{m'vc,m'vc'} = \frac{(p - q)(p + q + 1)(k - 1)}{2} = 6 \quad .$$

From Table 9,

$$\rho_{mvc,m'vc} = \frac{2N - p - q - k - 2}{2N} = \frac{590}{600} = .98333 \quad ,$$

$$\phi_{mvc,m'vc} = \frac{(p - q)(k - 1)[(p - q)^2 + (k - 1)^2 - 5]}{12(2N - p - q - k - 1)^2} = 0 \quad ,$$

$$\rho_{m'vc,m'vc'} = 1 - \left(\sum_{g=1}^{3} \frac{1}{m_g} - \frac{1}{m} \right)\left[\frac{q(p - q)^2 + 3(p - q) - 1}{6(k - 1)(p + q + 1)(p - q)} \right] - \frac{q(p - qk - 1)}{(p + q + 1)m}$$

$$= .99351 \quad .$$

and

$$\phi_{m'vc,m'vc'} = \frac{(p - q)}{48\rho^2_{m'vc,m'vc'}} \left\{ \sum_{g=1}^{3} \left(\frac{1}{m_g} \right)^2 - \frac{1}{m^2} \quad [(p - q)^2 - 1][p - q + 2] \right.$$

$$- 6(k - 1)(p + q + 1)(1 - \rho_{m'vc,m'vc'})^2$$

$$\left. + \frac{q(k - 1)}{m^2} [3(p + 1 - kq)^2 + (p - q)^2 + (k - 1)^2 q^2 - 5] \right\}$$

$$= \frac{1873}{(12)(2621)^2} = .00002 \quad .$$

Thus, from Table 9 and (4.17),

$$\rho_{mvc,m'vc'} = \frac{(\Sigma \frac{1}{m_g} - \frac{1}{m})[q(p-q)^2 + 3(p-q) - 1]}{6(p+q+3)(k-1)}$$

$$+ \frac{(q+1)(p+q-qk-k)}{(p+q+3)m}$$

$$= 1.00020 \quad .$$

From Table 9 and (4.18),

$$\phi_{mvc,m'vc'} = \Phi(mvc,m'vc,m'vc'; \frac{N}{m})$$

$$= \frac{1}{48\rho^2_{mvc,m'vc'}} \{(\frac{N}{m})^2 \rho^2_{mvc,m'vc}\phi_{mvc,m'vc} + \rho^2_{m'vc,m'vc'}\phi_{m'vc,m'vc'}$$

$$+ \frac{f_{mvc,m'vc}f_{m'vc,m'vc'}}{4(f_{mvc,m'vc} + f_{m'vc,m'vc'})}(\frac{N}{m}\rho_{mvc,m'vc} - \rho_{m'vc,m'vc'})^2\}$$

$$= \frac{1}{48(1.00444)^2} = \{(\frac{300}{291})^2(.98500)^2(0)^2 + (1.00076)^2(.00002)$$

$$+ \frac{(2)(6)}{(4)(8)}[\frac{300}{291}(.98500) - 1.00076]^2\}$$

$$= .0000C? \quad .$$

In the above example, $\rho_{mvc,m'vc}$, $\rho_{m'vc,m'vc'}$, and $\rho_{mvc,m'vc}$ were all close to 1, and $\phi_{mvc,m'vc}$, $\phi_{m'vc,m'vc'}$, and $\phi_{mvc,m'vc'}$ were all very close to 0. This result is not exceptional. For every test of hypothesis represented in Table 9, it can be shown that $\rho_{a,b} \to 1$ and $\phi_{a,b} \to 0$ as $N_1, N_2, \ldots, N_k$ all tend to ∞. (This fact follows since the limiting null distribution of $-2\log T_{a,b}$ is a chi-square distribution with $f_{a,b}$ degrees of freedom.) In general, $\phi_{a,b}$ is closer to 0 in large samples than $\rho_{a,b}$ is close to 1. For example, in the case considered above, the three ϕ -values were all 0 to four decimal places, while the ρ -values were .98333, .99351, and 1.00020, respectively. When $\phi_{a,b}$ is very close to 0, but $\rho_{a,b}$ is not so close to 1 that we can set $\rho = 1$ without loss of accuracy, $t_{a,b}(\gamma)$ may be found by setting $t_{a,b}(\gamma) = \chi(f_{a,b},\gamma)$, as can be seen by setting $\phi_{a,b} = 0$ in (4.15), and T^* may be found from (4.16). That is, when $\phi_{a,b}$ is very close to 0,

$$T^* = \exp -\tfrac{1}{2}\left[\frac{\chi^2(f_{a,b},\gamma)}{\rho_{a,b}} \right] \quad .$$

Of course, when $N_1, N_2, \ldots, N_k$ are so large that $\rho_{a,b} \doteq 1$ and $\phi_{a,b} \doteq 0$ (to several decimal place accuracy), then T^* may be found from the formula,

$$T^* = \exp -\tfrac{1}{2}\, \chi^2(f_{a,b},\gamma) \quad .$$

Table 8. Recommended Test Statistic for Each of the 14 Tests of Hypotheses

Hypothesis H_a	H_b	Recommended Statistic	Degrees of Freedom $f_{a,b}$ for Asymptotic Null Distribution	Comments
H_{mvc}	$H_{m'vc}$	$\lambda_{mvc,m'vc} = \left(\frac{\lvert W_{00} + A_{00} \rvert \lvert W_{11\cdot 0} \rvert}{\lvert W + A \rvert} \right)^{\frac{1}{2}N}$	$(p - q)(k - 1)$	1/
$H_{m'vc}$	H_{vc}	$\lambda_{m'vc,vc} = \left(\frac{\lvert W_{00} \rvert}{\lvert W_{00} + A_{00} \rvert} \right)^{\frac{1}{2}N}$	$q(k - 1)$	2/ 3/
H_{mvc}	H_{vc}	$\lambda_{mvc,vc} = \left(\frac{\lvert W \rvert}{\lvert W + A \rvert} \right)^{\frac{1}{2}N}$	$p(k - 1)$	2/
$H_{m'vc}$	$H_{m'vc'}$	$L_{m'vc,m'vc'} = \prod_{g=1}^{k} \left(\frac{\lvert \frac{1}{m_g} V_{11\cdot 0}^{(g)} \rvert}{\lvert \frac{1}{m} W_{11\cdot 0} \rvert} \right)^{\frac{1}{2}m_g}$	$\frac{(p - q)(p + q + 1)(k - 1)}{2}$	
H_{mvc}	$H_{m'vc'}$	$L_{mvc,m'vc'} = \prod_{g=1}^{k} \left(\frac{\lvert \frac{1}{m_g} V_{11\cdot 0}^{(g)} \rvert \lvert \frac{1}{m} (W_{00} + A_{00}) \rvert}{\lvert \frac{1}{m} (W + A) \rvert} \right)^{\frac{1}{2}m_g}$	$\frac{(p - q)(p + q + 3)(k - 1)}{2}$	

Table 8 (Continued)

$H_{m'vc'}$	$H_{vc'}$	$\lambda_{m'vc',vc'} = \left(\frac{\lvert W_{00}\rvert}{\lvert W_{00} + A_{00}\rvert}\right)^{\frac{1}{2}N} = \lambda_{m'vc,vc}$	$q(k - 1)$	2/ 3/
H_{vc}	$H_{vc'}$	$L_{vc,vc'} = \prod_{g=1}^{k}\left(\frac{\lvert\frac{1}{m_g} V_{11\cdot 0}^{(g)}\rvert}{\lvert\frac{1}{m} W_{11\cdot 0}\rvert}\right)^{\frac{1}{2}m_g} = L_{m'vc,m'vc'}$	$\frac{(p - q)(p + q + 1)(k - 1)}{2}$	
$H_{m'vc}$	$H_{vc'}$	$L_{m'vc,vc'} = \prod_{g=1}^{k}\left(\frac{\lvert\frac{1}{m_g} V_{11\cdot 0}^{(g)}\rvert \lvert W_{00}\rvert}{\lvert\frac{1}{m} W_{11\cdot 0}\rvert \lvert W_{00} + A_{00}\rvert}\right)^{\frac{1}{2}m_g}$	$\frac{(p + q)(p - q + 1)(k - 1)}{2}$	
H_{mvc}	$H_{vc'}$	$L_{mvc,vc'} = \prod_{g=1}^{k}\left(\frac{\lvert\frac{1}{m_g} V_{11\cdot 0}^{(g)}\rvert \lvert\frac{1}{m} W_{00}\rvert}{\lvert\frac{1}{m}(W + A)\rvert}\right)^{\frac{1}{2}m_g}$	$\frac{(p^2 - q^2 + 3p - q)(k - 1)}{2}$	
H_{vc}	H_t	$L_{vc',t} = \prod_{g=1}^{k}\left(\frac{\lvert\frac{1}{n_g} V_{00}^{(g)}\rvert}{\lvert\frac{1}{n} W_{00}\rvert}\right)^{\frac{1}{2}n_g}$	$\frac{q(q + 1)(k - 1)}{2}$	3/ 4/

Table 8 (Continued)

$H_{m'vc'}$	H_t	$L_{m'vc',t} = \prod_{g=1}^{k} \left(\frac{\lvert \frac{1}{n_g} V_{00}^{(g)} \rvert}{\lvert \frac{1}{n} (W_{00} + A_{00}) \rvert} \right)^{\frac{1}{2}n_g}$	$\frac{q(q+3)(k-1)}{2}$	3/ 5/
H_{vc}	H_t	$L_{vc,t} = \prod_{g=1}^{k} \left(\frac{\lvert \frac{1}{n_g} V^{(g)} \rvert}{\lvert \frac{1}{n} W \rvert} \right)^{\frac{1}{2}n_g}$	$\frac{p(p+1)(k-1)}{2}$	4/
$H_{m'vc}$	H_t	$L_{m'vc,t} = \prod_{g=1}^{k} \left(\frac{\lvert \frac{1}{n_g} V^{(g)} \rvert}{\lvert \frac{1}{n} W_{11\cdot 0} \rvert \lvert \frac{1}{n} (W_{00} + A_{00}) \rvert} \right)^{\frac{1}{2}n_g}$	$\frac{[p(p+1)+2q](k-1)}{2}$	
H_{mvc}	H_t	$L_{mvc,t} = \prod_{g=1}^{k} \left(\frac{\lvert \frac{1}{n_g} V^{(g)} \rvert}{\lvert \frac{1}{n} (W + A) \rvert} \right)^{\frac{1}{2}n_g}$	$\frac{p(p+3)(k-1)}{2}$	5/

1/ See Gleser and Olkin (1966). Note that their q is our p - q and vice versa.

2/ Standard MANOVA test. See Chapter 8 of Anderson (1958). Note that his q is our k .

Table 8 (Continued)

3/ Remember that this test is q -dimensional so that Anderson's p is equal to q here.

4/ Standard Test of Equality of Covariance Matrices. See Chapter 10, Section 2 of Anderson (1958). Note that his q is our k .

5/ Standard Simultaneous Test of Equality of Mean Vectors and Equality of Covariance Matrices. See Chapter 10, Section 3 of Anderson (1958). Note that his q is our k .

Note also that $m_g = N_g - q - 1 = n_g - q$, $g = 1,2,\ldots,k$,

and that $m = N - qk - k = n - qk$.

Table 9. Constants Needed to Apply the Approximation (4.13) to the Null Distribution of the Recommended Test Statistic

Test Statistic	$1 - \rho_{a,b}$	$\phi_{a,b}$
$\lambda_{mvc,m'vc}$	$\frac{p + q + k + 2}{2N}$	$\frac{(p - q)(k - 1)[(p - q)^2 + (k - 1)^2 - 5]}{12(2N - p - q - k - 1)^2}$
$\lambda_{m'vc,vc}$	$\frac{q + k + 2}{2N}$	$\frac{q(k - 1)[q^2 + (k - 1)^2 - 5]}{12(2N - p - k - 1)^2}$
$\lambda_{mvc,vc}$	$\frac{p + k + 2}{2N}$	$\frac{p(k - 1)[p^2 + (k - 1)^2 - 5]}{12(2N - p - k - 1)^2}$
$L_{m'vc,m'vc'}$	$\left[\frac{\Delta}{12mf_{m'vc,m'vc'}}\right] + \left[\frac{q(p - qk + 1)}{(p + q + 1)m}\right]$	$\frac{(p - q)}{48\rho^2_{m'vc,m'vc'}} \{[(p - q)^2 - 1][(p - q) + 2][\sum_{g=1}^{k} (\frac{1}{m_g})^2 - \frac{1}{m^2}] - 6(k - 1)(p + q + 1)(1 - \rho_{m'vc,m'vc'})^2 + \frac{q(k - 1)}{m^2}[3(p + 1 - kq)^2 + (p - q)^2 + (k - 1)^2 q^2 - 5]\}$

Table 9 (Continued)

$L_{mvc,m'vc'}$	$\left[\frac{\Delta}{12mf_{mvc,m'vc'}}\right] + \left[\frac{(q+1)(p+2-qk-k)}{(p+q+3)m}\right]$	$\Phi(mvc, m'vc, m'vc'; \frac{N}{m})$
$\lambda_{m'vc',vc'}$	$\frac{q+k+1}{2N}$	$\frac{q(k-1)[q^2+(k-1)^2-5]}{12(2N-q-k-1)^2}$
$L_{vc,vc'}$	$\left[\frac{\Delta}{12mf_{m'vc,vc'}}\right] + \left[\frac{q(p-qk+1)}{(p+q+1)m}\right]$	same as $\Phi_{m'vc,m'vc'}$
$L_{m'vc,vc'}$	$\left[\frac{\Delta}{12mf_{m'vc,vc'}}\right] + \left[\frac{(p-qk)(p-q)+p-k+2-2qk}{(p-q+1)(p+q)m}\right]$	$\Phi(m'vc, vc, vc'; \frac{N}{m})$
$L_{mvc,vc'}$	$\left[\frac{\Delta}{12mf_{mvc,vc'}}\right] + \left[\frac{(p-q)(q)(p-qk+1)+p(p-k+2-2qk)}{(p^2-q^2+3p-q)m}\right]$	$\Phi(mvc, vc, vc'; \frac{N}{m})$
$L_{vc',t}$	$\left[\sum_{g=1}^{k}\frac{1}{n_g}-\frac{1}{N}\right]\left[\frac{2q^2+3q-1}{6(q+1)(k-1)}\right]$	$\frac{q(q+1)}{48(\rho_{vc',t})^2}\{(q-1)(q+2)[\sum_{g=1}^{1}(\frac{1}{n_g})^2-\frac{1}{n^2}]$ $-6(k-1)(1-\rho_{vc',t})^2\}$

Table 9 (Continued)

$L_{m'vc',t}$	$\left[\sum_{g=1}^{k} \frac{1}{n_g} - \frac{1}{n}\right]\left[\frac{2q^2 + 3q - 1}{6(q+3)(k-1)}\right] + \frac{(q - k + 2)}{(q+3)n}$	$\Phi(m'vc', vc', t; \frac{N}{n})$
$L_{vc,t}$	$\left[\sum_{g=1}^{k} \frac{1}{n_g} - \frac{1}{n}\right]\left[\frac{2p^2 + 3p - 1}{6(p+1)(k-1)}\right]$	$\frac{p(p+1)}{48(\rho_{vc,t})^2}\{(p-1)(p+2)[\sum_{g=1}^{k}(\frac{1}{n_g})^2 - \frac{1}{n^2}] - 6(k-1)(1-\rho_{vc,t})^2\}$
$L_{m'vc,t}$	$\left[\sum_{g=1}^{k} \frac{1}{n_g} - \frac{1}{n}\right]\left[\frac{p(2p^2 + 3p - 1)}{12f_{m'vc,t}}\right] + \frac{q(q - k + 2)}{n[p(p+1)\ 2q]}$	$\Phi(m'vc, vc, t; \frac{N}{n})$
$L_{mvc,t}$	$\left[\sum_{g=1}^{k} \frac{1}{n_g} - \frac{1}{n}\right]\left[\frac{2p^2 + 3p - 1}{6(p+3)(k-1)}\right] + \frac{(p - k + 2)}{(p+3)n}$	$\Phi(mvc, vc, t; \frac{N}{n})$

Note: $\Delta = [\sum_{g=1}^{k} \frac{m}{m_g} - 1][p - q][q(p - q)^2 + 3(p - q) - 1]$.

5. An Application

In previous sections, we have described the implications of the hypotheses $H_{vc'}$, $H_{m'vc'}$, H_{vc}, $H_{m'vc}$, and H_{mvc} for the psychological testing situation in which k tests $(T_0,T_1),(T_0,T_2),\ldots,(T_0,T_k)$ are given to k separate groups of individuals. If the k groups of individuals taking the k tests can be regarded as k random samples of individuals from a certain population of individuals, and if the environments in which the k tests are given are homogeneous, then we would expect the distributions of the scores of individuals on the k forms $(T_0,T_1),(T_0,T_2),\ldots,$ (T_0,T_k) to have parameters related by hypothesis H_{mvc} if the k forms are parallel, and by hypothesis $H_{m'vc'}$ whether the k forms are parallel or not.

If the assignment of individuals to forms, or individual-form pairs to testing environments, has not been performed in such a way that differences among the parameters of the k test score distributions can be attributed solely to differences in the forms $(T_0,T_1),(T_0,T_2),\ldots,(T_0,T_k)$, then any of the hypotheses $H_{vc'}$, H_{vc}, or $H_{m'vc}$ may relate the parameters, or there may be no relationships among the parameters (H_t).

In this psychological testing context, an experimenter who believes that his experimental design has made adequate allowance for individual differences and environmental effects upon testing performance would usually start testing hypotheses by comparing the hypothesis H_{mvc} (parallel forms) to the hypothesis $H_{m'vc'}$. In this section, we illustrate the test of these two hypotheses in the context of the example described in Subsection 3.0. There, 3 forms (T_0,T_1), (T_0,T_2), (T_0,T_3) were each given to 100 individuals. The subtest T_0 common to all 3 forms has 2 parts, while the forms as a whole each have 3 parts. Hence, $k = 3$, $q = 2$, $p = 3$; $N_1 = N_2 = N_3 = 100$.

From Table 8, the recommended test statistic for testing H_{mvc} versus $H_{m'vc'}$ is

$$(5.1) \qquad L_{mvc,m'vc'} = \prod_{g=1}^{3} \left(\frac{\left|\frac{1}{m_g} V_{11\cdot 0}^{(g)}\right| \left|\frac{1}{m}(W_{00} + A_{00})\right|}{\left|\frac{1}{m}(W + A)\right|} \right)^{\frac{1}{2}m_g} ,$$

where $m_1 = m_2 = m_3 = 97$, $m = m_1 + m_2 + m_3 = 291$. Because $m_1 = m_2 = m_3$, and

$$\left|\frac{1}{m}(W + A)\right| / \left|\frac{1}{m}(W_{00} + A_{00})\right| = \left|\frac{1}{m} Q_{11\cdot 0}\right| ,$$

where

$$Q_{11\cdot 0} = W_{11} + A_{11} - (W_{10} + A_{10})(W_{00} + A_{00})^{-1}(W_{01} + A_{01}) ,$$

we can rewrite $L_{mvc,m'vc'}$ in the form

$$(5.2) \qquad L_{mvc,m'vc'} = \left(\frac{\prod_{g=1}^{3} \left|V_{11\cdot 0}^{(g)}\right|}{\left|\frac{1}{3} Q_{11\cdot 0}\right|^{3}} \right)^{\frac{1}{2}(97)} .$$

From the data given in Table 1, we find that

$$V_{11\cdot 0}^{(1)} = 1981.93688 , \quad V_{11\cdot 0}^{(2)} = 1540.40294 , \quad V_{11\cdot 0}^{(3)} = 1189.67254 ,$$

and that

$$Q_{11\cdot 0} = 4759.8120 \quad .$$

Note that $|V^{(g)}_{11\cdot 0}| = V^{(g)}_{11\cdot 0}$, $g = 1, 2, 3$, $|\frac{1}{3} Q_{11\cdot 0}| = \frac{1}{3} Q_{11\cdot 0}$ since these quantities are scalars. Thus, from (5.2)

$$(5.3) \qquad L_{mvc,m'vc'} = .00998 \quad .$$

In Subsection 4.3, we indicated that a test of H_{mvc} versus $H_{m'vc'}$ of level of significance approximately equal to γ rejects H_{mvc} when

$$L^* = \exp -\tfrac{1}{2} [\frac{t_{mvc,m'vc'}(\gamma)}{\rho_{mvc,m'vc'}}] \quad ,$$

and $t_{mvc,m'vc'}$ is obtained from (4.15). Since

$$f_{mvc,m'vc'} = 8 \quad , \quad \rho_{mvc,m'vc'} = 1.00444 \quad , \quad \phi_{mvc,m'vc'} = 0.000002 \quad ,$$

we see from (4.15) that

$$t_{mvc,m'vc'}(\gamma) \doteq \chi^2(8,\gamma) \quad ,$$

and that

$$L^* = \exp -\tfrac{1}{2} [\frac{\chi^2(8,\gamma)}{1.00401}] \quad .$$

If we wish to test H_{mvc} versus $H_{m'vc'}$ at level of significance $\gamma = 0.05$, then

$$\chi^2(8,\ 0.05) = 15.507 \quad ,$$

and

$$L^* = \exp -\tfrac{1}{2} \left[\frac{15.507}{1.00401} \right] = .00044 \quad .$$

Since $L_{mvc,m'vc'}$ is greater than L^*, we cannot reject H_{mvc} at the 0.05 level of significance. Thus, the three forms of the SAT can (tentatively) be regarded as parallel forms of the same test.

6. Acknowledgement

We would like to express our thanks to William H. Angoff and June Stern for their help in obtaining the data of Table 1, and to Henrietta Gallagher for her help in performing the computations.

Appendix. Derivation of the Maximum Likelihood Estimators

In Section 2 we noted that the joint density function $p(\bar{x},V)$ of the sufficient statistic $(\bar{x},V)$ has the form

$$p(\bar{x},V) = \prod_{g=1}^{k} p(\bar{x}^{(g)})p(V^{(g)}) \quad ,$$

where $p(\bar{x}^{(g)})$ is given by Equation (2.6), and $p(V^{(g)})$ is given by Equation (2.7), $g = 1,2,\ldots,k$. Adopting the approach to the derivation of the maximum likelihood estimators which was mentioned in Section 3, we break the likelihood $p(\bar{x},V)$ into two factors: (i) the marginal density function of $\bar{x}_0 = (\bar{x}_0^{(1)},\bar{x}_0^{(2)},\ldots,\bar{x}_0^{(k)})$ and $V_{00} = (V_{00}^{(1)},V_{00}^{(2)},\ldots, V_{00}^{(k)})$; and (ii) the conditional density function $p(\bar{x},V|\bar{x}_0,V_{00})$ of $(\bar{x},V)$ given $(\bar{x}_0,V_{00})$.

From (2.6), (2.7) and Theorems 2.4.3 and 7.3.3 of Anderson (1958), the marginal density function $p(\bar{x}_0,V_{00})$ is

$$p(\bar{x}_0,V_{00}) = \prod_{g=1}^{k} p(\bar{x}_0^{(g)})p(V_{00}^{(g)})$$

(A.1) $$= C_0 \prod_{g=1}^{k} \{ |V_{00}^{(g)}|^{\frac{1}{2}(n_g-q-1)} |\Sigma_{00}^{(g)}|^{-\frac{1}{2}N_g}$$

$$\exp -\tfrac{1}{2}[N_g(\bar{x}_0^{(g)} - \mu_0^{(g)})(\Sigma_{00}^{(g)})^{-1}(\bar{x}_0^{(g)} - \mu_0^{(g)})' + \mathrm{tr}(\Sigma_{00}^{(g)})^{-1}V_{00}^{(g)}]\} \quad ,$$

where C_0 is a certain constant. Since the conditional density function $p(\bar{x},V|\bar{x}_0,V_{00})$ of $(\bar{x},V)$ given $(\bar{x}_0,V_{00})$ is equal to $p(\bar{x},V)/p(\bar{x}_0,V_{00})$, it can be shown that

$$p(\bar{x},V|\bar{x}_0,V_{00}) = C_1 \prod_{g=1}^{k} \{|V_{00}^{(g)}|^{-\frac{1}{2}(p-q)}|V_{11\cdot 0}^{(g)}|^{\frac{1}{2}(n_g-p-1)}|\Sigma_{11\cdot 0}^{(g)}|^{-\frac{1}{2}N_g}$$

$$\exp -\tfrac{1}{2}[N_g(\bar{x}_1^{(g)} - \alpha^{(g)} - \bar{x}_0^{(g)}\beta^{(g)})(\Sigma_{11\cdot 0}^{(g)})^{-1}(\bar{x}_1^{(g)} - \bar{x}_0^{(g)}\beta^{(g)})'$$

(A.2)

$$+ \operatorname{tr} V_{00}^{(g)}(B^{(g)} - \beta^{(g)})(\Sigma_{11\cdot 0}^{(g)})^{-1}(B^{(g)} - \beta^{(g)})'$$

$$+ \operatorname{tr}(\Sigma_{11\cdot 0}^{(g)})^{-1}V_{11\cdot 0}^{(g)}]\} \quad ,$$

where C_1 is a certain constant, the parameters $\alpha^{(g)}$, $\beta^{(g)}$, and $\Sigma_{11\cdot 0}^{(g)}$, $g = 1,2,\ldots,k$, are defined by Equation (3.1), and the sample quantities $B^{(g)}$ and $V_{11\cdot 0}^{(g)}$, $g = 1,2,\ldots,k$, are defined by Equation (3.3).

The joint density function $p(\bar{x},V)$ of the sufficient statistic $(\bar{x},V)$ is the product

(A.3) $$p(\bar{x},V) = p(\bar{x},V|\bar{x}_0,V_{00})p(\bar{x}_0,V_{00})$$

of (A.1) and (A.2). In Section 3, we have shown that the parameterization of $p(\bar{x},V)$ in terms of $\mu_0^{(g)}$, $\Sigma_{00}^{(g)}$, $\alpha^{(g)}$, $\beta^{(g)}$, and $\Sigma_{11\cdot 0}^{(g)}$, $g = 1,2,\ldots,k$ is equivalent to the original parameterization of $p(\bar{x},V)$ in terms of $\mu^{(g)}$ and $\Sigma^{(g)}$, $g = 1,2,\ldots,k$. In this Appendix, we find maximum likelihood estimators (MLE) of the parameters $\mu_0^{(g)}$, $\Sigma_{00}^{(g)}$, $\alpha^{(g)}$, $\beta^{(g)}$, and $\Sigma_{11\cdot 0}^{(g)}$, $g = 1,2,\ldots,k$, under each of the hypotheses $H_{vc'}$, $H_{m'vc'}$, H_{vc} , $H_{m'vc}$, and H_{mvc} . These MLE can then be transformed (see Section 3) to obtain the MLE of the original parameters $\mu^{(g)}$ and $\Sigma^{(g)}$, $g = 1,2,\ldots,k$.

To obtain the MLE of the parameters under the various hypotheses, we make repeated use of the following lemmas.

Lemma A.1. Let Z_j be given s x t matrices and H_j be given s x s nonnegative definite matrices, $j = 1,2,\ldots,r$. Assume that $\sum_{j=1}^{r} H_j$ is nonsingular. Then for all s x t matrices Ξ,

$$\text{(A.4)} \quad \sum_{j=1}^{r} (Z_j - \Xi)'H_j(Z_j - \Xi) = \sum_{j=1}^{r} (Z_j - \hat{\Xi})'H_j(Z_j - \hat{\Xi}) + (\hat{\Xi} - \Xi)'(\sum_{j=1}^{r} H_j)(\hat{\Xi} - \Xi) \quad ,$$

where $\hat{\Xi} = (\sum_{j=1}^{r} H_j)^{-1} \sum_{j=1}^{r} H_jZ_j$. Hence for all s x t matrices Ξ and any t x t nonnegative definite matrix Δ,

$$\text{(A.5)} \quad \sum_{j=1}^{r} \text{tr } H_j(Z_j - \Xi)\Delta(Z_j - \Xi)' \geq \sum_{j=1}^{r} \text{tr } H_j(Z_j - \hat{\Xi})\Delta(Z_j - \hat{\Xi})' \quad ,$$

with equality in (A.5) if $\Xi = \hat{\Xi}$.

Proof. Note that

$$\text{(A.6)} \quad (Z_j - \Xi)'H_j(Z_j - \Xi) = (Z_j - \hat{\Xi})'H_j(Z_j - \hat{\Xi}) + (Z_j - \hat{\Xi})'H_j(\hat{\Xi} - \Xi) + (\hat{\Xi} - \Xi)'H_j(Z_j - \hat{\Xi}) + (\hat{\Xi} - \Xi)'H_j(\hat{\Xi} - \Xi) \quad ,$$

and that $\sum_{j=1}^{r} (Z_j - \hat{\Xi})'H_j = \sum_{j=1}^{r} H_j(Z_j - \hat{\Xi}) = 0$. From these two facts, (A.4) follows. Since for all Ξ (including the case when $\Xi = \hat{\Xi}$),

$$\sum_{j=1}^{r} \text{tr } H_j(Z_j - \Xi)\Delta(Z_j - \Xi)' = \text{tr} \sum_{j=1}^{r} (Z_j - \Xi)'H_j(Z_j - \Xi)\Delta \quad ,$$

and since $\text{tr}(\hat{\Xi} - \Xi)'(\sum_{j=1}^{r} H_j)(\hat{\Xi} - \Xi)\Delta \geq 0$, (A.5) follows directly from (A.4). For future use, note that

$$(A.7)\qquad \sum_{j=1}^{r} \operatorname{tr} H_j(Z_j - \hat{\Xi})\Delta(Z_j - \hat{\Xi})' = \operatorname{tr} \Delta[\sum_{j=1}^{r} Z_j' H_j Z_j - \hat{\Xi}'(\sum_{j=1}^{r} H_j)\hat{\Xi}] \quad .$$

Lemma A.2. Let U be a given $t \times t$ nonnegative definite matrix, and let ℓ be a positive integer. Then for all $t \times t$ positive definite matrices Δ,

$$(A.8)\qquad |\Delta|^{-\frac{1}{2}\ell} \exp[-\tfrac{1}{2}\operatorname{tr} \Delta^{-1}U] \leq |\tfrac{1}{\ell} U|^{-\frac{1}{2}\ell} e^{-\frac{1}{2}t\ell} \quad ,$$

with equality holding in (A.8) if $\Delta = (1/\ell)U$.

Proof. This lemma is a direct consequence of Lemma 3.2.2 of Anderson (1958).

A.1 Maximum Likelihood Estimators under General Alternatives

From Equation (A.1),

$$(A.9)\qquad p(\bar{x}_0, V_{00}) = C_0[\prod_{g=1}^{k} |V_{00}^{(g)}|^{\frac{1}{2}(n_g - q - 1)}] f(\underset{\sim}{\mu}_0, \underset{\sim}{\Sigma}_{00}) \quad ,$$

where $\underset{\sim}{\mu}_0 = (\mu_0^{(1)}, \mu_0^{(2)}, \ldots, \mu_0^{(k)})$, $\underset{\sim}{\Sigma}_{00} = (\Sigma_{00}^{(1)}, \Sigma_{00}^{(2)}, \ldots, \Sigma_{00}^{(k)})$,

$$f(\underset{\sim}{\mu}_0, \underset{\sim}{\Sigma}_{00}) = \{\exp[-\tfrac{1}{2}A(\underset{\sim}{\mu}_0, \underset{\sim}{\Sigma}_{00})]\} \prod_{g=1}^{k} \{|\Sigma_{00}^{(g)}|^{-\frac{1}{2}N_g} \exp[-\operatorname{tr}(\Sigma_{00}^{(g)})^{-1} V_{00}^{(g)}]\} \quad ,$$

(A.10)

and

$$A(\underset{\sim}{\mu}_0,\underset{\sim}{\Sigma}_{00}) = \sum_{g=1}^{k} N_g(\bar{x}_0^{(g)} - \mu_0^{(g)})(\Sigma_{00}^{(g)})^{-1}(\bar{x}_0^{(g)} - \mu_0^{(g)})'$$

(A.11)

$$= \sum_{g=1}^{k} \mathrm{tr}\, N_g(\bar{x}_0^{(g)} - \mu_0^{(g)})(\Sigma_0^{(g)})^{-1}(\bar{x}_0^{(g)} - \mu_0^{(g)})' \quad .$$

Note that $A(\underset{\sim}{\mu}_0,\underset{\sim}{\Sigma}_{00}) \geq 0$, with $A(\underset{\sim}{\mu}_0,\underset{\sim}{\Sigma}_{00}) = 0$ if $\mu_0^{(g)} = \bar{x}_0^{(g)}$, $g = 1,2,\ldots,k$. Hence

(A.12) $$f(\underset{\sim}{\mu}_0,\underset{\sim}{\Sigma}_{00}) \leq \prod_{g=1}^{k} \{|\Sigma_{00}|^{-\frac{1}{2}N_g} \exp -\tfrac{1}{2}\mathrm{tr}(\Sigma_{00}^{(g)})^{-1}V_{00}^{(g)}\}$$

with equality when $\mu_0^{(g)} = \bar{x}_0^{(g)}$, $g = 1,2,\ldots,k$. Applying Lemma A.2 to each term in the product on the right-hand side of (A.12), we conclude that

(A.13) $$f(\underset{\sim}{\mu}_0,\underset{\sim}{\Sigma}_{00}) \leq \prod_{g=1}^{k} \{|\frac{1}{N_g} V_{00}^{(g)}|^{-\frac{1}{2}N_g} e^{-\frac{1}{2}qN_g}\} = e^{-\frac{1}{2}qN} \prod_{g=1}^{k} |\frac{1}{N_g} V_{00}^{(g)}|^{-\frac{1}{2}N_g} ,$$

with equality in (A.13) holding if $\mu_0^{(g)} = \bar{x}_0^{(g)}$, $\Sigma_{00}^{(g)} = (N_g)^{-1}V_{00}^{(g)}$, $g = 1,2,\ldots,k$.

Let us turn now to $p(\bar{x},V|\bar{x}_0,V_{00})$. Let $\underset{\sim}{\alpha} = (\alpha^{(1)},\alpha^{(2)},\ldots,\alpha^{(k)})$, $\underset{\sim}{\beta} = (\beta^{(1)},\beta^{(2)},\ldots,\beta^{(k)})$, and $\underset{\sim}{\Sigma}_{11\cdot 0} = (\Sigma_{11\cdot 0}^{(1)},\Sigma_{11\cdot 0}^{(2)},\ldots,\Sigma_{11\cdot 0}^{(k)})$. Define

$$D(\underset{\sim}{\alpha},\underset{\sim}{\beta},\underset{\sim}{\Sigma}_{11\cdot 0})$$

(A.14)

$$= \sum_{g=1}^{k} \mathrm{tr}[N_g(\bar{x}_1^{(g)} - \alpha^{(g)} - \bar{x}_0^{(g)}\beta^{(g)})(\Sigma_{11\cdot 0}^{(g)})^{-1}(\bar{x}_1^{(g)} - \alpha^{(g)} - \bar{x}_0^{(g)}\beta^{(g)})'] \quad ,$$

and

$$\text{(A.15)} \qquad E(\beta,\Sigma_{11\cdot 0}) = \sum_{g=1}^{k} \mathrm{tr}[V_{11\cdot 0}^{(g)}(B^{(g)} - \beta^{(g)})(\Sigma_{11\cdot 0}^{(g)})^{-1}(B^{(g)} - \beta^{(g)})']$$

Then from Equation (A.2),

$$\text{(A.16)} \qquad p(\bar{x},V|\bar{x}_0,V_{00}) = C_1 \prod_{g=1}^{k} \{|V_{00}^{(g)}|^{-\frac{1}{2}(p-q)} |V_{11\cdot 0}^{(g)}|\}^{\frac{1}{2}(n_g-p-1)} h(\alpha,\beta,\Sigma_{11\cdot 0}) \quad ,$$

where

$$\text{(A.17)} \qquad h(\alpha,\beta,\Sigma_{11\cdot 0}) = [\exp -\tfrac{1}{2} D(\alpha,\beta,\Sigma_{11\cdot 0})][\exp -\tfrac{1}{2} E(\beta,\Sigma_{11\cdot 0})] \prod_{g=1}^{k} \{|\Sigma_{11\cdot 0}^{(g)}|^{-\frac{1}{2}N_g} \exp[-\tfrac{1}{2}\mathrm{tr}(\Sigma_{11\cdot 0}^{(g)})^{-1}V_{11\cdot 0}^{(g)}]\} \quad .$$

It is clear by inspection of (A.14) and (A.15) that $D(\alpha,\beta,\Sigma_{11\cdot 0}) \geq 0$ and $E(\beta,\Sigma_{11\cdot 0}) \geq 0$ for all α , β , $\Sigma_{11\cdot 0}$; and that $D(\alpha,\beta,\Sigma_{11\cdot 0}) = 0$ and $E(\beta,\Sigma_{11\cdot 0}) = 0$ if $\alpha^{(g)} = a^{(g)} = \bar{x}_1^{(g)} - \bar{x}_0^{(g)}B^{(g)}$ and $\beta^{(g)} = B^{(g)}$, $g = 1,2,\ldots,k$. Further, from Lemma A.2,

$$|\Sigma_{11\cdot 0}^{(g)}|^{-\frac{1}{2}N_g} \exp[-\tfrac{1}{2}\mathrm{tr}(\Sigma_{11\cdot 0}^{(g)})^{-1}V_{11\cdot 0}^{(g)}] \leq |\frac{1}{N_g} V_{11\cdot 0}^{(g)}|^{-\frac{1}{2}N_g} e^{-\frac{1}{2}(p-q)N_g} \quad ,$$

with equality if $\Sigma_{11\cdot 0}^{(g)} = (N_g)^{-1}V_{11\cdot 0}^{(g)}$, $g = 1,2,\ldots,k$. Hence, we conclude from (A.17) that for all α , β , $\Sigma_{11\cdot 0}$,

$$\text{(A.18)} \qquad h(\alpha,\beta,\Sigma_{11\cdot 0}) \leq e^{-(p-q)N} \prod_{g=1}^{k} |\frac{1}{N_g} V_{11\cdot 0}^{(g)}|^{-\frac{1}{2}N_g} \quad ,$$

with equality holding in (A.18) if $\alpha^{(g)} = a^{(g)}$, $\beta^{(g)} = B^{(g)}$, and $\Sigma_{11\cdot 0}^{(g)} = (N_g)^{-1}V_{11\cdot 0}^{(g)}$, $g = 1,2,\ldots,k$.

Let $\Theta = (\underset{\sim}{\mu}_0, \underset{\sim}{\Sigma}_{00}, \underset{\sim}{\alpha}, \underset{\sim}{\beta}, \underset{\sim}{\Sigma}_{11\cdot 0})$, $\Theta_0 = (\underset{\sim}{\mu}_0, \underset{\sim}{\Sigma}_{00})$, $\Theta_1 = (\underset{\sim}{\alpha}, \underset{\sim}{\beta}, \underset{\sim}{\Sigma}_{11\cdot 0})$. Then $\Theta = (\Theta_0, \Theta_1)$. The largest possible space of possible values for Θ is ω_t , where

$$\omega_t = \{\Theta\colon \Theta_0 \in \omega_{0,t}, \Theta \in \omega_{1,t}\} ,$$

and

$$\omega_{0,t} = \{\Theta_0 = (\underset{\sim}{\mu}_0, \underset{\sim}{\Sigma}_{00}) : \mu_0^{(g)} \text{ an arbitrary } 1 \times q \text{ vector and } \Sigma_{00}^{(g)} \text{ an arbitrary } q \times q \text{ positive definite matrix, } g = 1,2,\ldots,k\} ,$$

$$\omega_{1,t} = \{\Theta_1 = (\underset{\sim}{\alpha}, \underset{\sim}{\beta}, \underset{\sim}{\Sigma}_{11\cdot 0}) : \alpha^{(g)} \text{ an arbitrary } 1 \times (p - q) \text{ vector, } \beta^{(g)} \text{ an arbitrary } q \times (p - q) \text{ vector, and } \Sigma_{11\cdot 0}^{(g)} \text{ an arbitrary } (p - q) \times (p - q) \text{ positive definite matrix, } g = 1,2,\ldots,k\} .$$

Let H_a be a hypothesis which restricts the parameters $\Theta = (\underset{\sim}{\mu}_0, \underset{\sim}{\Sigma}_{00}, \underset{\sim}{\alpha}, \underset{\sim}{\beta}, \underset{\sim}{\Sigma}_{11\cdot 0})$ to a subspace ω_a of ω_t of the form

$$\omega_a = \{(\underset{\sim}{\mu}_0, \underset{\sim}{\Sigma}_{00}, \underset{\sim}{\alpha}, \underset{\sim}{\beta}, \underset{\sim}{\Sigma}_{11\cdot 0})\colon (\underset{\sim}{\mu}_0, \underset{\sim}{\Sigma}_{00}) \in \omega_{0,a} \text{ and } (\underset{\sim}{\alpha}, \underset{\sim}{\beta}, \underset{\sim}{\Sigma}_{11\cdot 0}) \in \omega_{1,a}\} ,$$

where $\omega_{0,a}$ is a subset of $\omega_{0,t}$ and $\omega_{1,a}$ is a subset of $\omega_{1,t}$. It then follows from (A.3), (A.9), and (A.16) that

$$\sup_{\Theta\epsilon\omega_a} p(\bar{x},V) = [C_0C_1 \prod_{g=1}^{k} |V^{(g)}|^{\frac{1}{2}(n_g-p-1)}][\sup_{\Theta_0\epsilon\omega_{0,a}} f(\underset{\sim}{\mu}_0,\underset{\sim}{\Sigma}_{00})]$$

(A.19)

$$[\sup_{\Theta_1\epsilon\omega_{1,a}} h(\underset{\sim}{\alpha},\underset{\sim}{\beta},\underset{\sim}{\Sigma}_{11\cdot 0})] \quad .$$

In particular, it follows from (A.13) and (A.18), and from (A.19), that

$$\text{(A.20)} \qquad \max_{\Theta\epsilon\omega_t} p(\bar{x},V) = H(V) \prod_{g=1}^{k} |\frac{1}{N_g} V^{(g)}|^{-\frac{1}{2}N_g} \quad ,$$

where

$$\text{(A.21)} \qquad H(V) = C_0C_1e^{-\frac{1}{2}pN} \prod_{g=1}^{k} |V^{(g)}|^{\frac{1}{2}(n_g-p-1)} \quad .$$

The maximum in (A.20) is achieved when $\mu_0^{(g)} = \bar{x}_0^{(g)}$, $\Sigma_{00}^{(g)} = (N_g)^{-1}V_{00}^{(g)}$, $\alpha^{(g)} = a^{(g)}$, $\beta^{(g)} = B^{(g)}$, and $\Sigma_{11\cdot 0}^{(g)} = (N_g)^{-1}V_{11\cdot 0}^{(g)}$, $g = 1,2,\ldots,k$, since (as shown above) equality in (A.13) and (A.18) is achieved for these values of the parameters. Thus, the MLE of the parameters under general alternatives are:

$$\hat{\mu}_0^{(g)} = \bar{x}_0^{(g)} \quad , \quad \hat{\Sigma}_{00}^{(g)} = \frac{1}{N_g} V_{00}^{(g)} \quad , \quad \hat{\alpha}^{(g)} = a^{(g)} \quad , \quad \hat{\beta}^{(g)} = B^{(g)} \quad ,$$

$$\hat{\Sigma}_{11\cdot 0}^{(g)} = \frac{1}{N_g} V_{11\cdot 0}^{(g)} \quad ,$$

for $g = 1,2,\ldots,k$, and the maximum of the likelihood is given by (A.20).

A.2 Maximum Likelihood Estimations under $H_{vc'}$

The hypothesis $H_{vc'}$ restricts the parameters $\Theta = (\underset{\sim}{\mu}_0, \underset{\sim}{\Sigma}_{00}, \underset{\sim}{\alpha}, \underset{\sim}{\beta}, \underset{\sim}{\Sigma}_{11\cdot 0})$ to the parametric subspace $\omega_{vc'}$ described in Subsection 3.1. Note that

$$\text{(A.22)} \qquad \omega_{vc'} = \{\Theta\colon \Theta_0 \in \omega_{0,vc'}, \Theta_1 \in \omega_{1,t}\} \quad ,$$

where

$$\omega_{0,vc'} = \{(\underset{\sim}{\mu}_0, \underset{\sim}{\Sigma}_0)\colon (\underset{\sim}{\mu}_0, \underset{\sim}{\Sigma}_0) \in \omega_{0,t}, \Sigma_{00}^{(1)} = \Sigma_{00}^{(2)} = \cdots = \Sigma_{00}^{(k)}\} \quad .$$

Let Σ_{00} represent the common value of $\Sigma_{00}^{(1)}, \Sigma_{00}^{(2)}, \ldots, \Sigma_{00}^{(k)}$ under $H_{vc'}$. Note that when $\Sigma_{00}^{(1)} = \Sigma_{00}^{(2)} = \cdots = \Sigma_{00}^{(k)} = \Sigma_{00}$,

$$\text{(A.23)} \qquad f(\underset{\sim}{\mu}_0, \underset{\sim}{\Sigma}_{00}) = \{\exp[-\tfrac{1}{2}A(\underset{\sim}{\mu}_0, \underset{\sim}{\Sigma}_{00})]\} |\Sigma_{00}|^{-\frac{1}{2}N} \exp -\tfrac{1}{2}\mathrm{tr}\, \Sigma_{00}^{-1} W_{00} \quad ,$$

where $W_{00} = \sum_{g=1}^{k} V_{00}^{(g)}$. Since $A(\underset{\sim}{\mu}_0, \underset{\sim}{\Sigma}_{00}) \geq 0$, with $A(\underset{\sim}{\mu}_0, \underset{\sim}{\Sigma}_{00}) = 0$ if $\mu_0^{(g)} = \bar{x}_0^{(g)}$, $g = 1, 2, \ldots, k$, and by an application of Lemma A.2, we conclude that when $\Sigma_{00}^{(1)} = \Sigma_{00}^{(2)} = \cdots = \Sigma_{00}^{(k)} = \Sigma_{00}$,

$$\text{(A.24)} \qquad f(\underset{\sim}{\mu}_0, \underset{\sim}{\Sigma}_{00}) \leq |\tfrac{1}{N} W_{00}|^{-\frac{1}{2}N} e^{-\frac{1}{2}qN}$$

for all $\underset{\sim}{\mu}_0$, Σ_{00}, with equality achieved in (A.24) if $\mu_0^{(g)} = \bar{x}_0^{(g)}$, $g = 1, 2, \ldots, k$ and $\Sigma_{00} = (N)^{-1} W_{00}$.

We conclude from (A.22), (A.19), (A.18), and (A.24) that

$$\text{(A.25)} \qquad \max_{\Theta \in \omega_{vc'}} p(\bar{x}, V) = H(V) |\tfrac{1}{N} W_{00}|^{-\frac{1}{2}N} \prod_{g=1}^{k} |\tfrac{1}{N_g} V_{11\cdot 0}^{(g)}|^{-\frac{1}{2}N_g} \quad .$$

As shown above, equality is achieved in (A.18) and (A.24) when $\mu_0^{(g)} = \bar{x}_0^{(g)}$, $\alpha^{(g)} = a^{(g)}$, $\beta^{(g)} = B^{(g)}$, and $\Sigma_{11\cdot 0}^{(g)} = (N_g)^{-1} V_{11\cdot 0}^{(g)}$, and when $\Sigma_{00}^{(1)} = \Sigma_{00}^{(2)} = \cdots = \Sigma_{00}^{(k)} = (N)^{-1} W_{00}$. Hence, it follows that under $H_{vc'}$ the MLE of $\underset{\sim}{\mu}_0$, $\Sigma_{00}^{(1)} = \Sigma_{00}^{(2)} = \cdots = \Sigma_{00}^{(k)} = \Sigma_{00}$, $\underset{\sim}{\alpha}$, $\underset{\sim}{\beta}$, and $\underset{\sim}{\Sigma}_{11\cdot 0}$ are given by Equation (3.5). The maximum of the likelihood under $H_{vc'}$ is given by Equation (A.25).

A.3 Maximum Likelihood Estimators under $H_{m'vc'}$

The hypothesis $H_{m'vc'}$ restricts the parameters $\Theta = (\underset{\sim}{\mu}_0, \underset{\sim}{\Sigma}_{00}, \underset{\sim}{\alpha}, \underset{\sim}{\beta}, \underset{\sim}{\Sigma}_{11\cdot 0})$ to the parametric subspace $\omega_{m'vc'}$ described in Subsection 3.2. Note that

(A.26) $$\omega_{m'vc'} = \{\Theta:\ \Theta_0 \in \omega_{0,m'vc'},\ \Theta_1 \in \omega_{1,t}\} \quad ,$$

where

$$\omega_{0,m'vc'} = \{(\underset{\sim}{\mu}_0, \Sigma_{00}):\ (\underset{\sim}{\mu}_0, \Sigma_{00}) \in \omega_{0,t}\ ,\ \mu_0^{(1)} = \mu_0^{(2)} = \cdots = \mu_0^{(k)},$$
$$\text{and } \Sigma_{00}^{(1)} = \Sigma_{00}^{(2)} = \cdots = \Sigma_{00}^{(k)}\} \quad .$$

Let μ_0 represent the common value of $\mu_0^{(1)}, \mu_0^{(2)}, \ldots, \mu_0^{(k)}$ and Σ_{00} represent the common value of $\Sigma_{00}^{(1)}, \Sigma_{00}^{(2)}, \ldots \Sigma_{00}^{(k)}$. Note that when $\mu_0^{(1)} = \mu_0^{(2)} = \cdots = \mu_0^{(k)} = \mu_0$, $\Sigma_{00}^{(1)} = \Sigma_{00}^{(2)} = \cdots = \Sigma_{00}^{(k)} = \Sigma_{00}$,

(A.27) $$f(\underset{\sim}{\mu}_0, \Sigma_{00}) = \{\exp[-\tfrac{1}{2}\, A(\underset{\sim}{\mu}_0, \underset{\sim}{\Sigma}_{00})]\}\, |\Sigma_{00}|^{-\frac{1}{2}N} \exp -\tfrac{1}{2}\mathrm{tr}\ \Sigma_{00}^{-1} W_{00} \quad ,$$

where in this case

$$A(\underset{\sim}{\mu}_0, \underset{\sim}{\Sigma}_{00}) = \sum_{g=1}^{k} N_g\ \mathrm{tr}(\bar{x}_0^{(g)} - \mu_0)\Sigma_{00}^{-1}(\bar{x}_0^{(g)} - \mu_0)' \quad .$$

Applying Lemma A.1, we find that for all $(\underset{\sim}{\mu}_0, \underset{\sim}{\Sigma}_{00}) \in \omega_{0,m'vc'}$,

$$A(\underset{\sim}{\mu}_0, \underset{\sim}{\Sigma}_{00}) \geq \sum_{g=1}^{k} \operatorname{tr} N_g(\bar{x}_0^{(g)} - \bar{\bar{x}}_0)(\Sigma_{00})^{-1}(\bar{x}_0^{(g)} - \bar{\bar{x}}_0)'$$

$$= \operatorname{tr}(\Sigma_{00})^{-1}[\sum_{g=1}^{k} N_g(\bar{x}_0^{(g)} - \bar{\bar{x}}_0)'(\bar{x}_0^{(g)} - \bar{\bar{x}}_0)]$$

$$= \operatorname{tr}(\Sigma_{00})^{-1}A_{00} \quad ,$$

where $\bar{\bar{x}}_0 = (\sum_{g=1}^{k} N_g)^{-1} \sum_{g=1}^{k} N_g \bar{x}_0^{(g)} = (N)^{-1} \sum_{g=1}^{k} N_g \bar{x}_0^{(g)}$. Thus for all $(\underset{\sim}{\mu}_0, \underset{\sim}{\Sigma}_{00}) \in \omega_{0,m'vc'}$,

$$\text{(A.28)} \qquad f(\underset{\sim}{\mu}_0, \Sigma_{00}) \leq |\Sigma_{00}|^{-\frac{1}{2}N} \exp -\tfrac{1}{2}\operatorname{tr}(\Sigma_{00})^{-1}(W_{00} + A_{00}) \quad .$$

An application of Lemma A.2 to the right-hand side of (A.28) yields

$$\text{(A.29)} \qquad f(\underset{\sim}{\mu}_0, \Sigma_{00}) \leq |\tfrac{1}{N}(W_{00} + A_{00})|^{-\frac{1}{2}N} e^{-\frac{1}{2}qN}$$

for all $(\underset{\sim}{\mu}_0, \Sigma_{00}) \in \omega_{0,m'vc'}$, with equality when $\mu_0^{(1)} = \mu_0^{(2)} = \cdots = \mu_0^{(k)} = \bar{\bar{x}}_0$ and $\Sigma_{00}^{(1)} = \Sigma_{00}^{(2)} = \cdots = \Sigma_{00}^{(k)} = (N)^{-1}(W_{00} + A_{00})$.

We conclude from (A.20), (A.19), (A.18), and (A.29) that

$$\text{(A.30)} \qquad \max_{\Theta \in \omega_{m'vc'}} p(\bar{x}, V) = H(V)|\tfrac{1}{N}(W_{00} + A_{00})|^{-\frac{1}{2}N} \prod_{g=1}^{k} |\tfrac{1}{N_g} V_{11\cdot 0}^{(g)}|^{-\frac{1}{2}N_g} \quad .$$

Since, as shown above, equality is achieved in (A.18) and (A.29) when $\mu_0^{(1)} = \mu_0^{(2)} = \cdots = \mu_0^{(k)} = \bar{\bar{x}}_0$, $\Sigma_{00}^{(1)} = \Sigma_{00}^{(2)} = \cdots = \Sigma_{00}^{(k)} = (N)^{-1}(W_{00} + A_{00})$, $\alpha^{(g)} = a^{(g)}$, $\beta^{(g)} = B^{(g)}$, $\Sigma_{11\cdot 0}^{(g)} = (N_g)^{-1}V_{11\cdot 0}^{(g)}$, $g = 1,2,\ldots,k$, it follows that under $H_{m'vc'}$ the MLE of $\mu_0^{(1)} = \mu_0^{(2)} = \cdots = \mu_0^{(k)} = \mu_0$,

$\Sigma_{00}^{(1)} = \Sigma_{00}^{(2)} = \cdots = \Sigma_{00}^{(k)} = \Sigma_{00}$, $\underset{\sim}{\alpha}$, $\underset{\sim}{\beta}$, $\underset{\sim}{\Sigma}_{11\cdot 0}$ are given by Equation (3.9). The maximum of the likelihood under $H_{m'vc}$, is given by Equation (A.30).

A.4 Maximum Likelihood Estimators under H_{vc} and $H_{m'vc}$

The hypothesis H_{vc} restricts the parameters Θ to the parametric subspace:

(A.31) $$\omega_{vc} = \{\Theta: \Theta_0 \in \omega_{0,vc'}, \Theta_1 \in \omega_{1,vc''}\} ,$$

and the hypothesis $H_{m'vc}$ restricts the parameters Θ to the parametric subspace

(A.32) $$\omega_{m'vc} = \{\Theta: \Theta_0 \in \omega_{0,m'vc'}, \Theta_1 \in \omega_{1,vc''}\} ,$$

where

$$\omega_{1,vc''} = \{(\underset{\sim}{\alpha},\underset{\sim}{\beta},\underset{\sim}{\Sigma}_{11\cdot 0}): (\underset{\sim}{\alpha},\underset{\sim}{\beta},\underset{\sim}{\Sigma}_{11\cdot 0}) \in \omega_{1,t} , \beta^{(1)} = \beta^{(2)} = \cdots = \beta^{(k)},$$
$$\text{and } \Sigma_{11\cdot 0}^{(1)} = \Sigma_{11\cdot 0}^{(2)} = \cdots = \Sigma_{11\cdot 0}^{(k)}\} .$$

Let β be the common value of $\beta^{(1)},\beta^{(2)},\ldots,\beta^{(k)}$, and let $\Sigma_{11\cdot 0}$ be the common value of $\Sigma_{11\cdot 0}^{(1)},\Sigma_{11\cdot 0}^{(2)},\ldots,\Sigma_{11\cdot 0}^{(k)}$. Note that when $\beta^{(1)} = \beta^{(2)} = \cdots = \beta^{(k)} = \beta$, $\Sigma_{11\cdot 0}^{(1)} = \Sigma_{11\cdot 0}^{(2)} = \cdots = \Sigma_{11\cdot 0}^{(k)} = \Sigma_{11\cdot 0}$,

(A.33) $$D(\underset{\sim}{\alpha},\underset{\sim}{\beta},\underset{\sim}{\Sigma}_{11\cdot 0}) \geq 0 ,$$

with equality achieved if

$$\alpha^{(g)} = \bar{x}_1^{(g)} - \bar{x}_0^{(g)}\beta , \qquad g = 1,2,\ldots,k .$$

Also,

$$(A.34)\quad E(\underset{\sim}{\beta},\underset{\sim}{\Sigma}_{11\cdot 0}) = E((\beta,\beta,\ldots,\beta),(\Sigma_{11\cdot 0},\Sigma_{11\cdot 0},\ldots,\Sigma_{11\cdot 0}))$$

$$= \sum_{g=1}^{k} \operatorname{tr} V_{00}^{(g)}(B^{(g)} - \beta)(\Sigma_{11\cdot 0})^{-1}(\beta^{(g)} - \beta)' \quad ..$$

Applying Lemma A.1 (and Equation (A.7)) to (A.34), we find that when $(\underset{\sim}{\alpha},\underset{\sim}{\beta},\underset{\sim}{\Sigma}_{11\cdot 0}) \in \omega_{1,vc}''$,

$$(A.35)\quad E(\underset{\sim}{\beta},\underset{\sim}{\Sigma}_{11\cdot 0}) \geq \operatorname{tr}(\Sigma_{11\cdot 0})^{-1}[\sum_{g=1}^{k} V_{10}^{(g)}(V_{00}^{(g)})^{-1}V_{01}^{(g)} - W_{10}W_{00}^{-1}W_{01}] \quad ,$$

with equality achieved for $\beta^{(1)} = \beta^{(2)} = \ldots = \beta^{(k)} = (\sum_{g=1}^{k} V_{00}^{(g)})^{-1}\sum_{g=1}^{k} V_{00}^{(g)}B^{(g)} = W_{00}^{-1}W_{01}$. Hence, for all $(\underset{\sim}{\mu},\underset{\sim}{\beta},\underset{\sim}{\Sigma}_{11\cdot 0}) \in \omega_{1,vc}''$, it follows from (A.33) and (A.35) that

$$(A.36)\quad h(\underset{\sim}{\alpha},\underset{\sim}{\beta},\underset{\sim}{\Sigma}_{11\cdot 0}) \leq |\Sigma_{11\cdot 0}|^{-\frac{1}{2}N}[\exp -\tfrac{1}{2}\operatorname{tr}(\Sigma_{11\cdot 0}^{(g)})^{-1}W_{11\cdot 0}] \quad ,$$

with equality achieved in (A.36) if $\alpha^{(g)} = \bar{x}_1^{(g)} - \bar{x}_0^{(g)}W_{00}^{-1}W_{01}$, $g = 1,2,\ldots,k$, and $\beta = W_{00}^{-1}W_{01}$. Applying Lemma A.2 to the right-hand side of (A.36), we conclude that for all $(\underset{\sim}{\alpha},\underset{\sim}{\beta},\underset{\sim}{\Sigma}_{11\cdot 0}) \in \omega_{1,vc}''$,

$$(A.37)\quad h(\underset{\sim}{\alpha},\underset{\sim}{\beta},\underset{\sim}{\Sigma}_{11\cdot 0}) \leq |\tfrac{1}{N}W_{11\cdot 0}|^{-\frac{1}{2}N}\, e^{-\frac{1}{2}(p-q)N} \quad ,$$

with equality achieved in (A.37) if $\alpha^{(g)} = \bar{x}_1^{(g)} - \bar{x}_0^{(g)}W_{00}^{-1}W_{01}$, $\beta^{(1)} = \beta^{(2)} = \ldots = \beta^{(k)} = W_{00}^{-1}W_{01}$, and $\Sigma_{11\cdot 0}^{(1)} = \Sigma_{11\cdot 0}^{(2)} = \ldots = \Sigma_{11\cdot 0}^{(k)} = (N)^{-1}W_{11\cdot 0}$.

We conclude from (A.31), (A.19), (A.24), and (A.37) that

$$(A.38)\quad \max_{\Theta\in\omega_{vc}} p(\bar{x},V) = H(V)\,|\tfrac{1}{N}W_{00}|^{-\frac{1}{2}N}|\tfrac{1}{N}W_{11\cdot 0}|^{-\frac{1}{2}N} \quad .$$

Since equality is achieved in (A.24) and (A.37) for $\mu_0^{(g)} = \bar{x}_0^{(g)}$, $\alpha^{(g)} = \bar{x}_1^{(g)} - \bar{x}_0^{(g)} W_{00}^{-1} W_{01}$, $g = 1,2,\dots,k$, and for $\Sigma_{00}^{(1)} = \Sigma_{00}^{(2)} = \dots = \Sigma_{00}^{(k)} = (N)^{-1} W_{00}$, $\beta^{(1)} = \beta^{(2)} = \dots = \beta^{(k)} = W_{00}^{-1} W_{01}$, $\Sigma_{11\cdot 0}^{(1)} = \Sigma_{11\cdot 0}^{(2)} = \dots = \Sigma_{11\cdot 0}^{(k)} = (N)^{-1} W_{11\cdot 0}$, it follows that under H_{vc} the MLE are given by Equation (3.12). The maximum of the likelihood under H_{vc} is given by Equation (A.38).

Similarly, from (A.32), (A.19), (A.29), and (A.37), it follows that

$$\text{(A.39)} \quad \max_{\Theta \in \omega_{m'vc}} p(\bar{x}, V) = H(V) \left|\frac{1}{N}(W_{00} + A_{00})\right|^{-\frac{1}{2}N} \left|\frac{1}{N} W_{11\cdot 0}\right|^{-\frac{1}{2}N} .$$

The maximum likelihood estimators of the parameters under $H_{m'vc}$ are given by Equation (3.14), as can be seen from the sufficient conditions given above for equality to be achieved in (A.29) and (A.37). The maximum of the likelihood under $H_{m'vc}$ is given by Equation (A.39).

A.5 Maximum of the Likelihood under H_{mvc}

The hypothesis H_{mvc} restricts the parameters Θ to the parametric subspace

$$\text{(A.40)} \quad \omega_{mvc} = \{\Theta\colon \Theta_0 \in \omega_{0,m'vc'}, \ \Theta_1 \in \omega_{1,m''vc''}\}$$

where

$$\omega_{1,m''vc''} = \{(\alpha, \beta, \Sigma_{11\cdot 0})\colon (\alpha, \beta, \Sigma_{11\cdot 0}) \in \omega_{1,t},$$
$$\alpha^{(1)} = \alpha^{(2)} = \dots = \alpha^{(k)}, \ \beta^{(1)} = \beta^{(2)} = \dots = \beta^{(k)},$$
$$\Sigma_{11\cdot 0}^{(1)} = \Sigma_{11\cdot 0}^{(2)} = \dots = \Sigma_{11\cdot 0}^{(k)}\} .$$

By applying Lemma A.1 separately to $D(\underset{\sim}{\alpha},\underset{\sim}{\beta},\underset{\sim}{\Sigma}_{11\cdot 0})$ and to $E(\underset{\sim}{\beta},\underset{\sim}{\Sigma}_{11\cdot 0})$ when $(\underset{\sim}{\alpha},\underset{\sim}{\beta},\underset{\sim}{\Sigma}_{11\cdot 0}) \in \omega_{1,m''vc''}$, it can be shown that

$$D(\underset{\sim}{\alpha},\underset{\sim}{\beta},\underset{\sim}{\Sigma}_{11\cdot 0}) + E(\underset{\sim}{\beta},\underset{\sim}{\Sigma}_{11\cdot 0}) \geq \mathrm{tr}\{(\Sigma_{11\cdot 0})^{-1}[\sum_{g=1}^{k} V_{10}^{(g)}(V_{00}^{(g)})^{-1}V_{01}^{(g)} + A_{11} - (W_{10} + A_{10})(W_{00} + A_{00})^{-1}(W_{01} + A_{01})]\}$$

with equality achieved when $\alpha^{(1)} = \alpha^{(2)} = \ldots \alpha^{(k)} = \bar{\bar{x}}_1 - \bar{\bar{x}}_0(W_{00} + A_{00})^{-1}(W_{01} + A_{01})$, and $\beta^{(1)} = \beta^{(2)} = \ldots = \beta^{(k)} = (W_{00} + A_{00})^{-1}(W_{01} + A_{01})$. Thus, when $(\underset{\sim}{\alpha},\underset{\sim}{\beta},\underset{\sim}{\Sigma}_{11\cdot 0}) \in \omega_{1,m''vc''}$,

$$\text{(A.41)} \qquad h(\underset{\sim}{\alpha},\underset{\sim}{\beta},\underset{\sim}{\Sigma}_{11\cdot 0}) \leq |\Sigma_{11\cdot 0}|^{-\frac{1}{2}N} \exp -\tfrac{1}{2}\mathrm{tr}(\Sigma_{11\cdot 0})^{-1}Q_{11\cdot 0}$$

where

$$Q_{11\cdot 0} = W_{11} + A_{11} - (W_{10} + A_{10})(W_{00} + A_{00})^{-1}(W_{01} + A_{01}) \quad .$$

Applying Lemma A.2 to the right-hand side of (A.41) yields the result that

$$\text{(A.42)} \qquad h(\underset{\sim}{\alpha},\underset{\sim}{\beta},\underset{\sim}{\Sigma}_{11\cdot 0}) \leq |\tfrac{1}{N} Q_{11\cdot 0}|^{-\frac{1}{2}N} e^{-\frac{1}{2}(p-q)N}$$

for all $(\underset{\sim}{\alpha},\underset{\sim}{\beta},\underset{\sim}{\Sigma}_{11\cdot 0}) \in \omega_{1,m''vc''}$, with equality when $\alpha^{(1)} = \alpha^{(2)} = \ldots = \alpha^{(k)} = \bar{\bar{x}}_1 - \bar{\bar{x}}_0(W_{00} + A_{00})^{-1}(W_{01} + A_{01})$, $\beta^{(1)} = \beta^{(2)} = \ldots = \beta^{(k)} = (W_{00} + A_{00})^{-1}(W_{01} + A_{01})$, and $\Sigma_{11\cdot 0}^{(1)} = \Sigma_{11\cdot 0}^{(2)} = \ldots = \Sigma_{11\cdot 0}^{(k)} = (N)^{-1}Q_{11\cdot 0}$.

From (A.40), (A.19), (A.29), and (A.42) it follows that

$$\text{(A.43)} \qquad \max_{\Theta\in\omega_{mvc}} p(\bar{x},V) = H(V)\,|\tfrac{1}{N}(W_{00} + A_{00})|^{-\frac{1}{2}N}\,|\tfrac{1}{N} Q_{11\cdot 0}|^{-\frac{1}{2}N}$$

$$= H(V)\,|\tfrac{1}{N}(W + A)|^{-\frac{1}{2}N} \quad .$$

The maximum likelihood estimators of the parameters under H_{mvc} are given implicitly in Section 3.5 (and explicitly above by the conditions for equality in (A.29) and (A.42)). The maximum of the likelihood under H_{mvc} is given by Equation (A.43).

References

Anderson, T. W. (1957). Maximum likelihood estimates for a multivariate normal distribution when some observations are missing. Journal of the American Statistical Association 52, 200-203.

Anderson, T. W. (1958). An Introduction to Multivariate Statistical Analysis. John Wiley and Sons, Inc., New York.

Bartlett, M. S. (1937). Properties of sufficiency and statistical tests. Proceedings of the Royal Society A, 160, 268-282.

Bhargava, R. P. (1962). Multivariate tests of hypotheses with incomplete data. Technical Report No. 3. Stanford University.

Box, G. E. P. (1949) A general distribution theory for a class of likelihood criteria. Biometrika 36, 317-346.

Das Gupta, S. (1969). Properties of power functions of some tests concerning dispersion matrices of multivariate normal distributions. Annals of Mathematical Statistics 40, 697-701.

Gleser, L. J. and Olkin, I. (1966). A k -sample regression model with covariance. In Multivariate Analysis (ed. P. R. Krishnaiah), Academic Press, Inc., New York, 59-72.

Gleser, L. J. and Olkin, I. (1969). Linear models in multivariate analysis. In Essays in Probability and Statistics (ed. R. C. Bose et al.), University of North Carolina Press, Chapel Hill, 267-292.

Gleser, L. J. and Olkin, I. (1972a). Estimation for a regression model with an unknown covariance matrix. Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, I, University of California Press, Berkeley, 541-568.

Gleser, L. J. and Olkin, I. (1972b). A note on Box's general method of approximation for the null distributions of likelihood criteria. To appear.

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

Sugiura, N. and Nagao, H. (1968). Unbiasedness of some test criteria for the equality of one or two covariance matrices. Annals of Mathematical Statistics 39, 1686-1692.