XVI

A Survey of Replacement Variate Generators

1. *Introduction*

In this section, the replacement variate rationale that characterizes the employment of component and latent variate generators is illustrated through a consideration of a number of well-known generators ("models"). All replacement variate generators considered are analyzed with respect their population characteristics. In practice, the parameters of generators have to be estimated on the basis of a sample drawn from the population P under study. The sole aim of the current treatment is the reorientation of traditional thinking on latent variable and component "models", and, in particular, the recognition of these "models" as replacement variate generators. The aim is *not* to rehash in great detail the mathematical details of particular generators, there already being in existence many accounts of this subject matter.

2. *The first principal component (pc1) generator*

Let $\underline{\mathbf{Y}}$ contain a set of p input variates $\mathbf{Y}_j$, j=1..p, jointly distributed in a population P under study. The scores that comprise the distributions of these variates are produced by following some particular set of rules of score production $\{r_1,r_2,...,r_p\}$. To begin, let $\underline{\mathbf{X}}=\underline{\mathbf{Y}}-\underline{\mu}_Y$, so that $E(\underline{\mathbf{X}})=\underline{0}$ and $E\underline{\mathbf{X}}\underline{\mathbf{X}}'=\Sigma$.

PC1 replacement variates

A replacement variate $\mathbf{c}$ is sought such that

<u>ri</u>) $\mathbf{c}=\underline{t}'\underline{\mathbf{X}}$

<u>rii</u>) $V(\mathbf{c})=\underline{t}'\Sigma\underline{t}$ is a maximum over all unit length $\underline{t}$, i.e., $\underline{t}$ s.t. $\underline{t}'\underline{t}=1$.

From (ri) it follows that any distributional requirements made of $\mathbf{c}$ can be brought about only if $\underline{\mathbf{X}}$ satisfies complementary distributional requirements. While, for purposes of hypothesis testing, it is common to treat $\underline{\mathbf{X}}$ as multivariate normal, which, if correct, induces $\mathbf{c}$ to be univariate normal, no such distributional requirements are here stipulated. Note also that the symbol $\mathbf{c}$ stands for *each* of the constituents of set C, i.e., any random variate that satisfies (ri)-(rii).

Existence

A p-vector $\underline{t}$ must be found that satisfies (ri) and (rii). The solution is well known. Let $\varphi=\underline{t}'\Sigma\underline{t}-\gamma(\underline{t}'\underline{t}-1)$. The partial derivatives of φ with respect $\underline{t}$ and γ are

(15.1) $$\frac{\partial\varphi}{\partial \underline{t}}\underline{t}'\Sigma\underline{t}-\gamma(\underline{t}'\underline{t}-1)=2\underline{t}'\Sigma-2\gamma\underline{t}'$$

(15.2) $$\frac{\partial\varphi}{\partial \gamma}\underline{t}'\Sigma\underline{t}-\gamma(\underline{t}'\underline{t}-1)=\underline{t}'\underline{t}-1$$

From (15.1), it follows that $\underline{t}$ is the first eigenvector $\underline{v}_1$, and γ the first eigenvalue λ_1, of Σ. Obviously, then, $\mathbf{c}=\underline{v}_1'\underline{\mathbf{X}}$ is the first principal component of the distribution of $\underline{\mathbf{X}}$. The replacement variate $\mathbf{c}$ will, therefore, be labeled $\mathbf{c}_1$. Requirements (ri)-(rii) define the *pc1-generator*, and, because $\mathbf{c}_1$ always exists, the input variates $\underline{\mathbf{X}}$ are always replaceable under the pc1-generator, i.e., $\underline{\mathbf{X}}$ is always *pc1-replaceable*.

Cardinality of replacement

Because Σ is gramian and of rank $t\leq p$, it possesses p eigenvalue/eigenvector pairs, $\{[\lambda_1,\underline{v}_1], [\lambda_2,\underline{v}_2],\ldots,[\lambda_p,\underline{v}_p]\}$, but only t of the eigenvalues are non-zero. From well known theory (Jolliffe,), the pc1 replacement is unique, i.e., Card(C)=1, as long as $\lambda_1>\lambda_2$, for then there exists no ambiguity as to the choice of $\underline{t}$, i.e., it will be the eigenvector of Σ associated with λ_1. And, of course, one could continue on and discuss the conditional principal component generators which yield the 2nd through tth principal component variates.

Construction formula

The construction formula is

(15.3) $\mathbf{c}_1=\underline{v}_1'\underline{\mathbf{X}}$.

<u>Consequences:</u>

The following are quantitative consequences of the pc1 replacement, or, in other words, consequences of (ri)-(rii).

<u>Ci</u>) $E(\mathbf{c}_1)=E(\underline{v}_1'\underline{\mathbf{X}})=0$ (from (15.1) and (ri))

<u>Cii</u>) $V(\mathbf{c}_1)=\underline{t}'\Sigma\underline{t}=\underline{v}_1'\Sigma\underline{v}_1=\lambda_1$ (from (15.2))

i.e., the variance of the replacement variate $\mathbf{c}$ is equal to the first eigenvalue of Σ.

<u>Ciii</u>) $\underline{\mathbf{X}}=\underline{v}_1\mathbf{c}_1+\underline{\mathbf{r}}$, in which $\underline{\mathbf{r}}=(I-\underline{v}_1\underline{v}_1')\underline{\mathbf{X}}$.

That is, the pc1 replacement can be given a model-like presentation, the phenomena represented by the $\mathbf{X}_j$ made to look as if they are related to phenomena represented by the terms $\mathbf{c}_1$ and $\underline{\mathbf{r}}$. The fact that replacement variate generators can often be expressed in this way has helped to maintain the false picture that latent variate generators represent relations between unobservable and observable entities. But note that, prior to analysis, i.e., prior to the employment of Σ to derive $\underline{v}_1$, there existed no rule for the production of scores comprising the distributions of $\mathbf{c}_1$ and $\underline{\mathbf{r}}$. These terms are not linked to constituents of natural reality by antecedently stipulated rules of correspondence. Consequence (Ciii) is simply a useful decomposition of the input variates in terms of constructed replacement variates.

Civ) $E(\mathbf{c}_1\underline{\mathbf{r}}')=E(\underline{v}_1'\underline{\mathbf{XX}}'(I-\underline{v}_1\underline{v}_1'))=\underline{v}_1\Sigma-\underline{v}_1\Sigma\underline{v}_1\underline{v}_1'=\lambda_1\underline{v}_1'-\lambda_1\underline{v}_1'=\underline{0}'$ (from (Cii) and (15.2))

Cv) $E(\mathbf{c}_1\underline{\mathbf{X}}')=E(\underline{v}_1'\underline{\mathbf{XX}}')=\underline{v}_1'\Sigma= \lambda_1\underline{v}_1'$ (from (15.2))

Cvi) $E(\underline{\mathbf{r}})=E(I-\underline{v}_1\underline{v}_1')\underline{\mathbf{X}}=\underline{0}$

Cviii) $E(\underline{\mathbf{X}}\underline{\mathbf{r}}')=E(\underline{\mathbf{XX}}'(I-\underline{v}_1\underline{v}_1')')=\Sigma-\Sigma\underline{v}_1\underline{v}_1'=\Sigma-\underline{v}_1\underline{v}_1'\lambda_1$ (from (15.2))

Cix) $E(\underline{\mathbf{rr}}')=E((I-\underline{v}_1\underline{v}_1')\underline{\mathbf{XX}}'(I-\underline{v}_1\underline{v}_1')')=\Sigma-\underline{v}_1\underline{v}_1'\lambda_1=\sum_{i=2}^{p}\lambda_i\underline{v}_i\underline{v}_i$,

in which $\underline{v}_i$ is the ith eigenvector, and λ_i, the ith eigenvalue, of Σ. Note that, since $(I-\underline{v}_1\underline{v}_1')$ is indempotent and of rank (p-1), rank($E(\underline{\mathbf{rr}}')$)=(p-1). Thus, Σ is decomposable as follows:

Cx) $\Sigma=C(E(\underline{\mathbf{X}}|\mathbf{c}_1))+E(C(\underline{\mathbf{X}}|\mathbf{c}_1))=C(E(\underline{\mathbf{X}}|\mathbf{c}_1))+E(\underline{\mathbf{rr}}')=A+B$,

in which A is the rank unity matrix $\underline{v}_1\underline{v}_1'\lambda_1$, and B is the rank (p-1) matrix $\sum_{i=2}^{p}\lambda_i\underline{v}_i\underline{v}_i$.

Optimality properties

As is clear from (ri)-(rii), the PC1 replacement finds a linear combination of the input variates, $\mathbf{c}=\underline{t}'\underline{\mathbf{X}}$, that has maximum variance among all of those linear combinations for which $\underline{t}'\underline{t}=1$. This linear combination is $\mathbf{c}_1$, the first principal component of the distribution of $\underline{\mathbf{X}}$. The replacement of the $\underline{\mathbf{X}}_j$ by $\mathbf{c}_1$ is optimal in a variety of other important, and well known, senses, several of which are, herein, reviewed:

i. Consider the vector of residuals, $\underline{\mathbf{X}}-(\underline{t}'\Sigma\underline{t})^{-1}\Sigma\underline{t}\underline{t}'\underline{\mathbf{X}}$, of the linear regression of the input variates $\mathbf{X}_j$, j=1..p, on a linear combination $\underline{t}'\underline{\mathbf{X}}$ ($\underline{t}'\underline{t}=1$) of the input variates. The replacement variate $\mathbf{c}_1=\underline{v}_1'\underline{\mathbf{X}}$ is that choice of linear combination of the $\mathbf{X}_j$ which minimizes $Etr(\underline{\mathbf{X}}-(\underline{t}'\Sigma\underline{t})^{-1}\Sigma\underline{t}\underline{t}'\underline{\mathbf{X}})(\underline{\mathbf{X}}-(\underline{t}'\Sigma\underline{t})^{-1}\Sigma\underline{t}\underline{t}'\underline{\mathbf{X}})')$, i.e., the expected value of the sum of squared residuals. In other words, the replacement variate $\mathbf{c}_1$ is that variate that is most "similar" to the p input variates when the latter are taken as a set. Knowledge of the p parameters of the regressions of the $\mathbf{X}_j$ on $\mathbf{c}_1$ allows for the calculation of

$\frac{1}{2}p(p+1)$ implied covariances and variances, $\hat{\Sigma}$, and these are as close as possible to being perfect reproductions of the elements of Σ.

Proof. The aim is to find the $\underline{t}$ that minimizes

$$trE(\underline{\mathbf{X}}-(\underline{t}'\Sigma\underline{t})^{-1}\Sigma\underline{t}\underline{t}'\underline{\mathbf{X}})(\underline{\mathbf{X}}-(\underline{t}'\Sigma\underline{t})^{-1}\Sigma\underline{t}\underline{t}'\underline{\mathbf{X}})'=tr(\Sigma)-tr(\Sigma\underline{t}(\underline{t}'\Sigma\underline{t})^{-1}\underline{t}\Sigma),$$

subject to $\underline{t}'\underline{t}=1$. Since tr(Σ) is not a function of $\underline{t}$, the task becomes one of maximizing $tr\Sigma\underline{t}(\underline{t}'\Sigma\underline{t})^{-1}\underline{t}'\Sigma=\underline{t}'\Sigma\underline{t}$ subject to $\underline{t}'\underline{t}=1$. But this is precisely the same eigenproblem as (15.1)-(15.2), and yields the same solution, i.e., $\underline{t}=\underline{v}_1$, in which $\underline{v}_1$ is the first eigenvector of Σ.

To put this another way, to possess replacement variate $\mathbf{c}_1$ in conjunction with $\underline{v}_1$, yields a set of linear conditional expectations (linear predictions) of the $\mathbf{X}_j$,

$$\hat{\underline{\mathbf{X}}}=E(\underline{\mathbf{X}}|\mathbf{c})=\underline{v}_1\underline{v}_1'\underline{\mathbf{X}},$$

that reproduces the original input variates $\underline{\mathbf{X}}$ better than any other single linear combination of the $\mathbf{X}_j$.

Replacement Loss

Since the aim of the replacement is to derive a variate $\mathbf{c}_1$ such that knowledge of the p parameters of the regressions of the $\mathbf{X}_j$ on $\mathbf{c}_1$ can be used to reproduce the $\frac{1}{2}p(p+1)$ covariances and variances of Σ, it is natural to quantify the degree to which this aim has been achieved. The vector of linear regressions of the $\mathbf{X}_j$ on the replacement variate $\mathbf{c}_1$ is $\underline{v}_1\mathbf{c}_1$. The loss inherent to the replacement of the p input variates by $\mathbf{c}_1$ is, therefore, given by the elements of the matrix $\Sigma-\hat{\Sigma}=\Sigma-E(\underline{v}_1\mathbf{c}_1\underline{v}_1'\mathbf{c}_1)=$ $\Sigma-\underline{v}_1\underline{v}_1'\lambda_1=\sum_{i=2}^{p}\lambda_i\underline{v}_i\underline{v}_i'=V_r\Lambda_rV_r'$, in which the columns of V_r contain eigenvectors 2 through t of Σ, and diagonal matrix Λ_r contains the corresponding eigenvalues. A natural measure of replacement loss is then $tr(V_r\Lambda_rV_r'V_r\Lambda_rV_r')=\sum_{i=2}^{p}\lambda_i^2$, or, better yet, the ratio of this quantity to $tr(\Sigma\Sigma)=\sum_{i=1}^{p}\lambda_i^2$, which is equal to the sum of the squared elements of Σ.

Characteristics of C

Under the condition that $\lambda_1>\lambda_2$, the set *C* contains but one variate, and, hence, a description of the properties of *C* is a description of the properties of $\mathbf{c}_1$. Therefore, one need not consider the similarity of the elements of *C*, but certainly might consider what the pc1 replacement implies about the relationship between the single pc1 replacement variate $\mathbf{c}_1$ and external variates (variates not part of the set of input variates). Let $\underline{\mathbf{Z}}$ contain a set of external variates. As is well known,

(15.4) $$E(\mathbf{c}_1\underline{\mathbf{Z}}')=E(\underline{v}_1'\underline{\mathbf{XZ}}')=\underline{v}_1'\Sigma_{XZ}$$

and

(15.5) $$\rho(\mathbf{c}_1,\underline{\mathbf{Z}}')=\frac{1}{\lambda_1}\underline{v}_1'\Sigma_{XZ}D_Z^{-1/2}$$

in which $D_Y^{-1/2}$ is the diagonal matrix containing the reciprocals of the standard deviations of the $\mathbf{Z}_j$.

Testability

Because $\underline{\mathbf{X}}$ is always pc1-replaceable, there is no need for a test of replaceability. However, since the essence of the replacement is variance and covariance explanation, it is perfectly reasonable to consider hypotheses about the value of λ_1 in some population *P*. For example, in light of optimality property (i), it might be tested whether the vector of residuals of the linear regressions of the input variates $\mathbf{X}_j$, j=1..p, on the replacement variate $\mathbf{c}_1$ has a null covariance matrix, i.e., whether λ_1=tr(Σ), or, equivalently, whether it is the case that rank(Σ)=1. This hypothesis will, in practice, always be false, but, as will be discussed later, so will be the hypotheses standardly tested in the employment of the linear factor generator.

2. *The unidimensional linear common factor (ulcf) generator*

The mathematics of factor analysis are well known. In this section, these mathematics are presented as they should always have been presented, i.e., without reference to the Central Account. In a comparison of linear factor and component analysis, Mulaik (1990) states that "...the essential feature of a component model is that a component variable is defined as some specific linear composite of a set of observed variables chosen in some particular way to optimize some quantitative criterion" (p.53). But, as will be seen, if the sentence part "linear composite of a set of observed variables" were to be replaced by "constructed random variate", this would, as McDonald's 1975 analysis managed to show, constitute a perfectly accurate description of both the ulcf and pc1 generators. For, as with the pc1 generator, the ulcf generator is a recipe for the construction of variates that must satisfy a set of quantitative requirements that are uniquely those of the linear factor generator. Let $\underline{\mathbf{Y}}$ contain a set of p input variates $\mathbf{Y}_j$, j=1..p, jointly distributed in a population *P* under study. The scores that comprise the distributions of these variates are produced by following the some particular set of antecedently stipulated rules of score production $\{r_1,r_2,...,r_p\}$. Once again, let $\underline{\mathbf{X}}=\underline{\mathbf{Y}}-\underline{\mu}_Y$ so that $E(\underline{\mathbf{X}})=\underline{0}$ and $E\underline{\mathbf{XX}}'=\Sigma$.

ulcf replacement variates

A continuous variate $\boldsymbol{\theta}$ is sought such that

ri) $E(\boldsymbol{\theta})=0$

rii) $V(\boldsymbol{\theta})=E\boldsymbol{\theta}^2=1$,

and

riii) The p residuals, $\underline{\mathbf{l}}$, of the linear regressions of the input variates, $\mathbf{X}_j$, j=1..p, on $\boldsymbol{\theta}$ have a covariance matrix that is diagonal and positive definite.

Note: In this particular formulation of the ulcf generator, no further requirements are placed on the distribution of $\boldsymbol{\theta}$.

Consequences of (ri)-(riii)

Ci) The linear conditional expecation of $\underline{\mathbf{X}}$ given $\boldsymbol{\theta}=\theta_o$ is equal to

$E(\underline{\mathbf{X}}|\boldsymbol{\theta}=\theta_o)_{lin}=E(\underline{\mathbf{X}}\boldsymbol{\theta})\theta_o=\underline{\sigma}_{x,\theta}\theta_o$ (from (rii))

To square the treatment with standard notation, symbolize $\underline{\sigma}_{x,\theta}$ as $\underline{\Lambda}$, in which case $E(\underline{\mathbf{X}}|\boldsymbol{\theta}=\theta_o)_{lin}=\underline{\Lambda}\theta_o$. The following are, then, consequences of the ulcf replacement:

Cii) $\underline{\mathbf{l}}=\underline{\mathbf{X}}-\underline{\Lambda}\boldsymbol{\theta}$

Ciii) $\underline{\mathbf{X}}=\underline{\Lambda}\boldsymbol{\theta}+\underline{\mathbf{l}}$

This is the ulcf decomposition of $\underline{\mathbf{X}}$. As alluded to earlier, the appearance of this decomposition helped foster the misconception that the ulcf generator represents structural relations between phenomena represented by the input variates, on the one hand, and an unobservable property (cause) represented by $\boldsymbol{\theta}$, on the other. This is, of course, an illusion, for while the scores that comprise the distribution of each $\mathbf{X}_j$ were indeed produced by an antecedently stipulated rule of score production r_j, and, occasionally, can be correctly claimed to be signified by an ordinary language concept, "φ", this is not the case for the scores that comprise the distributions of each of $\boldsymbol{\theta}$ and $\underline{\mathbf{l}}$. Without such antecedently stated rules of score production to establish correspondence relations, the symbols $\boldsymbol{\theta}$ and $\underline{\mathbf{l}}$ cannot rightly be said to *stand for* any features of natural reality. They are simply place-holders for any random variates (vectors) that satisfy the requirements of the ulcf generator.

Civ) $E(\boldsymbol{\theta}\underline{\mathbf{l}}')=E\boldsymbol{\theta}(\underline{\mathbf{X}}-\underline{\Lambda}\boldsymbol{\theta})'=\underline{\Lambda}'-\underline{\Lambda}'=\underline{0}$. (from (rii) and $E\boldsymbol{\theta}\underline{\mathbf{X}}=\underline{\sigma}_{x,\theta}=\underline{\Lambda}$)

Cv) $E(\underline{\mathbf{l}})=E(\underline{\mathbf{X}}-\underline{\Lambda}\boldsymbol{\theta})=\underline{0}$ (from $E(\underline{\mathbf{X}})=\underline{0}$ and (ri))

Cvi) $\Sigma = E(\underline{\mathbf{XX}}')= C(E(\underline{\mathbf{X}}|\boldsymbol{\theta}))+E(C(E(\underline{\mathbf{X}}|\boldsymbol{\theta}))=E((\underline{\Lambda}\boldsymbol{\theta}+\underline{\mathbf{l}})(\underline{\Lambda}\boldsymbol{\theta}+\underline{\mathbf{l}})')=\underline{\Lambda\Lambda}'+\Psi$.

(from (rii), (riii), and (Cii))

Cvii) $E(\underline{\mathbf{X}}\underline{\mathbf{l}}')=E((\underline{\Lambda}\boldsymbol{\theta}+\underline{\mathbf{l}})\underline{\mathbf{l}}')=\Psi$ (from (Cii) and (riii))

Cviii) $\boldsymbol{\theta}$ cannot be constructed as a linear combination of the $\mathbf{X}_j$. That is, there does not exist a vector $\underline{t}$ such that $\boldsymbol{\theta}=\underline{t}'\underline{\mathbf{X}}$ and also that (ri)-(riii) are satisfied.

Proof By (riii), $C(\underline{\mathbf{l}})=\Sigma-\underline{\sigma}_{x,\theta}\underline{\sigma}_{x,\theta}'=\Psi$, with Ψ diagonal and positive definite, and, hence, of rank p. Now, if $\boldsymbol{\theta}=\underline{t}'\underline{\mathbf{X}}$, for some $\underline{t}$, then $\underline{\sigma}_{x,\theta}=\Sigma\underline{t}$ and $C(\underline{\mathbf{l}})=\Sigma-\Sigma\underline{t}\underline{t}'\Sigma=\Sigma-\Sigma\underline{t}(\underline{t}'\Sigma\underline{t})^{-1}\underline{t}'\Sigma=\Sigma^{1/2}(I-\Sigma^{1/2}\underline{t}(\underline{t}'\Sigma\underline{t})^{-1}\underline{t}'\Sigma^{1/2})\Sigma^{1/2}$. But $(I-\Sigma^{1/2}\underline{t}(\underline{t}'\Sigma\underline{t})^{-1}\underline{t}'\Sigma^{1/2})$ is idempotent and of rank (p-1). Hence, $C(\underline{\mathbf{l}})$ is of rank (p-1), which contradicts (riii).

Existence

The implication [(ri)-(riii)]$\rightarrow$[$\Sigma=\underline{\Lambda\Lambda}'+\Psi$, Ψ diagonal and positive definite] is true. Therefore, ~[$\Sigma=\underline{\Lambda\Lambda}'+\Psi$, Ψ diagonal, positive definite]$\rightarrow$~[(ri)-(riii)] is also true. Hence, if Σ cannot be decomposed as $\underline{\Lambda\Lambda}'+\Psi$, Ψ diagonal and positive definite, then [(ri)-(riii)] cannot obtain, i.e., $\underline{\mathbf{X}}$ cannot be ulcf replaced. Similarly, [$\Sigma=\underline{\Lambda\Lambda}'+\Psi$, Ψ diagonal and positive definite]$\rightarrow$[(ri)-(riii)] is true (Wilson, 1928a; Guttman, 1955), hence, ~[(ri)-(riii)]$\rightarrow$~[$\Sigma=\underline{\Lambda\Lambda}'+\Psi$, Ψ diagonal and positive definite] is true. A necessary and sufficient condition that $\underline{\mathbf{X}}$ be ulcf-replaceable is, then, that $\Sigma=\underline{\Lambda\Lambda}'+\Psi$, in which Ψ is diagonal and positive definite. That is, there exists at least one random variate $\boldsymbol{\theta}$ that satisfies (ri)-(riii) iff $\Sigma=\underline{\Lambda\Lambda}'+\Psi$, Ψ diagonal and positive definite. Piaggio (1931), Kestelman (1952), and Guttman (1955) derived general formulas for the construction of $\boldsymbol{\theta}$.

The factor analytic decomposition $\Sigma=\Lambda_t\Lambda_t'+\Psi$, in which Ψ is diagonal and positive definite, for t>1, is not sufficient for a t-dimensional lcf replacement. The unidimensional quadratic replacement, for example, also yields the decomposition $\Sigma=\Lambda_2\Lambda_2'+\Psi$, Ψ diagonal and positive definite.

Cardinality of replacement

As will be recalled from Chapter IV, if $\underline{\mathbf{X}}$ is ulcf-replaceable, then Card(C)=∞. (Wilson, 1928a; Piaggio, 1931). That is, if a particular $\underline{\mathbf{X}}$ is ulcf-replaceable, then its replacement is not unique, there being constructible an infinity of random variates each of which satisfies (ri)-(riii).

Construction formula

Recall from Chapter IV, that, if $\underline{\mathbf{X}}$ is ulcf-replaceable, the construction formula for ulcf replacement variates (elements of set C), is $\boldsymbol{\theta}_i=\underline{\Lambda}'\Sigma^{-1}\underline{\mathbf{X}}+w^{1/2}\mathbf{s}_i$, in which $w=(1-\underline{\Lambda}'\Sigma^{-1}\underline{\Lambda})$ and $\mathbf{s}_i$ is an arbitrary random variate for which $C(\mathbf{s}_i,\underline{\mathbf{X}})=\underline{0}$, $E(\mathbf{s}_i)=0$, and $V(\mathbf{s}_i)=1$ (Piaggio, 1931; Kestelman, 1952; Guttman, 1955). It follows then that $\underline{\mathbf{l}}_i=\underline{\mathbf{X}}-\underline{\Lambda}\boldsymbol{\theta}_i=\underline{\mathbf{X}}-\underline{\Lambda}(\underline{\Lambda}'\Sigma^{-1}\underline{\mathbf{X}}+w^{1/2}\mathbf{s}_i)=\Psi\Sigma^{-1}\underline{\mathbf{X}}-w^{1/2}\underline{\Lambda}\mathbf{s}_i$. If additional distributional requirements are placed on the $\boldsymbol{\theta}_i$, then, because the distribution of $\underline{\mathbf{X}}$ is an empirical matter, these requirements must ultimately be realized through requirements imposed on the $\mathbf{s}_i$.

Note: In multidimensional linear factor replacements, because $\Sigma=\Lambda TT'\Lambda'+\Psi=\Lambda^*\Lambda^{*\prime}+\Psi$, for any T such that $TT'=T'T=I_m$, the matrix Λ is also indeterminate. A choice must then be made as to a favoured matrix of regression weights, before the construction formula is defined. The role played by Thurstone's principal of parsimony is to provide a rationale for choosing a favored Λ.

As was seen in Chapter V, some have claimed that, because ulcf replacement variates ("common factors") are constructed in accord with a formula that involves an arbitrary random variate $\mathbf{s}_i$, these variates must be absurdities. McDonald and Mulaik, in particular, have taken this apparent absurdity as a prima facie argument that ulcf replacement variates (i.e., random variates constructed so as to satisfy (ri)-(riii)) are not *really* common factors to $\underline{\mathbf{X}}$. However, they are wishing for something that latent variable models cannot deliver, namely the Central Account. What the employment of the ulcf generator *does* allow is a test of replaceability, and, if a particular $\underline{\mathbf{X}}$ is deemed to be ulcf replaceable, the right to produce replacement variates that satisfy precisely requirements (ri)-(riii) that are unambiguously imposed by the ulcf generator. Because requirements (ri)-(riii) are the *defining* requirements of the ulcf generator, the variates that satisfy these requirements cannot help but be ulcf replacement variates. The "paradox" mentioned in Chapter IV, that, when one partials from $\boldsymbol{\theta}$ all of the information about $\boldsymbol{\theta}$ that is contained in $\underline{\mathbf{X}}$, and yet is left with a positive residual variance, $E(V(\boldsymbol{\theta}|\underline{\mathbf{X}}))$, is indicative of the fact that one must go beyond the input variates to construct a new variate that satisfies (ri)-(riii). It is a fact that variates that satisfy requirements (ri)-(riii) cannot also be functions of the input random variates (this was proven very nicely in McDonald, 1975). It is also a consequence of the defining requirements of the ulcf generator that if a particular $\underline{\mathbf{X}}$ is ulcf-replaceable, then it is ulcf-replaceable by an infinity of variates, each not a function of $\underline{\mathbf{X}}$. There cannot be anything mysterious or controversial about this, for this consequence arises from the requirements unambiguously imposed by the ulcf generator. If the replacement variates generated under the ulcf generator are "absurd", the problem lies in the choice of defining requirements, i.e., in the very formulation of the linear factor generator.

In passing judgment on the ulcf replacement several points must be borne in mind:

i) Every replacement variate generator produces its particular brand of optimal replacement variates at a cost. This can mean that a replacement variate constructed by a generator to satisfy requirement A cannot then have property B. Two costs of the brand of optimality required of a ulcf replacement variate are that $\mathrm{Card}(C)=\infty$ and that it is not possible that such variates are functions of the input random variates. Once again, there can be nothing controversial about this, because it follows directly from requirements (ri)-(riii) that unambiguously characterize the ulcf-generator, and that were stipulated by psychometricians when they invented linear factor analysis. No replacement generator can deliver on all senses of optimality. If a replacement *is* desired in which Card(C)=1, then the cost will be the sacrificing of certain of the senses of optimality upon which the ulcf generator currently *does* deliver.

ii) It is true that there is a third cost of ulcf replacement optimality, for it does not seem that a $\boldsymbol{\theta}_i$ whose construction includes an arbitrary random variate, $\mathbf{s}_i$, could be used to scale individuals, this, occasionally, being an aim of a factor analysis. But current practice, which has the researcher retaining the mythology of the Central Account, and insisting that preference should be given to "factor score estimators" over the variates constructed so as to satisfy the defining requirements of the ulcf generator, must be overhauled. If it is deemed essential to produce a variate that can be used to scale individuals, then the linear factor generator is the wrong choice, because its replacement variates do not satisfy this requirement. They, *uniquely*, insist upon (ri)-(riii), and replacement variates that satisfy (ri)-(riii) possess undesirable properties with respect the aim of scaling individuals;

iii) The foregoing analysis makes it sound as if the development of linear factor analysis featured a careful planning in regard the optimality/cost profile of the replacement variates produced under the ulcf generator. Techniques are not usually developed in so careful a fashion. In fact, Spearman, busy as he was deducing social policy implications of the employment of linear factor analysis, and positing various mental energies, simply did not understand what his "model" entailed. First, he wrongly believed that Garnett had proven that the factors of factor analysis were "unique", a proof to which he attached great importance:

> There is another particularly important limitation to the divisibility of the variables into factors. It is that the division into general and specific factors all mutually independent can be effected *in one way only*; in other words, it is unique. For the proof of this momentous theorem, we have to thank Garnett (Spearman, 1927, p.v11).

Dodd's review of factor analysis reinforced this view: "The converse proposition that, if (I) is true, 'g' and 's' alone determine the observed variables was assumed in this [Hart and Spearman, 1912] article and not proved until later (12, 18, 23)" (1928, p.213). It is unreasonable to expect Spearman and his contemporaries to have possessed a flawless grasp of the logic of his newly introduced technique. Mistakes, unwarranted suppositions, and uncertainty are very often a part of the early history of novel technology. However, when Wilson *corrected* his misconceptions about factor analysis, Spearman chose to misrepresent the facts, rather than, perhaps, rethink the formulation of his generator. The rapid evolution of the Central Account mythology, and the consequent need to protect it, meant that the dispassionate analysis of optimality/cost tradeoffs in regard the ulcf-replacement was effectively foregone. In the end, the only thing for the discipline to do was to fabricate stories (e.g., the estimation of unobservable measurements) to explain away the "absurdity" of the ulcf replacement.

Characteristics of set C

Set *C*, which contains the ulcf replacement variates (those variates constructed as $\boldsymbol{\theta}_i=\underline{\Lambda}'\Sigma^{-1}\underline{\mathbf{X}}+w^{1/2}\mathbf{s}_i$, in which $w=(1-\underline{\Lambda}'\Sigma^{-1}\underline{\Lambda})$, and $\mathbf{s}_i$ is an arbitrary random variate for which $C(\mathbf{s}_i,\underline{\mathbf{X}})=\underline{0}$, $E(\mathbf{s}_i)=0$, and $V(\mathbf{s}_i)=1$) is of infinite cardinality. In addition to describing the properties that the $\boldsymbol{\theta}_i$ must possess for inclusion in *C*, it is clearly of interest to describe other general features of *C*, notably the distinctness of the replacement variates it contains. A number of these features, including the smallest correlation between members of *C* (Guttman's ρ^*), and the range of correlations between members of *C* and an "external variate", were considered in Chapter IV.

Testability

If the distribution of $\underline{\mathbf{X}}$ in particular population *P* happens to be multivariate normal, then there exist well known tests (see, e.g., Lawley & Maxwell, 1963) of the hypothesis H_o: [$\Sigma=\underline{\Lambda\Lambda}'+\Psi$, Ψ diagonal and positive definite] against the alternative H_1: [Σ is any gramian matrix]. These tests are tests of the hypothesis that a set of input variates, $\mathbf{X}_j$, j=1..p, are ulcf-replaceable.

Optimality criteria

The primary sense of optimality to which answers a ulcf replacement variate, $\boldsymbol{\theta}_i$, is that which is clear from (ri)-(riii), these the defining requirements of the ulcf generator:

$\boldsymbol{\theta}_i$ is a variate that replaces the input variates, $\mathbf{X}_j$, j=1..p, in the sense that the p slope parameters, λ_j, of the linear regressions of the $\mathbf{X}_j$ on $\boldsymbol{\theta}_i$ can be used to reproduce the $\frac{1}{2}p(p-1)$ unique covariances contained in Σ.

That is, if $\underline{\mathbf{X}}$ is ulcf-replaceable,

(15.6) $$\sigma_{jk}=\sigma_{xj,\theta}\sigma_{xk,\theta}=\lambda_j\lambda_k,\ j\neq k.$$

Hence, the $\frac{1}{2}p(p-1)$ pair-wise linear dependencies of the joint distribution of the $\mathbf{X}_j$ can be accounted for by knowledge of the linear relationships between the $\mathbf{X}_j$ and a single constructed random variate, i.e., by a single dimension. In particular, each of the variates $\boldsymbol{\theta}_i$ contains all of the information in regard the pairwise linear relationships among the $\mathbf{X}_j$ in the sense that, once the $\mathbf{X}_j$ have been conditioned on any of the $\boldsymbol{\theta}_i$, there exist no further pairwise linear dependencies among the $\mathbf{X}_j$. In addition to this primary sense of optimality, ulcf replacement variates have been alleged to be optimal in a number of consequent, or secondary, senses.

i. Invariance properties

The ulcf replacement of a particular set of input variates, say $\underline{\mathbf{X}}_A$, possesses a number of attractive invariance properties. Let $\underline{\mathbf{X}}_A$ be ulcf replaceable, with ulcf representation $\underline{\mathbf{X}}_A=\underline{\Lambda}_A\boldsymbol{\theta}+\underline{\mathbf{l}}$. The replacement variates to $\underline{\mathbf{X}}_A$ are, then, constructed as $\boldsymbol{\theta}_i=\underline{\Lambda}_A'\Sigma_A^{-1}\underline{\mathbf{X}}+w_A^{1/2}\mathbf{s}_i$, in which $w_A=(1-\underline{\Lambda}_A'\Sigma_A^{-1}\underline{\Lambda}_A)$, and $\mathbf{s}_i$ is an arbitrary random variate for which $C(\mathbf{s}_i,\underline{\mathbf{X}}_A)=\underline{0}$, $E(\mathbf{s}_i)=0$, and $V(\mathbf{s}_i)=1$. Set C_A contains the $\boldsymbol{\theta}_i$. Now, consider which features of the replacement of $\underline{\mathbf{X}}_A$ survive if $\underline{\mathbf{X}}_A$ is rescaled by a diagonal, positive definite matrix D (this matrix might, for example, contain the reciprocals of the standard deviations of the $\mathbf{X}_{Aj}$), producing a new set of input variates $\underline{\mathbf{Z}}=D\underline{\mathbf{X}}_A$. It turns out that:

a. The input variates $\underline{\mathbf{Z}}$ are still ulcf replaceable, and, in fact, are replaceable by the variates contained in C_A. The vector of regression parameters (of $\underline{\mathbf{Z}}$ on each of the variates contained in C_A) is equal to $\underline{\Lambda}_Z=D\underline{\Lambda}_A$.

b. The ulcf decomposition $\Sigma_Z=D\Sigma_AD$ is simply $D\Sigma_AD=D\underline{\Lambda}_A\underline{\Lambda}_A'D+D\Psi_AD=\underline{\Lambda}_Z\underline{\Lambda}_Z'+\Psi_Z$.

It can be seen, then, that if $\underline{\mathbf{X}}_A$ is ulcf replaceable, and the parameters of this replacement are known, $\underline{\mathbf{Z}}=D\underline{\mathbf{X}}_A$ is also ulcf replaceable, and the parameters of its ulcf replacement are simply rescaled versions of the parameters of the ulcf replacement of $\underline{\mathbf{X}}_A$.

ii. Generalizability

This topic will be considered in detail in a later section on the variate domain formulation of the ulcf replacement. The issue pertains to the conditions under which a subset of variates that ulcf-replace a particular set of variates $\underline{\mathbf{X}}^*$, i.e., a subset of the variates contained in $C_{\underline{X}*}$, ulcf-replace, not only $\underline{\mathbf{X}}^*$, but additional variates. In such a case, the products of a given factor analysis (ulcf-replacement) may rightly be said to "reach beyond the original set of input variates."

Replacement Loss

Because the aim of the ulcf replacement is to derive a variate $\boldsymbol{\theta}$ such that knowledge of the p slope parameters of the regressions of a set of input variates $\mathbf{X}_j$, j=1..p, on $\boldsymbol{\theta}$, can be used to reproduce the $\frac{1}{2}p(p-1)$ unique covariances contained in Σ, and since, in most applications, this aim will not be fully realized, it is only natural to quantify the degree to which this aim has been achieved. Because $\underline{\mathbf{X}}$ is ulcf-replaceable only if $\Sigma_{\underline{X}}=\underline{\Lambda}\underline{\Lambda}'+\Psi$ for some particular choice of $\underline{\Lambda}$ and Ψ, in which Ψ is diagonal and positive definite, the quantification of replacement loss can be carried out by quantifying the discrepancy between $\Sigma_{\underline{X}}$ and $\tilde{\Sigma}=\underline{\tilde{\Lambda}}\underline{\tilde{\Lambda}}'+\tilde{\Psi}$, in which $\underline{\tilde{\Lambda}}$ and $\tilde{\Psi}$ are chosen so as to make $\tilde{\Sigma}$ as close as possible to $\Sigma_{\underline{X}}$. The estimation of this discrepancy is at root of RMSEA measures of fit now employed within the field of structural equation modeling.

Additional restrictions

A variety of additional requirements or restrictions have, for a variety of reasons, been considered by psychometricians. Several examples are, herein, discussed.

Distributional requirements for $\boldsymbol{\theta}$: The ulcf generator (2.7)-(2.8) does not require of ulcf replacement variates that they have some particular distribution, whereas generator (2.11)-(2.12) requires that each has a normal distribution. Distributional requirements made of ulcf replacement variates have, traditionally, been misportrayed as "assumptions." They are, in fact, additional restrictions that must be satisfied by variates in order for them to be ulcf-replacement variates. Now, distributional claims about $\underline{\mathbf{X}}$ must "make sense", in that they must square with evidence about the distribution of $\underline{\mathbf{X}}$ in *P*, culled from samples drawn from *P*. The moment requirements that must be satisfied by the $\mathbf{s}_i$ are that $C(\mathbf{s}_i,\underline{\mathbf{X}})=\underline{0}$, $E(\mathbf{s}_i)=0$, and $V(\mathbf{s}_i)=1$. Given that $\underline{\mathbf{X}}$ has some particular density $g_{\underline{X}}$, the distribution of $\mathbf{s}_i$ can always be chosen so as to yield some desired continuous distribution for the $\boldsymbol{\theta}_i$, because $\boldsymbol{\theta}_i=\underline{\Lambda}_X'\Sigma^{-1}\underline{\mathbf{X}}+w^{1/2}\mathbf{s}_i$, and, as a result,

$$f_{\theta_i}=\left|J\left(\begin{bmatrix}\underline{X}\\ s_i\end{bmatrix}\to\begin{bmatrix}\theta_i\\ \underline{t}\end{bmatrix}\right)\right|^{-1}\int_{\underline{t}} f_{(\underline{X},s_i)}\left(\frac{(\theta_i-\underline{\Lambda}'\Sigma^{-1}\underline{X})}{w^{1/2}},\underline{t}^{-1}(\underline{X},s_i)\right)d\underline{t}$$

, in which $\underline{t}$ is any vector chosen so that the transformation of $\begin{bmatrix}\underline{X}\\ s_i\end{bmatrix}$ to $\begin{bmatrix}\theta_i\\ \underline{t}\end{bmatrix}$ is non-singular, and $J(.\to.)$ is the jacobian of the transformation. Bartholomew has repeatedly stated (e.g., 1980, p.295) that, because the distribution of $\boldsymbol{\theta}_i$ is "unknown", it may be chosen "to suit our convenience." Clearly, this latitude in regard the distribution of $\boldsymbol{\theta}_i$ does not come about as a result of its being "unknown" (if it was truly "unknown", the task would then be one of discovery, and he would have no business *choosing* this distribution "to suit his convenience"), but as a result of the requirements imposed by the ulcf generator on random variates in order that they can rightly be called latent variates to $\underline{\mathbf{X}}$.

Additional requirements to yield Card(*C*)=1: As was evident from Chapters IV and V, the fact that, under the ulcf replacement, Card(*C*) is equal to infinity, has been the cause of consternation amongst psychometricians. It has been suggested herein that the true source of this concern was not the non-uniqueness of the ulcf replacement per se, but rather the idea that common factors are constructed random variates, and the threat this poses to the Central Account. Regardless, various suggestions, some of these dating back to Spearman's responses to E.B. Wilson, have been made as to how determinate linear factor replacements might be achieved. These suggestions, being as they were entries in the indeterminacy debate, have typically been offered up as "ways to save the factor model." The factor "model" is, however, a replacement variate generator, and, hence, the issue is a very general one, pertaining as it does to conditions under which a replacement generator can be made to yield a replacement in which Card(*C*)=1. Several of these conditions will now be reviewed and critiqued.

i. Extra-factor analytic research. In a number of his papers, Mulaik has recommended that further research could be undertaken to resolve the "indeterminacy impasse" in a given application of factor analysis. His ideas were discussed in some detail in Chapters V and VI. To recount, he mentions two possibilities. In the first, parameter estimates generated in an application of a factor analytic generator are employed as a clue in a subsequent search for some natural phenomenon that, on non-factor analytic grounds could be judged as the cause of the phenomena represented by the manifest variates. As was noted, this program of research would then effectively change the nature of latent variable modeling, and would lead to the discarding of the CA mythology. In the second, Mulaik envisions extra-factor analytic research undertaken to facilitate the selection of one preferred *variate* from among those in *C*. It was argued that this suggestion suffers from Mulaik's mistaken view that variates are the objects of scientific inquiry, and that, additional knowledge about some psychological phenomena, say, that of self-esteem, might be used, in a given application, to pick out from set *C* a "true" common factor. As argued in Part II, contrary to the confusions inherent to the practice of latent variate interpretation, the variates contained in *C* are not signified by concepts from ordinary language. They are not, e.g., various types of self-esteem or dominance. They are just variates, differentiated according to their statistical properties and means of production. The only way to put Mulaik's second suggestion into operation would be by formulating further *quantitative* stipulations that would single out an element of *C* as *the* ulcf replacement of a given **X**. In the next section, one such stipulation is considered.

ii. Variate domain foundation. The variate domain, or "behaviour domain", treatment of latent variable modeling has been a topic of some controversy. The idea is that the set of p input variates to be analyzed in a given application are but a subset of the variates contained in a population (domain) of variates that could potentially have been analyzed. The aim of factor analyses based on subsets drawn from this domain is then to allow for the making of inferences about the domain, notably about the identities and natures of the factors of the domain. There is no question that both Spearman and Guttman came to view it as essential that factor analysis be founded on the variate domain conception. McDonald (e.g., 1996a, p.598) has further taken his abstractive property conception of the referent of the concept *latent variate to* ***X*** as implying a variate domain foundation to latent variable modeling. The variate domain conception has been controversial because, as was seen in Chapter V, many have seen it as the key to ridding factor analytic practice of the difficulties caused by the indeterminacy property of the ulcf replacement. Moreover, its treatment has been burdened by non-mathematical obscurities, notably in regard the concept of *behaviour domain* itself, and the means by which to decide what is, and is not, contained within a given domain.

The current treatment will: i) Review the variate domain treatment from the perspective of the logic of replacement variate generators. The mathematics reviewed draws heavily upon the papers of Mulaik and McDonald (1978) and McDonald and Mulaik (1979), these papers representing the most satisfactory integration of variate domain thinking into practical application.[1]; ii) Show that McDonald's use of the variate domain as foundation for his abstractive property interpretation of latent variable modeling is a covert attempt to assert the truth of the Central Account. The following conclusions will be reached:

a) The mathematics of the variate domain treatment of latent variable modeling establish that, under certain highly restrictive conditions, the number of replacement variates of a subset of p input variates drawn from a domain of variates that, additionally, replace the remaining (k-p) variates contained in the domain, decreases as k increases, until, in the limit, i.e., as $k \rightarrow \infty$, it is equal to unity. That is, if certain very restrictive conditions hold, the latitude inherent to the replacement of a given set of variates, A, can be reduced by insisting that the variates that replace set A, also replace an increasingly large number of additional variates;

b) The conditions that must be satisfied in order for this limiting uniqueness to obtain are prohibitively severe;

c) McDonald is mistaken that "difficulties in the traditional use of the linear factor model", and, in particular, those attendant to the practice of latent variate interpretation, are eliminated by employing a behaviour domain foundation. In fact, his abstractive property notion is just the CAM dressed in fancy clothing, and the fact that a replacement is unique does not save the CAM. Regardless of the cardinality of a replacement, latent variable models are not detectors of properties/attributes, and the practice of latent variate interpretation, resting, as it does, on a claim of the coherence of non-normative, conceptual signification, makes no sense;

d) The conditions that must be satisfied so as to support the claim that, in the limit, the cardinality of a ulcf replacement is unity, render equivalent a variety of distinct replacement generators, including the ulcf, pc1, and image generators. That is to say, the replacement variates produced by these generators happen to coincide under the special condition of limiting uniqueness of the ulcf replacement.

Basic results

As was seen in Chapter V, the indeterminacy property of ulcf representations, a property that has to do with the latitude inherent to ulcf replacements, has frequently been misportrayed as an issue of "measurement precision" or "predictive efficacy." Not surprisingly, then, the attempt to remedy problems perceived to arise from the indeterminacy property through the use of limiting results based on variate domain formulations have standardly been misportrayed as solutions to measurement and predictive problems. But the issue of determinacy in the limit bears neither on measurement, nor prediction (although it has a mathematical kinship with the latter), but,

1 An earlier (1996) account, Maraun (1996), was alleged by Professor Mulaik to constitute a misrepresentation of Mulaik(1978). I must acknowledge the correctness of Mulaik's charge, and hope, this time, to do justice to his fine paper.

rather, on the reduction in the cardinality of the set of variates that ulcf replace some particular $\underline{\mathbf{X}}^*$.

Consider a simple example of the linear prediction of a random variate $\mathbf{Z}$ by a set of p random variates $\{\mathbf{X}_1,\mathbf{X}_2,...,\mathbf{X}_p\}$. For convenience, let all variates have expectations of zero. One seeks a p-vector of real coefficients, $\underline{\alpha}$, such that the squared Pearson Product Moment Correlation of $\mathbf{Z}$ and $\underline{\mathbf{X}}'\underline{\alpha}$, $\rho^2(\mathbf{Z},\underline{\mathbf{X}}'\underline{\alpha})$, is a maximum. Maximizing $\rho^2(\mathbf{Z},\underline{\mathbf{X}}'\underline{\alpha})$ is equivalent to minimizing the quantity

$$(15.7) \qquad E(\mathbf{Z}-\underline{\mathbf{X}}'\underline{\alpha})^2$$

and the solution, $\underline{\alpha}=\Sigma_X^{-1}\underline{\sigma}_{XZ}$, exists and is unique so long as Σ_X^{-1} exists. The squared Pearson Product Moment Correlation of $\mathbf{Z}$ and the optimal linear predictor $\underline{\mathbf{X}}'\Sigma_X^{-1}\underline{\sigma}_{XZ}$ is called the *squared multiple correlation coefficient*, or *coefficient of determination*, and is symbolized as $R^2_{\mathbf{Z}.\underline{\mathbf{X}}}$. The quantity $(1-R^2_{\mathbf{Z}.\underline{\mathbf{X}}})$ corresponds to the proportion of variation in $\mathbf{Z}$ not linearly associated with $\underline{\mathbf{X}}$. Now, $R^2_{\mathbf{Z}.\underline{\mathbf{X}}}=1$ only if $\mathbf{Z}$ lies in the span of $\{\mathbf{X}_1,\mathbf{X}_2,...,\mathbf{X}_p\}$, for this is precisely the condition under which there exists a vector of real coefficients $\underline{\alpha}$ such that $E(\mathbf{Z}-\underline{\mathbf{X}}'\underline{\alpha})^2=0$. If this condition does not obtain, as is typically the case, the linear prediction of $\mathbf{Z}$ by$\{\mathbf{X}_1,\mathbf{X}_2,...,\mathbf{X}_p\}$ will be less than perfect. In marked contrast to the determinacy of ulcf replacements (the case in which $\rho^*=1$), $R^2_{\mathbf{Z}.\underline{\mathbf{X}}}$ can attain the value of unity when the number of predictors is finite, its value determined by the "naturally occuring" joint distribution of $\mathbf{Z}$ and $\underline{\mathbf{X}}$ in the population *P* under study. That is, unless the variates themselves are constructed in some artificial manner, the joint distribution of $\mathbf{Z}$ and $\underline{\mathbf{X}}$ in a population *P* under study can be viewed as a constituent of natural reality, and $R^2_{\mathbf{Z}.\underline{\mathbf{X}}}$, a desciption of certain of its properties.

Now, consider adding a (p+1)th variate, $\mathbf{Y}$, to the set of predictors. The (1+p+1)-element joint distribution of $(\mathbf{Z},\underline{\mathbf{X}},\mathbf{Y})$ in population *P* is now of interest. Because $R^2_{\mathbf{Z}.(\underline{\mathbf{X}}:\mathbf{Y})}=R^2_{\mathbf{Z}.\underline{\mathbf{X}}}+R^2_{\mathbf{Z}.(\mathbf{Y}|\underline{\mathbf{X}})}$, in which $R^2_{\mathbf{Z}.(\mathbf{Y}|\underline{\mathbf{X}})}$ is the squared semi-partial correlation of $\mathbf{Z}$ and $\mathbf{Y}$, i.e., the squared correlation of $\mathbf{Z}$ and $(\mathbf{Y}-\underline{\mathbf{X}}'\Sigma_X^{-1}\underline{\sigma}_{XY})$, and $R^2_{\mathbf{Z}.(\mathbf{Y}|\underline{\mathbf{X}})}\geq 0$, it follows that $R^2_{\mathbf{Z}.(\underline{\mathbf{X}}:\mathbf{Y})}\geq R^2_{\mathbf{Z}.\underline{\mathbf{X}}}$. That is, adding a (p+1)th variate, $\mathbf{Y}$, to the predictive set will improve the linear prediction of $\mathbf{Z}$ so long as $\mathbf{Z}$ is linearly associated with the residual from the linear regression of $\mathbf{Y}$ on $\underline{\mathbf{X}}$. Note, however, the important fact that $\mathbf{Z}$, the variate to be predicted, *does not change as a function of the variates contained in the set of predictors*. Variate $\mathbf{Z}$ is precisely the same variate after the addition of $\mathbf{Y}$ as it was before the addition of $\mathbf{Y}$. Thus, regardless of which predictors are used, or how many predictors are used, it is always clear *what* is to be predicted, and how predictive efficacy can be partitioned with respect the predictors.

This general issue of prediction, and the complimentary issue of the adding of predictor variates to improve prediction, does *not* describe the determinacy issue of ulcf replacements. The distinction between the two scenarios is clear when it is noted that *the variate $\boldsymbol{\theta}$ that appears in the equations of the ulcf generator is not fixed prior to analysis*. Prior to analysis, no rule exists by which scores on $\boldsymbol{\theta}$ are produced. Hence, there does not exist a joint distribution of scores on $\boldsymbol{\theta}$ and those on $\{\mathbf{X}_1,\mathbf{X}_2,...,\mathbf{X}_p\}$, in population *P*, whose properties can be studied in the determination of a predictive relationship between

$\boldsymbol{\theta}$ and $\{\mathbf{X}_1,\mathbf{X}_2,...,\mathbf{X}_p\}$. The symbol $\boldsymbol{\theta}$ represents any random variate that satisfies the requirements imposed by the ulcf generator in the replacement of $\{\mathbf{X}_1,\mathbf{X}_2,...,\mathbf{X}_p\}$. The distributional specifications that are sometimes a part of the ulcf generator, e.g., that $\boldsymbol{\theta}$ and $\{\mathbf{X}_1,\mathbf{X}_2,...,\mathbf{X}_p\}$ have a joint normal distribution, are *requirements* imposed on the distribution of constructed random variate $\boldsymbol{\theta}$ by the ulcf generator. Supplementing $\{\mathbf{X}_1,\mathbf{X}_2,...,\mathbf{X}_p\}$ with additional variates turns out, not surprisingly, to generate additional requirements that must be satisfied by a variate that is to qualify as a ulcf replacement variate. Determinacy in the limit, then, is not a predictive issue, but is, rather, about nested sets of replacement variates, and the possibility that, by increasing the number of input variates that must be replaced, enough restrictions will eventually be imposed to bring about a replacement in which Card(C)=1. To motivate analysis of this topic, consider a simple example in which many of the key issues are evident.

Let there be a particular set of p input variates $\underline{\mathbf{X}}_o$, and let this set be ulcf replaceable with ulcf representation $\underline{\mathbf{X}}_o=\underline{\Lambda}_o\boldsymbol{\theta}+\Psi_o^{1/2}\underline{\boldsymbol{\delta}}$. Let the set of replacement variates to $\underline{\mathbf{X}}_o$ be $C_{[\underline{\mathbf{X}}o]}$, and note that this set has minimum correlation equal to $\rho^*_{(\underline{\mathbf{X}}o)}=2R^2_{\boldsymbol{\theta}.(\underline{\mathbf{X}}o)}-1=2\underline{\Lambda}_o'\Sigma_{\underline{\mathbf{X}}o}^{-1}\underline{\Lambda}_o-1$ (see 4.47). The variates, $\boldsymbol{\theta}_i$, contained in $C_{[\underline{\mathbf{X}}o]}$, and constructed as $\boldsymbol{\theta}_i=\underline{\Lambda}_o'\Sigma_{\underline{\mathbf{X}}o}^{-1}\underline{\mathbf{X}}_o+w_o^{1/2}\mathbf{s}_i$ ($w_o=(1-\underline{\Lambda}_o'\Sigma_o^{-1}\underline{\Lambda}_o)$, $C(\mathbf{s}_i,\underline{\mathbf{X}}_o)=\underline{0}$, $E(\mathbf{s}_i)=0$, and $V(\mathbf{s}_i)=1$), each have the properties $E(\boldsymbol{\theta}_i)=0$, $V(\boldsymbol{\theta}_i)=1$, and $E(\underline{\mathbf{X}}_o\boldsymbol{\theta}_i)=\underline{\Lambda}_o$. Now, add a (p+1)th variate $\mathbf{V}$ to the set, $\underline{\mathbf{X}}_o$, of input variates.

Theorem (Steiger, 1996): A sufficient condition for the set $[\underline{\mathbf{X}}_o:\mathbf{V}]$ to be ulcf-replaceable by a subset $S(C_{[\underline{\mathbf{X}}o]})$ of $C_{[\underline{\mathbf{X}}o]}$ is that $E(\underline{\mathbf{X}}_o\mathbf{V})=\kappa\underline{\Lambda}_o$, for some real number $\kappa\neq0$.

Proof

If $E(\underline{\mathbf{X}}_o\mathbf{V})=\kappa\underline{\Lambda}_o$, for some real number $\kappa\neq0$, then $[\underline{\mathbf{X}}_o:\mathbf{V}]$ is ulcf-replaceable, with ulcf representation

(15.8) $$\begin{pmatrix}\underline{\mathbf{X}}_o\\ \mathbf{V}\end{pmatrix}=\begin{pmatrix}\Lambda_o\\ \kappa\end{pmatrix}\boldsymbol{\varphi}+\Omega^{\frac{1}{2}}\underline{\boldsymbol{\varepsilon}},$$

in which $V(\boldsymbol{\varphi})=1$, $C(\underline{\boldsymbol{\varepsilon}})=I$, $C(\boldsymbol{\varphi},\underline{\boldsymbol{\varepsilon}})=\underline{0}$, and $\Omega=\begin{pmatrix}\Psi_o & \circ\\ \circ & (\sigma^2_V-\kappa^2)\end{pmatrix}$. Because $E(\underline{\mathbf{X}}_o\boldsymbol{\varphi})=\underline{\Lambda}_o$, the variates $\boldsymbol{\varphi}_j$ contained in set $C_{[\underline{\mathbf{X}}o:\mathbf{V}]}$, i.e., the variates that replace $[\underline{\mathbf{X}}_o:\mathbf{V}]$, are also contained in $C_{[\underline{\mathbf{X}}o]}$. That is, there exists a non-empty subset of $C_{[\underline{\mathbf{X}}o]}$ that ulcf-replaces $[\underline{\mathbf{X}}_o:\mathbf{V}]$□

Note that the replacement variates to $[\underline{\mathbf{X}}_o:\mathbf{V}]$ contained in set $C_{[\underline{\mathbf{X}}o:\mathbf{V}]}$ are constructed as $\boldsymbol{\varphi}_j=[\underline{\Lambda}_o:\kappa]'\Sigma_{(\underline{\mathbf{X}}o:\mathbf{V})}^{-1}\begin{pmatrix}\underline{\mathbf{X}}_o\\ \mathbf{V}\end{pmatrix}+W_{(\underline{\mathbf{X}}o:\mathbf{V})}^{1/2}\mathbf{s}_j$, and that $C_{[\underline{\mathbf{X}}o:\mathbf{V}]}$ has minimum correlation

$\rho^*_{(\underline{\mathbf{X}}o:V)}=2R^2_{\varphi.(\underline{\mathbf{X}}o:V)}-1=2[\underline{\Lambda}_o:\kappa]'\Sigma_{(\underline{\mathbf{X}}o:V)}^{-1}\begin{pmatrix}\Lambda_o\\ \kappa\end{pmatrix}-1$. Because Ω is positive definite, it follows that

(15.9)

$$t_{(\underline{\mathbf{X}}o:V)}=[\underline{\Lambda}_o':\kappa]\begin{pmatrix}\Psi_o^{-1} & \circ\\ \circ & (\sigma^2_V-\kappa^2)^{-1}\end{pmatrix}\begin{bmatrix}\Lambda_o\\ \kappa\end{bmatrix}=\underline{\Lambda}_o'\Psi_o^{-1}\underline{\Lambda}_o+\kappa^2(\sigma^2_V-\kappa^2)^{-1}>\underline{\Lambda}_o'\Psi_o^{-1}\underline{\Lambda}_o=t_{(\underline{\mathbf{X}}o)}.$$

Hence, from (4.48), it follows that

(15.10) $$\rho^*_{(\underline{\mathbf{X}}o:V)}=\frac{(t_{(\underline{X}_o:V)}-1)}{(t_{(\underline{X}_o:V)}+1)}>\frac{(t_{(\underline{X}_o)}-1)}{(t_{(\underline{X}_o)}+1)}=\rho^*_{(\underline{\mathbf{X}}o)}$$

That is, the minimum correlation for set $C_{[\underline{\mathbf{X}}o:V]}$ is larger than that for set $C_{[\underline{\mathbf{X}}o]}$. Thus, if $\underline{\mathbf{X}}_o$ is ulcf replaceable, insisting that a subset of its replacement variates replace, additionally, a (p+1)th input variate **V** has the effect of imposing additional requirements on the replacement, the result being a reduction in the latitude inherent to the replacement. The implication is that, if $\underline{\mathbf{X}}$ is ulcf replaceable, and it is insisted that a subset of its replacement variates replace an increasingly larger number of additional variates, then, if this replaceability continues to obtain, the number of imposed requirements will eventually render the cardinality of this subset unity.

However, the conditions under which this convergence to unique replacement will take place are severe, as will later be seen. It is suggestive to note that if just a single variate is added to set $\underline{\mathbf{X}}_o$, a reduction in the cardinality of the replacement will occur only if $E(\underline{\mathbf{X}}_o\mathbf{V})$ is equal to $\kappa\underline{\Lambda}_o$, $\kappa\neq0$. As Steiger (1996a, Theorem 2) shows, this is equivalent to the requirement that the linear conditional expectation of **V** given $\underline{\mathbf{X}}_o$ is a linear function of the determinate parts of the replacement variates, $\boldsymbol{\theta}_i$. The proof runs as follows:

$E(\mathbf{V}|\underline{\mathbf{X}}_o=\underline{X}_o)=\underline{\sigma}_{X_oV}'\Sigma_{\underline{X}_o}^{-1}\underline{X}_o$. Because $\underline{\sigma}_{X_oV}=E(\underline{\mathbf{X}}_o\mathbf{V})$ must be equal to $\kappa\underline{\Lambda}_o$, $\kappa\neq0$, $E(\mathbf{V}|\underline{\mathbf{X}}_o=\underline{X}_o)$ must then equal $\kappa\underline{\Lambda}_o'\Sigma_{\underline{\mathbf{X}}o}^{-1}\underline{\mathbf{X}}_o=\kappa\mathbf{D}_{\boldsymbol{\theta}}$ (see 4.2iii). Hence, the additional variate **V** must only be linearly related to a single linear composite, $\kappa\mathbf{D}_{\boldsymbol{\theta}}$, of the original input variates, $\underline{\mathbf{X}}_o$.

Let it be the case that $E(\underline{\mathbf{X}}_o\mathbf{V})=\kappa\underline{\Lambda}_o$, $\kappa\neq0$. It follows then that all of the replacement variates contained in $C_{[\underline{\mathbf{X}}o:V]}$ possess the properties required for inclusion in $C_{[\underline{\mathbf{X}}o]}$, but not all of the variates contained in $C_{[\underline{\mathbf{X}}o]}$ have the covariance of κ with **V** required for inclusion in $C_{[\underline{\mathbf{X}}o:V]}$. Only a subset do. Hence, $C_{[\underline{\mathbf{X}}o:V]}\subset C_{[\underline{\mathbf{X}}o]}$. What are the additional restrictions imposed on the replacement variates $\boldsymbol{\theta}_i=\underline{\Lambda}_o'\Sigma_{\underline{\mathbf{X}}o}^{-1}\underline{\mathbf{X}}_o+w_o^{1/2}\mathbf{s}_i$, the satisfaction of which admits them to $S(C_{[\underline{\mathbf{X}}o]})$? To put it another way, what distinguishes subset $S(C_{[\underline{\mathbf{X}}o]})$ from the other $\boldsymbol{\theta}_i$ contained in $C_{[\underline{\mathbf{X}}o]}$? Clearly, as for any variate contained in $C_{[\underline{\mathbf{X}}o]}$, these variates must satisfy: $E(\boldsymbol{\theta}_i\underline{\mathbf{X}}_o)=\underline{\Lambda}_o$, $E(\boldsymbol{\theta}_i)=0$, and $V(\boldsymbol{\theta}_i)=1$. These properties are realized by the imposition of moment constraints on $\mathbf{s}_i$, namely that $E(\mathbf{s}_i\underline{\mathbf{X}}_o)=\underline{0}$, $E(\mathbf{s}_i)=0$, and $V(\mathbf{s}_i)=1$. Variates in $S(C_{[\underline{\mathbf{X}}o]})$ additionally satisfy the requirement that $E(\boldsymbol{\theta}_i\mathbf{V})=\kappa$. This requirement is equivalent to the requirement that

(15.11) $$E((\underline{\Lambda}_o'\Sigma_{\underline{\mathbf{X}}o}^{-1}\underline{\mathbf{X}}_o+w_o^{1/2}\mathbf{s}_i)\mathbf{V})=\kappa,$$

which obtains only if

(15.12) $$\sigma_{\mathbf{V}\mathbf{s}i}=\frac{(\kappa-\underline{\Lambda}_o'\Sigma_{\underline{X}_o}^{-1}\underline{\sigma}_{\underline{X}_oV})}{w_o^{1/2}}$$

Hence, the elements of the subset $S(C_{[\underline{\mathbf{X}}o]})$ of variates that replace both the original set of input variates $\underline{\mathbf{X}}_o$ and, additionally, the variate $\mathbf{V}$, are constructed by choosing $\mathbf{s}_i$ so as to satisfy one further restriction, that $\sigma_{\mathbf{V}\mathbf{s}i}=\dfrac{(\kappa-\underline{\Lambda}_o'\Sigma_{\underline{X}_o}^{-1}\underline{\sigma}_{\underline{X}_oV})}{w_o^{1/2}}$.

Now, let the situation of a "domain of variates" be considered.

Definition (domain of input variates). A domain, D, of input variates is a set of input variates, $\mathbf{X}_j$, j=1..k, jointly distributed in some population, P_T, from which subsets of variates can be drawn as input to particular analyses.

Note, that with the notion of "domain of variates" comes the implication that the variates represent phenomena of a "similar type". McDonald and Mulaik, for example speak of variates that "...have been selected on the basis of certain common attributes defined in advance..." (1979, p.305). The idea is that there exists a rule r_σ for the production of a potentially unlimited number of "indicators" of some construct σ of interest. Mulaik (1996, p.582) makes this explicit: "I would now emphasize that having a *prior definition* of a domain in general, whether with a determinate common factor or not, means distinguishing the variables of the domain in some way external to the model by, say, some mark, some behavioral attribute, some operation and context of measurement that is an observable distinguishing feature of each variable in the domain." The researcher can then envision constructing a "universe of content or behaviour domain...defined in advance of any statistical analysis" (McDonald & Mulaik, 1979, p.302). It was argued in Chapter IX that Mulaik and McDonald's grasp of concept meaning and signification is confused, and that their program is problematic on many fronts. The issue of interest at the moment, however, is the mathematics of limiting determinacy.

Let there be a domain, D, containing k input variates, $\mathbf{X}_i$, i=1..k. For convenience, let these variates have expectations of zero, and imagine drawing p<k of these variates for analysis. The p variates sampled are placed in random vector $\underline{\mathbf{X}}_A$, and the remaining m=(k-p) of the variates of D in vector $\underline{\mathbf{X}}_B$. Let $\underline{\mathbf{X}}_A$ be ulcf-replaceable with ulcf representation $\underline{\mathbf{X}}_A=\underline{\Lambda}_A\boldsymbol{\theta}_A+\Psi_A^{1/2}\underline{\boldsymbol{\delta}}_A$, and let $C_{[A]}$ stand for the set of ulcf replacements to $\underline{\mathbf{X}}_A$. The minimum correlation of the replacement is equal to $\rho^*_{(A)}=2R^2_{\boldsymbol{\theta}.(A)}-1=2\underline{\Lambda}_A'\Sigma_{(A)}^{-1}\underline{\Lambda}_A-1$, and the variates contained in $C_{[A]}$ are constructed as $\boldsymbol{\theta}_i=\underline{\Lambda}_A'\Sigma_{(A)}^{-1}\underline{\mathbf{X}}_A+w_A^{1/2}\mathbf{s}_i$, in which $w_A=(1-\underline{\Lambda}_A'\Sigma^{-1}\underline{\Lambda}_A)$, $C(\mathbf{s}_i,\underline{\mathbf{X}})=\underline{0}$, $E(\mathbf{s}_i)=0$, and $V(\mathbf{s}_i)=1$. Each of the variates contained in $C_{[A]}$ has the properties $E(\boldsymbol{\theta}_i)=0$, $V(\boldsymbol{\theta}_i)=1$, and $E(\underline{\mathbf{X}}\boldsymbol{\theta}_i)=\underline{\Lambda}_A$.

Definition (A-replaceability): Domain D will be said to be A-replaceable if there exists a subset of $C_{[A]}$, $S_D(C_{[A]})$, whose elements each ulcf-replace the variates contained in D. If such a subset exists, its elements will be called A-replacements to D.

Theorem (A-replaceability): Let $\underline{\mathbf{X}}_A$ be ulcf-replaceable with ulcf representation $\underline{\mathbf{X}}_A=\underline{\Lambda}_A\boldsymbol{\theta}_A+\Psi_A^{1/2}\underline{\boldsymbol{\delta}}_A$. Domain D is, then, A-replaceable if and only if $[\underline{\mathbf{X}}_A:\underline{\mathbf{X}}_B]$ is ulcf-replaceable.

Proof

→

Assume that D is A-replaceable: There exists a subset of $C_{[A]}$, $S_D(C_{[A]})$, whose elements, $\boldsymbol{\varphi}_i$, ulcf-replace the variates of D. Because the variates $\boldsymbol{\varphi}_i$ are contained in $C_{[A]}$, it must be the case that $E(\underline{\mathbf{X}}_A\boldsymbol{\varphi}_i)=\underline{\Lambda}_A$, and, hence, that the ulcf representation of the variates contained in D is

(15.13) $$\begin{pmatrix}\underline{\mathbf{X}}_A\\ \underline{\mathbf{X}}_B\end{pmatrix}=\begin{pmatrix}\underline{\Lambda}_A\\ \underline{\Lambda}_B\end{pmatrix}\boldsymbol{\varphi}+\Omega^{\frac{1}{2}}\underline{\boldsymbol{\varepsilon}},$$

in which $V(\boldsymbol{\varphi})=1$, $C(\underline{\boldsymbol{\varepsilon}})=I$, $C(\boldsymbol{\varphi},\underline{\boldsymbol{\varepsilon}})=\underline{0}$, and $\Omega^{1/2}$ is a diagonal, positive definite matrix. That is, $[\underline{\mathbf{X}}_A:\underline{\mathbf{X}}_B]$ is ulcf-replaceable.

←

Assume that $[\underline{\mathbf{X}}_A:\underline{\mathbf{X}}_B]$ is ulcf-replaceable, with ulcf representation

(15.14) $$\begin{pmatrix}\underline{\mathbf{X}}_A\\ \underline{\mathbf{X}}_B\end{pmatrix}=\begin{pmatrix}\underline{\gamma}\\ \underline{\Lambda}_B\end{pmatrix}\boldsymbol{\varphi}+\Omega^{\frac{1}{2}}\underline{\boldsymbol{\varepsilon}},$$

in which $V(\boldsymbol{\varphi})=1$, $C(\underline{\boldsymbol{\varepsilon}})=I$, $C(\boldsymbol{\varphi},\underline{\boldsymbol{\varepsilon}})=\underline{0}$, and $\Omega^{1/2}$ is a diagonal, positive definite matrix. Then

$$(15.15) \qquad C[\underline{\mathbf{X}}_A:\underline{\mathbf{X}}_B]=E\begin{pmatrix}\underline{\mathbf{X}}_A\underline{\mathbf{X}}_A' & \underline{\mathbf{X}}_A\underline{\mathbf{X}}_B' \\ \underline{\mathbf{X}}_B\underline{\mathbf{X}}_A' & \underline{\mathbf{X}}_B\underline{\mathbf{X}}_B'\end{pmatrix}=\begin{pmatrix}\underline{\gamma}\underline{\gamma}+U_A & \underline{\gamma}\underline{\Lambda}_B' \\ \underline{\Lambda}_B\underline{\gamma}' & \underline{\Lambda}_B\underline{\Lambda}_B'+U_B\end{pmatrix}$$

and, since $E(\underline{\mathbf{X}}_A\underline{\mathbf{X}}_A')=\underline{\gamma}\underline{\gamma}'+U_A$, it must be that $\underline{\gamma}=\underline{\Lambda}_A$, and, hence, that $E(\underline{\mathbf{X}}_A\boldsymbol{\varphi})=\underline{\Lambda}_A$. Thus, the variates $\boldsymbol{\varphi}_i$ that ulcf replace $[\underline{\mathbf{X}}_A:\underline{\mathbf{X}}_B]$, those contained in $C_{[A:B]}$, are also contained in $C_{[A]}$. In other words, there exists a subset of C_A that ulcf-replaces $[\underline{\mathbf{X}}_A:\underline{\mathbf{X}}_B]$, and domain D is A-replaceable$_\square$

Note that, if domain D is A-replaceable, $U_A=E(\underline{\mathbf{X}}_A\underline{\mathbf{X}}_A')-\underline{\Lambda}_A\underline{\Lambda}_A'=\Psi_A$. Note also the severity of the restrictions that must be satisfied by the distribution of $[\underline{\mathbf{X}}_A:\underline{\mathbf{X}}_B]$ in order that D be A-replaceable. From (7.15), it follows, in analogy to the single variate case, that it must be possible to choose a vector of real coefficients $\underline{\Lambda}_B$ such that $E(\underline{\mathbf{X}}_A\underline{\mathbf{X}}_B')=\underline{\Lambda}_A\underline{\Lambda}_B'$ and $E(\underline{\mathbf{X}}_B\underline{\mathbf{X}}_B')=\underline{\Lambda}_B\underline{\Lambda}_B'+U_B$, in which U_B is diagonal and positive definite. The number of such restrictions is an increasing function of m, the number of variates contained in $\underline{\mathbf{X}}_B$. If k is taken as being very large, then the number of restrictions that must be satisfied by the distribution of $[\underline{\mathbf{X}}_A:\underline{\mathbf{X}}_B]$ will also be very large. Given that domain D is A-replaceable, let the set $C_{[D]}$ contain its replacement variates. Elements of $C_{[D]}$ are constructed as $\boldsymbol{\varphi}_i=[\underline{\Lambda}_A:\underline{\Lambda}_B]'\Sigma_{(A:B)}^{-1}\begin{pmatrix}\underline{\mathbf{X}}_A\\ \underline{\mathbf{X}}_B\end{pmatrix}+W_{(A:B)}^{1/2}\mathbf{s}_i$. Set $C_{[D]}$ has minimum correlation equal to $\rho^*_{(A:B)}=2R^2_{\boldsymbol{\varphi}.(A:B)}-1=2[\underline{\Lambda}_A:\underline{\Lambda}_B]'\Sigma_{(A:B)}^{-1}\begin{pmatrix}\underline{\Lambda}_A\\ \underline{\Lambda}_B\end{pmatrix}-1$.

Theorem (relationship between sets $C_{[D]}$ and $C_{[A]}$ under A-replaceability): If domain D is A-replaceable, then every replacement variate $\boldsymbol{\varphi}_i$ contained in $C_{[D]}$ is also contained in $C_{[A]}$, i.e., $C_{[D]}\subset C_{[A]}$, but not every variate, $\boldsymbol{\theta}$, contained in $C_{[A]}$ is contained in $C_{[D]}$. That is, $C_{[A]}\not\subset C_{[D]}$, and the set of A-replacements to D is a proper subset of $C_{[A]}$.

Proof (Mulaik & McDonald, 1978)

Set $C_{[D]}$ contains those variates $\boldsymbol{\varphi}_i$ that ulcf replace $[\underline{\mathbf{X}}_A:\underline{\mathbf{X}}_B]$, and, because D is A-replaceable, it must be the case that $E(\underline{\mathbf{X}}_A\boldsymbol{\varphi}_i)=\underline{\Lambda}_A$. That is, $C_{[D]}\subset C_{[A]}$. Thus, given that D is A-replaceable, $C_{[D]}$ is precisely that set of variates, $S_D(C_{[A]})$, that is both contained in $C_{[A]}$ and replaces the variates contained in set B. What remains to be shown is that not every variate contained in $C_{[A]}$ is also contained in $C_{[D]}$. The correlation, $\rho_{ij[A]}$, between any two replacement variates contained in $C_{[A]}$ must lie within the bounds $\rho^*_{(A)}\le\rho_{ij[A]}\le 1$, while the correlation $\rho_{ij[D]}$, between any two replacement variates contained in $C_{[D]}$ must lie within the bounds $\rho^*_{(A:B)}\le\rho_{ij[D]}\le 1$. Because

$$(15.16) \qquad C[\underline{\mathbf{X}}_A:\underline{\mathbf{X}}_B]=\begin{pmatrix}\underline{\Lambda}_A\underline{\Lambda}_A+\Psi_A & \underline{\Lambda}_A\underline{\Lambda}_B' \\ \underline{\Lambda}_B\underline{\Lambda}_A' & \underline{\Lambda}_B\underline{\Lambda}_B'+U_B\end{pmatrix},$$

with both Ψ_A and U_B diagonal and positive definite,

(15.17)

$$t_{(A:B)}=[\underline{\Lambda}_A':\underline{\Lambda}_B']\begin{pmatrix}\Psi_A^{-1} & \circ \\ \circ & U_B^{-1}\end{pmatrix}\begin{bmatrix}\underline{\Lambda}_A \\ \underline{\Lambda}_B\end{bmatrix}=\underline{\Lambda}_A'\Psi_A^{-1}\underline{\Lambda}_A+\underline{\Lambda}_B'U_B^{-1}\underline{\Lambda}_B>\underline{\Lambda}_A'\Psi_A^{-1}\underline{\Lambda}_A=t_{(A)}.$$

Hence,

(15.18) $$\rho^*_{(A:B)}=\frac{(t_{(A:B)}-1)}{(t_{(A:B)}+1)}>\frac{(t_{(A)}-1)}{(t_{(A)}+1)}=\rho^*_{(A)}.$$

Let particular variate $\boldsymbol{\theta}^*$ be an element of set $C_{[D]}$, and hence, also an element of $C_{[A]}$. Then also contained in $C_{[A]}$ is a subset of replacement variates, $\min(C_{[A]},\boldsymbol{\theta}^*)$, each of which has minimum correlation $\rho^*_{(A)}$ with $\boldsymbol{\theta}^*$. Since the minimum correlation for $C_{[D]}$ is equal to $\rho^*_{(A:B)}$, which is greater than $\rho^*_{(A)}$, the variates in $\min(C_{[A]},\boldsymbol{\theta}^*)$ cannot be elements of $C_{[D]}$□

Given the A-replaceability of D, all of the variates $\boldsymbol{\varphi}_i$ contained in $C_{[D]}$ have the property that $E(\underline{\mathbf{X}}_A\boldsymbol{\varphi}_i)=\underline{\Lambda}_A$. That is, these variates are also contained in $C_{[A]}$. However, the variates $\boldsymbol{\theta}_i$ that are contained in $C_{[A]}$ do not necessarily have the property that $E(\underline{\mathbf{X}}_B\boldsymbol{\theta}_i)=\underline{\Lambda}_B$, this required for inclusion in $C_{[D]}$. Only a subset of the variates contained in $C_{[A]}$ have this property. Hence, $C_{[D]}\subset C_{[A]}$. Once again, it can be asked what additional restrictions are imposed on the replacement variates $\boldsymbol{\theta}_i=\underline{\Lambda}_A'\Sigma_{(A)}^{-1}\underline{\mathbf{X}}_A+w_A^{1/2}\mathbf{s}_i$, the satisfaction of which admits them to $S_D(C_{[A]})$. Such variates must satisfy the requirements that $E(\boldsymbol{\theta}_i\underline{\mathbf{X}}_A)=\underline{\Lambda}_A$, $E(\boldsymbol{\theta}_i)=0$, $V(\boldsymbol{\theta}_i)=1$, and the additional requirement that $E(\boldsymbol{\theta}_i\underline{\mathbf{X}}_B)=\underline{\Lambda}_B$. The first three properties are realized by the imposition of the standard moment constraints on $\mathbf{s}_i$, namely that $E(\mathbf{s}_i\underline{\mathbf{X}}_A)=\underline{0}$, $E(\mathbf{s}_i)=0$, $V(\mathbf{s}_i)=1$. To satisfy the additional requirement that $E(\boldsymbol{\theta}_i\underline{\mathbf{X}}_B)=\underline{\Lambda}_B$, it must be the case that

(15.19) $$E((\underline{\Lambda}_A'\Sigma_{(A)}^{-1}\underline{\mathbf{X}}_A+w_A^{1/2}\mathbf{s}_i)\underline{\mathbf{X}}_B')=\underline{\Lambda}_B',$$

and, hence, that

(15.20) $$\underline{\sigma}_{\underline{B}si}=\frac{(\underline{\Lambda}_B-\Sigma_{BA}\Sigma_A^{-1}\underline{\Lambda}_A)}{w_A^{1/2}}$$

Thus, if D is A-replaceable, the elements of the subset $S_D(C_{[\underline{\mathbf{X}}]})$ of variates which both replace the original set of input variates, $\underline{\mathbf{X}}_A$, and, additionally, the variates $\underline{\mathbf{X}}_B$, are constructed by choosing $\mathbf{s}_i$ variates which satisfy the m additional moment constraints of (15.20). Hence, further specifications are added to the standard construction formula, or recipe, of the ulcf replacement generator.

Theorem (convergence to uniqueness of the A-replacements to D): Let domain D be A-replaceable. Then, as m, the number of variates contained in $\underline{\mathbf{X}}_B$ grows indefinitely large, i.e., $k\rightarrow\infty$, the cardinality of the set of A-replacements to D converges to unity, i.e., $\mathrm{Card}(S_D(C_{[A]}))\rightarrow 1$.

Proof

If domain D is A-replaceable, then, as $k\rightarrow\infty$, $t_{(A:B)}$ of (15.17) becomes increasingly large, and $\rho^*_{(A:B)}=\dfrac{(t_{(A:B)}-1)}{(t_{(A:B)}+1)}\rightarrow 1$. Hence, in the limit, $C_{[D]}$ contains but one variate, i.e., $\mathrm{Card}(S_D(C_{[A]})\rightarrow 1$□

Now, this, and analogous results, is the raison d'etre of the variate domain response to the indeterminacy problem. For it shows one way in which a single replacement variate can be singled out as preferred from among the elements of $C_{[A]}$: The preferred variate is that variate that is contained in $C_{[A]}$ and, additionally, replaces a very large number of additional variates. It must be noted, however, that:

i) Such a variate does not necessarily exist. Within the framework discussed herein, its existence requires that, as $m\rightarrow\infty$, D continues to be ulcf-replaceable. As mentioned earlier, this will only occur given the satisfaction of an enormous number of restrictions on the elements of the covariance matrix of the variates to be replaced.

ii) Contrary to the views of Spearman, Piaggio, and, in certain installments, McDonald and Mulaik, the issue of limiting determinacy can be subsumed neither within the logic of standard prediction, nor reliability. Determinacy in the limit is just the limiting uniqueness of a replacement. It is the special situation in which, by adding variates to the set that must be replaced, so many restrictions are eventually imposed that only one replacement variate to $\underline{\mathbf{X}}_A$, one variate constructed as $\boldsymbol{\theta}_i=\underline{\Lambda}_A'\Sigma_{(A)}^{-1}\underline{\mathbf{X}}_A+w_A^{1/2}\mathbf{s}_i$, satisfies them all. This means of achieving uniqueness of replacement is akin to that encountered in non-metric multidimensional scaling, wherein a t-dimensional solution configuration is "tightened" through the addition of variates to be represented in the configuration.

Theorem (relationships between A-replaceable domains): Let there be three sets of variates, $\underline{\mathbf{X}}_A$ consisting of p variates, $\underline{\mathbf{X}}_B$ of m variates, and $\underline{\mathbf{X}}_C$ of n variates, and let there be no overlap in these sets. Let $\underline{\mathbf{X}}_A$ be ulcf replaceable with ulcf representation $\underline{\mathbf{X}}_A=\underline{\Lambda}_A\boldsymbol{\theta}_A+\Psi_A^{1/2}\boldsymbol{\delta}_A$, $C_{[A]}$ be the set of replacement variates to $\underline{\mathbf{X}}_A$, and

$\rho^*_{(A)}=2R^2_{\boldsymbol{\theta}.(A)}-1=2\underline{\Lambda}_A'\Sigma_{(A)}^{-1}\underline{\Lambda}_A-1$ be its minimum correlation. The following statements regarding the variate domains [A:B], [A:C], and [A:B:C] are true:

i. Domain [A:B:C] is A-replaceable if and only if $[\underline{\mathbf{X}}_A:\underline{\mathbf{X}}_B:\underline{\mathbf{X}}_C]$ is ulcf-replaceable;

ii. If domains [A:B], [A:C], and [A:B:C] are each A-replaceable, and $S_1(C_{[A]})$, $S_2(C_{[A]})$, and $S_3(C_{[A]})$ are those subsets of $C_{[A]}$ that A-replace [A:B], [A:C], and [A:B:C], respectively, then $S_1(C_{[A]})\equiv C_{[A:B]}$, $S_2(C_{[A]})\equiv C_{[A:C]}$, and $S_3(C_{[A]})\equiv C_{[A:B:C]}$.

iii. If domains [A:B] and [A:C] are each A-replaceable, then $C_{[A:B]},C_{[A:C]}\subset C_{[A]}$

iv. If domains [A:B], [A:C], and [A:B:C] are each A-replaceable, then $C_{[A:B:C]}\subset C_{[A:B]},C_{[A:C]}$

v. If domain [A:B:C] is A-replaceable, then, as $(m+n)\rightarrow\infty$, the cardinality of the set of A-replacements to [A:B:C] converges to unity, i.e., $\mathrm{Card}(S_3(C_{[A]}))\rightarrow 1$.

vi. If domains [A:B] and [A:C] are each A-replaceable, $\rho(\boldsymbol{\varphi}_i,\boldsymbol{\omega}_j)$, the correlation between any member of $C_{[A:B]}$ and any member of $C_{[A:C]}$, satisfies the bounds $\rho^*_{(A)}=2\underline{\Lambda}_A'\Sigma_{(A)}^{-1}\underline{\Lambda}_A-1\leq\rho(\boldsymbol{\varphi}_i,\boldsymbol{\omega}_j)\leq 1$.

vii. The upper bound of (vi) is obtained if, in addition to the conditions stated in (vi), [A:B:C] is A-replaceable and m and n both go to infinity.

Proof

i. The proof is analogous to that of the eariler theorem on A-replaceability□

ii. Let:

Set [A:B] be A-replaceable and $[\underline{\mathbf{X}}_A:\underline{\mathbf{X}}_B]$ have ulcf representation

(15.21) $$\begin{pmatrix}\underline{\mathbf{X}}_A\\ \underline{\mathbf{X}}_B\end{pmatrix}=\begin{pmatrix}\underline{\Lambda}_A\\ \underline{\Lambda}_B\end{pmatrix}\boldsymbol{\varphi}+\begin{pmatrix}\Psi_A^{1/2} & \circ\\ \circ & \Psi_B^{1/2}\end{pmatrix}\begin{pmatrix}\underline{\boldsymbol{\varepsilon}}_A\\ \underline{\boldsymbol{\varepsilon}}_B\end{pmatrix};$$

Set $C_{[A:B]}$ contain the replacement variates to $[\underline{\mathbf{X}}_A:\underline{\mathbf{X}}_B]$, this set having minimum correlation equal to $\rho^*_{(A:B)}=2R^2_{\boldsymbol{\varphi}.(A:B)}-1=2[\underline{\Lambda}_A:\underline{\Lambda}_B]'\Sigma_{(A:B)}^{-1}\begin{pmatrix}\underline{\Lambda}_A\\ \underline{\Lambda}_B\end{pmatrix}-1$;

Set [A:C] be A-replaceable and $[\underline{\mathbf{X}}_A:\underline{\mathbf{X}}_C]$ have ulcf representation

(15.22) $$\begin{pmatrix}\underline{\mathbf{X}}_A\\ \underline{\mathbf{X}}_C\end{pmatrix}=\begin{pmatrix}\underline{\Lambda}_A\\ \underline{\Lambda}_C\end{pmatrix}\boldsymbol{\omega}+\begin{pmatrix}\Psi_A^{1/2} & \circ\\ \circ & \Psi_C^{1/2}\end{pmatrix}\begin{pmatrix}\underline{\boldsymbol{\varepsilon}}_A\\ \underline{\boldsymbol{\varepsilon}}_C\end{pmatrix};$$

Set $C_{[A:C]}$ contain the replacement variates to $[\underline{\mathbf{X}}_A:\underline{\mathbf{X}}_C]$, this set having minimum correlation equal to $\rho^*_{(A:C)}=2R^2_{\boldsymbol{\omega}.(A:C)}-1=2[\underline{\Lambda}_A:\underline{\Lambda}_C]'\Sigma_{(A:C)}^{-1}\begin{pmatrix}\underline{\Lambda}_A\\ \underline{\Lambda}_C\end{pmatrix}-1$;

Set [A:B:C] be A-replaceable, and $[\underline{\mathbf{X}}_A:\underline{\mathbf{X}}_B:\underline{\mathbf{X}}_C]$ have ulcf representation

(15.23) $$\begin{pmatrix}\underline{\mathbf{X}}_A\\ \underline{\mathbf{X}}_B\\ \underline{\mathbf{X}}_C\end{pmatrix}=\begin{pmatrix}\underline{\Lambda}_A\\ \underline{\Lambda}_B\\ \underline{\Lambda}_C\end{pmatrix}\boldsymbol{\nu}+\begin{pmatrix}\Psi_A^{1/2} & \circ & \circ\\ \circ & \Psi_B^{1/2} & \circ\\ \circ & \circ & \Psi_C^{1/2}\end{pmatrix}\begin{pmatrix}\underline{\boldsymbol{\varepsilon}}_A\\ \underline{\boldsymbol{\varepsilon}}_B\\ \underline{\boldsymbol{\varepsilon}}_C\end{pmatrix};$$

Set $C_{[A:B:C]}$ contain the replacement variates to $[\underline{\mathbf{X}}_A:\underline{\mathbf{X}}_B:\underline{\mathbf{X}}_C]$, this set having minimum correlation equal to $\rho^*_{(A:B:C)}=2R^2\boldsymbol{\nu}_{.(A:C)}-1=2[\underline{\Lambda}_A:\underline{\Lambda}_B:\underline{\Lambda}_C]'\Sigma_{(A:B:C)}^{-1}\begin{pmatrix}\underline{\Lambda}_A\\ \underline{\Lambda}_B\\ \underline{\Lambda}_C\end{pmatrix}-1$.

Consider the first set equality, $S_1(C_{[A]})\equiv C_{[A:B]}$. $C_{[A:B]}$ contains those variates $\boldsymbol{\varphi}_i$ that replace $[\underline{\mathbf{X}}_A:\underline{\mathbf{X}}_B]$, and, because [A:B] is A-replaceable, have the property that $E(\underline{\mathbf{X}}_A\boldsymbol{\varphi}_i)=\underline{\Lambda}_A$. That is, they are also contained in $C_{[A]}$. Hence, $C_{[A:B]}$ is precisely that set of variates contained in $C_{[A]}$ that replaces, in addition to the variates of set A, the variates of set B. Thus, $S_1(C_{[A]})\equiv C_{[A:B]}$. The two other equalities are proven in analogous fashion□

iii. Because $E(\boldsymbol{\omega}\underline{\mathbf{X}}_A]=\underline{\Lambda}_A$, the variates $\boldsymbol{\omega}_i$ that are contained in $C_{[A:C]}$ are also contained in $C_{[A]}$. Analogously, since $E(\boldsymbol{\varphi}\underline{\mathbf{X}}_A]=\underline{\Lambda}_A$, the variates, $\boldsymbol{\varphi}_i$, contained in $C_{[A:B]}$ are also contained in $C_{[A]}$. That there exist variates in $C_{[A]}$ that are not contained in $C_{[A:B]}$, and also variates contained in $C_{[A]}$ that are not contained in $C_{[A:C]}$ follows from an argument involving minimum correlations analogous to that of the proof of Theorem (relationship between sets $C_{[D]}$ and $C_{[A]}$ under A-replaceability)□

iv. Because, from (7.23), $E(\boldsymbol{\nu}\underline{\mathbf{X}}_A)=\underline{\Lambda}_A$, $E(\boldsymbol{\nu}\underline{\mathbf{X}}_B)=\underline{\Lambda}_B$, and $E(\boldsymbol{\nu}\underline{\mathbf{X}}_C)=\underline{\Lambda}_C$, the variates, $\boldsymbol{\nu}_i$, contained in $C_{[A:B:C]}$ are also contained in each of $C_{[A:B]}$ and $C_{[A:C]}$. That there exist variates in each of $C_{[A:B]}$ and $C_{[A:C]}$ that are not contained in $C_{[A:B:C]}$ follows from an argument involving minimum correlations analogous to that of the proof of Theorem (relationship between sets $C_{[D]}$ and $C_{[A]}$ under A-replaceability)□

v. $\rho^*_{(A:B:C)}=\dfrac{(t_{(A:B:C)}-1)}{(t_{(A:B:C)}+1)}$. As $(m+n)\rightarrow\infty$, $t_{(A:B:C)}\rightarrow\infty$, and $\rho^*_{(A:B:C)}\rightarrow 1$□

vi. Because $\rho(\boldsymbol{\varphi}_i,\boldsymbol{\omega}_j)$ is a correlation defined on population *P*, the upper bound is obvious. For the lower bound, note that, because [A:B] and [A:C} are, at least, individually A-replaceable, the variates $\boldsymbol{\varphi}_i$ contained in $C_{[A:B]}$ and the variates $\boldsymbol{\omega}_j$ contained in $C_{[A:B]}$ are also contained in $C_{[A]}$. The elements of set $C_{[A]}$ cannot be correlated less than $\rho^*_{(A)}=2\underline{\Delta}_A'\Sigma_{(A)}^{-1}\underline{\Delta}_A-1$□

vii. If $m\rightarrow\infty$ and $n\rightarrow\infty$, then $Card(S_1(C_{[A]}))\rightarrow 1$ and $Card(S_2(C_{[A]}))\rightarrow 1$. Hence, $Card(C_{[A:B]})\rightarrow 1$ and $Card(C_{[A:C]})\rightarrow 1$. If [A:B:C] is, in addition, A-replaceable, then it has ulcf representation (7.22) with the replacement variates $\mathbf{v}$ contained in set $C_{[A:B:C]}$. Now, from (iv), $C_{[A:B:C]}\subset C_{[A:B]},C_{[A:C]}$. That is, the variates $\mathbf{v}_i$ are contained in both $C_{[A:B]}$ and $C_{[A:C]}$. But since the cardinalities of $C_{[A:B]}$ and $C_{[A:C]}$ are both unity, it must also be that $Card(C_{[A:B:C]})=1$, and this single replacement variate is contained in both $C_{[A:B]}$ and $C_{[A:C]}$. Obviously then, $\rho(\boldsymbol{\varphi}_i,\boldsymbol{\omega}_j)=1$□

Theorem (convergence to uniqueness of the A-replacements to [A:B:C]): Let there be two domains of variates, D_1 consisting of the p variates contained in $\underline{\mathbf{X}}_A$, and a remaining m variates contained in $\underline{\mathbf{X}}_B$, and D_2 consisting of the variates contained in $\underline{\mathbf{X}}_A$, and a remaining n variates contained in $\underline{\mathbf{X}}_C$, and let there be no overlap between $\underline{\mathbf{X}}_A$, $\underline{\mathbf{X}}_B$, and $\underline{\mathbf{X}}_C$. Finally, let domain D_3 consist of the union of the variates contained in $\underline{\mathbf{X}}_A$, $\underline{\mathbf{X}}_B$, and $\underline{\mathbf{X}}_C$. If D_3 is A-replaceable, then, as $(m+n)\rightarrow\infty$, $Card(S_{D3}(C_{[A]})\rightarrow 1$, in which $S_{D3}(C_{[A]})$ is the subset of $C_{[A]}$ that A-replaces D_3.

Proof

If domain D_3 is A-replaceable, then, as $(m+n)\rightarrow\infty$, $t_{(A:B:C)}$ becomes increasingly large, and $\rho^*_{(A:B:C)}=\dfrac{(t_{(A:B:C)}-1)}{(t_{(A:B:C)}+1)}\rightarrow 1$. Hence, in the limit, $C_{[D3]}$ contains but one variate. That is, $Card(S_{D3}(C_{[A]})\rightarrow 1$□

In groundbreaking work, Guttman (e.g., 1955) had earlier considered the problem of limiting determinacy. His treatment was based on a consideration of a domain of variates, $\{\mathbf{X}_1,\mathbf{X}_2,...\mathbf{X}_k\}$, and the behaviour of the sequence, $\Sigma_k=\underline{\Delta}_k\underline{\Delta}_k'+\Psi_k$, formed by drawing successively more variates from this domain. That is, he was interested in the conditions under which the sequence remains ulcf-representable as the number of variates to be represented becomes large. Guttman defined limiting determinacy (here given for the case of a unidimensional replacement) as when $\lim_{n\rightarrow\infty}\underline{\Delta}_n'\Sigma^{-1}{}_n\underline{\Delta}_n=1$, and identified conditions under which this determinacy obtains. One necessary condition is given in the next theorem.

Theorem (convergence to diagonal matrix of Σ_k^{-1}; Guttman, 1955): If a sequence of domains D_k is determinate in the limit, then, as $k\rightarrow\infty$, Σ_k^{-1} converges to a diagonal matrix.

Proof

If the ulcf replacement of D_k is determinate in the limit, then $\lim_{k \to \infty} \underline{\Lambda}_k{}'\Sigma^{-1}{}_k \underline{\Lambda}_k = 1$. Because, for all k, $\Sigma_k = \underline{\Lambda}_k \underline{\Lambda}_k{}' + \Psi_k$, in which Ψ_k is diagonal and positive definite, it follows that

$\Sigma_k^{-1} = \Psi_k^{-1} - \Psi_k^{-1} \underline{\Lambda}_k \underline{\Lambda}_k{}' \Psi_k^{-1} a_k$, in which $a_k = (1 + \underline{\Lambda}_k{}' \Psi_k^{-1} \underline{\Lambda}_k)^{-1}$. As $\lim_{k \to \infty} \underline{\Lambda}_k{}'\Sigma^{-1}{}_k \underline{\Lambda}_k = 1$,

$\lim_{k \to \infty} a_k = \lim_{k \to \infty} (1 + \underline{\Lambda}_k{}'\Psi_k^{-1} \underline{\Lambda}_k) = \infty$, and, Σ_k^{-1} converges to Ψ_k^{-1}, a diagonal matrix□

If one could design a study in which a core set of p variates is augmented by additional variates, then Guttman's criterion could be used to gain insight into whether limiting uniqueness of the ulcf replacement is, at all, possible. What would be required is a demonstration that, as more variates were added, the sequence of sets of variates remained ulcf representable, and S_k^{-1} moved towards diagonality. If this can be demonstrated, then it is possible that an essentially unique ulcf replacement can be obtained with relatively few variates, especially if the magnitudes of the regression parameters contained in $\underline{\Lambda}_k$ are large relative to the magnitudes of the parameters contained in Ψ_k^{-1}.

As shown in McDonald and Mulaik (1979), what has been called, herein, A-replaceability is not the only means by which the limiting uniqueness of a replacement can be achieved. A-replaceability requires that a subset of the ulcf-replacements of a core set of variates ulcf-replace, not only these original variates, but a set of additional variates drawn from the same variate domain. One might consider relaxing this requirement somewhat by requiring that a subset not only ulcf-replace the core set of variates, but *partially* replace a set of additional variates. The sense of *partial replaceability* one insists upon is open to negotiation, with McDonald and Mulaik (1979) allowing the residuals of the linear regressions of certain of the input variates (but not those of the original set) on the replacement variates to correlate.

Now, as the mathematics suggest, insisting that some of the replacement variates of a set of input variates replace an additional m variates, imposes additional requirements on the replacement, and effectively reduces the range of variates that satisfy the full set of imposed requirements. This might then appear to be a promising means to achieve a replacement with a cardinality of unity. However, there seems to have arisen within psychometrics the general attitude that the mere *possibility* of unique replacements within variate domains settles concern in regard non-unique replacements (indeterminacy). This view will have to be brought to heal, for variate domain technology bears fruit in regard the production of unique replacements only if researchers actually follow the paths of action suggested by it, and only when these paths are *available*.

If the user of latent variate technology were in the habit of employing the insights available from variate domain treatments to reduce the cardinality of replacements that suffer from indeterminacy, his use of these generators would have a very particular look.

He would, for example, test the conditions that must be fulfilled in order for the limiting uniqueness to obtain, report upon aborted attempts to achieve a unique replacement when the conditions were not met, calculate measures such as Guttman's ρ^*, carefully describe the rules for inclusion of variates in domains of interest, and, importantly, describe how this domain was sampled. This is all made clear in Mulaik and McDonald (1978) in which it is stated that: "...we believe that in practice researchers should initially attempt a strict definition of the domain of empirical measures to be represented in a study and should set up a well-defined sampling procedure for selecting variables from the domain." (1978, p.191); "We now see that if we possess a prior definition of a domain of variables as Guttman has said is necessary, then the assumption that the domain possesses a determinate factor space has the status of a strong falsifiable empirical hypothesis. The hypothesis may be falsified in principle by observing that a reasonably large number of variables from the domain do not yield a correlation matrix whose inverse tends to a diagonal matrix [or]...if two overlapping sets of variables [$\boldsymbol{\eta},\mathbf{v}_1$] and [$\boldsymbol{\eta},\mathbf{v}_2$] chosen from the domain in such a way that each conforms to the *g*-factor law [is A-replaceable], jointly do not form a set [$\boldsymbol{\eta},\mathbf{v}_1,\mathbf{v}_2$] conforming to the *g*-factor law. Such a falsification of the hypothesis could occur with a relatively small number of variables. The hypothesis may also be falsified with as few as three variables from a domain if they form a Heywood case" (1978, p.188); "The hypothesis that an infinite domain has a determinate factor space may be falsified but not confirmed with less than all variables from the domain" (1978, p.190).

In fact, there is no sign that the current practice of employing latent variate generators is founded on variate domain technology. Those who have seen in the variate domain treatment a solution to the problems of indeterminacy seem to have taken the mere *possibility* of such a treatment, and the mere possibility that the required conditions hold, as being sufficient to rid practice equivalent to a general solution to the problems that they perceive as arising from non-unique replacements (indeterminacy). If it were only this simple. Steiger (1996, p.546) puts it well: "...Clearly, one can always *assume* there are more variables out there that fit the factors that fit the variables one has now. This is rather like an experimenter "solving" the problem of an overly wide confidence interval for the correlation between two variables by assuming that he/she has sampled 45 observations from a larger domain with the same correlation as the present observations! Indeed, many very difficult problems in statistics could be "solved" with the aid of such an approach."

Lacking any support for the claim that the employments of latent variate generators are founded on variate domain technology, McDonald has turned to claiming that this technology is, nevertheless, "implicit" in applied work. Thus he states that "I believe a well conceptualized domain is at least implicit in all well conducted applications of FA/IRT to the estimation (with S.E.s) of abilities, attitudes, beliefs etcetera." (1996a, p.599); "Guttman (1953) conjectured as a rule of thumb that in the common factor model about ten to fifteen variables approximate an infinity of them for empirical applications of behaviour domain concepts..." (1996a, p.598); "Behaviour domain theory is not merely one way to create a correspondence between the factor equations and test data, but I claim that it is indeed the implicit theory governing the practice of factor analysis" (1996b, p.669). But much more is needed than mere allegations of "implicit belief" in the variate domain conception. As McDonald (1996b) himself points out, given that all of the conditions required for uniqueness of replacement in a variate domain are, in fact, met, the issue then becomes the practical one of sampling enough input variates so as to make the replacement, at the least, virtually unique. This practical problem is addressed by the taking of practical, non-controversial steps, by existing practitioners of latent variable modeling (e.g., the discussion of the problem in their published work). The apparent inability of McDonald and others to acknowledge the difference between the actual *taking* of these steps, and their mere possibility, is what justifies Steiger's comment.[2]

The wishful thinking that inevitably attends the discussion of indeterminacy and variate domain technology is often accompanied by bouts of amnesia when it comes time to mention the requirements that must be satisfied in order to bring about uniqueness of replacement, or to present a sober assessment as to the ease with which these requirements can actually be satisfied. There is a strong historical precedent for this amnesia. As will be recalled, Piaggio (1933) recommended that to remedy the indeterminacy problem, one merely needs to "take the number of variates to be infinity." Irwin (1935), soon after, offered the same "solution". McDonald and Mulaik state that "For finite p, the squared multiple correlation of the manifest variates with the common factor is strictly less than unity. If one can conceivably find infinitely many variables with nonzero loadings on *X*, then in the limit as [the number of variates] approaches infinity...the squared multiple correlation approaches unity" (1979, p.299). As their later analysis shows, determinacy in the limit is not quite so assured as these words would suggest. In reminiscinces about his own career as a statistical consultant, McDonald (1996a, p.599) states "I have advised on dozens of such analyses, and my first task as a psychologist is to understand my client's implicit or explicit conceptual domain. If they do not have an abstract concept capable of further extension with congeneric indicators, I recommend the use of composites." This is a commendable sentiment, but why does McDonald mention the *capability* to create further congeneric indicators as the key to the wholly grail of latent variate generator use, when, on his own account, what must happen is that *an existing domain of variates must actually satisfy all of the multitude of requirements necessary to bring about limiting determinacy*? In particular, the researcher must have already created enough variates to allow for a test in regard the hypothesis of limiting uniqueness of replacement. And, finally, if the whole endeavour, according to McDonald, rests on the estimation of the unique, limiting replacement variate $\boldsymbol{\theta}_D$ (which, as will be seen, turns out to be none other than the first principal component variate), then what does the latent variable modeller do when, inevitably, the requirements for the limiting uniqueness are not met? One cannot know, because published latent variable analyses make no mention of limiting determinacy, and offer up no evidence to support claims of limiting determinacy.

Discussions of variate domain treatments have very often carried with them the strong implication that these treatments allow for work to progress in latent variable modeling undeterred by the supposed implications of the indeterminacy property. This

2 McDonald is, of course, entitled to suggest that employers of latent variate replacement technology *should* carry out their work within a variate domain framework.

implication is nicely illustrated by McDonald (1996b, p.663) who deduces from the possibility of a variate domain founding for latent variable modeling that "...behavioral scientists- at ETS, ACT, and elsewhere, can go on doing what they wish to do." But this is mere rhetorical ploy, for the requirements that must be satisfied in order to bring about a limiting determinacy are sufficiently difficult to realize that, if they *were* taken seriously in applied work, they would likely grind the latent variable industry to a halt. To quote Steiger, "Testing the infinite domain *model* would require gathering more variables, to somehow test the notion that one is in fact sampling from an infinite behaviour domain. Of course factor analysts often have absolutely no intention of doing that, with good reason. First, many have already exhausted their efforts gathering the variables for the first factor analysis. Second, all the evidence seems to suggest that, generally, the number of factors required to obtain an adequate fit to the data tends to go up as a relatively constant function of the number of variables" (1996, p.547).

Instead of an honest accounting of the implications of variate domain technology, what has been assembled is a list of "ifs":

"...if one worries about the relationship of factor-score estimates to the variables being estimated, it is because one is able, under certain conditions, to define the variables being estimated on the basis of a behaviour domain that fits the common factor model consistently with the given variables." (McDonald & Mulaik, 1979, p.305)

"If a behaviour domain can be factored so that one of its common factors has the same loadings on a core set of variables as when the latter are factored separately, it is then meanigful to consider the relationship between the possible factors of the core and the possible factors of the behaviour domain. In such a case, the loadings of the core set uniquely mark a factor variable in the domain, and the addition of further variables with nonzero loadings on the factor will determine it as precisely as one pleases, ultimately yielding an infinite sequence of variables that determine the factor exactly" (1979, p.304)

But such promissory notes are largely irredeamable. Consider the phrase "one is able, under certain conditions, to define the variables being estimated on the basis of a behaviour domain that fits the common factor model consistently with the given variables". How exactly can a variate be "defined" (presumably this is the mathematical sense of definition, for this variate is, after all, a limit of an infinite sequence of variates) on the basis of a behaviour domain that fits the common factor model, when the consituents of the domain are not individually defined, and there is no way to know whether such imaginary variates do, in fact, "fit the common factor model"?

2a. McDonald's abstractive property position

Some have talked of the "management of indeterminacy", meaning by this the computation of indeterminacy indices and the use of large numbers of input variates in order to avoid the difficulties that arise from the indeterminacy property of the ulcf replacement. In considering the possibility of managing indeterminacy, Steiger (1996, p.620) states that "...I too do not think we necessarily have to *discard* factor analysis or

IRT theory, or ETS as a consequence of factor indeterminacy". If the issue of behaviour domain theory were simply about the conditions under which a unique ulcf replacement could be achieved, then, indeed, attention could be turned to "managing indeterminacy" through the taking of various practical steps. But, of course, variate domain treatments were seized upon by psychometricians during the indeterminacy debate and the indeterminacy debate did not occur as a result of concern over the cardinality of ulcf replacements, but, rather, of perceived threats to the Central Account. Variate domain positions were championed in an attempt to protect the Central Account from threats posed by the indeterminacy property. The indeterminacy property suggested that the Central Account was a mythology (which, indeed, it was). It is not surprising, then, to find that McDonald takes the adoption of a behaviour domain foundation for latent variable modeling as placing on firm logical footing his abstractive property position. For, as will be shown, McDonald's abstractive property position is nothing but a fancily clad version of the CAM, and, hence, suffers from the same defects documented in Part II. In particular, it presumes the mistaken belief that latent variable models are detectors/discoverers of properties/attributes, and rests on the badly confused conceptions of concept meaning and signification that were documented in Chapters IX and XII.

For the unidimensional case, McDonald's abstractive property position can be paraphrased as follows:

i. There is a set of k variates, D:$\{\mathbf{X}_1,\mathbf{X}_2,...,\mathbf{X}_p,\mathbf{X}_{(p+1)},\mathbf{X}_{(p+2),...,}\mathbf{X}_{(k)}\}$, comprising a "universe of content" or "behaviour domain". This domain is divided into two subsets, A, of p variates, and B, of (k-p) variates.
ii. Let subset A be ulcf representable, with vector of regression weights $\underline{\Lambda}_A$.
iii. Consider the sequence of sets $\{S_1,S_2,...,S_{(k-p)}\}$ formed by augmenting subset A with, first, one variate sampled from B, then two, then three, etc.
iv. Imagine that each of the S_i are ulcf representable. That is, $\underline{\mathbf{X}}_i=\underline{\Lambda}_i\boldsymbol{\theta}_i+\Psi_i^{1/2}\underline{\boldsymbol{\delta}}_i$, i=1..(k-p), so that $\Sigma_i=\underline{\Lambda}_i\underline{\Lambda}_i'+\Psi_i$, i=1..(k-p), in which $V(\boldsymbol{\theta}_i)=1$, $C(\underline{\boldsymbol{\delta}}_i)=I$, $C(\boldsymbol{\theta}_i,\underline{\boldsymbol{\delta}}_i)=\underline{0}$, and Ψ_i is diagonal and positive definite.
v. Let it be the case that $\underline{\Lambda}_i=\begin{pmatrix}\Lambda_A\\ \underline{\gamma}\end{pmatrix}$, i=1..(k-p), in which $\underline{\gamma}$ is an i×1 vector.
vi. The concept *latent variate to D* signifies the variate $\boldsymbol{\theta}_D$ associated with the representation to which $\underline{\mathbf{Y}}_i=\underline{\Lambda}_i\boldsymbol{\theta}_i+\Psi_i^{1/2}\underline{\boldsymbol{\delta}}_i$ converges as $(k-p)\rightarrow\infty$: "...the loadings of the core set uniquely mark a factor variable in the domain, and the addition of further variables with nonzero loadings on the factor will determine it as precisely as one pleases, ultimately yielding an infinite sequence of variables that determines the factor exactly" (McDonald & Mulaik, 1979, p.304)

vii. The latent variate to D is the common property, say κ, of the items contained within D.

"It is the generic character of the common property that implies the behavior domain idealization" (McDonald, 1996b, p.670)

"...I am describing the rule of correspondence which I both recommend as the normative rule of correspondence, and conjecture to be the rule as a matter of fact followed in most applications, namely: In an application, the common factor of a set of tests/items corresponds to their common property" (McDonald, 1996b, p.670).

"The problem remains, however, of a range of ambiguity with respect to the ways in which the results may be extended by the use of further variables to measure the discovered attribute more precisely" (McDonald & Mulaik, 1979, p.305).

"But in the absence of an agreed domain of variables based on prespecified common attributes, two investigators may seize upon distinct sets of common attributes and proceed to build distinct extended batteries of variables on the basis of the same core set, in terms of these distinct attributes. Ultimately, in principle, they can create two distinct test batteries of infinite length, built on the same original test and corresponding to two distinct attributes and to two distinct random variables, either of which might have been the common factor of the original variables. It is factor indeterminacy that supplies a mathematical latitude for such sets of alternative common attributes to be found." (McDonald & Mulaik, 1979, p.305)

"Their hope is that when these variables are subjected to a factor analysis, the psychological attributes that determine the correlations among the variables will reveal themselves" (McDonald & Mulaik, 1979, p.305)

viii. The items that comprise domain D are "indicators" of κ.

ix. Because κ is "defined" on an infinity of variates, it is not observable. The latent variable modeller cannot know that the common property of the items is, in fact, κ. Instead, the identity of the common property of the items must be inferred on the basis of $\underline{\Lambda}_A$. This is the task engaged in by the researcher in a factor/latent variate interpretation.

"..the widely accepted aim of factor analysis, namely, the interpretation of a common factor in terms of the common attribute of the tests that have high loadings on it..."; "what attribute of the individuals the factor variable represents" (McDonald & Mulaik, 1979, p.298).

x. The scores that comprise the distribution of the random variate $\boldsymbol{\theta}_D$ are measurements of the individuals in population *P* with respect property κ. These scores are signified by concept "κ". Unfortunately, they are not observable (κ is not directly measurable) and, hence, must be estimated on the basis of a sample of items drawn from D.

"If I estimate the attitude of one freshly drawn examinee by Bartlett's (1937) ML formula and attach an S.E., the S.E. will be large using 19 items, smaller with 40, and very small but nonzero with 95. What is being estimated by any of these is the common property of the items in a test lengthened further, by drawing items from a clearly conceptualized domain" (McDonald, 1996a, p.599)

Clearly, this is simply a jazzed-up version of the CAM. The difference between the standard CAM and McDonald's abstractive property version is that, in the latter: i) The criterion for existence of a common property of the phenomena represented by $\underline{\mathbf{X}}$ is that a domain of variates from which the variates $\underline{\mathbf{X}}$ are drawn is either A-replaceable or partially A-replaceable; ii) the common property is taken to be a variate that either A-replaces, or partially A-replaces, the domain of variates. Hence, in all essential details the abstractive property position is just the CAM, and, thus, suffers from precisely the same defects.

2b. The ulcf and pc1 generators: A comparison

The comparison of linear factor analysis and principal component analysis is old sport. Latent variable modelers have traditionally believed factor analysis to be superior to component analysis. It has been argued, herein, that these beliefs have largely been the product of commitment to a mythology, the Central Account, and not reasoned argument. Both linear factor analysis and principal component analysis involve the application of replacement variate generators. As is clear from the preceding sections, the replacement variates produced under these generators answer to different senses of optimality and entail different costs. Two generators are properly compared by considering their optimality/cost profiles in regard the research tasks to which they will be put. There is simply no point in talking about which is "better" in some absolute sense. It will, therefore, be worthwhile to examine the principal component analysis versus linear factor analysis question afresh and without distraction from the mystical elements of the CA.

The "special case" refrain. When component and factor generators are discussed, it has been a common refrain that "the family of factor models contains the family of component models." The basis for this refrain is that one may derive various component generators by placing restrictions on various factor generators. For example, the covariance structure of the pc1 replacement is produced by setting to null the matrix Ψ in the ulcf covariance structure, $\Sigma=\underline{\Lambda\Lambda}'+\Psi$. This fact is taken as evidence of the *precedence* of the ulcf generator. Gorsuch (1990, p.33), for example, states that "Common factor analysis is the general case of which component analysis is a special case." Indeed. Employing the same style of argument, one should conclude that CO is inferior to CO_2 on the grounds that it is "derivable" from CO_2 through the dropping of an oxygen molecule. The fact that the covariance structure of the pc1 replacement is derivable from that of the ulcf replacement by the dropping of a term is of nothing more than notational interest. What is important is that the pc1 replacement has a profoundly different optimality/cost profile than the ulcf replacement. The construction formulas according to which each type of replacement variate (ulcf and pc1 replacement variates) are constructed, answer to the optimality/cost profiles characteristic of the ulcf and pc1 generators. Hence, these formulas produce replacement variates that possess precisely the properties that they were designed to possess. Neither type of variate can be said to be, in general terms, superior or inferior to the other. Moreover, in the same way that it would be a mistake to believe that all of the properties of CO can be deduced from CO_2 because "CO can be written as CO_2 with one less O", so too is it foolish to believe that the fact that the pc1 covariance structure is a special case of the ulcf covariance structure, provides a definitive account of the relationship between the ulcf and pc1 generators, and the variates produced under each.

The "testability" and "model" refrains. Another common refrain is that "the linear factor model is testable, while the principal component model is not." Recall from chapter III, Bentler and Kano's (1990) assessement:

> This is one of its virtues, as compared to components analysis, because the ability to reject a model is a fundamental aspect of data analysis using structural models (Bentler, 1989; Joreskog & Sorbom, 1988). In fact, the component *model*, or class of models as described by Velicer and Jackson, is not a model at all. As noted by Dunteman (1989, p.56), "Principal components analysis is a procedure to decompose the correlation matrix without regard to an underlying model. There is no hypothesis tested with the model. It is a nonfalsifiable procedure for analyzing data that can always be applied with any data set. Of course, there is a sampling theory for eigenvalues of a sample covariance matrix that can be used to test hypotheses about population roots of a covariance matrix (e.g., Anderson, 1963), but such hypotheses are typically hard to frame and not very informative about the structure of a correlation matrix...

McDonald (e.g., 1975, p.143) also seems impressed by the fact that linear factor analysis yields a testable hypothesis: "That is, the fundamental theorem of factor analysis yields a testable hypothesis, while the fundamental theorem of image analysis is merely tautological."

This sort of commentary badly mishandles the related issues of "grounds for model-hood" and "testability". The quote from Bentler and Kano implies that a necessary condition for model-hood is testability, but fails to make explicit the sense of *model* in play. The result is the failure to grasp that neither the principal component, nor the linear factor, "model" are models in the classical sense of "representer of a state of affairs". As discussed in Chapter X, the formulation of a classical representational model requires the antecedent laying down of rules of correspondence, which, in turn, presupposes the ability to antecedently identify the relata of the model/modelled relationship. In carrying out a linear factor, or principal component, analysis, the researcher does not lay down rules of correspondence to link the replacement variate terms that appear in the equations of these generators, and particular features of natural reality for which they are to stand. The sense of the concept *testability* that Bentler and Kano (1990) wish to invoke is, similarly, left hanging. What is certain is that statistical testability itself is not sufficient to establish a basis of distinction between the two types of generator. Both principal component and factor analysis are frameworks within which endless statistical hypotheses can be generated and tested. For example, within the framework of principal components analysis, one could easily enough test the hypotheses H_o:$(\kappa_1-\kappa_4)=.013$ and H_o:$\kappa_1/2\kappa_3=1$, in which κ_i is the ith eigenvalue of the covariance matrix of some set of variates. At the risk of stating the obvious, what makes an hypothesis test of interest to science is that it is a test of an important empirical claim.

The reason that latent variable modellers harp on the "testability" of the factor analytic "model" is simply that they wrongly believe that the standard test of H_o: $\Sigma=\underline{\Lambda\Lambda}'+\Psi$ is a test of the picture described by the Central Account. But the Central Account is mythology masquerading as empirical claim. The point is that Bentler and Kano's belief in the superiority of linear factor analysis does not arise from mathematical analysis, but rather commitment to the Central Account, as is clear from their characterizing component based hypotheses as "..typically hard to frame". What has, all along, made linear factor analytic hypotheses "easy to frame" is the fact that they are accompanied by a familiar story, the Central Account. In reality, a latent variable model is a replacement variate generator, and, hence, allows for the testing of various hypotheses of replaceability. It is CA fueled nonsense to suggest that component generators, and, notably, the pc1 generator, do not have precisely the same status.

The status of the pc1 and ulcf generators with respect the issue of model-hood and testability can be summarized as follows:

i. Neither linear factor, nor principal component, analysis involve the employment, or testing, of models, at least if the term *model* is employed in a classical sense to denote a representation of features of natural reality.[3] Both are, rather, replacement variate generators, and the tests one can carry out in employing each are tests of replaceability.

ii. In regard the issue of testability, consider the following chart.

	ulcf generator	pc1 generator
Hypothesis	H_o: $\Sigma=\underline{\Lambda\Lambda}'+\Psi$, Ψ diagonal, positive definite	H_o: $\Sigma=\lambda_1\underline{v}_1\underline{v}_1'$
Paraphrase	(A)	(B)

(A) Does there exist a variate, $\boldsymbol{\theta}$, that replaces the input variates, $\mathbf{X}_j$, j=1..p, in the sense that the p parameters from the linear regressions of the $\mathbf{X}_j$ on $\boldsymbol{\theta}$ reproduce the $\frac{1}{2}p(p-1)$ unique covariances contained in Σ.

(B) Does there exist a variate, $\boldsymbol{\theta}$, that replaces the input variates, $\mathbf{X}_j$, j=1..p, in the sense that the p parameters from the linear regressions of the $\mathbf{X}_j$ on $\boldsymbol{\theta}$ reproduce the $\frac{1}{2}p(p+1)$ unique variances and covariances contained in Σ

3 They are both models when *model* is used in the currently popular, but trivial, sense to denote "a statement in regard the joint distribution of a set of random variates."

It is hoped that Bentler would not disagree with the claim that principal component replaceability hypothesis (B) is no less trivial than ulcf replaceability hypothesis (A). Bentler might, of course, scoff at the hypothesis H_o: $\Sigma=\lambda_1\underline{v}_1\underline{v}_1'$ on the oft-cited grounds that "it never holds in any naturally occurring population *P*." But, if he were inclined to do so, he should recall that the same can be said of the ulcf hypothesis of replaceability, H_o: $\Sigma=\underline{\Lambda\Lambda}'+\Psi$ versus H_1:~H_o. This is why, *in practice*, the factor analyst departs from a program of strict hypothesis testing, and turns to looking for a ulcf replacement variate, $\boldsymbol{\theta}$, for which the p residuals of the linear regressions of the input variates, $\mathbf{X}_j$, j=1..p, on $\boldsymbol{\theta}$, have a covariance matrix that is *as close as possible* to being diagonal. Thus, one must add a third level to this chart:

Loss minimization	$\min(\sum_{i=1}^{p-1}\sum_{j=i+1}^{p}(\sigma_{ij}-\lambda_i\lambda_j)^2)$	$\min(\sum_{i=1}^{p-1}\sum_{j=i}^{p}(\sigma_{ij}-\lambda_1 v_{i1}v_{j1})^2)$
Paraphrase	(C)	(D)

(C) Find a variate, $\boldsymbol{\theta}$, that replaces the input variates, $\mathbf{X}_j$, j=1..p, in the sense that the p parameters from the linear regressions of the $\mathbf{X}_j$ on $\boldsymbol{\theta}$ reproduce the $\frac{1}{2}p(p-1)$ covariances contained in Σ with as little mean-square loss as possible.

(D) Find a variate, $\boldsymbol{\theta}$, that replaces the input variates, $\mathbf{X}_j$, j=1..p, in the sense that the p parameters from the linear regressions of the $\mathbf{X}_j$ on $\boldsymbol{\theta}$ reproduce the $\frac{1}{2}p(p+1)$ unique variances and covariances contained in Σ with as little mean-square loss as possible.

It is the inexplicable comparison of (A) (the fact that a linear factor analysis standardly begins with the testing of an hypothesis that is always incorrect) with (D) (the fact that principal component analysis usually foregoes an initial hypothesis test and proceeds immediately to the approximation of $\underline{\mathbf{X}}$) that has, historically, been seized upon, along with the Central Account's misportrayal of latent variate generators as models in the classical sense of the term, to create the illusion that, in contrast to component generators, the employment of linear factor generators involve a special brand of testability. This also suggests the dubious nature of the claim that "principal component analysis is variance oriented while linear factor analysis is covariance oriented". In fact, principal component analysis is variance and covariance oriented while linear factor analysis is covariance oriented. The testable hypotheses (A) and (B) each claim that a certain subset of the elements of Σ can be reproduced on the basis of knowledge of the parameters of the regression functions of the input variates on a (p+1)th constructed replacement variate. In analogy to linear factor analysis, one could easily begin a principal component analysis by formally testing (B), before moving on to (D). To test an hypothesis within the context of either ulcf or pc1 is to test an hypothesis of replaceability.

iii. *Generalizability and variate domains*. McDonald and Mulaik (e.g., 1979, p.305) have harped on the theme that one of the most important characteristics of latent variable models is that, under certain conditions, they can be generalized to other variates not included in an original analysis. Mulaik has suggested that indeterminacy is a virtue of latent variable models because it is this property which allows such "generalizability": "...trying to confine the basis for making a generalization from experience to specific, determinate phenomena already observed and defined, as some do by prescribing the use of *operational definitions* or by urging the use of specific linear combinations of a set of observed variables (component factors) to stand for what is common to them, may actually get in the way of formulating creative generalizations and syntheses that go beyond what is already known or observed but which nevertheless are eventually tied to experience by the efforts to test such generalizations empirically with extensions to new data" (Mulaik, 1996, p.54). Mulaik is, here, guilty of conflating conceptual and empirical issues. Ambiguity in regard the correct employment of the concept that is to signify the phenomena of interest *is* ambiguity in regard the phenomena of interest (i.e., *what* constituent of natural reality is to be studied). No creative generalizations can arise from such conceptual confusion.

Mulaik is mistaken in believing that providing definitions for concept that signify phenomena constrains the generalizations that can be made in regard what is known about the phenomena itself. Providing a definition of denotative concept "ψ" settles *what* "ψ" denotes, and, hence, *what* is to be studied (ψ-things). The ability to make various empirical generalizations about ψ-things is an entirely different issue that *presupposes* clarity in the employment of concept "ψ". There is, however, a grain of truth to what Mulaik claims, for indeterminacy is the property that Card(C)>1 and is a consequence of the latitude inherent to the construction formula $\boldsymbol{\theta}_i=\underline{\Lambda}'\Sigma^{-1}\underline{\mathbf{X}}+w^{1/2}\mathbf{s}_i$. The variate $\mathbf{s}_i$ must only satisfy the moment restrictions $E(\mathbf{s}_i)=0$, $V(\mathbf{s}_i)=1$, and $E(\underline{\mathbf{X}}\mathbf{s}_i)=0$. As was seen in the previous section on variate domains, *if certain requirements are satisfied* this very latitude in the choice of $\mathbf{s}_i$ can be used to choose an $\mathbf{s}_i$ in such a way that the resulting replacement variates $\boldsymbol{\theta}_i=\underline{\Lambda}'\Sigma^{-1}\underline{\mathbf{X}}+w^{1/2}\mathbf{s}_i$ now possess the required properties to replace not only $\underline{\mathbf{X}}$, but also additional variates. The requirements that pc1 replacement variates must satisfy result in their being linear combinations of the set of variates they must replace. This effectively removes from a pc1 replacement the latitude possessed by a ulcf replacement. The question that must be asked is "are the costs of the latitude inherent to the ulcf replacement worth the possibility of this particular brand of generalizability?" The answer perhaps turns on whether one believes that the object of science is to work with ambiguous notions of domains of variates or to describe and explain phenomena.

iv. *Asymptotic equivalence of ulcf and pc1 replacements*. Supposing the variate domain formulation of the ulcf replacement, if the replacement of a domain of input variates is A-replaceable, and the representation has a cardinality of unity, then the limiting variate $\boldsymbol{\theta}_D$ that McDonald claims is a common property of the items turns out to be the pc1 replacement variate of the variates. Moreover, the vector of regression coefficients from the ulcf replacement is, in the limit, equivalent to the first eigenvector from the pc1 replacement. A number of different proofs of these asymptotic equivalences, based on various premises, have been given. The simple proof of Schneeweiss (1989) and Bentler and Kano (1990), in which asymptotic equivalence obtains under very mild assumptions is, herein, restated.

Theorem (asymptotic equivalence of pc1 and ulcf replacements; Schneeweiss, 1989; Bentler & Kano, 1990). Let there be a domain, D, of variates $\{\mathbf{X}_1,\mathbf{X}_2,...\mathbf{X}_n\}$, a sequence of sets of variates $\{S_3,S_4,...,S_k\}$, each set formed by drawing S_i variates from D, and let each set S_i, i=3..k, be ulcf representable, $\underline{\mathbf{Y}}_i=\underline{\Lambda}_i\boldsymbol{\theta}_i+\Psi_i^{1/2}\underline{\boldsymbol{\delta}}_i$, in which Ψ_i is diagonal and positive definite, $V(\boldsymbol{\theta}_i)=1$, $C(\underline{\boldsymbol{\delta}}_i)=I$, and $C(\boldsymbol{\theta}_i,\underline{\boldsymbol{\delta}}_i)=\underline{0}$, so that $\Sigma_i=\underline{\Lambda}_i\underline{\Lambda}_i'+\Psi_i$, i=3..k. That is, let the sets S_i remain ulcf representable as k becomes progressively larger. Consider also the sequence $\Sigma_i\underline{v}_{1i}=\kappa_{1i}\underline{v}_{1i}$, $\underline{v}_{1i}'\underline{v}_{1i}=1$, in which $\underline{v}_{1i}$ and κ_{1i} are the first eigenvector and eigenvalue of Σ_i. The variate $\mathbf{c}_{1i}=\underline{v}_{1i}'\underline{\mathbf{X}}_i$ is the first principal component of the distribution of S_i. If, as $k\rightarrow\infty$, $\underline{\Lambda}_i'\underline{\Lambda}_i\rightarrow\infty$ and there exists a real number $\zeta_o>0$ such that, for all j, $\zeta_o>\zeta_{ij}$, in which ζ_{ij} is the jth element of Ψ_i, then $\rho(\mathbf{c}_{1i},\boldsymbol{\theta}_i)\rightarrow 1$.

Proof

a) Because $\boldsymbol{\theta}_i=\underline{\Lambda}_i'\Sigma_i^{-1}\underline{\mathbf{X}}_i+w_i^{1/2}\mathbf{s}_{ij}$ are the variates contained in set C_i that contains the common factors of set S_i, $\rho(\mathbf{c}_{1i},\boldsymbol{\theta}_i)=\rho(\underline{v}_{1i}'\underline{\mathbf{X}}_i,\underline{\Lambda}_i'\Sigma_i^{-1}\underline{\mathbf{X}}_i+w_i^{1/2}\mathbf{s}_{ij})=\frac{(\underline{v}_{1i}'\underline{\Lambda}_i)}{\sqrt{\kappa_{1i}}}$. Now, $\Sigma_i\underline{v}_{1i}=(\underline{\Lambda}_i\underline{\Lambda}_i'+\Psi_i)\underline{v}_{1i}$, so that $\underline{v}_{1i}'\Sigma_i\underline{v}_{1i}=(\underline{v}_{1i}'\underline{\Lambda}_i)^2+\underline{v}_{1i}'\Psi_i\underline{v}_{1i}$. Hence, $\kappa_{1i}-\underline{v}_{1i}'\Psi_i\underline{v}_1=(\underline{v}_{1i}'\underline{\Lambda}_i)^2=\kappa_{1i}\rho(\mathbf{c}_{1i},\boldsymbol{\theta}_i)^2$. If it is the case that there exists $\zeta_o>0$ such that, for all j, $\zeta_o>\zeta_{ij}$, then, because $\underline{v}_{1i}'\underline{v}_{1i}=1$, $\rho(\mathbf{c}_{1i},\boldsymbol{\theta}_i)^2>(1-\frac{\zeta_o}{\kappa_{1i}})$. Finally, as $k\rightarrow\infty$, $\kappa_{1i}\rightarrow\infty$, and $\rho(\mathbf{c}_{1i},\boldsymbol{\theta}_i)^2\rightarrow 1$□

Bentler and Kano (1990) also prove that, as $k\rightarrow\infty$, $(\sqrt{\kappa_{1i}}\,\underline{v}_{1i}-\underline{\Lambda}_i)'(\sqrt{\kappa_{1i}}\,\underline{v}_{1i}-\underline{\Lambda}_i)\rightarrow 0$, i.e. that the vector of regression weights from a ulcf representation are asymptotically equivalent to the scaled first eigenvector.

These results make it abundantly clear that both the so-called common factor and the first principal component are constructed random variates (replacement variates), *the very same constructed random variate under the limiting conditions that bring about limiting uniqueness in the ulcf replacement*. Yet psychometricians continue to assert the Central Account, painting factor analysis, *in distinction to principal component analysis*, as the process of employing a detector/discoverer of unobservable properties/attributes. The truth about these matters is manifestly clear, and can be easily grasped by scholars of psychometrics if they are willing to dispense with the Central Account. The idea is this:

i. To paraphrase McDonald and Mulaik (1979), unidimensional linear factor analysis is concerned with the existence of a *random variate* $\boldsymbol{X}$ (an additional variable defined on the population, *P*, of objects under study) that has the property that, when a set of input variates $\boldsymbol{Y}_1,...,\boldsymbol{Y}_n$ are linearly regressed on it, the resulting residuals are uncorrelated.

ii. The replacement variates of unidimensional linear factor analysis are asymptotically equivalent to the first principal component. The primary sense, and a number of the secondary senses, of optimality of replacement delivered by the ulcf generator have already been reviewed. Principal component analysis has, traditionally, been recognized as involving the construction of new or synthetic variates that stand in certain optimal relations to an original set of variates. That is, principal component analysis has traditionally, and correctly, been characterized as a replacement variate generator.

iii. The ulcf and pc1 replacements are each optimal in their own defining senses, and come with their own specific costs. If it is insisted that a replacement variate, $\boldsymbol{\theta}$, not only renders conditionally uncorrelated a set of input variates, $\{\mathbf{Y}_1,\mathbf{Y}_2,\ldots,\mathbf{Y}_p\}$, but also that the residual covariance matrix be diagonal and positive definite, then $\boldsymbol{\theta}$ cannot be a linear function of the input $\mathbf{Y}_j$. If the $\{\mathbf{Y}_1,\mathbf{Y}_2,\ldots,\mathbf{Y}_p\}$ are ulcf-replaceable, it turns out that the set of ulcf replacements is of infinite cardinality (i.e., that Card(C)>1) and this means that the ulcf replacement is not unique. In contrast, under very mild conditions, the pc1 replacement is both unique and the pc1 replacement variate is a linear combination of the $\mathbf{Y}_j$. Which brand of replacement should be preferred would depend, presumably, on the tasks to which it will be put. Certainly, however, the Central Account, with its misportrayals and incoherences, cannot be legitimately employed to adjudicate the case.

iv. The diagonal and positive definite property of the residual covariance matrix of the ulcf has often been taken as the foundation of the CAM, in that the elements of this matrix are seen as "measurement error variances" or "unreliabilities". But the presence of a diagonal and positive definite matrix in a set of equations does *not* magically transform the equations into a "measurement model." As argued in Chapter IX, the interpretation of factor analysis as bearing on measurement issues springs from adherence to the Central Account mythology, and represents a blatant misportrayal of measurement and conceptual signification.

v. For finite p, Card(C_{pc1})=1 while Card(C_{ulcf})>1. But if, as more and more variates are added to the set of input variates, this set continues to remain ulcf-replaceable, then, in the limit, enough restrictions will be imposed to make Card(C_{ulcf}) equal to unity. If this limiting uniqueness does obtain, the single variate contained in C_{ulcf} is just the single replacement variate contained in C_{pc1}, and, certainly, is no more the common property/attribute of the phenomena represented by the variates than is the pc1 variate (or any other statistical function of the input variates).

vi. The difference between the finite p and variate domain foundations for ulcf and pc1 generators is akin to that between the data analytic (finite N) and inferential (infinite population) cases of general statistics. As with the latter distinction, the former member of the pair is akin to fact, while the latter is, at best, a useful abstraction. Moreover, the fact that, if the limiting requirements do hold (and it can never be known whether they do, or even what precisely is *meant* by an infinite domain of variables), the ulcf and pc1-replacements are equivalent, does not mean that the differences that exist at finite p are unimportant. Limiting equivalence of the two generators is interesting and highlights the fact that they are both just replacement variate generators, but in the all-important cases of finite p analyses, those analyses that are actually carried out by researchers, the ulcf and pc1 generators behave markedly differently.

3. *A general class of component generators*

Schonemann and Steiger (1976) have considered a class of replacement variate generators that they call regression component models (herein called regression component (rc) generators). This family of generators contains, as a special case, the pc1 replacement. Let $\underline{\mathbf{X}}$ contain a set of p input variates $\mathbf{X}_j$, j=1..p, jointly distributed in a population *P* under study. The scores that comprise the distributions of these variates are produced by following the rules $\{r_1,r_2,...,r_p\}$. Assume that $E(\underline{\mathbf{X}})=\underline{0}$ so that $E\underline{\mathbf{X}}\underline{\mathbf{X}}'=\Sigma$, and that rank($\Sigma$)=p, so that Σ is nonsingular.

Regression component(rc) replacement variates

A replacement variate **c** is sought such that

ri) $\mathbf{c}=\underline{t}'\underline{\mathbf{X}}$

rii) The p residuals $\underline{\mathbf{l}}=\underline{\mathbf{X}}-\underline{\sigma}_{\mathbf{X},\mathbf{c}}\sigma^{-2}_{\mathbf{c}}\mathbf{c}$ of the linear regressions of the input variates $\mathbf{X}_j$, j=1..p, on the replacement variate **c** have a covariance matrix $\Sigma_{\underline{\mathbf{l}}}$ with certain prescribed properties, the properties chosen defining a particular generator.

The term $\underline{\sigma}_{\mathbf{X},\mathbf{c}}\sigma^{-2}_{\mathbf{c}}\mathbf{c}$ is the linear conditional expectation of $\underline{\mathbf{X}}$ given the replacement variate **c**: $E(\underline{\mathbf{X}}|\mathbf{c}=c_o)=\underline{\sigma}_{\mathbf{X},\mathbf{c}}\sigma^{-2}_{\mathbf{c}}c_o$. Particular rc generators are derived by imposing further requirements on the covariance matrix of the residuals. For certain rc-generators, replacement only occurs given that the distribution of $\underline{\mathbf{X}}$ satisfies certain requirements, and, for these generators, rc-replaceability is a testable hypothesis. As Schonemann and Steiger (1976) emphasize, it is very easy indeed to generate testable hypotheses within the context of component analysis.

Consequences of (ri)-(rii)

The following are consequent properties of rc-replacements.

Ci) $E(\mathbf{c})=E(\underline{t}'\underline{\mathbf{X}})=0$ (from the fact that $E(\underline{\mathbf{X}})=\underline{0}$ and (ri))

Cii) $V(\mathbf{c})=\underline{t}'\Sigma\underline{t}$ (from (ri))

Ciii) $\underline{\sigma}_{\mathbf{X},\mathbf{c}}\sigma^{-2}_{\mathbf{c}}=(\underline{t}'\Sigma\underline{t})^{-1}\Sigma\underline{t}$

Civ) $\underline{\mathbf{l}}=\underline{\mathbf{X}}-\underline{\sigma}_{\mathbf{X},\mathbf{c}}\sigma^{-2}_{\mathbf{c}}\mathbf{c}=\underline{\mathbf{X}}-\underline{\sigma}_{\mathbf{X},\mathbf{c}}\sigma^{-2}_{\mathbf{c}}\mathbf{c}=(I-\Sigma\underline{t}(\underline{t}'\Sigma\underline{t})^{-1}\underline{t}')\underline{\mathbf{X}}=(I-\underline{a}\underline{t}')\underline{\mathbf{X}}$, in which $\underline{a}=\Sigma\underline{t}(\underline{t}'\Sigma\underline{t})^{-1}$ is the vector of linear regression weights for the regression of $\underline{\mathbf{X}}$ on **c**.

Cv) $\underline{\mathbf{X}}=\Sigma\underline{t}(\underline{t}'\Sigma\underline{t})^{-1}\underline{t}'\underline{\mathbf{X}}+(I-\Sigma\underline{t}(\underline{t}'\Sigma\underline{t})^{-1}\underline{t}')\underline{\mathbf{X}}=\underline{a}\mathbf{c}+\underline{\mathbf{l}}$.

Cvi) $E(\underline{\mathbf{l}})=E(I-\underline{a}\underline{t}')\underline{\mathbf{X}}=\underline{0}$

Cvii) $\Sigma_{\underline{l}}=E(\underline{\mathbf{l}}\underline{\mathbf{l}}')=E((I-\underline{a}\underline{t}')\underline{\mathbf{X}}\underline{\mathbf{X}}'(I-\underline{a}\underline{t}')')=\Sigma-\Sigma\underline{t}(\underline{t}'\Sigma\underline{t})^{-1}\underline{t}'\Sigma$

Note that, because $\Sigma_{\underline{l}}=\Sigma-\Sigma\underline{t}(\underline{t}'\Sigma\underline{t})^{-1}\underline{t}'\Sigma=\Sigma^{1/2}(I-\Sigma^{1/2}\underline{t}(\underline{t}'\Sigma\underline{t})^{-1}\underline{t}'\Sigma^{1/2})\Sigma^{1/2}$, and the middle term $(I-\Sigma^{1/2}\underline{t}(\underline{t}'\Sigma\underline{t})^{-1}\underline{t}'\Sigma^{1/2})$ is idempotent, it follows that $\text{rank}(\Sigma_{\underline{l}})=\text{rank}(I-\Sigma^{1/2}\underline{t}(\underline{t}'\Sigma\underline{t})^{-1}\underline{t}'\Sigma^{1/2})=\text{tr}(I-\Sigma^{1/2}\underline{t}(\underline{t}'\Sigma\underline{t})^{-1}\underline{t}'\Sigma^{1/2})=(p-1)$. That is, the replacement represents one dimension of variability in the distribution of the input variates, (p-1) dimensions being consigned to residual.

Cviii) $E(\mathbf{c}\underline{\mathbf{l}}')=E(\underline{t}'\underline{\mathbf{X}}\underline{\mathbf{X}}'(I-\underline{a}\underline{t}'))=\underline{t}'\Sigma-\underline{t}'\Sigma\underline{t}(\underline{t}'\Sigma(\underline{t}'\Sigma\underline{t})^{-1})=\underline{0}'$ (from (Cii) and (Civ))

That is, the replacement variate and residual variates are uncorrelated.

Cix) $\Sigma=C(E(\underline{\mathbf{X}}|\mathbf{c}))+E(C(\underline{\mathbf{X}}|\mathbf{c}))=\underline{b}\underline{b}'+\Sigma_{\underline{l}}$

in which $\underline{b}=\underline{a}(\underline{t}'\Sigma\underline{t})^{1/2}$. Hence, Σ has, under an rc-replacement, a decomposition of the same form as that brought about by the ulcf replacement (Schonemann & Steiger, 1976). The difference between the two lies in the fact that, in the ulcf representation, $\text{rank}(E(C(\underline{\mathbf{X}}|\mathbf{c}))=\text{rank}(\Sigma_{\underline{l}})=p$, while in unidimensional rc-analysis, $\text{rank}(E(C(\underline{\mathbf{X}}|\mathbf{c}))=\text{rank}(\Sigma_{\underline{l}})=(p-1)$. As Schonemann and Steiger note, and as was noted earlier in this chapter, the $\text{rank}(E(C(\underline{\mathbf{X}}|\mathbf{c}))=p$ requirement of the ulcf replacement is the source of its $\text{Card}(C)=\infty$ (indeterminacy) property.

Cx) $\underline{t}=\Sigma^{-1}\underline{a}(\underline{a}'\Sigma^{-1}\underline{a})^{-1}$ (from (Civ))

Cxi) $\underline{a}\underline{t}'=(\underline{t}'\Sigma\underline{t})^{-1}\Sigma\underline{t}\underline{t}'$ is idempotent

Cxii) $\underline{t}'\underline{a}=\underline{t}'(\underline{t}'\Sigma\underline{t})^{-1}\Sigma\underline{t}=1$

Cxiii) $E(\underline{\mathbf{X}}\mathbf{c})=E(\underline{\mathbf{X}}\underline{\mathbf{X}}'\underline{t})=\Sigma\underline{t}=\underline{a}(\underline{a}'\Sigma^{-1}\underline{a})^{-1}$

Cxiv) $E(\underline{\mathbf{X}}\underline{\mathbf{l}}')=E(\underline{\mathbf{X}}\underline{\mathbf{X}}'(I-\underline{a}\underline{t}')')=\Sigma-\Sigma\underline{t}\underline{a}'=\Sigma(I-\underline{t}(\underline{t}'\Sigma\underline{t})^{-1}\underline{t}'\Sigma)$ (from (Civ))

Special cases and existence

a. If **c** must be chosen so as to satisfy (ri), (rii), and (riii): $(\text{tr}(\Sigma_{\underline{l}}'\Sigma_{\underline{l}})$ a minimum), then $\underline{t}=\underline{v}_1=\underline{a}$, in which $\underline{v}_1$ is the first eigenvector of Σ, and the rc-generator is just the pc1-generator. As was seen earlier, this replacement always exists.

b. If **c** must be chosen so as to satisfy (ri), (rii), and (riii): $(\text{tr}(\Sigma_{\underline{l}}'\Sigma_{\underline{l}})=0)$, then **c** only exists if $\text{rank}(\Sigma)=1$, in which case $\underline{t}$ is the first eigenvector of Σ.

c. If, for rescaled variates $\Psi^{-1/2}\underline{\mathbf{X}}$, in which Ψ is diagonal and positive definite, **c** must be chosen so as to satisfy (ri), (rii), and (riii): ($\Sigma_{\underline{\mathbf{l}}}$ possesses (p-1) non-zero eigenvalues each equal to unity), then **c** does not necessarily exist. In fact, as the following theorem shows, $\Psi^{-1/2}\underline{\mathbf{X}}$ is rc-replaceable in this particular sense if and only if $\underline{\mathbf{X}}$ is ulcf replaceable.

Theorem (equivalence of ulcf and case (c) rc-replaceability; Schonemann & Steiger, 1976). A set of input random variates $\underline{\mathbf{X}}$ is ulcf replaceable with ulf representation $\underline{\mathbf{X}}=\underline{\Lambda}\boldsymbol{\theta}+\Psi^{1/2}\underline{\boldsymbol{\delta}}$ if and only if $\Psi^{-1/2}\underline{\mathbf{X}}$ is rc-replaceable with $\Sigma_{\underline{\mathbf{l}}}$ possessing (p-1) non-zero eigenvalues equal to unity.

Proof

$\rightarrow$

Let $\underline{\mathbf{X}}$ be ulcf replaceable with ulcf representation $\underline{\mathbf{X}}=\underline{\Lambda}\boldsymbol{\theta}+\Psi^{1/2}\underline{\boldsymbol{\delta}}$, from which it follows that $\Sigma_{\underline{X}}=\underline{\Lambda}\underline{\Lambda}'+\Psi$. Then, from (4.2vi), $\Psi^{-1/2}\Sigma\Psi^{-1/2}$ has a single eigenvalue λ_1 of greater than unity and (p-1) eigenvalues of unity. Hence, $\Psi^{-1/2}\Sigma\Psi^{-1/2}=\lambda_1\underline{v}_1\underline{v}_1+LL'$, in which the columns of $(\underline{v}_1{:}L)$ contain the p unit normalized eigenvectors of $\Psi^{-1/2}\Sigma\Psi^{-1/2}$. Define $\underline{a}$ in (Cv) to be $\sqrt{\lambda_1}\,\underline{v}_1$, from which it follows from (Cx) that $\underline{t}=\frac{1}{\sqrt{\lambda_1}}\underline{v}_1$, and, from (Civ), that $\underline{\mathbf{l}}=(I-\underline{a}\underline{t}')\Psi^{-1/2}\underline{\mathbf{X}}$. Then

$$C(\underline{\mathbf{l}})=\Psi^{-1/2}\Sigma\Psi^{-1/2}-\Psi^{-1/2}\Sigma\Psi^{-1/2}\frac{1}{\sqrt{\lambda_1}}\underline{v}_1\left[\frac{1}{\sqrt{\lambda_1}}\underline{v}_1{}'\Psi^{-1/2}\Sigma\Psi^{-1/2}\frac{1}{\sqrt{\lambda_1}}\underline{v}_1\right]^{-1}\frac{1}{\sqrt{\lambda_1}}\underline{v}_1\Psi^{-1/2}\Sigma\Psi^{-1/2}=LL'.$$

Since $C(\underline{\mathbf{l}})$ is then symmetric and idempotent, $\mathrm{rank}(C(\underline{\mathbf{l}}))=\mathrm{tr}(LL')=\mathrm{tr}(L'L)=(p-1)$, and the (p-1) non-zero eigenvalues are equal to unity.

$\leftarrow$

Let $\Psi^{-1/2}\underline{\mathbf{X}}$ be rc-replaceable with $\Sigma_{\underline{\mathbf{l}}}$ possessing (p-1) non-zero eigenvalues equal to unity. Then
$\Psi^{-1/2}\underline{\mathbf{X}}=\underline{a}\underline{t}'\Psi^{-1/2}\underline{\mathbf{X}}+(I-\underline{a}\underline{t}')\Psi^{-1/2}\underline{\mathbf{X}}=\underline{a}\underline{t}'\Psi^{-1/2}\underline{\mathbf{X}}+\underline{\mathbf{l}}$, from which it follows that $\Psi^{-1/2}\Sigma\Psi^{-1/2}=\underline{a}^*\underline{a}^{*}{}'+\Sigma_{\underline{\mathbf{l}}}$. Because $\Sigma_{\underline{\mathbf{l}}}$ possesses (p-1) non-zero eigenvalues of unity and is symmetric, $\Psi^{-1/2}\Sigma\Psi^{-1/2}=\underline{a}^*\underline{a}^{*}{}'+L_2L_2'$, in which $L_2'L_2=I_{(p-1)}$. Because it must be that $\underline{a}^{*}{}'\Sigma_{\underline{\mathbf{l}}}=\underline{a}^{*}{}'L_2L_2'=\underline{0}'$, let $\underline{a}^{*}=\varphi\underline{v}$, in which $\underline{v}'\underline{v}=1$ and $\underline{v}'L_2=\underline{0}'$ ($\underline{v}$ is the orthonormal complement of L_2). It follows then that $\Psi^{-1/2}\Sigma\Psi^{-1/2}=\underline{a}^*\underline{a}^{*}{}'+\Sigma_{\underline{\mathbf{l}}}=\varphi^2\underline{v}\underline{v}'+L_2L_2'=\underline{v}(\varphi^2-1)\underline{v}'+\underline{v}\underline{v}'+L_2L_2'=\underline{b}\underline{b}'+I$. Thus, $\Sigma=\underline{b}^*\underline{b}^{*}{}'+\Psi$, in which Ψ is diagonal and positive definite□

If **c** exists for the case (c) rc replacement, then $\underline{a}=\sqrt{\lambda_1}\,\underline{v}_1$ and $\underline{t}=\frac{1}{\sqrt{\lambda_1}}\underline{v}_1$, in which λ_1 and $\underline{v}_1$ are the first eigenvalue and eigenvector, respectively, of the matrix $\Psi^{-1/2}\Sigma\Psi^{-1/2}$. The essence of this rc-generator centres on one sense of what might be called *essential unidimensionality of replacement*: While the residuals of the regression of the input variates on the rc-replacement variate do not have a diagonal covariance matrix (unlike in the ulcf replacement), their distribution is spherical. Hence, any further replacement variates that could be constructed following the first (to bring about better reproduction of the covariances of the $\mathbf{X}_j$) are indistinguishable in regard the amount of variation that they would explain. Schonemann and Steiger (1976) describe this as discarding replacement variates two to p as correlated "error".

Cardinality of replacement

The cardinality of the replacement will depend upon the particulars of requirement (riii). Certainly, for rc-generators (a), (b), and (c), if the replacement exists, then it is unique, and *C* contains but one variate.

Construction formula

The construction formula is $\mathbf{c}=\underline{t}'\underline{\mathbf{X}}$.

Optimality properties

The primary sense of optimality possessed by rc-replacements lies in the fact that they are variance-covariance reproducers. They differ in terms of the stipulations they make with regard the residual covariance matrix $\Sigma_{\underline{l}}$. Generator (a), for example, insists that $(\mathrm{tr}(\Sigma_{\underline{l}}'\Sigma_{\underline{l}})$ be a minimum. That is, that knowledge of the parameters of the regressions of the $\mathbf{X}_j$ on $\mathbf{c}$ allows for the *best possible* reproduction of the parameters contained in Σ. Generator (b), on the other hand, insists that $\mathrm{tr}(\Sigma_{\underline{l}}'\Sigma_{\underline{l}})=0$, thereby calling for a replacement that perfectly reproduces the parameters of Σ. Generator (c) calls for a best possible reproduction in which, additionally, the variance and covariance not reproduced is associated uniformly with a further (p-1) replacement variates.

Characteristics of C

The rc-replacements considered herein each have a cardinality of unity. Hence, the internal characteristics of set *C* need not be considered. With regard the relation of the single replacement variate contained in *C* and external variates, $\underline{\mathbf{Y}}$,

$$(15.24)\qquad (\mathbf{c}\underline{\mathbf{Y}}')=E(\underline{t}'\underline{\mathbf{X}}\underline{\mathbf{Y}}')=\underline{t}'\Sigma_{XY},$$

and

$$(15.25)\qquad \rho(\mathbf{c}_1,\underline{\mathbf{Y}}')=\frac{1}{(\underline{t}'\Sigma\underline{t})^{\frac{1}{2}}}\underline{t}'\Sigma_{XY}D_Y^{-1/2},$$

in which $D_Y^{-1/2}$ is the diagonal matrix containing the reciprocals of the standard deviations of the $\mathbf{Y}_j$.

Testability

Certain rc-replacements always exist, e.g., rc-replacement (a), in which case there is no need for a test of replaceability. Others, e.g., replacements (b) and (c), do not necessarily exist, and in these cases a test of rc-replaceability makes sense.

4. *The partial image replacement generator*

Let $\underline{\mathbf{X}}$ contain a set of p input variates $\mathbf{X}_j$, j=1..p, jointly distributed in a population *P* under study. The scores that comprise the distributions of these variates are produced by following the rules $\{r_1,r_2,...,r_p\}$. Assume that $E(\underline{\mathbf{X}})=\underline{0}$ so that $E\underline{\mathbf{XX}}'=\Sigma$, and that $\text{rank}(\Sigma)=p$. The partial image (pi) replacement generator yields a multidimensional replacement of $\underline{\mathbf{X}}$.

Partial image (pi) replacement variates

A replacement p-vector $\underline{\mathbf{p}}$ is sought such that

ri) $\underline{\mathbf{p}}=F\underline{\mathbf{X}}$, so that $\underline{\mathbf{X}}=\underline{\mathbf{p}}+(I-F)\underline{\mathbf{X}}=\underline{\mathbf{p}}+\underline{\mathbf{a}}$

rii) $E(\underline{\mathbf{Xa}}')$ is diagonal

riii) $E(\text{diag}(\underline{\mathbf{pa}}'))=E(\text{diag}(F\underline{\mathbf{XX}}(I-F')))$ is comprised of zeros

Requirement (ri) states that the replacement vector $\underline{\mathbf{p}}$ must be a linear transformation of the input variates $\underline{\mathbf{X}}$, (rii) that, for j=1..p, $\mathbf{X}_j$ must only have non-zero correlation with the jth element of $\underline{\mathbf{a}}$, and (riii) that, for j=1..p, $E(\mathbf{p}_j\mathbf{a}_j)=0$.

Existence

$E(\underline{\mathbf{Xa}}')=E(\underline{\mathbf{XX}}'(I-F)')=\Sigma(I-F)'$, which, from (rii), is a diagonal matrix D. Because $\text{rank}(\Sigma)=p$, $(I-F)=D\Sigma^{-1}$. From (riii), $E(\text{diag}(F\underline{\mathbf{XX}}(I-F')))=E(\text{diag}((I-D\Sigma^{-1})\underline{\mathbf{XX}}\Sigma^{-1}D))=\text{diag}(D-D\Sigma^{-1}D)$ must be comprised of zeros. Thus, $D=[\text{diag}(\Sigma^{-1})]^{-1}$. Let D be resymbolized as Σ_d. It follows, then, that $\underline{\mathbf{p}}=F\underline{\mathbf{X}}=(I-\Sigma_d\Sigma^{-1})\underline{\mathbf{X}}$ and $\underline{\mathbf{a}}=\Sigma_d\Sigma^{-1}\underline{\mathbf{X}}$. The jth element of $\underline{\mathbf{p}}$ (called by Guttman (1953) the partial image of $\mathbf{X}_j$) is the linear conditional expectation of $\mathbf{X}_j$ on the remaining (p-1) input variates, while the jth element of $\underline{\mathbf{a}}$ (called by Guttman (1953) the partial anti-image of $\mathbf{X}_j$) is the corresponding residual variate. Note also that setting the jth diagonal element of $(I-\Sigma_d\Sigma^{-1})$ equal to zero in the construction of $\underline{\mathbf{p}}$ as $(I-\Sigma_d\Sigma^{-1})\underline{\mathbf{X}}$, ensures that each variate plays no role in the replacement of itself.

Consequences of (ri)-(riii)

The following are consequences of the replacement:

Ci) $E(\underline{\mathbf{p}})=E(F\underline{\mathbf{X}})=0$ (from the fact that $E(\underline{\mathbf{X}})=\underline{0}$)

Cii) $E(\underline{\mathbf{pp}}')=(I-\Sigma_d\Sigma^{-1})E\underline{\mathbf{XX}}'(I-\Sigma_d\Sigma^{-1})'=(I-\Sigma_d\Sigma^{-1})\Sigma(I-\Sigma_d\Sigma^{-1})'=(\Sigma-\Sigma_d)\Sigma^{-1}(\Sigma-\Sigma_d)$

Ciii) $E(\underline{\mathbf{a}})=E(I-F)\underline{\mathbf{X}}=\underline{0}$ (from the fact that $E(\underline{\mathbf{X}})=\underline{0}$)

Civ) $E(\underline{\mathbf{aa}}')=E(\Sigma_d\Sigma^{-1}\underline{\mathbf{XX}}'\Sigma^{-1}\Sigma_d)=\Sigma_d\Sigma^{-1}\Sigma_d$

Cv) $E(\underline{\mathbf{pa}}')=E((I-\Sigma_d\Sigma^{-1})E\underline{\mathbf{XX}}'\Sigma^{-1}\Sigma_d)=\Sigma_d-\Sigma_d\Sigma^{-1}\Sigma_d$

That is, the replacement vector (containing the partial images) and the anti-image variates are not, in general, uncorrelated, but, from (riii), $E(\mathbf{p}_j\mathbf{a}_j)=0$, j=1..p. That is, the diagonal of $\Sigma_d-\Sigma_d\Sigma^{-1}\Sigma_d$ is comprised of zeros.

Cvi) $\Sigma=E\underline{\mathbf{pp}}'+2E\underline{\mathbf{pa}}'+E\underline{\mathbf{aa}}'=[(\Sigma-\Sigma_d)\Sigma^{-1}(\Sigma-\Sigma_d)]+[\Sigma_d-\Sigma_d\Sigma^{-1}\Sigma_d]+\Sigma_d\Sigma^{-1}\Sigma_d$

This is a decomposition that differs from a linear factor decomposition by the non-vanishing of the matrix $E\underline{\mathbf{pa}}'$, and by the non-diagonality (for finite p) of the matrix $\Sigma_d\Sigma^{-1}\Sigma_d$.

Cvii) Letting $\underline{\mathbf{X}}_{(-j)}$ stand for the set of all input variates *except* the jth:

$R^2_{a_j\cdot\underline{X}_{(-j)}}=0$ for j=1..p (from (rii))

$R^2_{p_j\cdot\underline{X}_{(-j)}}=1$ for j=1..p

These properties follow from the construction formula for $\underline{\mathbf{p}}$, $\mathbf{p}_j=\underline{t}'\underline{\mathbf{x}}_{(-j)}$, in which $\underline{t}$ is the (p-1)-vector obtained by eliminating from the jth row of $(I-\Sigma_d\Sigma^{-1})$, the jth element.

Cviii) The jth diagonal element of $E(\underline{\mathbf{aa}}')$ is the variance of the jth residual variate, $\mathbf{a}_j$, i.e., $\sigma^2_{ej}=E(\mathbf{X}_j-\mathbf{p}_j)^2=(1-R^2_{X_j\cdot\underline{X}_{(-j)}})$. A necessary and sufficient condition that σ^2_{ej} be a minimum is that $\mathbf{a}_j$ be uncorrelated with all of the $\mathbf{X}_i$ except $\mathbf{X}_j$ (Guttman, 1956). It follows, then, from requirement (rii), that σ^2_{ej}, j=1..p, is a minimum.

Optimality properties

The pi-replacement is optimal in the sense that it yields the p-dimensional replacement vector, $\underline{\mathbf{p}}=F\underline{\mathbf{X}}$, that minimizes $Etr(\underline{\mathbf{X}}-\underline{\mathbf{p}})(\underline{\mathbf{X}}-\underline{\mathbf{p}})'=trE\underline{\mathbf{aa}}'=tr\Sigma_d\Sigma^{-1}\Sigma_d=\sum_{j=1}^{p}\sigma^2_{e_j}$. This is truly a variance oriented replacement, in the sense that a separate replacement variate is generated to reproduce the variance parameter of each input variate. That is, a single replacement variate $\underline{t}'\underline{\mathbf{X}}$ is sought that minimizes $Etr(\underline{\mathbf{X}}-(\underline{t}'\Sigma\underline{t})^{-1}\Sigma\underline{t}\underline{t}'\underline{\mathbf{X}})(\underline{\mathbf{X}}-(\underline{t}'\Sigma\underline{t})^{-1}\Sigma\underline{t}\underline{t}'\underline{\mathbf{X}})'$. In the rc-replacement, on the other hand, a single replacement variate must reproduce, as well as possible p variances and $\frac{1}{2}p(p-1)$ covariances.

Asymptotic equivalence of pi and lf replacements.

Guttman (1953, 1955) recommended using Σ_d as a proxy for Ψ in the linear factor analytic representation $\underline{\mathbf{X}}=\Lambda\underline{\boldsymbol{\theta}}+\Psi^{1/2}\underline{\boldsymbol{\delta}}$, because, if $\underline{\mathbf{X}}$ continues to be linear factor representable as $p\to\infty$, then $\Sigma^{-1}\to\Psi^{-1}$. But because $diag(\Sigma^{-1})=\Sigma_d$, with jth element $(1-R^2_{x_j\cdot\underline{x}_{(-j)}})^{-1}$, it follows that $(1-R^2_{x_j\cdot\underline{x}_{(-j)}})\to\zeta_j$, the jth element of Ψ, and $R^2_{x_j\cdot\underline{x}_{(-j)}}\to h^2_j$, the "communality" of the jth input variate. Furthermore, if, as $p\to\infty$, $\Sigma^{-1}\to\Psi^{-1}$, then, from (Cvi), $\Sigma\to[(\Sigma-\Sigma_d)\Sigma^{-1}(\Sigma-\Sigma_d)]+\Psi=\Lambda\Lambda'+\Psi$, in which Λ is a $p\times rank(E(\underline{\mathbf{pp}}'))$ matrix. Thus, if $\underline{\mathbf{X}}$ remains linear factor representable as $p\to\infty$, then, in the limit, image and common factor replacement variates are identical.

5. *The LISREL replacement generator*

Consider now a second family of multidimensional replacement generators. Let $\underline{\mathbf{z}}'=[\underline{\mathbf{y}}',\underline{\mathbf{x}}']$ contain a set of (p+q) input variates, jointly distributed in a population *P* under study. The scores that comprise the distributions of these variates are produced by following the rules $\{r_1,r_2,...,r_{p+q}\}$. Assume that $E(\underline{\mathbf{z}})=\underline{0}$ so that $E\underline{\mathbf{zz}}'=\Sigma_z$.

LISREL replacement variates

LISREL is said to be a confirmatory technique. But the view that it can be used to test theories about phenomena is, once again, the Central Account speaking. The latent variate terms of LISREL models are not linked to phenomena under study through the antecedent laying down of rules of correspondence. The confirmatory claim is properly interpreted to mean that a researcher must himself specify, in a given application, certain of the requirements inherent to the replacement brought about under the LISREL generator. As will be seen, the control that the researcher exerts over the replacement manifests itself in particular requirements that a random variate must satisfy in order to qualify as a LISREL replacement variate. Definitions of the matrices that appear in this section are as provided in Chapter 2, section 3b. We will only consider one of the simpler LISREL replacements and will assume that the parameters of any particular replacement are identified (see Bollen, 1989, for a review of this topic).

In the employment of the LISREL generator, an (n+m) element random vector $\underline{\boldsymbol{\rho}}'=[\boldsymbol{\xi}',\boldsymbol{\zeta}']$ is sought such that

ri) $E(\underline{\boldsymbol{\rho}})=\underline{0}$

rii) $C(\underline{\boldsymbol{\rho}})$ is block diagonal, i.e., $E(\boldsymbol{\xi\zeta}')$ is a null matrix

riii) $\underline{\boldsymbol{\eta}}=B\underline{\boldsymbol{\eta}}+\Gamma\boldsymbol{\xi}+\boldsymbol{\zeta}=(I-B)^{-1}(\Gamma\boldsymbol{\xi}+\boldsymbol{\zeta})$

and

riv) The (p+q) residuals $\underline{\mathbf{l}}$ of the linear regressions of the input variates $\underline{\mathbf{z}}$ on the (m+n) replacement vector $\underline{\boldsymbol{\kappa}}=[\underline{\boldsymbol{\eta}}',\underline{\boldsymbol{\xi}}']$ have a covariance matrix $C(\underline{\mathbf{l}})=\Theta$ that is positive definite and of some particular form (not necessarily diagonal).

Consequences

Ci) The linear conditional expectation of $\underline{\mathbf{z}}$ on $\underline{\boldsymbol{\kappa}}$, $E(\underline{\mathbf{z}}|\underline{\boldsymbol{\kappa}})_{lin}$, is equal to $\Sigma_{\underline{z},\underline{\kappa}}\Sigma_{\underline{\kappa}}^{-1}\underline{\boldsymbol{\kappa}}$. Now,

(15.26) $$\Sigma_{\underline{z},\underline{\kappa}}=E\begin{pmatrix} E(\underline{\mathbf{y}}\underline{\boldsymbol{\eta}}') & E(\underline{\mathbf{y}}\underline{\boldsymbol{\xi}}') \\ E(\underline{\mathbf{x}}\underline{\boldsymbol{\eta}}') & E(\underline{\mathbf{x}}\underline{\boldsymbol{\xi}}') \end{pmatrix}=\Lambda\Sigma_{\underline{\kappa}}$$

in which $\Lambda=\begin{pmatrix} \Lambda_y & \circ \\ \circ & \Lambda_x \end{pmatrix}$, and $\Sigma_{\underline{\kappa}}=\begin{pmatrix} (I-B)^{-1}(\Gamma\Phi\Gamma'+\Psi)(I-B)^{-1\prime} & (I-B)^{-1}\Gamma\Phi \\ \Phi\Gamma'(I-B)^{-1\prime} & \Phi \end{pmatrix}$.

Hence, $E(\underline{\mathbf{z}}|\underline{\boldsymbol{\kappa}})_{lin}=\Lambda\underline{\boldsymbol{\kappa}}$.

Cii) $\underline{\mathbf{l}}=\underline{\mathbf{z}}-\Lambda\underline{\boldsymbol{\kappa}}$

Ciii) Because $C(E(\underline{\mathbf{z}}|\underline{\boldsymbol{\kappa}}))=C(\Lambda\underline{\boldsymbol{\kappa}})=\Lambda\Sigma_{\underline{\kappa}}\Lambda'$, and $E(C(\underline{\mathbf{l}}))=\Sigma_z-\Lambda\Sigma_{\underline{\kappa}}\Lambda'=\Theta$,

$$\Sigma_z=\Lambda\Sigma_{\underline{\kappa}}\Lambda'+\Theta$$

That is,

(15.27)
$$\Sigma_z=\begin{pmatrix} \Lambda_y & \circ \\ \circ & \Lambda_x \end{pmatrix}\begin{pmatrix} (I-B)^{-1}(\Gamma\Phi\Gamma'+\Psi)(I-B)^{-1\prime} & (I-B)^{-1}\Gamma\Phi \\ \Phi\Gamma'(I-B)^{-1\prime} & \Phi \end{pmatrix}\begin{pmatrix} \Lambda_y & \circ \\ \circ & \Lambda_x \end{pmatrix}'+\Theta,$$

which is the well known LISREL covariance structure.

Existence

The LISREL replacement is just a special linear factor analytic replacement. Hence, the implication [(ri)-(riv)]$\rightarrow$[$\Sigma_z=\Lambda\Sigma_{\underline{\kappa}}\Lambda'+\Theta$, Θ positive definite and of some particular form] is true. Therefore, ~[$\Sigma_z=\Lambda\Sigma_{\underline{\kappa}}\Lambda'+\Theta$, Θ positive definite, of some particular form]$\rightarrow$~[(ri)-(riv)] is also true. Hence, if Σ_z cannot be decomposed as $\Lambda\Sigma_{\underline{\kappa}}\Lambda'+\Theta$, Θ a positive definite matrix of some particular form, then [(ri)-(riv)] cannot obtain. Vittadini (1989) has proven that, if Σ_z has the LISREL covariance structure (7.30), then there exists at least one vector $\underline{\boldsymbol{\rho}}'=[\boldsymbol{\xi}',\boldsymbol{\zeta}']$ that satisfies (ri)-(riv). Thus, under this condition, there will also exist at least one replacement vector $\underline{\boldsymbol{\kappa}}=[\boldsymbol{\xi}',\underline{\boldsymbol{\eta}}']$. If a set of input variates $\underline{\mathbf{z}}'=[\underline{\mathbf{y}}',\underline{\mathbf{x}}']$ are replaceable under the LISREL generator, they will be said to be L-replaceable, and $\underline{\boldsymbol{\kappa}}$ will be called an L-replacement to $\underline{\mathbf{z}}$.

Cardinality of replacement

Given that particular $\underline{\mathbf{z}}$ is L-replaceable, let C^* contain all vectors $\underline{\boldsymbol{\rho}}$ that satisfy (ri)-(riv). Vittadini (1989) has proven that, for finite (p+q), Card(C^*)=∞. Hence, if C stands for the set of replacement vectors, i.e., the set containing the vectors $\underline{\boldsymbol{\kappa}}$ that L-replace the input variates, it follows that Card(C)=∞. That is, the L-replacement of $\underline{\mathbf{z}}$ is not unique, there being constructible an infinity of random vectors each of which satisfies (ri)-(riv) if, in fact, $\underline{\mathbf{z}}$ is L-replaceable.

Construction formula

The construction formula for the vectors contained in C^* is

(15.28)
$$\begin{pmatrix}\underline{\zeta}\\ \underline{\xi}\end{pmatrix}_i=\begin{pmatrix}\Psi & \circ\\ \circ & \Phi\end{pmatrix}\begin{pmatrix}\Lambda_y(I-B)^{-1} & \Lambda_y(I-B)^{-1}\Gamma\\ \circ & \Lambda_X\end{pmatrix}'\Sigma_z^{-1}\underline{\mathbf{z}}+W^{1/2}\underline{\mathbf{s}}_i$$

$$=\Omega\Lambda^{*\prime}\Sigma_z^{-1}\underline{\mathbf{z}}+W^{1/2}\underline{\mathbf{s}}_i,$$

in which $E(\underline{\mathbf{s}}_i)=\underline{0}$, $C(\underline{\mathbf{s}}_i)=I$, $E(\underline{\mathbf{z}}\underline{\mathbf{s}}_i')=0_{(p+q)\times(m+n)}$, $W=(\Omega-\Omega\Lambda^{*\prime}\Sigma_z^{-1}\Lambda^*\Omega)$,

$$\Omega=C\begin{pmatrix}\underline{\zeta}\\ \underline{\xi}\end{pmatrix}_i=\begin{pmatrix}\Psi & \circ\\ \circ & \Phi\end{pmatrix},\text{ and }\Lambda^*=\begin{pmatrix}\Lambda_y(I-B)^{-1} & \Lambda_y(I-B)^{-1}\Gamma\\ \circ & \Lambda_X\end{pmatrix}.$$

Because $\underline{\boldsymbol{\eta}}=(I-B)^{-1}(\Gamma\boldsymbol{\xi}+\boldsymbol{\zeta})$, it then follows that the construction formula for the vectors contained in C is

(15.29)

$$\kappa_i = \begin{pmatrix} \underline{\eta} \\ \underline{\xi} \end{pmatrix}_i = \begin{pmatrix} (I-B)^{-1} & (I-B)^{-1}\Gamma \\ \circ & I \end{pmatrix} \left[\begin{pmatrix} \Psi & \circ \\ \circ & \Phi \end{pmatrix} \begin{pmatrix} \Lambda_y(I-B)^{-1} & \Lambda_y(I-B)^{-1}\Gamma \\ \circ & \Lambda_X \end{pmatrix}' \Sigma_z^{-1}\underline{z} + W^{1/2}\underline{s}_i \right]$$

$$=F\Omega\Lambda^{*\prime}\Sigma_z^{-1}\underline{\mathbf{z}}+FW^{1/2}\underline{\mathbf{s}}_i,$$

in which $F = \begin{pmatrix} (I-B)^{-1} & (I-B)^{-1}\Gamma \\ \circ & I \end{pmatrix}$.

Characteristics of C

Vittadini (1989) derived Guttman's minimum correlation measure for the case of the L-replacement. While one could consider the minimum correlation for any of the (2m+n+p+q) sets of constructed variates $C_{\eta i}$, i=1..m, $C_{\xi j}$, j=1..n, $C_{\zeta k}$, k=1..m, $C_{\delta s}$, s=1..q, $C_{\varepsilon t}$, t=1..p, attention will, herein, be focussed on the two vectors of replacement variates $\underline{\boldsymbol{\eta}}$ and $\underline{\boldsymbol{\xi}}$. To begin, the matrix of covariances between construction $\{\underline{\boldsymbol{\zeta}}_i,\underline{\boldsymbol{\xi}}_i\}$ and distinct construction $\{\underline{\boldsymbol{\zeta}}_j,\underline{\boldsymbol{\xi}}_j\}$is

(15.30)

$$\Sigma_{\{\underline{\zeta},\underline{\xi}\}_{ij}} =$$

$$\begin{pmatrix} \Psi & \circ \\ \circ & \Phi \end{pmatrix} \begin{pmatrix} \Lambda_y(I-B)^{-1} & \Lambda_y(I-B)^{-1}\Gamma \\ \circ & \Lambda_X \end{pmatrix}' \Sigma_Z^{-1} \begin{pmatrix} \Lambda_y(I-B)^{-1} & \Lambda_y(I-B)^{-1}\Gamma \\ \circ & \Lambda_X \end{pmatrix} \begin{pmatrix} \Psi & \circ \\ \circ & \Phi \end{pmatrix} + W^{1/2}\Sigma_{\underline{s}_i,\underline{s}_j}W^{1/2}$$

$$=\Omega\Lambda^{*\prime}\Sigma_z^{-1}\Lambda^{*}\Omega+W^{1/2}\Sigma_{\underline{s}i,\underline{s}j}W^{1/2}$$

in which $\Sigma_{\underline{s}i,\underline{s}j}$ is the (m+n) by (m+n) correlation matrix $E(\underline{\mathbf{s}}_i\underline{\mathbf{s}}_j')$. The covariance matrix between any replacement vector $\{\underline{\boldsymbol{\eta}}_i,\underline{\boldsymbol{\xi}}_i\}$ and distinct replacement vector $\{\underline{\boldsymbol{\eta}}_j,\underline{\boldsymbol{\xi}}_j\}$ is then

(15.31) $$\Sigma_{\{\underline{\eta},\underline{\xi}\}_{ij}} = \begin{pmatrix} (I-B)^{-1} & (I-B)^{-1}\Gamma \\ \circ & I \end{pmatrix} \Sigma_{\{\underline{\zeta},\underline{\xi}\}_{ij}} \begin{pmatrix} (I-B)^{-1} & (I-B)^{-1}\Gamma \\ \circ & I \end{pmatrix}'$$

$$=F\Omega\Lambda^{*\prime}\Sigma_z^{-1}\Lambda^{*}\Omega F'+FW^{1/2}\Sigma_{\underline{s}i,\underline{s}j}W^{1/2}F'$$

The diagonal elements of $\Sigma_{\{\underline{\eta},\underline{\xi}\}_{ij}}$ are minimized by taking $\underline{\mathbf{s}}_j$ to be $-\underline{\mathbf{s}}_i$, in which case $\Sigma_{\underline{s}i,\underline{s}j}$=-I, and

$$\Sigma_{\{\underline{\zeta},\underline{\xi}\}_{ij}\,min} = 2\begin{pmatrix}\Psi & \circ \\ \circ & \Phi\end{pmatrix}\begin{pmatrix}\Lambda_y(I-B)^{-1} & \Lambda_y(I-B)^{-1}\Gamma \\ \circ & \Lambda_X\end{pmatrix}'\Sigma_z^{-1}\begin{pmatrix}\Lambda_y(I-B)^{-1} & \Lambda_y(I-B)^{-1}\Gamma \\ \circ & \Lambda_X\end{pmatrix}\begin{pmatrix}\Psi & \circ \\ \circ & \Phi\end{pmatrix} - \begin{pmatrix}\Psi & \circ \\ \circ & \Phi\end{pmatrix}$$

$$=2\Omega\Lambda^{*}{}'\Sigma_z^{-1}\Lambda^{*}\Omega-\Omega.$$

It then follows that

(15.32)
$$R_{\{\underline{\eta},\underline{\xi}\}_{ij\,min}} = D^{-1/2}\begin{pmatrix}(I-B)^{-1} & (I-B)^{-1}\Gamma \\ \circ & I\end{pmatrix}\Sigma_{\{\underline{\zeta},\underline{\xi}\}_{ij\,min}}\begin{pmatrix}(I-B)^{-1} & (I-B)^{-1}\Gamma \\ \circ & I\end{pmatrix}' D^{-1/2}$$

$$=D^{-1/2}F2\Omega\Lambda^{*}{}'\Sigma_z^{-1}\Lambda^{*}\Omega F'D^{-1/2}-D^{-1/2}F\Omega F'D^{-1/2}$$

in which D=Diag($\Sigma_{\{\underline{\boldsymbol{\eta}}i,\underline{\boldsymbol{\xi}}i\}}$). The (m+n) diagonal elements of this matrix contain the minimum correlations for the sets $C_{\eta i}$, i=1..m, and $C_{\xi j}$, j=1..n.

In view of the fact that $\Sigma_z^{-1}=\Theta^{-1}-\Theta^{-1}\Lambda^{*}(\Omega+\Lambda^{*}{}'\Theta^{-1}\Lambda^{*})^{-1}\Lambda^{*}{}'\Theta^{-1}$, it follows that

(15.33) $$\Sigma_{\{\underline{\zeta},\underline{\xi}\}_{ij}\,min}=2\Omega[B-B(\Omega+B)^{-1}B]\Omega-\Sigma_{\{\underline{\zeta},\underline{\xi}\}_{ij}},$$

in which $B=\Lambda^{*}{}'\Theta^{-1}\Lambda^{*}$,

(15.34) $$\Sigma_{\{\underline{\eta},\underline{\xi}\}_{ij}\,min}=2FB\Omega BF'-2F\Omega B(\Omega+B)^{-1}B\Omega F'-F\Sigma_{\{\underline{\zeta},\underline{\xi}\}_{ij}}F',$$

and

(15.39) $$R_{\{\underline{\eta},\underline{\xi}\}_{ij\,min}}=2D^{-1/2}FB\Omega BF'D^{-1/2}-2D^{-1/2}F\Omega B(\Omega+B)^{-1}B\Omega F'D^{-1/2}-R_{\{\underline{\zeta},\underline{\xi}\}_{ij}}$$

If only the replacement variates $\underline{\eta}$ and $\underline{\xi}$ are considered, the average minimum correlation is equal to

(15.40) $$\frac{1}{(m+n)}\,\mathrm{tr}(R_{\{\underline{\eta},\underline{\xi}\}_{ij_{min}}}).$$

If all of the (m+n+p+q) constructed variates contained in the vectors $\underline{\xi}$, $\underline{\zeta}$, $\underline{\delta}$, $\underline{\varepsilon}$, are considered, and are required to have variances of unity and be mutually uncorrelated, then the average minimum correlation is equal to (p+q-m-n)/(p+q+m+n). Vittadini (1989) gave bounds for the average minimum correlation in the general case.

Example

Consider the case in which p=4, q=4, m=2, n=2, and, in the population,

$$\Sigma_z=\begin{pmatrix}\Lambda_y(I-B)^{-1} & \Lambda_y(I-B)^{-1}\Gamma\\ \circ & \Lambda_X\end{pmatrix}\begin{pmatrix}\Psi & \circ\\ \circ & \Phi\end{pmatrix}\begin{pmatrix}\Lambda_y(I-B)^{-1} & \Lambda_y(I-B)^{-1}\Gamma\\ \circ & \Lambda_X\end{pmatrix}'+\Theta,$$

with

$$\Lambda y:=\begin{pmatrix}.72 & 0\\ .76 & 0\\ 0 & .5\\ 0 & .5\end{pmatrix}\quad \Lambda x:=\begin{pmatrix}.95 & 0\\ .72 & 0\\ 0 & .39\\ 0 & .11\end{pmatrix}\quad B:=\begin{pmatrix}0 & 0\\ .18 & 0\end{pmatrix}\quad \Gamma:=\begin{pmatrix}.45 & 0\\ 0 & .78\end{pmatrix}\quad \Psi:=\begin{pmatrix}.26 & 0\\ 0 & .07\end{pmatrix}\quad \Phi:=\begin{pmatrix}1 & .34\\ .34 & 1\end{pmatrix}$$

$$\Theta=\begin{pmatrix}0.7602 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 0.7329 & 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0.8159 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0.8159 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0.0975 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 0.4816 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 0 & 0.8479 & 0\\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.9879\end{pmatrix}$$

In this case,

$$\begin{pmatrix}\underline{\boldsymbol{\eta}}\\ \underline{\xi}\end{pmatrix}_i = \begin{pmatrix}(I-B)^{-1} & (I-B)^{-1}\Gamma\\ \circ & I\end{pmatrix}\begin{pmatrix}\underline{\zeta}\\ \underline{\xi}\end{pmatrix}_i$$

$$= \begin{pmatrix}(I-B)^{-1} & (I-B)^{-1}\Gamma\\ \circ & I\end{pmatrix}\left[\begin{pmatrix}\Psi & \circ\\ \circ & \Phi\end{pmatrix}\begin{pmatrix}\Lambda_y(I-B)^{-1} & \Lambda_y(I-B)^{-1}\Gamma\\ \circ & \Lambda_X\end{pmatrix}'\Sigma_z^{-1}\underline{\mathbf{z}} + W^{1/2}\underline{\mathbf{s}}_i\right]$$

$$= \begin{pmatrix} .186 & .204 & 019 & .019 & .27 & .041 & -.002 & .000\\ .030 & .033 & .250 & .250 & .180 & .028 & .206 & .050\\ .026 & .029 & .011 & .011 & .840 & .129 & .008 & .002\\ -.003 & -.003 & .275 & .275 & .179 & .028 & .276 & .067\end{pmatrix}\underline{\mathbf{z}} + \begin{pmatrix} .435 & -.002 & .106 & .000\\ .078 & .260 & .016 & .525\\ 0 & 0 & .294 & .060\\ 0 & 0 & 0 & .768\end{pmatrix}\underline{\mathbf{s}}i$$

The matrix of minimum correlations for this replacement is equal to

$$R_{\{\underline{\eta},\underline{\xi}\}_{ij_{min}}} = \begin{pmatrix} .1504 & .2394 & .5803 & .2347\\ .2394 & -.107 & .3603 & -.1029\\ .5803 & .3603 & .8277 & .3033\\ .2347 & -.1029 & .3033 & -.1993\end{pmatrix}$$

Hence, the minimum correlations for the sets $C_{\eta 1}$, $C_{\eta 2}$, $C_{\xi 1}$, and $C_{\xi 2}$ which contain, respectively, the replacement variates $\boldsymbol{\eta}_1$, $\boldsymbol{\eta}_2$, ξ_1, and ξ_2, range from a low of -.1993 for $C_{\xi 2}$, to a high of .8277 for $C_{\xi 1}$. Evidently, there exists a great deal of latitude in the ξ_2 replacement (i.e., radically different random variates are contained within $C_{\xi 2}$), and much less so in regard the $\boldsymbol{\xi}_1$ replacement (the variates contained in $C_{\xi 1}$ are relatively similar: the most dissimilar pairs of $\boldsymbol{\xi}_1$ replacement variates have in common $100\times.83^2$=69% of their variance).

Optimality criteria

The L-replacement is just a special case of the usual linear factor analytic replacement, there being two chief differences:

i. In the case of the L-replacement, the covariance matrix of the residuals, Θ, is allowed to be non-diagonal. If it is insisted that Θ be diagonal and positive definite, and a particular set of input variates $\underline{\mathbf{z}}_o$ is L-replaceable, then the replacement vector $\underline{\boldsymbol{\kappa}}$ replaces the $\mathbf{z}_{oj}$ in the usual factor analytic sense that the $\frac{1}{2}(p+q)(p+q-1)$ unique covariances contained in Σ_z can be reproduced via knowledge of the values of the parameters contained in the matrices Λ_X, Λ_Y, B, Γ, Ψ, and Φ. If, on the other hand, it is insisted that Θ be of some specific non-diagonal form, and a particular set of input variates, $\underline{\mathbf{z}}_o$, is L-replaceable, then the replacement vector $\underline{\boldsymbol{\kappa}}$ *partially* replaces the $\mathbf{z}_{oj}$ in this sense. That is, certain of the covariances in Σ_z are allowed to be imperfectly reproduced by the parameters of the p regressions of $\underline{\mathbf{z}}$ on $\underline{\boldsymbol{\kappa}}$.

ii. The discipline's immersion in the Central Account has fostered the view that, in employing LISREL, and covariance structure analytic techniques in general, the researcher is testing hypotheses about the relationships amongst observable and unobservable entities, and, in particular, hypotheses as to the existence of the latter class of entity. In many applications, versions of the CAC are invoked, the employment of the LISREL generator is mislabeled as "causal modeling", and such testing is viewed as the testing of "causal hypotheses." In other applications, the CAM is invoked and LISREL based tests are portrayed as tests of measurement claims. But this is fantasy. The LISREL generator is confirmatory in the sense that, in contrast to the standard linear factor generators, the analyst must specify in a fair bit of detail the properties that variates must have in order that they can rightly be said to L-replace a set of input variates. Some of these requirements are substantive in nature, while others are pragmatic restrictions imposed to identify the replacement being an example of the latter.

Consider the oft-tested hypotheses that a set of p input variates is either parallel, tau-equivalent, or congeneric (as in, e.g., Joreskog & Sorbom, 1993, p.115). On the mythology of the Central Account, these hypotheses are portrayed as measurement hypotheses. In fact, they are simply hypotheses of L-replaceability. First, consider the simple hypothesis that a set of input variates $\underline{\mathbf{X}}_o$ is ulcf-replaceable, or in the parlance of psychometrics, that these variates constitute a set of congeneric measures. To hypothesize ulcf-replaceability is to hypothesize that there exist a parameter $\underline{\Lambda}_o$ such that the residuals of the linear regressions of the $\mathbf{X}_{oj}$ on a replacement variate have a diagonal, positive definite covariance matrix Ψ_o. Recall from Chapter IV that, if $\underline{\mathbf{X}}_o$ is ulcf-replaceable, the construction formula of the ulcf replacement variates can be expressed as

$$(15.37) \qquad \boldsymbol{\theta}_i=\underline{\Lambda}_o'\Sigma_o^{-1}\underline{\mathbf{X}}_o+w_o^{1/2}\mathbf{s}_i=w_o\underline{\Lambda}_o'\Psi_o^{-1}\underline{\mathbf{X}}+w_o^{1/2}\mathbf{s}_i= w_o\sum_{j=1}^{p}\frac{\lambda_{oj}}{\sigma^2_{o\zeta_j}}\mathbf{X}_{oj}+w_o^{1/2}\mathbf{s}_i$$

Note that the contribution of each input variate $\mathbf{X}_{oj}$ to $\boldsymbol{\theta}_i$ is governed by the ratio $\frac{\lambda_{oj}}{\sigma^2_{o\zeta_j}}$.

Thus, a variate whose regression weight is large, and residual variance small, will have relatively large input into the constructed replacement variate $\boldsymbol{\theta}_i$.

Now consider the hypothesis of parallelism, i.e., that $\lambda_j=\lambda\ \forall j$, and $\sigma^2_{o\zeta_j}=\sigma^2_{o\zeta}\ \forall j$. If a set of input variates is ulcf replaceable with the added restrictions that that $\lambda_j=\lambda\ \forall j$ and $\sigma^2_{o\zeta_j}=\sigma^2_{o\zeta}\ \forall j$, then the resulting construction formula is

(15.38) $$\boldsymbol{\theta}_i=\frac{\lambda}{(\sigma^2_{o\zeta}+p\lambda^2)}\mathbf{t}+w_o^{1/2}\mathbf{s}_i$$

in which $\mathbf{t}=\sum_{j=1}^{p}\mathbf{X}_{oj}$. Hence, to hypothesize parallelism is to hypothesize that there exist parameters λ and $\sigma^2_{o\zeta}$ such that the residuals of the linear regressions of the $\mathbf{X}_{oj}$ on a replacement variate $\boldsymbol{\theta}_i=\frac{\lambda}{(\sigma^2_{o\zeta}+p\lambda^2)}\mathbf{t}+w_o^{1/2}\mathbf{s}_i$ have a diagonal, positive definite covariance matrix. For this hypothesis of replaceability to be true, each input variate must make an identical contribution to the replacement variate $\boldsymbol{\theta}_i$. Hence, to hypothesize parallelism is to hypothesize the existence of a ulcf replacement variate to which the input variates make equal contributions.

Finally, consider the hypothesis of tau-equivalence i.e., that $\lambda_j=\lambda\ \forall j$. If this hypothesis is correct, then the associated construction formula is

(15.39) $$\boldsymbol{\theta}_i=\lambda w\sum_{j=1}^{p}\frac{1}{\sigma^2_{o\zeta_j}}\mathbf{X}_{oj}+w_o^{1/2}\mathbf{s}_i .$$

Hence, to hypothesize tau-equivalence is to hypothesize that there exists (p+1) parameters λ and $\sigma^2_{o\zeta_j}$, j=1..p, such that the residuals of the linear regressions of the $\mathbf{X}_{oj}$ on a replacement variate $\boldsymbol{\theta}_i=\lambda w\sum_{j=1}^{p}\frac{1}{\sigma^2_{o\zeta_j}}\mathbf{X}_j+w_o^{1/2}\mathbf{s}_i$ have a diagonal, positive definite covariance matrix. Under this hypothesis of replaceability, the contribution of each input variate to the constructed replacement variate varies only through the residual variance of the variate.

Testability

A great deal of attention has been given to the topic of testing the general hypothesis

$$H_o\text{: }\Sigma_z=\begin{pmatrix}\Lambda_y(I-B)^{-1} & \Lambda_y(I-B)^{-1}\Gamma\\ \circ & \Lambda_X\end{pmatrix}\begin{pmatrix}\Psi & \circ\\ \circ & \Phi\end{pmatrix}\begin{pmatrix}\Lambda_y(I-B)^{-1} & \Lambda_y(I-B)^{-1}\Gamma\\ \circ & \Lambda_X\end{pmatrix}'+\Theta$$

versus the alternative H_1: [Σ_z is any gramian matrix] (see Bollen, 1989, for a review). Under the supposition that $\underline{\mathbf{z}}$ has a multivariate normal distribution, a test may be derived according to either maximum likelihood or generalized least squares principles. Obviously, a test of H_o versus H_1 is a test of whether or not $\underline{\mathbf{z}}$ is L-replaceable.

6. *Unidimensional quadratic factor (uqf) generator*

Let $\underline{\mathbf{X}}$ contain a set of p input variates, jointly distributed in a population *P* under study. The scores that comprise the distributions of these variates are produced by following some particular set of rules $\{r_1, r_2, ..., r_{p+q}\}$. Assume that $E(\underline{\mathbf{X}})=\underline{0}$ so that $E\underline{\mathbf{XX}}'=\Sigma$.

uqf replacement variates

A replacement variate $\boldsymbol{\theta}$ is sought with the properties that

ri) $\boldsymbol{\theta}\sim N(0,1)$,

and

rii) The residuals $\underline{\mathbf{l}}$ of the linear regressions of the input variates $\mathbf{X}_j$, j=1..p, on $\underline{\mathbf{f}}(\boldsymbol{\theta})=\begin{bmatrix}\boldsymbol{\theta}\\ \frac{(\boldsymbol{\theta}^2-1)}{\sqrt{2}}\end{bmatrix}$ have a covariance matrix Ψ that is diagonal and positive definite.

Comment: This is a unidimensional replacement in the sense that but one variate $\boldsymbol{\theta}$ is sought, but, if such a variate exists, the input variates have a quadratic regression on it. The requirement that $C(\underline{\mathbf{l}})=\Sigma-\Lambda\Omega\Lambda'=\Psi$, in which Ψ is diagonal and positive definite, $\Omega=C(\underline{\mathbf{f}}(\boldsymbol{\theta}))$ and $\Lambda=E(\underline{\mathbf{Xf}}(\boldsymbol{\theta})')$, insists upon the usual linear factor analytic brand of reproducibility. That is, knowledge of a set of regression parameters (in this case those involving first and second degree polynomials in $\boldsymbol{\theta}$) allows one to reproduce the covariances contained in Σ.

Consequences

Ci) $$E\underline{\mathbf{f}}(\boldsymbol{\theta})=E\begin{bmatrix}\boldsymbol{\theta}\\ \frac{(\boldsymbol{\theta}^2-1)}{\sqrt{2}}\end{bmatrix}=\begin{bmatrix}E\boldsymbol{\theta}\\ \frac{E\boldsymbol{\theta}^2-1}{\sqrt{2}}\end{bmatrix}=\underline{0}$$

Cii) $\Omega=E\underline{\mathbf{f}}(\boldsymbol{\theta})\underline{\mathbf{f}}(\boldsymbol{\theta})'=$

$$\begin{bmatrix}E\boldsymbol{\theta}^2 & E\boldsymbol{\theta}\frac{(\boldsymbol{\theta}^2-1)}{\sqrt{2}}\\ E\boldsymbol{\theta}\frac{(\boldsymbol{\theta}^2-1)}{\sqrt{2}} & E\frac{(\boldsymbol{\theta}^2-1)}{\sqrt{2}}\frac{(\boldsymbol{\theta}^2-1)}{\sqrt{2}}\end{bmatrix}=\begin{bmatrix}1 & \frac{E\boldsymbol{\theta}^3-E\boldsymbol{\theta}}{\sqrt{2}}\\ \frac{E\boldsymbol{\theta}^3-E\boldsymbol{\theta}}{\sqrt{2}} & \frac{E\boldsymbol{\theta}^4-2E\boldsymbol{\theta}^2+1}{\sqrt{2}}\end{bmatrix}=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}=I$$

since, from (ri), $E(\boldsymbol{\theta})=0$, $E(\boldsymbol{\theta}^2)=1$, $E(\boldsymbol{\theta}^3)=0$, and $E(\boldsymbol{\theta}^4)=3$.

Ciii) From (rii) and (Cii), $E(\underline{\mathbf{X}}|\underline{\mathbf{f}}(\boldsymbol{\theta})=\underline{f}(\theta_o))=E(\underline{\mathbf{X}}\underline{\mathbf{f}}(\boldsymbol{\theta})')\Omega^{-1}\underline{f}(\theta_o)=\Lambda\underline{f}(\theta_o)$ is the linear conditional expectation of $\underline{\mathbf{X}}$ given $\underline{\mathbf{f}}(\boldsymbol{\theta})=\underline{f}(\theta_o)$, and, hence, $(\underline{\mathbf{l}}|\underline{\mathbf{f}}(\boldsymbol{\theta})=\underline{f}(\theta_o))=((\underline{\mathbf{X}}-\Lambda\underline{f}(\theta_o))|\underline{\mathbf{f}}(\boldsymbol{\theta})=\underline{f}(\theta_o))$.

Civ) $E(\underline{\mathbf{l}}|\underline{\mathbf{f}}(\boldsymbol{\theta})=\underline{f}(\theta_o))=E((\underline{\mathbf{X}}-\Lambda\underline{f}(\theta_o))|\underline{\mathbf{f}}(\boldsymbol{\theta})=\underline{f}(\theta_o))=E(\underline{\mathbf{X}}|\underline{\mathbf{f}}(\boldsymbol{\theta})=\underline{f}(\theta_o))-\Lambda\underline{f}(\theta_o)=\underline{0}$. Thus, $E(\underline{\mathbf{l}})=EE(\underline{\mathbf{l}}|\underline{\mathbf{f}}(\boldsymbol{\theta}))=\underline{0}$

Cv) $E(\underline{\mathbf{X}}\underline{\mathbf{l}}')=E(\underline{\mathbf{X}}(\underline{\mathbf{X}}-\Lambda\underline{\mathbf{f}}(\boldsymbol{\theta}))')=\Sigma-\Lambda\Lambda'$.

Cvi) $E(\underline{\mathbf{f}}(\boldsymbol{\theta})\underline{\mathbf{l}}')=E(\underline{\mathbf{f}}(\boldsymbol{\theta})(\underline{\mathbf{X}}-\Lambda\underline{\mathbf{f}}(\boldsymbol{\theta}))')=\underline{0}\underline{0}'_{(2\times p)}$

Cvii) $\underline{\mathbf{X}}=\Lambda\underline{\mathbf{f}}(\boldsymbol{\theta})+\underline{\mathbf{l}}$

Cviii) $\Sigma=C(E(\underline{\mathbf{X}}|\underline{\mathbf{f}}(\boldsymbol{\theta}))+E(C((\underline{\mathbf{X}}|\underline{\mathbf{f}}(\boldsymbol{\theta}))=\Lambda\Lambda'+\Psi$,

in which Λ is a (p×2) matrix of real coefficients, and Ψ, diagonal and positive definite. Clearly, the uqf replacement has precisely the same covariance structure as the two-dimensional linear factor replacement.

Existence

Let $\Lambda=[\underline{\lambda}_1,\underline{\lambda}_2]$. Note that, if $\underline{\lambda}_2=\underline{0}$, (rii) would call for the usual ulcf-replaceability, and $\boldsymbol{\theta}$ would exist so long as $\Sigma=\underline{\lambda}_1\underline{\lambda}_1'+\Psi$, in which Ψ is diagonal and positive definite. The construction formula for $\boldsymbol{\theta}$ in this case would simply be $\boldsymbol{\theta}_i=\underline{\lambda}_1'\Sigma^{-1}\underline{\mathbf{X}}+w^{1/2}\mathbf{s}_i$, in which $\mathbf{s}_i$ is selected so that it is statistically independent of $\underline{\mathbf{X}}$ and $\boldsymbol{\theta}_i\sim N(0,1)$. Now, the second element of $\underline{\mathbf{f}}(\boldsymbol{\theta})$ can be obtained from the first by the formula $\frac{(\boldsymbol{\theta}_i^{\,2}-1)}{\sqrt{2}}$. Thus, it follows that two conditions must be met in order for $\underline{\mathbf{X}}$ to be uqf-replaceable:

i. $\Sigma=\Lambda^*\Lambda^{*\prime}+\Psi$, in which Ψ is diagonal and positive definite, and Λ^* is a (p×2) matrix of real coefficients.

ii. From (rii) and (Ciii), it must be the case that $E(\underline{\mathbf{X}}\underline{\mathbf{f}}(\boldsymbol{\theta})')=\Lambda^*$. If $\boldsymbol{\theta}_i$ is constructed as $\underline{\lambda}_1'\Sigma^{-1}\underline{\mathbf{X}}+w^{1/2}\mathbf{s}_i$, in which $\underline{\lambda}_1$ is the first column of Λ^*, and $\mathbf{s}_i$ is chosen so that it is statistically independent of $\underline{\mathbf{X}}$ and $\underline{\lambda}_1'\Sigma^{-1}\underline{\mathbf{X}}+w^{1/2}\mathbf{s}_i\sim N(0,1)$, then it will always be the case that

(15.40) $$E\underline{\mathbf{X}}\boldsymbol{\theta}_i=E(\underline{\mathbf{X}}(\underline{\lambda}_1'\Sigma^{-1}\underline{\mathbf{X}}+w^{1/2}\mathbf{s}_i)=\underline{\lambda}_1.$$

However, it must also be the case that

(15.41) $$E\left(\underline{\mathbf{X}}\frac{(\boldsymbol{\theta}_i^{\,2}-1)}{\sqrt{2}}\right)=E\left(\underline{\mathbf{X}}\frac{((\underline{\lambda}_1'\Sigma^{-1}\underline{\mathbf{X}}+w^{1/2}\mathbf{s}_i)^2-1)}{\sqrt{2}}\right)=\underline{\lambda}_2.$$

Now, the left member of this equality, say $\underline{\lambda}_2(\underline{\lambda}_1)$, is equal to

(15.42) $$E\left(\underline{\mathbf{X}}\frac{((\underline{\lambda}_1'\Sigma^{-1}\underline{\mathbf{X}}+w^{1/2}\mathbf{s}_i)^2-1)}{\sqrt{2}}\right)=$$
$$E\left(\underline{\mathbf{X}}\frac{((\underline{\lambda}_1'\Sigma^{-1}\underline{\mathbf{X}}\underline{\mathbf{X}}'\Sigma^{-1}\underline{\lambda}_1+2w^{1/2}\underline{\lambda}_1'\Sigma^{-1}\underline{\mathbf{X}}\mathbf{s}_i+w\mathbf{s}_i^{\,2})-1)}{\sqrt{2}}\right).$$

Because $\mathbf{s}_i$ is chosen to be statistically independent of $\underline{\mathbf{X}}$, the kth element of $\underline{\lambda}_2(\underline{\lambda}_1)$ is equal to

(15.43) $$\frac{(\underline{\lambda}_1'\Sigma^{-1}[E\mathbf{X}_k\underline{\mathbf{X}}\underline{\mathbf{X}}']\Sigma^{-1}\underline{\lambda}_1)}{\sqrt{2}}=\frac{\underline{t}'\Sigma_k\underline{t}}{\sqrt{2}},$$

in which $t=\underline{\lambda}_1'\Sigma^{-1}$ and $\Sigma_k=\begin{pmatrix} EX_kX_1^{\,2} & EX_kX_1X_2 & . & EX_kX_1X_p \\ EX_kX_2X_1 & EX_kX_2^{\,2} & . & EX_kX_2X_p \\ . & . & . & . \\ EX_kX_pX_1 & EX_kX_pX_2 & . & EX_kX_p^{\,2} \end{pmatrix}$

There are p such matrices Σ_k each containing product moments of the third order of the distribution of $\underline{\mathbf{X}}$. Thus, it may be concluded that $\underline{\mathbf{X}}$ is uqf-replaceable, with replacement variate $\boldsymbol{\theta}_i=\underline{\lambda}_1'\Sigma^{-1}\underline{\mathbf{X}}+w^{1/2}\mathbf{s}_i$, only if there exists a p-vector of real coefficients $\underline{\lambda}_1$ such that $\Sigma-\Lambda\Lambda'=\Psi$, in which Ψ is diagonal, and positive definite, and $\Lambda=[\underline{\lambda}_1:\underline{\lambda}_2(\underline{\lambda}_1)]$. Put another way, $\underline{\mathbf{X}}$ is uqf-replaceable only if $\Sigma=\Lambda^*\Lambda^{*\prime}+\Psi$, in which Ψ is diagonal and positive definite, Λ^* is a (p×2) matrix of real coefficients, and there exists an orthonormal matrix T, such that $\Lambda^*T=\Lambda=[\underline{\lambda}_1:\underline{\lambda}_2(\underline{\lambda}_1)]$.

Construction formula

The construction formula is $\boldsymbol{\theta}_i=\underline{\lambda}_1'\Sigma^{-1}\underline{\mathbf{X}}+w^{1/2}\mathbf{s}_i$, in which $\mathbf{s}_i$ is chosen so that it is statistically independent of $\underline{\mathbf{X}}$ and $\boldsymbol{\theta}_i$~N(0,1).

Cardinality of replacement

As is clear from the construction formula, so long as p, the number of input variates, is finite, the set *C* of replacement variates $\boldsymbol{\theta}_i=\underline{\lambda}_1'\Sigma^{-1}\underline{\mathbf{X}}+w^{1/2}\mathbf{s}_i$ has a cardinality of infinity.

Characteristics of C

The latitude inherent to the set *C* of replacement variates $\boldsymbol{\theta}_i$ can be quantified, as per usual, by Guttman's minimum correlation, $\rho^*=2\underline{\lambda}_1'\Sigma^{-1}\underline{\lambda}_1-1$.

Optimality criteria

If $\underline{\mathbf{X}}$ is uqf replaceable, then each replacement variate $\boldsymbol{\theta}_i$ replaces the input variates $\mathbf{X}_j$ in the sense that the 2p regression parameters contained in $\Lambda=[\underline{\lambda}_1:\underline{\lambda}_2(\underline{\lambda}_1)]$ contain all of the information necessary to reproduce the $\frac{1}{2}p(p-1)$ unique covariances contained of Σ. That is,

$$\sigma_{ij}=\lambda_{1i}\lambda_{1j}+\underline{\lambda}_{2i}(\underline{\lambda}_1)\underline{\lambda}_{2j}(\underline{\lambda}_1)\quad i\neq j \tag{15.44}$$

Put another way, $\underline{\mathbf{f}}(\boldsymbol{\theta})$ contains all of the information about the pairwise linear dependencies among the $\mathbf{X}_j$, in the sense that, following conditioning of $\underline{\mathbf{X}}$ on $\underline{\mathbf{f}}(\boldsymbol{\theta})$, no linear dependencies among the $\mathbf{X}_j$ remain.

Testability

If (ri)-(rii) hold, then, from (Cviii), $\Sigma=\Lambda\Lambda'+\Psi$ in which Ψ is diagonal and positive definite, and $\Lambda=[\underline{\lambda}_1:\underline{\lambda}_2(\underline{\lambda}_1)]$. Consider any decomposition $\Sigma=\Lambda^*\Lambda^{*\prime}+\Psi$, Ψ diagonal and positive definite (e.g., one derived from a two-dimensional linear factor analysis of $\underline{\mathbf{X}}$). Then $\Lambda=\Lambda^*T$, in which $TT'=T'T=I_2$. It thus follows that $\underline{\mathbf{X}}=\Lambda^*\underline{\mathbf{w}}+\underline{\boldsymbol{\delta}}=\Lambda T'\underline{\mathbf{w}}+\underline{\boldsymbol{\delta}}=\Lambda\underline{\mathbf{f}}(\boldsymbol{\theta})+\underline{\mathbf{l}}$, in which $C(\underline{\mathbf{w}})=I_2$, $C(\underline{\mathbf{w}},\underline{\boldsymbol{\delta}}')=0$, $C(\underline{\boldsymbol{\delta}})=\Psi=C(\underline{\mathbf{l}})$, and $\underline{\mathbf{f}}(\boldsymbol{\theta})$ is an orthonormal tranformation of $\underline{\mathbf{w}}$. Now,

$$\underline{\mathbf{b}}=(\Lambda^{*\prime}\Psi^{-1}\Lambda^*)^{-1}\Lambda^{*\prime}\Psi^{-1}\underline{\mathbf{X}}=(\Lambda^{*\prime}\Psi^{-1}\Lambda^*)^{-1}\Lambda^{*\prime}\Psi^{-1}\Lambda^*\underline{\mathbf{w}}+(\Lambda^{*\prime}\Psi^{-1}\Lambda^*)^{-1}\Lambda^{*\prime}\Psi^{-1}\underline{\boldsymbol{\delta}}=T\underline{\mathbf{f}}(\boldsymbol{\theta})+\underline{\boldsymbol{\delta}}^*, \tag{15.45}$$

with the variates contained in vector $\underline{\mathbf{b}}$ recognizable as Bartlett factor score predictors. Thus, the vector of Bartlett predictors is a perturbed orthornormal transformation of $\underline{\mathbf{f}}(\boldsymbol{\theta})$. The covariance matrix of these variates is

$$C(\underline{\mathbf{b}})=C(\underline{\mathbf{f}}(\boldsymbol{\theta}))+C(\underline{\boldsymbol{\delta}}^*)=I+(\Lambda^{*\prime}\Psi^{-1}\Lambda^*)^{-1}. \tag{15.46}$$

It follows, then, that if $\underline{\mathbf{X}}$ is uqf-replaceable, so long as the diagonals of the matrix $C(\underline{\boldsymbol{\delta}}^*)$ are "small", the joint distribution of the Bartlett predictors will be roughly the same as the joint distribution of $T\underline{\mathbf{f}}(\boldsymbol{\theta})$. Because $E(\frac{(\boldsymbol{\theta}^2-1)}{\sqrt{2}}|\boldsymbol{\theta}=\theta^*)=\frac{(\theta^{*2}-1)}{\sqrt{2}}$, the distribution of $\underline{\mathbf{f}}(\boldsymbol{\theta})$ will have a curvilinear shape. This fact has been used by McDonald (1962, 1967) as a means of distinguishing between the two cases of 2-dimensional linear factor, and uqf, replaceability of a given $\underline{\mathbf{X}}$: In the linear case, the joint distribution of the two replacement variates $(\boldsymbol{\theta}_1,\boldsymbol{\theta}_2)$ will not have a curvilinear shape. Hence, given that the hypothesis H_o: $[\Sigma=\Lambda^*\Lambda^{*\prime}+\Psi$, Ψ diagonal and positive definite] has been retained, a low-cost second step is to produce a bivariate plot of Bartlett factor score predictions. The lack of any noticeable curvilinearity is evidence against the hypothesis of uqf-replaceability. McDonald (1962, 1967) has provided techniques for the testing of hypotheses of nonlinear (quadratic and higher-order) replaceability which involve the minimization of a loss function in terms of T, and involves estimates of the higher order joint moments of $\underline{\mathbf{b}}$ and $\underline{\boldsymbol{\delta}}^*$, from which estimates of the higher-order joint moments of $\underline{\mathbf{w}}$ are deducible.

From (Cviii), the implication $[(ri)\text{-}(rii)]\rightarrow[\Sigma=\Lambda_2\Lambda_2'+\Psi$, Ψ diagonal, positive definite] is thus true, and, hence, so is the implication $\sim[\Sigma=\Lambda_2\Lambda_2'+\Psi$, Ψ diagonal, positive definite]$\rightarrow\sim[(ri)\text{-}(rii)]$. Hence, if Σ cannot be decomposed as $\Lambda_2\Lambda_2'+\Psi$, Ψ diagonal, positive definite, then [(ri)-(rii)] cannot obtain ($\underline{\mathbf{X}}$ is not uqf replaceable). But $[\Sigma=\Lambda_2\Lambda_2'+\Psi$, Ψ diagonal, positive definite] is not sufficient for uqf-replaceability. As was shown in the section on *existence*, Λ_2 must have the special form $[\underline{\lambda}_1:\underline{\lambda}_2(\underline{\lambda}_1)]$. Hence, the following sequential procedure could be employed:

i. Test if there exists Λ^* such that $\Sigma-\Lambda^*\Lambda^{*\prime}=\Psi$, Ψ diagonal, positive definite.

ii. If decision (i) is in the affirmative, estimate Σ_k, k=1..p, and Σ^{-1}.

iii. For arbitrary Λ^* satisfying (i), test to see if there exists a rotation T, T'T=TT'=I, such that $\Lambda^*T=[\underline{\lambda}_1:\underline{\lambda}_2(\underline{\lambda}_1)]$. If so, then make decision that $\underline{\mathbf{X}}$ is uqf-replaceable.

On the other hand, uqf-replaceability could, in principle, be tested as a covariance structure hypothesis: H_o: $[\Sigma=\underline{\lambda}_1\underline{\lambda}_1'+\underline{\lambda}_2(\underline{\lambda}_1)\underline{\lambda}_2(\underline{\lambda}_1)'+\Psi$, Ψ diagonal, positive definite].

This hypothesis can be unpacked as

H_o: $[\Sigma=\underline{\lambda}_1\underline{\lambda}_1'+\frac{1}{2}(I\otimes\underline{\lambda}_1')\Sigma_3\underline{\lambda}_1\underline{\lambda}_1'\Sigma_3(\underline{\lambda}_1\otimes I)+\Psi$, Ψ diagonal and positive definite], in which Σ_3 is the $(p^2 \times p)$ matrix $\begin{pmatrix} \Sigma^{-1}\Sigma_1\Sigma^{-1} \\ \Sigma^{-1}\Sigma_2\Sigma^{-1} \\ . \\ \Sigma^{-1}\Sigma_p\Sigma^{-1} \end{pmatrix}$. The vector $\underline{\mathbf{X}}$ cannot, of course, be reasonably asserted to be multivariate normal, which rules out employment of the usual likelihood-ratio approaches.

7. *Linear discriminant (ld) replacement generator*

Let $\underline{\mathbf{X}}$ contain a set of p continuous input variates $\mathbf{X}_j$, j=1..p, jointly distributed in a population *P* under study. The scores that comprise the distributions of these variates are produced in accord with rules $\{r_1,r_2,...,r_p\}$ of score production. Let *P* be divisible into two sub-populations, P_1 and P_2, and let there exist, prior to analysis, a rule, r_c, for the identification of those members of *P* that belong to each of P_1 and P_2. The rule will, in general, take the form:

> r_c: A member of population *P*, p_i, is a member of P_1, i.e., $p_i \in P_1$, if and only if it has characterics$\{c_1,c_2,...c_t\}$;
> else, $p_i \in P_2$.

Application of r_c to the members of *P* will yield a Bernouilli variate $\mathbf{Y}$, representing sub-population membership, and taking on values of, say, 1 (representing membership in P_1) with probability π_1, and 2 (representing membership in P_2) with probability $\pi_2=(1-\pi_1)$. The joint moments of $\underline{\mathbf{X}}$ conditional on specific values of $\mathbf{Y}$ can be considered and, in particular:

(15.47) $E(\underline{\mathbf{X}}|\mathbf{Y}=1)=\underline{\mu}_1$, $E(\underline{\mathbf{X}}|\mathbf{Y}=2)=\underline{\mu}_2$, and $C(\underline{\mathbf{X}}|\mathbf{Y}=1)=C(\underline{\mathbf{X}}|\mathbf{Y}=2)=\Sigma_W$.

The latter is a testable requirement for this particular version of the ld replacement variate generator. It follows, then, that

(15.48) $E(\underline{\mathbf{X}})=\pi_1\underline{\mu}_1+\pi_2\underline{\mu}_2$

and

(15.49) $C(\underline{\mathbf{X}}) = \Sigma = E(C(\underline{\mathbf{X}} \mid \mathbf{Y})) + C(E(\underline{\mathbf{X}} \mid \mathbf{Y})) = \Sigma_W + \pi_1 \pi_2 * (\underline{\mu}_1 - \underline{\mu}_2)(\underline{\mu}_1 - \underline{\mu}_2)'$

Linear discriminant (ld) replacement variates

A replacement variate **c** is sought such that

ri) $\mathbf{c}=\underline{t}'\underline{\mathbf{X}}$

and

rii) $(E(\mathbf{c}|P_1)-E(\mathbf{c}|P_2))^2$ is a maximum over all **c** for which $V(\mathbf{c}|P_1)=V(\mathbf{c}|P_2)=\underline{t}'\Sigma_W\underline{t}=1$.

Existence

Note that

(15.50)
$$(E(\mathbf{c}|P_1)-E(\mathbf{c}|P_2))^2=(E(\underline{t}'\underline{\mathbf{X}}|P_1)-E(\underline{t}'\underline{\mathbf{X}}|P_2))^2=\underline{t}'(\underline{\mu}_1-\underline{\mu}_2)(\underline{\mu}_1-\underline{\mu}_2)'\underline{t}=\underline{t}'\Sigma_B\underline{t}.$$

Hence, a p-vector $\underline{t}$ must be found such that $\underline{t}'\Sigma_B\underline{t}$ is a maximum over all $\underline{t}$ for which $\underline{t}'\Sigma_W\underline{t}=a$, in which a is a positive constant. Because the choice of a does not affect the construction of **c**, a will be chosen to be unity. Letting $\varphi=\underline{t}'\Sigma_B\underline{t}-\gamma(\underline{t}'\Sigma_W\underline{t}-1)$,

(15.51)
$$\frac{\partial\varphi}{\partial\underline{t}}\underline{t}'\Sigma_B\underline{t}-\gamma(\underline{t}'\Sigma_W\underline{t}-1)=2\underline{t}'\Sigma_B-2\gamma\underline{t}'\Sigma_W$$

(15.52)
$$\frac{\partial\varphi}{\partial\gamma}\underline{t}'\Sigma_B\underline{t}-\gamma(\underline{t}'\Sigma_W\underline{t}-1)=\underline{t}'\Sigma_W\underline{t}-1$$

Hence, $\underline{t}$ is the first eigenvector $\underline{w}_1$ of $\Sigma_W^{-1}\Sigma_B$, and the maximum of $(E(\mathbf{c}|P_1)-E(\mathbf{c}|P_2))^2$ is equal to γ_1, the first eigenvalue of $\Sigma_W^{-1}\Sigma_B$. It can, therefore, be concluded that $\underline{\mathbf{X}}$ is ld-replaceable if: i) $C(\underline{\mathbf{X}}|P_1)=C(\underline{\mathbf{X}}|P_2)=\Sigma_W$; ii) Σ_W is nonsingular.

Consequences

The following are consequences of the ld replacement.

Ci) $E(\mathbf{c})=E(\underline{w}_1'\underline{\mathbf{X}})=\pi_1\underline{w}_1'\underline{\mu}_1+\pi_2\underline{w}_1'\underline{\mu}_2$ (from (15.51))

Cii) $V(\mathbf{c})=\underline{w}_1'\Sigma\underline{w}_1=\underline{w}_1'\Sigma_W\underline{w}_1+\pi_1\pi_2\underline{w}_1'(\underline{\mu}_1-\underline{\mu}_2)(\underline{\mu}_1-\underline{\mu}_2)'\underline{w}_1=1+\pi_1\pi_2\gamma_1$ (from (15.52))

$\underline{\text{Ciii}}$) $E(\mathbf{c}\underline{\mathbf{X}}')=E(\underline{w}_1'\underline{\mathbf{X}}\underline{\mathbf{X}}')=\underline{w}_1'\Sigma=\underline{w}_1'\Sigma_W+\pi_1\pi_2\underline{w}_1'(\underline{\mu}_1-\underline{\mu}_2)(\underline{\mu}_1-\underline{\mu}_2)'$ (from (15.52) and (7.54))

$$=(\frac{1}{\gamma_1}+\pi_1\pi_2)\underline{w}_1'\Sigma_B$$

$\underline{\text{Civ}}$) $E(\underline{\mathbf{X}}|\mathbf{c}=c_o)=\Sigma\underline{w}_1(\underline{w}_1'\Sigma\underline{w}_1)^{-1}c_o$

$\underline{\text{Cv}}$) $\Sigma=C(E(\underline{\mathbf{X}}|\mathbf{c}))+E(C(\underline{\mathbf{X}}|\mathbf{c}))=\frac{1}{\gamma_1}\Sigma\underline{w}_1\underline{w}_1'\Sigma+\Sigma^{1/2}(I-\frac{1}{\gamma_1}\Sigma^{1/2}\underline{w}_1\underline{w}_1'\Sigma^{1/2})\Sigma^{1/2}$

Cardinality of replacement

If $\underline{\mathbf{X}}$ is ld-replaceable, i.e., $C(\underline{\mathbf{X}}|P_1)=C(\underline{\mathbf{X}}|P_2)=\Sigma_W$ and Σ_W is nonsingular, then the cardinality of set C that contains the replacement variates **c** is unity.

Construction formula

The construction formula is $\mathbf{c}=\underline{w}_1'\underline{\mathbf{X}}$, in which $\underline{w}_1$ is the first eigenvector of $\Sigma_W^{-1}\Sigma_B$.

Optimality criteria

The replacement variate **c** is that variate that contains the most information about normalized mean differences between sub-populations P_1 and P_2. In particular, of all one-dimensional projections of $\underline{\mathbf{X}}$, the projection of $\underline{\mathbf{X}}$ onto $\underline{w}_1$ (the projection that produces **c**) yields the greatest separation of the two p-dimensional point clusters in R^p defined by P_1 and P_2. Hence, if one were forced to give up the p input variates $\mathbf{X}_j$ as a basis for distinguishing quantitatively between the sub-populations P_1 and P_2, replacing these variates with a single new variate that is a linear combination of them, then the variate to be chosen would be **c** as defined above.

Characteristics of C

The cardinality of C is unity and the single variate it contains has properties

(15.53) $E(\mathbf{c}\underline{\mathbf{z}}')=E(\underline{t}'\underline{\mathbf{X}}\underline{\mathbf{z}}')=\underline{w}_1'\Sigma_{Xz}$,

and

(15.54) $\rho(\mathbf{c},\underline{\mathbf{z}}')=\frac{1}{\underline{w}_1'\Sigma\underline{w}_1}\underline{w}_1'\Sigma_{Xz}D_z^{-1/2}=\frac{1}{(1+\pi_1\pi_2\gamma_1)}\underline{w}_1'\Sigma_{Xz}D_z^{-1/2}$, in which $\underline{\mathbf{z}}$ is an external set of variates and $D_z^{-1/2}$ is the diagonal matrix containing the reciprocals of the standard deviations of the $\mathbf{z}_j$.

Testability

The ld-replacement always exists and, hence, in its standard usage there is no need for an hypothesis test. One could, of course, invent testable hypotheses within the ld framework. One example is the hypothesis that the separation between the two populations P_1 and P_2 is greater than some hypothesized value.

8. *Latent profile (lp) replacement generator*

Let $\underline{\mathbf{X}}$ contain a set of p continuous input variates $\mathbf{X}_j$, j=1..p, jointly distributed in a population *P* under study. The scores that comprise the distributions of these variates are produced in accord with rules $\{r_1,r_2,...,r_p\}$ of score production. As with the ld replacement generator, latent class analysis is concerned with two (or more) subpopulations of *P*. In an application of the the ld generator, the analyst does possess a rule for the identification of those members of *P* that belong to each of subpopulations P_1 and P_2, respectively. Application of this rule to the members of *P* generates the Bernoulli random variate **Y**, and allows for the calculation of moments of $\underline{\mathbf{X}}$ conditional on particular values of **Y**. A projection of $\underline{\mathbf{X}}$, i.e., a single replacement variate, is sought that yields the greatest separation of the subpopulations. In an application of the latent profile generator, on the other hand, the analyst does *not* possess, prior to analysis, a rule by which each member of population *P* can be assigned to one of the subpopulations referred to in the generator's equations. That is, these sub-populations do not exist prior to analysis. The aim is to *create* a rule r_{lc} whose application to the members of *P* will result in a discrete random variate $\boldsymbol{\theta}$ whose values represent sub-population membership, and with $\boldsymbol{\theta}$ replacing the input variates $\mathbf{X}_j$, j=1..p, in certain optimal ways. As will be seen, the standard formulation of the lp replacement is essentially equivalent to that of the ulcf replacement, the sole difference being that $\boldsymbol{\theta}$ must be discrete, rather than continuous. The number of values that $\boldsymbol{\theta}$ is allowed to assume is a defining feature of any lp replacement. The current treatment considers only the special case in which $\boldsymbol{\theta}$ must have a (two-valued) Bernoulli distribution.

Perhaps more than any other latent variable technology, the latent class and profile "models" have been infected with the metaphysics of the Central Account. They have frequently been portrayed as detectors of existing "latent classes", "natural kinds", or "true types" (see, e.g., Meehl, 1965, 1973, 1992; Meehl & Golden, 1982; Waller & Meehl, 1998). Correlatively, tests of the conformity of $\underline{\mathbf{X}}$s to latent class and profile models have routinely been portrayed as tests of whether there exist "latent classes" or "natural kinds" which "underlie the $\mathbf{X}_j$". Meehl, for example, developed his *taxometric* latent variable technology, which overlaps with latent class and profile analysis, in the hopes that it could be used by applied researchers to detect "latent taxa" (i.e., discrete *types* which are said to underlie, perhaps causally, responding to a set of indicator variables) when, in fact, they do exist.

(2-valued) latent profile (lp) replacement variates

A replacement variate $\boldsymbol{\theta}$ is sought such that

ri) $\boldsymbol{\theta}$~*Bernoulli*(π_1), taking on values θ_1 and θ_2 with probabilities π_1 and π_2=(1-π_1), respectively,

rii) E($\boldsymbol{\theta}$)=0 and E($\boldsymbol{\theta}^2$)=1,

and

riii) The vector of residuals $\underline{\mathbf{l}}$ of the linear regressions of the input variates $\mathbf{X}_j$, j=1..p, on $\boldsymbol{\theta}$ has a covariance matrix Ψ that is diagonal and positive definite.

Consequences

The following are consequences of the requirements of replacement specified under the lp generator ((ri)-(riii)).

Ci) From (ri) and (rii), E($\boldsymbol{\theta}$)= $\pi_1\theta_1+\pi_2\theta_2$=0 and E($\boldsymbol{\theta}^2$)=$\pi_1\theta_1^2+\pi_2\theta_2^2$=1. Hence,

$$\theta_1=\frac{-(1-\pi_1)\theta_2}{\pi_1} \text{ and } \frac{\pi_1(1-\pi_1)^2\theta_2^{\ 2}}{\pi_1^{\ 2}}+\pi_2\theta_2^{\ 2}=1,$$

from which it follows that

$$\theta_2^{\ 2}=\left(\frac{\pi_1(1-\pi_1)^2+\pi_2\pi_1^{\ 2}}{\pi_1^{\ 2}}\right)^{-1}=\frac{\pi_1}{\pi_2}.$$

Thus, it can be concluded that the two values assumed by $\boldsymbol{\theta}$ must be $\theta_2=\sqrt{\frac{\pi_1}{\pi_2}}$ and $\theta_1=-\sqrt{\frac{\pi_2}{\pi_1}}$.

Cii) The conditional expectations of $\underline{\mathbf{X}}$ given each of the two values of $\boldsymbol{\theta}$ are

$$E(\underline{\mathbf{X}}|\boldsymbol{\theta}=-\sqrt{\frac{\pi_2}{\pi_1}})=\underline{\mu}_1 \text{ and } E(\underline{\mathbf{X}}|\boldsymbol{\theta}=\sqrt{\frac{\pi_1}{\pi_2}})=\underline{\mu}_2,$$

and, hence, the two conditional residuals are

$(\underline{l}|\boldsymbol{\theta}=-\sqrt{\frac{\pi_2}{\pi_1}})=(\underline{\mathbf{X}}-\underline{\mu}_1|\boldsymbol{\theta}=-\sqrt{\frac{\pi_2}{\pi_1}})$ and $(\underline{l}|\boldsymbol{\theta}=\sqrt{\frac{\pi_1}{\pi_2}})=(\underline{\mathbf{X}}-\underline{\mu}_2|\boldsymbol{\theta}=\sqrt{\frac{\pi_1}{\pi_2}})$

Ciii) Let $E((\underline{\mathbf{X}}-\underline{\mu}_1)(\underline{\mathbf{X}}-\underline{\mu}_1)'|\boldsymbol{\theta}=-\sqrt{\frac{\pi_2}{\pi_1}})=\Psi_1$ and $E((\underline{\mathbf{X}}-\underline{\mu}_2)(\underline{\mathbf{X}}-\underline{\mu}_2)'|\boldsymbol{\theta}=\sqrt{\frac{\pi_1}{\pi_2}})=\Psi_2$ be the conditional covariance matrices of $\underline{\mathbf{X}}$ at each of the two values assumed by $\boldsymbol{\theta}$. Because, from (riii),

$$C(\underline{l})=\pi_1(E(\underline{\mathbf{X}}-\underline{\mu}_1)(\underline{\mathbf{X}}-\underline{\mu}_1)'|\boldsymbol{\theta}=-\sqrt{\frac{\pi_2}{\pi_1}})+\pi_2(E(\underline{\mathbf{X}}-\underline{\mu}_2)(\underline{\mathbf{X}}-\underline{\mu}_2)'|\boldsymbol{\theta}=\sqrt{\frac{\pi_1}{\pi_2}})=\pi_1\Psi_1+\pi_2\Psi_2,$$

is diagonal and positive definite, both of the conditional covariance matrices Ψ_1 and Ψ_2 must also be diagonal and positive definite.

Civ) Because, from (Ciii), $C(\underline{l})=E(C(\underline{\mathbf{X}}|\boldsymbol{\theta}))=\pi_1\Psi_1+\pi_2\Psi_2$, it follows that

$$\Sigma=E(C(\underline{\mathbf{X}}|\boldsymbol{\theta}))+C(E(\underline{\mathbf{X}}|\boldsymbol{\theta}))=[\pi_1\Psi_1+\pi_2\Psi_2]+\pi_1\pi_2(\underline{\mu}_1-\underline{\mu}_2)(\underline{\mu}_1-\underline{\mu}_2)'$$

Cv) $\sigma_{ij}=\pi_1\pi_2(\mu_{1i}-\mu_{2i})(\mu_{1j}-\mu_{2j})$ (from (Civ) and (riii))

Cvi) From (Civ), the covariance matrix of the input variates has the ulcf decomposition

$$\Sigma=\underline{\Lambda\Lambda}'+\Phi,$$

in which $\underline{\Lambda}=\sqrt{\pi_1\pi_2}\,(\underline{\mu}_1-\underline{\mu}_2)$, and $\Phi=\pi_1\Psi_1+\pi_2\Psi_2$ is a diagonal, positive definite matrix (McDonald, 1967; Bartholomew & Knott, 1999; Molenaar & von Eye, 1994).

Cvii) From (Cvi), it follows that $\underline{\mathbf{X}}$ is representable as $\underline{\mathbf{X}}=\underline{\Lambda}\boldsymbol{\theta}+\underline{\mathbf{e}}$, in which $E(\boldsymbol{\theta})=0$, $E(\boldsymbol{\theta}^2)=1$, $E(\underline{\mathbf{X}}|\boldsymbol{\theta}=t)=\underline{\Lambda}t$, $E(\underline{\mathbf{e}})=\underline{0}$, and $C(\underline{\mathbf{e}})=\Phi$. The variate $\boldsymbol{\theta}$ has, of course, a Bernoulli distribution, and, from (Ci), assumes the values $\theta_2=\sqrt{\frac{\pi_1}{\pi_2}}$, and $\theta_1=-\sqrt{\frac{\pi_2}{\pi_1}}$ with probabilities π_2 and π_1, respectively. It follows from (Cii) and the fact that $E(\underline{\mathbf{X}}|\boldsymbol{\theta}=t)=\underline{\Lambda}t$, that $\underline{\mu}_1=-\underline{\Lambda}\sqrt{\frac{\pi_2}{\pi_1}}$ and $\underline{\mu}_2=\underline{\Lambda}\sqrt{\frac{\pi_1}{\pi_2}}$. The matrix $\Sigma-\Phi$ is of rank one, and, hence, can be represented as $\Sigma-\Phi=\underline{mm}'\omega=\underline{\Lambda\Lambda}'$, in which $\underline{m}'\underline{m}=1$, $\underline{m}$ and ω are, respectively, the single eigenvector and eigenvalue of $\Sigma-\Phi$, and in which $\underline{\Lambda}=\sqrt{\omega}\underline{m}$. Hence,

$$\mathbf{v} = \frac{1}{\sqrt{\omega}}\underline{m}'(\underline{\mathbf{X}}-\underline{\mu}) = \frac{1}{\sqrt{\omega}}\underline{m}'\underline{\Lambda}\boldsymbol{\theta}+\frac{1}{\sqrt{\omega}}\underline{m}'\underline{\mathbf{e}} = \frac{1}{\omega}\underline{\Lambda}'\underline{\Lambda}\boldsymbol{\theta}+\frac{1}{\omega}\underline{\Lambda}'\underline{\mathbf{e}} = \mathbf{w}+\mathbf{d},$$

in which $\mathbf{v}$ is called by McDonald (1967, pp.31-32) a component variate. The variate $\mathbf{w}$ has the same distribution as $\boldsymbol{\theta}$, but has variance $\frac{1}{\omega^2}(\underline{\Lambda}'\underline{\Lambda})^2$. If σ^2_d is small, $\mathbf{w}$ will have roughly the same distribution as $\mathbf{v}$, and, hence, $\mathbf{v}$ will have roughly the same distribution as $\boldsymbol{\theta}$.

Existence

Consider the usual case, in which (riii) is further refined to insist that $(\underline{\mathbf{X}}|\boldsymbol{\theta}=i)\sim N_p(\underline{\mu}_i,\Psi_i)$ (call this (riiiN)). In this case, the lp replacement is, save for the requirement that $\boldsymbol{\theta}$ has a bernoulli distribution, identical to the linear factor replacement (2.14)-(2.15). It might, then, be thought that because of the close ties between the lp and ulcf generators, and, in particular, result (Cvi), that the bernoulli variate $\boldsymbol{\theta}$ satisfying (ri)-(riiiN) can be constructed by dichotomizing, in some particular way, a ulcf replacement variate $\boldsymbol{\theta}_i=\underline{\Lambda}_i'\Sigma_i^{-1}\underline{\mathbf{X}}_i+w_i^{1/2}\mathbf{s}_i$. The idea would be that, since $C(\underline{\mathbf{X}}|\boldsymbol{\theta}_i=\theta_o)=\Psi$, Ψ diagonal and positive definite, it might be possible to find a real number, τ, such that

if $\boldsymbol{\theta}_i<\tau$ then $\boldsymbol{\theta}=0$;

if $\boldsymbol{\theta}_i\geq\tau$ then $\boldsymbol{\theta}=1$;

$C(\underline{\mathbf{X}}|\boldsymbol{\theta}=0)=\Psi_o$, Ψ_o diagonal and positive definite;

$C(\underline{\mathbf{X}}|\boldsymbol{\theta}=1)=\Psi_1$, Ψ_1 diagonal and positive definite.

Unfortunately, this cannot work. It can be shown that, for any real number τ^*, $C(\underline{\mathbf{X}}|\boldsymbol{\theta}_i<\tau^*)=a\underline{\Lambda}\underline{\Lambda}'+\Psi$, and $C(\underline{\mathbf{X}}|\boldsymbol{\theta}_i\geq\tau^*)=b\underline{\Lambda}\underline{\Lambda}'+\Psi$, in which $a=V(\boldsymbol{\theta}_i|\boldsymbol{\theta}_i<\tau^*)$, and $b=V(\boldsymbol{\theta}_i|\boldsymbol{\theta}_i\geq\tau^*)$. Clearly, these conditional covariance matrices are not diagonal. Hence, no real number τ^* can be found such that dichotomization of a ulcf replacement variate $\boldsymbol{\theta}_i$ by τ^* will yield a bernoulli variate with properties (ri)-(riiiN).

In fact, it is unknown, at present, the conditions under which a random variate satisfying (ri)-(riii) (or (ri)-(riiiN)) is constructable, if, in fact, it is at all constructable. This state of affairs may seem puzzling, given that applied researchers have, for some time, been carrying out latent class and profile analyses. If it is unclear as to the conditions under which a dichotomous variate with the properties required by the lp replacement exists, then what is taking place in these analyses? A rough sketch of practice is as follows:

i. Given (ri)-(riiiN), the unconditional distribution of $\underline{\mathbf{X}}$ is mixture normal. That is,

$f_{\underline{\mathbf{X}}}=\pi_1 f_{\underline{\mathbf{X}}|1}+\pi_2 f_{\underline{\mathbf{X}}|2}$, in which $f_{\underline{\mathbf{X}}|i}=\dfrac{|\Psi_i|^{-\frac{1}{2}}}{(2\pi)^{\frac{p}{2}}}\exp(-\frac{1}{2}(\underline{\mathbf{X}}-\underline{\mu}_i)'\Psi_i(\underline{\mathbf{X}}-\underline{\mu}_i))$.

For a sample of size n drawn from a given population under study, the log-likelihood under the null hypothesis, H_o:$f_{\underline{\mathbf{X}}}=\pi_1 f_{\underline{\mathbf{X}}|1}+\pi_2 f_{\underline{\mathbf{X}}|2}$ is then

$$\ln(L_{Ho}) = \ln(\prod_{j=1}^{n} f_{\underline{X}_j}) = \sum_{j=1}^{n}\ln(\pi_1 f_{\underline{X}_j|1}+\pi_2 f_{\underline{X}_j|2}).$$

The log-likelihood under the alternative hypothesis that the sample was drawn from a general multivariate normal may be symbolized as $\ln(L_{H1})$. Using an approximation from Wolfe(1971), when H_o is true, $2\ln(L_{H1})-2\ln(L_{Ho})$ has an approximate χ^2_d distribution, with $d=3p^2+9p$.

ii. If H_o is accepted, then a rule r_{lp} is created to "allocate sample members to classes", and this rule is standardly based on the posterior distribution $f_{\theta|\underline{\mathbf{X}}}$.

iii. Following allocation, the classes are named, this being the latent class/profile equivalent of latent variate interpretation. As Bartholomew and Knott (1999, p.142) put it, "Having fitted the model we will normally wish to consider what can be learned from the classification at which we have arrived. In the case of Macready and Dayton's mastery model we would expect to find two classes, with one having the characteristics of 'masters' and the other of 'non-masters'...For many response patterns, for example 1111, 1011 and 0110, we would be extremely confident in allocating individuals to the 'master' category. Similarly, we would have little hesitation in classifying someone responding 0000 as a non-master, and scarcely more if it was 1000."

The question is, *what* exactly does the researcher who employs this strategy have the right to claim. Not surprisingly, latent class and profile analyses are seen through the lenses of the Central Account. Thus, when there is evidence in favour of H_o, this evidence is taken as support for the claim that two distinct types of object have been detected in population *P*. The task is then viewed as being to correctly infer what types these, in fact, are. But this portrayal is mythology. As argued in Part 2, to construct a tool of detection the researcher would need to antecedently possess a concept whose rule of employment warranted ascription of the concept to the members of some particular sub-class, C, of population *P*. Population *P* could, then, rightly be said to be comprised of those who are in C, and those who are in $\overline{C}$. The researcher would then have a type of individual, those in sub-class C, whose detection was of interest. If it could then be shown that the distributions of a particular set of variates $\underline{\mathbf{X}}$ in each of C and $\overline{C}$, were as described by (ri)-(riiiN), then (ri)-(riiiN) could, indeed, be employed in the development of a tool for the detection of those in C (and $\overline{C}$). But in this case, the researcher would no longer be conducting a latent profile/class analysis, but, rather, a discriminant analysis of some sort.

The quote from Bartholomew and Knott gives the wrong impression that they were antecedently in possession of a concept *master*, and, hence, that they employed a latent class generator as a detector, the data providing evidence that the population under study was *truly* comprised of masters and non-masters. The antecedent possession of a concept *master* (the possession of a rule that fixed the ascription of this concept to individuals), would have been required for there to have been a sense to Barthlomew and Knott's expression of "extreme confidence in allocating individuals." For if one cannot know which individuals are and are not masters, if there does not exist a clear sense as to what it *is* to be a master, then there can be no sense to the notion that one is *confident* that one has made a correct decision in calling a given individual a master. Similarly, unless one has studied masters and non-masters, one could not come to know their characteristics, and, in particular, the response patterns that they were likely to yield. However, to study masters presupposes the capacity to identify such individuals, which, in turn, presupposes the capacity to reliably ascribe the concept *master*. Once again, if Bartholomew and Knott had possessed a technical sense of *master* (had possessed a rule that fixed the grounds of ascription of a technical sense of the term to individuals) there would have been no need for the employment of a statistical procedure whose foundations are a mythology. Their talk about what one should expect from masters arises from the shadow cast by various ordinary language senses of the term *master*, senses that any competent language user has mastered (senses, unfortunately, that Bartholomew and Knott did not bother to tie into the formulation of the technique that they employed).

In fact, the modus operandi of Bartholomew and Knott's analysis in particular, and latent profile and class analysis in general, is not detection, but rather construction. The aim inherent to the employment of these generators is the construction of a classification rule that yields a random variate that satisfies (ri)-(riiiN), the very requirements insisted upon by the lp replacement. Now, the truth of H_o does imply that the unconditional density of $\underline{\mathbf{X}}$ can be factored as $\pi_1 f_{\underline{\mathbf{X}}|1} + \pi_2 f_{\underline{\mathbf{X}}|2}$. But whereas the truth of the linear factor analytic representation $\Sigma=\underline{\Lambda\Lambda}+\Psi$ implies the constructability of variates that satisfy the requirements for factor variate-hood, it is unclear whether the constructability of a variate satisfying (ri)-(riiiN) follows from the truth of the factorization $\pi_1 f_{\underline{\mathbf{X}}|1} + \pi_2 f_{\underline{\mathbf{X}}|2}$. That is, allocation rule r_{lp} does induce a random variate with a bernoulli distribution, but it is unclear whether this induced random variate satisfies (rii)-(riii) (or (rii)-(riiiN)), the very requirements the lp generator stipulates that it must satisfy.

9. *Item response (ir) replacement generators*

Let $\underline{\mathbf{X}}$ contain a set of p dichotomous input variates $\mathbf{X}_j$, j=1..p, jointly distributed in a population *P* under study. The scores that comprise the distributions of these variates are produced in accord with some set of rules $\{r_1,r_2,...,r_p\}$.

ir replacement variates

While the family of ir generators is a large one, the basic requirements imposed by ir generators are roughly as follows:

A replacement variate $\boldsymbol{\theta}$ is sought such that

ri) $\boldsymbol{\theta} \sim f_\theta$, in which f_θ is a continuous density,

rii) $E(\boldsymbol{\theta})=t$ and $E(\boldsymbol{\theta}^2)=s$, $s>0$,

and

riii) Conditional on any value θ_0 of $\boldsymbol{\theta}$, the input variates, $\mathbf{X}_j$, $j=1..p$, are statistically independent. That is,

$$P(\underline{\mathbf{X}}=\underline{x}|\boldsymbol{\theta}=\theta_o) = \prod_{j=1}^{p} P(\mathbf{X}_j = x_j \,|\, \boldsymbol{\theta} = \theta_o)$$

$$=\prod_{j=1}^{p} P(\mathbf{X}_j = 1 \,|\, \boldsymbol{\theta} = \theta_o)^{x_j} (1 - P(\mathbf{X}_j = 1 \,|\, \boldsymbol{\theta} = \theta_o))^{1-x_j}$$

Clearly, this recipe is not well enough specified to allow for a solution. The current treatment will focus on that version of the ir replacement in which f_θ is left unspecified, (rii) is further restricted so that $E(\boldsymbol{\theta})=0$ and $E(\boldsymbol{\theta}^2)=1$, and (riii) is further elaborated by requiring that $P(\mathbf{X}_j=1|\boldsymbol{\theta}=\theta)=\boldsymbol{\Phi}(a_j(\boldsymbol{\theta}-b_j))$, in which $\Phi(\cdot)$ is the cumulative probability function of a standard normal variate. This particular ir generator will be called the "two-parameter normal ogive" (2pno) generator. A number of special cases will also be considered.

Existence

From (2.29) and (2.30), the generator equations are as follows:

(15.55) $$P(\underline{\mathbf{X}}=\underline{x}) = \int_{-\infty}^{\infty} \prod_{j=1}^{p} \Phi(a_j(\theta - b_j))^{x_j} (1 - \Phi(a_j(\theta - b_j)))^{1-x_j} dF(\theta)$$

Reparametrize as

(15.56) $$\lambda_j = \frac{a_j}{\sqrt{1+a_j^2}},\ \sigma_j^2 = \frac{1}{1+a_j^2},\ \text{and}\ \gamma_j = \frac{a_j b_j}{\sqrt{1+a_j^2}},$$

so that

(15.57) $$\Phi(a_j(\theta_0-b_j))=\int_{\gamma_j}^{\infty}\frac{1}{(2\pi\sigma_j^2)^{1/2}}\exp\left(-\frac{(Y_j-\lambda_j\theta_o)^2}{2\sigma_j^2}\right)dY_j=\int_{\gamma_j}^{\infty}\varphi\left(\lambda_j\theta_o,\sigma_j^2\right)dY_j$$

in which $\varphi(\mu,\sigma^2)$ is the normal density function with mean μ and variance σ^2. It follows then that the generator equations can be restated as

(15.58)

$$P(\underline{\mathbf{X}}=\underline{x}) = \int_{-\infty}^{\infty}\prod_{j=1}^{p}\left(\int_{\gamma_j}^{\infty}\varphi\left(\lambda_j\theta,(1-\lambda_j^2)\right)dY_j\right)^{x_j}\left(1-\int_{\gamma_j}^{\infty}\varphi\left(\lambda_j\theta,(1-\lambda_j^2)\right)dY_j\right)^{1-x_j} dF(\theta)$$

$$=\int_{-\infty}^{\infty}\int_{\underline{\gamma}} f_{\underline{Y}|\theta}(\underline{\Lambda}\theta,\Psi)d\underline{Y}dF(\theta)=\int_{\underline{\gamma}(\underline{x})} f_{\underline{Y}}(\underline{0},\underline{\Lambda}\underline{\Lambda}'+\Psi)d\underline{Y}$$

in which the consitional density $f_{\underline{Y}|\theta}(\underline{\Lambda}\theta,\Psi)$ is multivariate normal, with mean vector $\underline{\Lambda}\theta$ and covariance matrix Ψ, diagonal and positive definite. The elements of Ψ are the σ_j^2, j=1..p, and, from (15.58), are equal to $(1-\lambda_j^2)$. The unconditional density $f_{\underline{Y}}(\underline{0},\underline{\Lambda}\underline{\Lambda}'+\Psi)$ is not necessarily multivariate normal. The parameter $\underline{\gamma}(\underline{x})$ is the multidimensional rectangle of integration defined by the p parameters γ_j, j=1..p, and particular response pattern $\underline{x}$. The final expression shows that the 2pno replaceability of the p variates $\mathbf{X}_j$ is equivalent to the ulcf replaceability of a set of p continuous counterparts, $\mathbf{Y}_j$, j=1..p, that, when dichotomized by the parameters γ_j, j=1..p, yield the $\mathbf{X}_j$. It follows then that the replacement variate $\boldsymbol{\theta}$ called for in (ri)-(riii) is produced in the usual ulcf manner as

(15.59) $$\boldsymbol{\theta}_i=\underline{\Lambda}'\Sigma_Y^{-1}\underline{\mathbf{Y}}+w^{1/2}\mathbf{s}_i$$

in which $E(\mathbf{s}_i)=0$, $V(\mathbf{s}_i)=1$, and $C(\underline{\mathbf{Y}},\mathbf{s}_i)=\underline{0}$. An $\underline{\mathbf{X}}$ described by the (2pno) generator will be called 2pno-replaceable. However, to date no construction formula has been derived for $\underline{\mathbf{Y}}$. Hence, the account, as it stands, remains incomplete. That is, there currently does not exist a construction formula that links the $\boldsymbol{\theta}_i$ to the input variates.

If a given $\underline{\mathbf{X}}$ is 2pno-replaceable, then it is also ir-replaceable in the following senses.

a. [f_θ unspecified, $E(\boldsymbol{\theta}^*)=t\neq 0$, $V(\boldsymbol{\theta}^*)=s\neq 1$, $s>0$, and $P(\mathbf{X}_j=1|\boldsymbol{\theta}^*=\theta^*)=\Phi(a_j(\boldsymbol{\theta}^*-b_j))$].

This is the 2pno generator modified by the added restrictions that $\boldsymbol{\theta}$ have a mean of $t\neq0$ and variance of $s\neq1$. This brand of ir-replaceability of a given set of input variates follows from the 2pno-replaceability of these variates by noting that the variate one seeks under the requirements of (a) can be expressed as $\boldsymbol{\theta}^*=s^{1/2}\boldsymbol{\theta}+t$, in which $\boldsymbol{\theta}$ is a 2pno replacement variate. Hence, $P(\mathbf{X}_j=1|\boldsymbol{\theta}=\theta)=\Phi(a_j(\theta-b_j))=\Phi(a_j^*(\theta^*-b_j^*))=P(\mathbf{X}_j=1|\boldsymbol{\theta}^*=\theta^*)$, in which $a_j^*=\frac{a_j}{s^{1/2}}$ and $b_j^*=s^{1/2}b_j+t$. Thus, from (15.56) and (15.58), this brand of ir replaceability is equivalent to the ulcf replaceability of a set of continuous counterparts, $\mathbf{Y}_j$, j=1..p, in which

$$\text{(15.60)} \qquad \lambda_j=\frac{a_j}{\sqrt{s+a_j^2}},\ \sigma_j^2=\frac{s}{s+a_j^2},\ \text{and}\ \gamma_j=\frac{a_j(s^{1/2}b_j+t)}{\sqrt{s+a_j^2}}$$

b. [f_θ unspecified, $E(\boldsymbol{\theta})=0$, $E(\boldsymbol{\theta}^2)=1$, and $P(\mathbf{X}_j=1|\boldsymbol{\theta}=\theta)=\frac{\exp(a_j(\theta-b_j))}{1+\exp(a_j(\theta-b_j))}=L(a_j(\theta-b_j))$]

This is the two-parameter logistic (2pl) ir-generator. Haley (1952) showed that, for all real numbers c, $|\Phi(c)-L(1.7c)|<.01$. Hence, the 2pl generator with parameters a_j, j=1..p, and b_j, j=1..p, is essentially equivalent to the 2pno generator with parameters $a_j^*=1.7a_j$, j=1..p, and $b_j^*=b_j$, j=1..p. It follows then that this brand of ir replaceability is, from (15.56) and (15.58), equivalent to ulcf replaceability in which

$$\text{(15.61)} \qquad \lambda_j=\frac{a_j}{\sqrt{1.72^2+a_j^2}},\ \sigma_j^2=\frac{1.7^2}{1.7^2+a_j^2},\ \text{and}\ \gamma_j=\frac{a_jb_j}{\sqrt{1.7^2+a_j^2}}.$$

Cardinality of replacement

Once again, since the choice of $\mathbf{s}_i$ in (15.59) is arbitrary save for the usual moment constraints, the set *C* containing the replacement variates to $\underline{\mathbf{X}}$ is of infinite cardinality.

Construction formulas

The construction formula is $\boldsymbol{\theta}_i=\underline{\Lambda}'\Sigma_Y^{-1}\underline{\mathbf{Y}}+w^{1/2}\mathbf{s}_i$, in which $E(\mathbf{s}_i)=0$, $V(\mathbf{s}_i)=1$, and $C(\underline{\mathbf{Y}},\mathbf{s}_i)=\underline{0}$.

Characteristics of C

To quantify the latitude inherent to a given ir-replacement Guttman's ρ^* can be employed. This measure is, herein, tailored to a number of particular ir-generators.

i. 2pno generator

To begin, note that, from (15.58),

(15.62) $R_{\underline{Y}}=\underline{\Lambda\Lambda}'+\Psi$

which, from (15.56), can be re-expressed as

(15.63) $R_{\underline{Y}}=D_a^{1/2}\underline{aa}'D_a^{1/2}+D_a$,

in which D_a is a diagonal matrix whose jjth element is $\frac{1}{(1+a_j^2)}$. Recall that, from (4.48), $\rho^*=\frac{(t-1)}{(t+1)}$, in which $t=\underline{\Lambda}'\Psi^{-1}\underline{\Lambda}$. It then follows from (15.63) and the parameter correspondences of (15.56) that

(15.64) $$\rho^*=\frac{(\sum_{j=1}^{p}a_j^2-1)}{(\sum_{j=1}^{p}a_j^2+1)}.$$

As would be expected, the degree of latitude in a 2pno-replacement is governed by the magnitudes of the discrimination parameters.

ii. 2pl generator

It follows from (15.61) that

(15.65) $R_{\underline{Y}}=D^{1/2}\underline{aa}'D^{1/2}+1.72^2D=\underline{cc}'+D^*$,

in which D is a diagonal matrix whose elements are $\frac{1}{(1.7^2+a_j^2)}$, $\underline{c}=D^{1/2}\underline{a}$, and $D^*=1.72^2D$. Hence, $t=\underline{c}'D^{*-1}\underline{c}=\frac{1}{1.7^2}\underline{a}'\underline{a}$, from which it follows that

(15.66) $$\rho^{*}=\frac{(\frac{1}{1.7^2}\sum_{j=1}^{p}a_j^{2}-1)}{(\frac{1}{1.7^2}\sum_{j=1}^{p}a_j^{2}+1)}.$$

iii. [f_θ unspecified, E($\boldsymbol{\theta}$)=t, V($\boldsymbol{\theta}$)=s, s>0, and P($\mathbf{X}_j$=x_j|$\boldsymbol{\theta}$=θ)= $\boldsymbol{\Phi}(a_j(\boldsymbol{\theta}-b_j))$]

From (15.60), $R_{\underline{Y}}=D^{1/2}\underline{a}\underline{a}'D^{1/2}+sD=\underline{c}\underline{c}'+D^{*}$, in which D is a diagonal matrix whose elements are $\frac{1}{(s+a_j^{2})}$, $\underline{c}=D^{1/2}\underline{a}$, and $D^{*}=sD$. It follows then that $t=\underline{c}'D^{*-1}\underline{c}=\frac{1}{s}\underline{a}'\underline{a}$, from which it follows that

(15.66) $$\rho^{*}=\frac{(\frac{1}{s}\sum_{j=1}^{p}a_j^{2}-1)}{(\frac{1}{s}\sum_{j=1}^{p}a_j^{2}+1)}.$$

The one-parameter normal ogive (1pno) replacement

This is simply the 2pno generator with the additional requirements that a_j=1, j=1..p, so that P($\mathbf{X}_j$=1|$\boldsymbol{\theta}$=θ)=Φ(θ-b_j). In the case of the 1pno generator, $\lambda_j=\frac{1}{\sqrt{2}}$, j=1..p, $\gamma_j=\frac{b_j}{\sqrt{2}}$, j=1..p, and $\sigma_j^2=(1-\lambda_j^2)=\frac{1}{2}$, j=1..p. It follows, then, that

(15.67) $$R_{\underline{Y}}=\frac{1}{2}\underline{1}\underline{1}'+\frac{1}{2}I,$$

and

(15.68) $$\rho^{*}=\frac{(p-1)}{(p+1)}.$$

Note, then, that the p parameters, b_j, j=1..p, of the 1pno generator play no role in governing the degree of latitude inherent to the replacement. Should a given set of input dichotomous variates be 1pno-replaceable, the minimum correlation of the replacement is a function of only the number of variates replaced.

The Guttman and independence replacements

Consider the case of the ir generator that requires that:

riG) $\boldsymbol{\theta}\sim f_\theta$, in which f_θ is continuous but unspecified,

and

riiG) $P(\underline{\mathbf{X}}=\underline{x}|\boldsymbol{\theta}=\theta_o)=\prod_{j=1}^{p}P(\mathbf{X}_j=1|\boldsymbol{\theta}=\theta_o)^{x_j}(1-P(\mathbf{X}_j=1|\boldsymbol{\theta}=\theta_o))^{1-x_j}$

in which

(15.69) $$P(\mathbf{X}_j=1|\boldsymbol{\theta}=\theta_o)=1 \quad \text{if } \theta_o\geq t_j$$
$$=0 \quad \text{if } \theta_o<t_j,\ \text{for } t_j\in R$$

Requirement (riiG) states that the form of the regression of each input variate on the replacement variate $\boldsymbol{\theta}$ must be a "step-function." Requirements (riG)-(riiG) will be called the Guttman ir (Gir) generator, for reasons that will shortly become clear.

If the set of input variates, $\mathbf{X}_j$, j=1..p, are Gir-replaceable in a population *P* of objects under study, then it follows from (15.69) that the proportion of objects in *P* who score unity on input variate $\mathbf{X}_j$ is equal to

(15.70) $$P(\mathbf{X}_j=1)=\int_{-\infty}^{\infty}P(X_j=1|\theta)dF(\theta)=\int_{t_j}^{\infty}dF(\theta)=1-F(t_j)$$

in which $F(s)=P(\boldsymbol{\theta}\leq s)$. Let $t_{j(k)}$ indicate that t_j is the kth largest in the ordering of the t_j from the largest to the smallest, and rename each of the $\mathbf{X}_j$ by its associated $t_{j(k)}$. For example, $\mathbf{X}_{t_{j(2)}}$ indicates that $\mathbf{X}_j$ has the 2nd largest value of t_j. It then follows from (15.70) that

(15.71) $$P(\mathbf{X}_{tj(1)}=1)\leq P(\mathbf{X}t_{k(2)}=1)\leq...\leq P(\mathbf{X}t_{l(p)}=1)$$

That is, the numerical order of the magnitudes of the p marginal proportions $P(\mathbf{X}_j=1)$ is the reverse of that of the t_j. Moreover, only the following (p+1) response patterns $P(\underline{\mathbf{X}}=\underline{x})$ have non-zero probabilities of occurence in population *P*:

(15.72)

$$P(\mathbf{X}_{t_{j(p)}}=0, \mathbf{X}_{t_{k(p-1)}}=0, ..., \mathbf{X}_{t_{m(2)}}=0, \mathbf{X}_{t_{n(1)}}=0)$$

$$=\int_{-\infty}^{\infty}\prod_{j=1}^{p}(1-P(X_{t_{j(k)}}=1|\theta))dF(\theta)$$

$$=\int_{-\infty}^{t_{j(p)}}1dF(\theta)+\int_{t_{j(p)}}^{\infty}0dF(\theta)=F(t_{j(p)})$$

$$P(\mathbf{X}_{t_{j(p)}}=1, \mathbf{X}_{t_{k(p-1)}}=0, ..., \mathbf{X}_{t_{m(2)}}=0, \mathbf{X}_{t_{n(1)}}=0)$$

$$=\int_{-\infty}^{\infty}P(X_{t_{j(p)}}=1|\theta)\prod_{j(k)\neq j(p)}(1-P(X_{t_{j(k)}}=1|\theta))dF(\theta)$$

$$=\int_{-\infty}^{t_{j(p)}}0dF(\theta)+\int_{t_{j(p)}}^{t_{j(p-1)}}1dF(\theta)+\int_{t_{j(p-1)}}^{t_{j(\infty)}}0dF(\theta)=F(t_{j(p-1)})-F(t_{j(p)})$$

$$P(\mathbf{X}_{t_{j(p)}}=1, \mathbf{X}_{t_{k(p-1)}}=1, ..., \mathbf{X}_{t_{m(2)}}=0, \mathbf{X}_{t_{n(1)}}=0)$$

$$=\int_{-\infty}^{\infty}P(X_{t_{j(p)}}=1|\theta)P(X_{t_{j(p-1)}}=1|\theta)\prod_{j(k)\neq j(p),j(p-1)}(1-P(X_{t_{j(k)}}=1|\theta))dF(\theta)$$

$$=\int_{-\infty}^{t_{j(p-1)}}0dF(\theta)+\int_{t_{j(p-1)}}^{t_{j(p-2)}}1dF(\theta)+\int_{t_{j(p-2)}}^{t_{j(\infty)}}0dF(\theta)=F(t_{j(p-2)})-F(t_{j(p-1)})$$

.....

.....

$$=P(\mathbf{X}_{t_{j(p)}}=1, \mathbf{X}_{t_{k(p-1)}}=1, ..., \mathbf{X}_{t_{m(2)}}=1, \mathbf{X}_{t_{n(1)}}=1)$$

$$=\int_{-\infty}^{\infty}\prod_{j(k)}P(X_{t_{j(k)}}=1|\theta)dF(\theta)$$

$$=\int_{-\infty}^{t_{j(1)}}0dF(\theta)+\int_{t_{j(1)}}^{\infty}1dF(\theta)=1-F(t_{j(1)})$$

That is, any response pattern for which $\mathbf{X}_{t_{j(r)}}=0$ and $\mathbf{X}_{t_{k(s)}}=1$, r>s, j≠k, has a probability of occurrence equal to zero in *P*. The set of (p+1) response patterns with non-zero probabilities of occurrence (i.e., those of (15.72)) comprises the well known parallelogram pattern investigated by Guttman (1950). Input variates conforming to (riG)-(riiiG) are said to form a "perfect scale".

If the only response patterns with non-zero probabilities of occurrence in *P* are the (p+1) response patterns that form the perfect scale, then the p dichotomous input variates are Gir-replaceable. If (riG) is particularized to require that $\boldsymbol{\theta}$ has a standard normal distribution, then the replacement variate has the construction formula

(15.73) $$\boldsymbol{\theta}=\sum_{i=0}^{p}\delta_i(\sum_{j=1}^{p}\mathbf{X}_j)\mathbf{Z}_i$$

in which

$$\delta_i(\sum_{j=1}^{p}\mathbf{X}_j)=1 \text{ if } \sum_{j=1}^{p}\mathbf{X}_j = i,$$
$$=0 \text{ otherwise,}$$

and

$$\mathbf{Z}_0\sim\frac{\exp\left(-\frac{\mathbf{q}^2}{2}\right)}{(1-P(\mathbf{X}_{t_{j(p)}}=1))\sqrt{2\pi}} \qquad \mathbf{q}\le t_{j(p)},$$

$$\mathbf{Z}_i\sim\frac{\exp\left(-\frac{\mathbf{q}^2}{2}\right)}{(P(\mathbf{X}_{t_{j(p-i+1)}}=1)-P(\mathbf{X}_{t_{j(p-i)}}=1))\sqrt{2\pi}} \qquad t_{j(p-i+1)}\le\mathbf{q}\le t_{j(p-i)},\ i=1..(p-1),$$

$$\mathbf{Z}_p\sim\frac{\exp\left(-\frac{\mathbf{q}^2}{2}\right)}{(P(\mathbf{X}_{t_{j(1)}}=1))\sqrt{2\pi}} \qquad t_{j(1)}\le\mathbf{q}.$$

Random variates $\mathbf{Z}_i$, i=0..p, have truncated standard normal densities, and the numerical values of the $t_{j(k)}$ are determined by (15.70).

<u>Example</u>

Let there be p=3 variates, and let these variates be Gir replaceable in a population *P*. In particular, let it be the case that:

$P(\mathbf{X}_3=0,\mathbf{X}_1=0,\mathbf{X}_2=0)=.07$
$P(\mathbf{X}_3=1,\mathbf{X}_1=0,\mathbf{X}_2=0)=.24$
$P(\mathbf{X}_3=1,\mathbf{X}_1=1,\mathbf{X}_2=0)=.30$

$P(\mathbf{X}_3=1,\mathbf{X}_1=1,\mathbf{X}_2=1)=.39$

These proportions imply the marginals $P(\mathbf{X}_3=1)=.93$, $P(\mathbf{X}_1=1)=.69$, $P(\mathbf{X}_2=1)=.39$, and parameter values $t_{3(3)}=-1.476$, $t_{1(2)}=-.496$, $t_{2(1)}=.279$. The construction formula for the standard normal replacement variate is as per (15.73), with

$$\mathbf{Z}_0\sim\frac{\exp\left(-\frac{\mathbf{q}^2}{2}\right)}{.07\sqrt{2\pi}} \qquad \mathbf{q}\leq-1.476,$$

$$\mathbf{Z}_1\sim\frac{\exp\left(-\frac{\mathbf{q}^2}{2}\right)}{.24\sqrt{2\pi}} \qquad -1.476\leq\mathbf{q}\leq-.496,$$

$$\mathbf{Z}_2\sim\frac{\exp\left(-\frac{\mathbf{q}^2}{2}\right)}{.3\sqrt{2\pi}} \qquad -.496\leq\mathbf{q}\leq.279,$$

$$\mathbf{Z}_3\sim\frac{\exp\left(-\frac{\mathbf{q}^2}{2}\right)}{.39\sqrt{2\pi}} \qquad .279\leq\mathbf{q}$$

The replaceability brought about by (riG)-(riiiG) can be stated in the standard fashion by noting that knowledge of the marginal proportions $P(\mathbf{X}_j=1)$ allows for the recovery of the higher order proportions $P(\mathbf{X}_{t_{j(p)}}=a,\mathbf{X}_{t_{k(p-1)}}=b,\ldots,\mathbf{X}_{t_{m(2)}}=c,\mathbf{X}_{t_{n(1)}}=m)$, and, hence, all association parameters of the joint distribution of the input variates. But (riG)-(riiG) also imply that knowledge of the θ-score of any object drawn from population *P* allows for the reproduction of the object's response pattern: Any object, k, has scores of unity on those variates, j, for which $\theta_k>t_{j(i)}$, and zeros on all other variates. Similarly, knowledge that the jth variate is kth in the ordering of the t_j allows one to deduce those objects in *P* who scored unity on it (those for which $\theta_k>t_{j(k)}$) and those who scored zero on it (every other object). It follows that the full object by variate response matrix in population *P* can be reproduced through knowledge of only the θ-scores of the objects and the t_j parameters of the variates.

Consider, on the other hand, the ir generator that requires that:

<u>riI)</u> $\boldsymbol{\theta}\sim f_\theta$, in which f_θ is continuous but unspecified,

and

<u>riiI)</u> $P(\underline{\mathbf{X}}=\underline{x}|\boldsymbol{\theta}=\theta_o)=\prod_{j=1}^{p}P(\mathbf{X}_j=1|\boldsymbol{\theta}=\theta_o)^{x_j}(1-P(\mathbf{X}_j=1|\boldsymbol{\theta}=\theta_o))^{1-x_j}$

in which

$$P(\mathbf{X}_j=1|\boldsymbol{\theta}=\theta_o)=P(\mathbf{X}_j=1) \qquad \forall\ \theta_o \text{ and } j=1..p$$

Requirement (riiI) states that the regression of each input variate on the replacement variate $\boldsymbol{\theta}$ must be a flat-line. That is, each input variate must be unrelated to the replacement variate. Requirements (riI)-(riiI) will be called the independence (Iir) generator (see Holland, 1990, for related discussion of the Guttman and Independence generators).

Consider, once again, the 2pno generator. Now,

(15.74) as $a_j \to \infty$,

$$\boldsymbol{\Phi}(a_j(\boldsymbol{\theta}-b_j))=P(\mathbf{X}_j=1|\boldsymbol{\theta}=\theta)\to 1 \text{ when } \theta\geq b_j$$
$$\to 0 \text{ when } \theta<b_j.$$

Hence, as the a_j parameters become indefinitely large, the regression functions of the 2pno replacement become step-functions, and the 2pno and Gir replacements converge to equivalent replacements. But it is also the case that if $a_j\to\infty$, j=1..p, then, from (15.64), $\rho^*\to 1$. Hence, the Gir replacement is a 2pno replacement in which, additionally, it is insisted that all replacement variates be linearly related. If (Gir) is refined by insisting that $\boldsymbol{\theta}$ have the standard normal distribution, then all latitude inherent to set *C* is removed. That is, the replacement can be made unique. In fact, the Gir replacement is the only finite p ir replacement in which Card(*C*)=1.

On the other hand, in the Iir replacement, the requirement that $P(\mathbf{X}_j=1|\boldsymbol{\theta}=\theta_o)=P(\mathbf{X}_j=1)\ \forall\ \theta_o$, j=1..p, can be re-stated as the requirement that $\frac{d}{d\boldsymbol{\theta}}P(\mathbf{X}_j=1|\boldsymbol{\theta}=\theta_o)=0\ \forall\ \theta_o$, j=1..p. Because, when $a_j=0$, $\frac{d}{d\boldsymbol{\theta}}\boldsymbol{\Phi}(a_j(\boldsymbol{\theta}-b_j))=0$ for all $\forall\ \theta_o$, the Iir can be represented as a 2pno replacement in which the regression functions must have the form $P(\mathbf{X}_j=1|\boldsymbol{\theta}=\theta_o)=c_j+\Phi(a_j(\theta_o-b_j))$, with $a_j=0$, j=1..p, and $c_j=P(\mathbf{X}_j=1)-\Phi(0)$. But if $a_j=0$, j=1..p, then, from (15.64), $\rho^*=-1$. That is, the latitude inherent to the Iir replacement is at a maximum. It can then be summarized that the Gir and Iir replacements are at opposite extremes in regard the latitude they allow with regard the set of replacement variates constructed under each. When the latent variable modeller tests an hypothesis of Gir replaceability of some particular set of input variates, he is testing whether the input variates are 2pno-replaceable by a single replacement variate. When he tests an hypothesis of ir-replaceability of a particular set of input variates against the standard alternative that the variates "are described by the independence model", he is then, in fact, testing an hypothesis that the latitude inherent to the replacement of these variates is less than the maximum allowed under the Iir replacement.

References

Allen, M.J. & Yen, W.M. (1979). *Introduction to measurement theory*. Monterey, CA: Brooks/Cole.
"alpha particle". *A Dictionary of Physics*. Ed. Alan Isaacs. Oxford University Press, 2000. *Oxford Reference Online*. Oxford University Press.
Anderson, T., & Rubin, M. (1956). Statistical inference in factor analysis. *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, *5*, 11-150.
Anderson, T. W. (1959) Some Scaling Methods and Estimation Procedures in the Latent Class Model. In U.Grenander (ed.). *Probability and Statistics: The Harold Cramer Volume*, New York: Wiley.
Anderson, T.W. (1963). Asymptotic theory for principal components analysis. An*nals of Mathematical Statistics, 34,* 122-134.
Aitkin, M. (1985). Comments on D.J. Bartholomew, Foundations of factor analysis: some practical implications. *British Journal of Mathematical and Statistical Psychology*, 38, 127-128.
Austin, J.L. (1962). *Sense and Sensibilia*, G.J. Warnock (Ed.). Oxford: Oxford University Press.
Baker, F. (1965). Origins of the item parameters X_{50} and β as a modern item analysis technique. *Journal of Educational Measurement, 2(2),* p.167-180.
Baker, G. & Hacker, P. (1980). *Wittgenstein: Meaning and Understanding.* Chicago: University of Chicago Press.
Baker, G. & Hacker, P. (1982). The Grammar of Psychology: Wittgenstein's Bemerkungen über die philosophie der psychologie. *Language and Communication, 2(3),* 227-244.
Bartholomew, D. (1980). Factor analysis for categorical data. *Journal of the Royal Statistical Society, 42(3),* p.293-321.
Bartholomew, D. (1981). Posterior analysis of the factor model. *British Journal of Mathematical and Statistical Psychology, 34,* 93-99.
Bartholomew, D. (1984). The foundations of factor analysis. *Biometrika, 71(2)*, p.221-232.
Bartholomew, D. (1995). Spearman and the origin and development of factor analysis. *British Journal of Mathematical and Statistical Psychology, 48*, 211-220.
Bartholomew, D. (1996). Metaphor taken as math: Indeterminacy in the factor analysis model: Comment. *Multivariate Behavioral Research*, *31(4*), 551-554.
Bartholomew, D., & Knott, M. (1999). *Latent variable models and factor analysis.* London: Arnold.
"baryon" *A Dictionary of Physics*. Ed. Alan Isaacs. Oxford University Press, 2000. *Oxford Reference Online*. Oxford University Press.
Beck, A. T. (1967). *Depression: Causes and treatment.* Philadelphia: University of Pennsylvania Press.

Benet, V., & Waller, N.G. (1995). The Big Seven factor model of personality description: Evidence for its cross-cultural generality in a Spanish sample. *Journal of Personality* & *Social Psychology*, *69(4),* 701-718.

Bennett, M.R., & Hacker, P.M.S (2003). *Philosophical Foundations of Neuroscience.* Oxford: Blackwell Publishing.

Bentler, P.M. (1980). Multivariate analysis with latent variables: Causal modeling. *Annual Review of Psychology, 31, 1980*, 419-456.

Bentler, P. (1986). Structural modeling and Psychometrika: An historical perspective on growth and achievements. *Psychometrika, 51(1),* 35-51.

Bentler, P.M. (1989). *EQS structural equations program manual.* Los Angeles: BMDP Statistical Software.

Bentler, P.M., & Kano, Y. (1990). On the equivalence of factors and components. *Multivariate Behavioural Research, 25(1),* 67-74.

Bickel, P., & Doksum, K. (1977*). Mathematical Statistics: Basic Ideas and Selected Topics.* Oakland: Holden-Day, Inc.

Bielby, W.T., & Hauser, R.M. (1977). Structural Equation Models. *Annual Review of Sociologoy, 3,* 137 – 161.

Bijleveld, C. (1989). *Exploratory Linear Dynamic Systems Analysis.* Leiden: DSWO Press.

Binet, A. & Simon, T. (1916). *The development of intelligence in children (The Binet-Simon scale)* Translations by E. S. Kite. Baltimore: Williams & Wilkins.

Birnbaum, A. (1968). Some Latent Trait Models. In F.M. Lord & M.R. Novick (Eds.) *Statistical Theories of Mental Test Scores.* Reading, Massachusetts: Addison-Wesley. pp. 397-424.

Black, M. (1962). *Models and Metaphors: Studies in Language and Philosophy.* New York: Cornell University Press.

Blalock, H. (1961). *Causal Inferences in Nonexperimental Research.* University of North Carolina Press, Chapel Hill, NC.

Blalock, H.M. (1963). Making causal inferences for unmeasured variables from correlations among indicators. *American Journal of Sociology, 69*, 53-62.

Blinkhorn, S. (1997). Past imperfect, future conditional: Fifty years of test theory. *British Journal of Mathematical and Statistical Psychology, 50*, 175-185.

Bock, R. D., & Aitkin, M. (1981). Marginal maximum likelihood estimation of item parameters: Application of an EM algorithm. *Psychometrika, 46(4),* 443-459.

Bock, R. D., & Lieberman, M. (1970). Fitting a response model for n dichotomously scored items. *Psychometrika, 35(2),* 179-197.

Bollen, K. A. (1989). *Structural equations with latent variables.* New York: Wiley.

Boltzmann, L.. (1974) [1905]. *Theoretical Physics and Philosophical Problems*. B. McGuinness, (Ed). P. Foulkes, (Trans.) Dordrecht: Reidel.

Bravais, A. (1846), Analyse mathematique sur les probabilites des erreurs de situation d'un point," *Memoires par divers Savans, 9,* 255-332.

Browne, M. (1985). Comments on D.J. Bartholomew, Foundations of factor analysis: some practical implications. *British Journal of Mathematical and Statistical Psychology, 38,* 132-133.

Bueche, F. (1972). *Principles of Physics, 2nd Edition.* New York: McGraw-Hill.
Burt, C. (1940). *The factors of the mind: An introduction to factor-analysis in psychology.* London: University of London Press.
Burt, C. (1966). The early history of multivariate techniques in psychological research. *Multivariate Behavioral Research, 1(1),* 24-42.
Bush, R., & Mosteller, F. (1955). *Stochastic Models for Learning.* New York: Wiley.
Bush, R., Luce, D., & Suppes, P. (1978). Mathematical models. In W. Kruskal and J. Tanur (Eds.), *The International Encyclopedia of Statistics*. New York: The Free Press.
Camp, BH. (1932). The converse of Spearman's two-factor theorem. *Biometrika, 24,* 260-261.
Camp, BH. (1934). Spearman's general factor again. *Biometrika, 26(1),* 260-261.
Campbell, N. (1952). *What is Science.* New York: Dover.
Carnap, R (1950). Empricism, semantics, and ontology. *Revue Internationale de Philosophie*, *4,* 20-40. Reprinted in the Supplement to *Meaning and Necessity: A Study in Semantics and Modal Logic,* 4th ed. (1956). Chicago:University of Chicago Press.
Carnap, R. (1963). Intellectual biography. In P.A. Schlipp (Ed.), *The Philosophy of Rudolf Carnap.* La Salle, ILL: Open Court.
Cattell, R.B. (1952). *Factor Analysis: An Introduction and Manual for the Psychologist and Social Scientist.* Oxford: Harper.
Chow, S.L. (1998). In defense of the construct validity approach. *Theory* & *Psychology*, *8(4),* 481-487.
Christoffersson, A. (1975). Factor analysis of dichotomized variables. *Psychometrika, Vol 40(1)*, 5-32.
Costner, H.L. (1969). Theory, deduction and rules of correspondence. *American Journal of Sociology 75,* 245-263.
Cooper, C. (1998). *Individual Differences.* London: Arnold.
Crocker, L. & Algina, J. (1986). *Introduction to Classical and Modern Test Theor*y. New York: Holt Reinhard and Winston.
Cronbach, L.J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52,* 281-302.
Cureton, E., & D'Agostino, R. (1983). *Factor analysis.* Hillsdale, N. J.: Erlbaum Associates.
Damarin, F. (1970). A latent-structure model for answering personal questions. *Psychological Bulletin, 73(1),* 23-40.
De Leeuw, J. (1988). Model selection in multinomial experiments. In T. Dijkstra (Ed.), *On Model Uncertainty and its Statistical Implications.* Berlin: Springer-Verlag.
De Leeuw, J. (1996). Series editors introduction. In T. Heinen, *Latent class and discrete latent latent trait models: Similarities and differences.* London: Sage Publications.
Dobson, A. J. (1990). *An Introduction to Generalized Linear Models.* New York: Chapman and Hall.
Dodd, S. (1928). The theory of factors, I. *Psychological Review, 35,* 211-234.
Dunteman, G.H. (1989). *Principle Component Analysis.* Newbury Park, CA: Sage.

Dwyer, J.H. (1983). *Statistical models for the social and behavioral sciences*. New York : Oxford University Press.
Edwards, A.L. (1957). *Techniques of Scale Construction.* New York: Appleton-Century-Crofts.
Everitt, B. S. (1984). *An Introduction to Latent Variable Models.* New York: Chapman and Hall.
Eysenck, H. J (1953). The logical basis of factor analysis. *American Psychologist*, *8*, 105-113.
Falmagne, J. (1985). *Elements of Psychophysical Theory*. Oxford: Clarendon Press.
Feigl, H. (1950). Existential hypotheses: realistic versus phenomenalistic interpretations. *Philosophy of Science, 17,* 36-62.
Feigl, H. (1953). The scientific outlook: naturalism and humanism. In H. Feigl & M. Brodbeck (Eds.), *Readings in the philosophy of science* (pp. 8-18). New York: Appleton-Century-Crofts.
Feigl, H. (1956). Some major issues and developments in the philosophy of science of logical empiricism. In H. Feigl & M. Scriven (Eds.), *Minnesota studies in the philosophy of science* (Vol. 1). Minneapolis: University of Minnesota Press. pp. 3-37.
Ferguson, G. (1942). Item selection by the constant process. *Psychometrika, 7,* 19-29.
Fielding, A., (1977). Latent structure models. In: C. A. O'Muircheartaigh & C. Payne (Eds.). *Exploring Data Structures*. New York: Wiley. pp. 125-158.
Final Report by the *Ad Hoc Committee to Consider Certain Charges* to the Editorial Counsel of the Psychometric Society. June 1, 1978.
Finney, D. (1944). The application of probit analysis to the results on mental tests. *Psychometrika, 19,* 31-39.
Freedman, D. (1985). Statistics and the scientific method. In W. Mason and S. Fienberg (Eds.), *Cohort Analysis in Social Research.* New York: Springer-Verlag.
Frisch, R. (1934). *Statistical Confluence Analysis by Means of Complete Regression Systems*, (Economics Institute Publication No. 5). Oslo: Oslo University.
Fruchter, B. (1954). *Introduction to Factor Analysis.* New York: D. Van Nostrand Company.
Gabriel, K. (1971). The biplot graphical display of matrices with application to principal component analysis. *Biometrika, 58,* 453-467.
Galton, F. (1875). History of twins. In F.Galton, *Inquiries into Human Faculty and its Development.* (pp. 155-173). London: Dent.
Galton, F. (1877). Typical laws of heredity. *Proceedings of the Royal Institution, 8,* 282-301.
Gangestad, S., & Snyder, M. (1985). To carve nature at its joints: On the existence of discrete classes in personality. *Psychological Review*, *92(3)*, 317-349.
Garner, W. R., & Hake, H. W. Eriksen, C. W. (1956). Operationism and the concept of perception. *Psychological Review*, *63*, 149-159.
Garnett, J. (1919). On certain independent factors in mental measurements. *Proceedings of the Royal Statistical Society, 96,* 91-111.
Goldstein, H.,& Wood, R. (1989). Five decades of item response modeling. *British Journal of Mathematical and Statistical Psychology, 42(2),* 139-167.

Gorsuch, R. (1983). *Factor analysis.* Hillsdale, NJ: Lawrence Erlbaum Associates.
Gorsuch, R. (1990). Common factor analysis versus component analysis: Some well and little known facts. *Multivariate Behavioral Research, 25(1),* 33-39.
Grosser, H. (1962). *The Discovery of Neptune.* Cambridge, Mass: Harvard University Press.
Green, B. F. (1954). Attitude measurement. In G. Lindzey (Ed.), *Handbook of social psychology*, *Vol. 1.* Reading, MA: Addison-Wesley. pp. 335 – 369.
Green, B.F. Jr. (1976). On the factor score controversy. *Psychometrika, 41(2),* 263 – 266.
Guilford, J.P. (1936) *Psychometric Methods*. New York: McGraw-Hill.
Guilford, J.P. (1954) *Psychometric Methods* (2nd Edition). New York: McGraw-Hill.
Guttman, L. 1940. Multiple rectilinear prediction and the resolution into components. *Psychometrika*, *5*, 75-99.
Guttman, L. (1945). A basis for analyzing test-retest reliability. *Psychometrika, 10,* 255-282.
Guttman, L. (1950). The basis for scalogram analysis. In S.A. Stouffer, L. Guttman, E.A. Suchman, P.F. Lazarsfeld, S.A.Star, and J.A. Clausen (Eds.), *Measurement and Prediction.* NewYork: John Wiley & Sons, Inc. pp. 60-90.
P.F., Lazarsfeld, S.A. Star, & J.A. Clausen (Eds.), *Measurement and Prediction*. Princeton: Princeton University Press.
Guttman, L. (1953). Image theory for the structure of quantitative variates. *Psychometrika, 18,* 277-296.
Guttman, L. (1955). The determinacy of factor score matrices with implications for five other basic problems of common-factor theory. *The British Journal of Statistical Psychology, 8(2)*, 65-81.
Guttman, L. (1956). "Best possible" systematic estimates of communalities. *Psychometrika*, *21*, 273-285.
Guttman, L. (1957). Simple proofs of relations between the communality problem and multiple correlation. *Psychometrika, 22(2),* 147 – 157.
Guttman, L. (1975). *The Green-McDonald Proof of Nonexistence of Factor Analysis.* Unpublished manuscript, Hebrew University of Jerusalem and the Israel Institute of Applied Social Research.
Guttman, L. (1977). What is not what in statistics*. The Statistician, 26,* 81-107.
Hacker, P. (1986). *Insight and Illusion: Themes in the Philosophy of Wittgenstein.* Oxford: Clarendon Press.
Hacker, P. (1987). Languages, minds and brains. In C. Blakemore, & S. Greenfield (Eds.), *Mindwaves.* Oxford: Basil Blackwell Ltd.
Hacker, P. (1988). Language, rules and pseudo-rules. *Language and Communication, 8,* 159-172.
Haley, D.C. (1952). Estimation of the dosage mortality relationship when the dose is subject to error. Stanford Applied Mathematics & Statistics Laboratory, Technical Report No. 15. Stanford University.
Technical ReportHamid, P. N., & Cheng, S. (1996). The development and validation of an index of emotional disposition and mood state: The Chinese Affect Scale. *Educational* & *Psychological Measurement*, *56(6),* 995-1014.

Harman, H.H. (1960). *Modern factor analysis.* Chicago: University of Chicago Press.
Harre, R. (1970). *The Principles of Scientific Thinking.* Chicago: The University of Chicago Press.
Hart, B., & Spearman, C. (1912). General ability: Its existence and nature. *British Journal of Psychology, 5,* 51-84.
Hartley, R. E. (1954). Two kinds of factor analysis. *Psychometrika*, *19*, 195-203.
Hauser, R.M. & Goldberger, A.S. (1971). The treatment of unobservable variables in path analysis. In H. Costner (Ed.), *Sociological Methodology.* Washington, D.C.: American Sociological Association, pp. 81-117.
Hayduck, L.A. 1987. *Structural Equation Modeling with LISREL: Essentials and Advances.* Baltimore: The John Hopkins University Press.
Heermann, E.F. (1966). The algebra of factorial indeterminancy. *Psychometrika*, *31,* 539 – 543.
Heermann, E.F. (1964). The geometry of factorial indeterminacy. *Psychometrika, 29,* 371 – 381.
Hempel, C. G. (1952). Fundamentals of concept formation in empirical science. *Int. Encyclopedia Unified Science, Vol. II. No. 7.* Chicago: University of Chicago Press.
Hempel, C. 1965. *Aspects of Scientific Explanation: And other essays in philosophical science*. New York: The Free Press. pp. 101-122.
Henrysson, S. (1957). *Applicability of factor analysis in the behavioral sciences: a methodological study.* Oxford, England: Almqvist & Wiksell.
Holland, P (1981). When are item response models consistent with observed data? *Psychometrika, 46 (1),* 79 – 92.
Holland, P. (1990). On the sampling theory foundations of item response theory models. *Psychometrika, 55,* 577-601.
Holland, P. W. and Rosenbaum, P. R. (1986). Conditional association and unidimensionality in monotone latent variable models. *The Annals of Statisitcs, 14* 1523-1543.
Holzinger K.J., & Harman, H.H. (1941). *Factor Analysis.* Chicago: University of Chicago Press.
Horst, P. (1965). *Factor Analysis of Data Matrices.* New York: Holt, Rinehart and Winston, Inc.
Holzinger, K.J. (1930). *Statistical Resume of the Spearman Two-factor Theory.* Chicago: University of Chicago Press.
Howe, W.G. (1955). Some contributions to factor analysis. Oak Ridge: Oak Ridge National Laboratory (Report no. ORNL-1919).
Hoyle, R.H., & Lennox, R.D. (1991). Latent structure of self-monitoring. *Multivariate Behavioral Research*, *26(3)*, 511-540.
Intriligator, M. (1983). Economic and econometric models. In Z. Griliches and M. Intriligator (Eds.), *Handbook of Econometrics, Vol. I.* Amsterdam: North-Holland.
Irwin, J.O. (1935). On the indeterminacy in the estimate of *g*. *British Journal of Psychology, 25,* 393 – 394.

Jensen, A. (1985). The nature of black-white differences on various psychometric tests: Spearman's hypothesis. *Behavioral and Brain Sciences*, *8,* 193-263.
Jensen, A. (1969). How much can we boost IQ and scholastic achievement? *Harvard Educational Review, 39,* 1-123.
Jolliffe, I.T. (1986). *Principal Component Analysis.* New York: Springer.
Joreskog, K. G. (1967). Some contributions to maximum likelihood factor analysis. *Psychometrika*, *32(4),* 443-482.
Joreskog, K. G (1969). A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika*, *34(2 Pt.1)*, 183-202.
Joreskog, K.G (1973). A general method for estimating a linear structural equation. In A.S. Goldberger and O.D. Duncan (Eds.), *Structural Equation Models in the Social Sciences.* New York: Seminar Press.
Joreskog, K.G. (1979). Basic ideas of factor and component analysis. In K.G Joreskog & D. Sorbom, *Advances in factor analysis and structural equation models*. Cambridge, MA: Abt Books.
Joreskog, K.G, & Sorbom, D. (1979). *Advances in factor analysis and structural equation models*. Cambridge, MA: Abt Books.
Joreskog, K.G, & Sorbom, D. (1985). Simultaneous analysis of longitudinal data from several cohorts. In W. Mason and S. Fienberg, *Cohort Analysis in Social Research*. New York: Springer-Verlag.
Joreskog, K. G.; Sorbom, D. (1993). *LISREL 8: Structural equation modeling with the SIMPLIS command language.* Chicago: Scientific Software International, Inc.Lawrence Erlbaum Associates, Inc.
Jost, J. T., & Gustafson, D.F. (1998). Wittgenstein's problem and the methods of psychology: How "grammar depends on facts”. *Theory* & *Psychology*, *8(4),* 463-479.
Kalman, R. (1974). Comments on the scientific aspects of modeling. In M. Marois (Ed.), *Towards a Plan of Action for Mankind.* Amsterdam: North Holland.
Kalman, R. (1983). Identifiability and modeling in econometrics. In P. Krishnaiah (Ed.), *Developments in Statistics, Vol.4*. New York:Academic Press.
Keesling, J.W. (1972). *Maximum Likelihood Approaches to Causal Analysis*. Ph.D. dissertation, Chicago: University of Chicago.
Kestelman, H. (1952). The fundamental equation of factor analysis. *British Journal of Psychology*, *5*, 1952. 1-6.
Koch, S. (1954). Clark L. Hull. In W.K. Estes et al (Eds.), *Modern Learning Theory.* New York: Appleton-Century-Crofts.
Krantz, D. H. (1991). From indices to mappings: The representational approach to measurement. In D. R. Brown & J. E. K. Smith (Eds.), *Frontiers of Mathematical Psychology*New York: Springer-Verlag. pp. 1-52.
Land, K. C. (1970). On the estimation of path coefficients for unmeasured variables from correlations among observed variables. *Social Forces*, *48* , 506-511.
Lawley, D. N. (1940). The estimation of factor loadings by the method of maximum likelihood. *Proceedings of The Royal Society of Edinburgh, 60,* 64-82.
Lawley, D.N. (1943). On problems connected with item selection and test construction. *Proceedings of the Royal Society of Edinburgh, 1943, 61,* 273-287.

Lawley, D. (1947). On problems connected with item selection and test construction. *Proceedings of the Royal Society of Edinburgh, 61,* 273-287.
Lawley, D.N. & Maxwell, A.E. (1963). *Factor Analysis as a Statistical Method.* London: Butterworth (2nd edition 1971).
Lazarsfeld, P.F. (1950). The logical and mathematical structure of latent structure analysis. In S.A. Stouffer, L. Guttman, E.A. Suchman, P.F. Lazarsfeld, S.A.Star, and J.A. Clausen (Eds.), *Measurement and Prediction.* NewYork: John Wiley & Sons, Inc.
Lazarsfeld, P. F. (1959). Latent structure analysis. In S. Koch (Ed.), *Psychology*: *A Study of a Science, Vol. 3*. New York: McGraw-Hill.
Leary, T. (1957). *Interpersonal Diagnosis of Personality.* New York: Ronald Press.
Ledermann, W (1937). On the rank of the reduced correlational matrix in multiple-factor analysis. *Psychometrika, 2,* 85-93.
Ledermann, W. (1938). The orthogonal transformation of a factorial matrix into itself. *Psychometrika, 3,* 181 – 187.
Levy, P. (1995). Charles Spearman's contributions to test theory. *British Journal of Mathematical and Statistical Psychology, 48,* 221-235.
Lewis, C. (1986). Test theory and psychometrika: The past twenty-five years.; *Psychometrika*, *Vol 51(1),* 11-22.
Ljung, L. (1987). *System Identification: Theory for the User.* Englewood Cliffs, NJ: Prentice-Hall.
Loehlin, J. C. (1990). Component analysis versus common factor analysis: A case of disputed authorship. *Multivariate Behavioral Research, Vol 25(1)*, 29-31.
Loehlin, J. C. (1998). *Latent Variable Models: An Introduction to Factor, Path and Structural Analysis (3rd Edition).* New Jersey: Lawrence Erlbaum Associates.
Lohmoller, J. (1989). *Latent variable path modeling with partial least squares.* Heidelberg: Physica-Verlag.
Lord, F. (1952). A theory of test scores. *Psychometric Monographs, 7.*
Lord, F., & Novick, M. (1968). *Statistical Theories of Mental Test Scores.* London: Addison-Wesley Publishing Company.
Lovie, P., & Lovie, A. D. (1995). The cold equations: Spearman and Wilson on factor indeterminacy. *British Journal of Mathematical & Statistical Psychology, 48(2),* 237-253.
Mach, E. (1960). *The Science of Mechanics,* (6th Ed.). Trans. J. McCormack. LaSalle, Illinois: Open Court.
Maraun, M. D. (1996).Metaphor taken as math: Indeterminacy in the factor analysis model. *Multivariate Behavioral Research*, *31(4),* 517-538.
McArdle, J. J (1990). Principles versus principals of structural factor analyses. *Multivariate Behavioral Research*, *25(1),* 81-87.
McDonald, R. P. (1999). *Test Theory: A Unified Treatment.* London: Lawrence Erlbaum Associates, Publishers.
McDonald, R. P. (1996a). Latent traits and the possibility of motion. *Multivariate Behavioral Research*, *31(4),* 593-601.

McDonald, R. P. (1996b). Consensus emergens: A matter of interpretation. *Multivariate Behavioral Research, 31(4),* 663-672.
McDonald, R. P. (1983). Exploratory and confirmatory nonlinear common factor analysis. In H. Wainer & S. Messick (Eds.) *Principals of Modern Psychological Measurement: A Festschrift for Frederic M. Lord.* Hillsdale, New Jersey: Lawrence Erlbaum Associates, Inc.
McDonald, R. P. (1981). The dimensionality of tests and items. *British Journal of Mathematical and Statistical Psychology, 34,* 100-117.
McDonald, R. P. (1979). The simultaneous estimation of factor loadings and scores. *British Journal of Mathematical and Statistical Psychology, 32,* 212-228.
McDonald, R. P. (1977). The indeterminacy of components and the definition of common factors. *British Journal of Mathematical and Statistical Psychology, 30,* 165-176.
McDonald, R. P. (1975). Descriptive axioms for common factor theory, image theory and component theory. *Psychometrika, 40(2),* 137 – 152.
McDonald, R. P. (1974). The measurement of factor indeterminacy, *Psychometrika, 39,* 203-222.
McDonald, R.P. (1972) Some uncertainties about factor indeterminacy. Unpublished paper.
McDonald, R.P. (1967). Nonlinear factor analysis. *Psychometric Monographs*, No. 15, 1967. pp. 167.
McDonald, R.P. (1962) A note on the derivation of the general latent class model. *Psychometrika, 27,* 203 –206.
McDonald, R., & Burr, E. (1967). A comparison of four methods of construction factor scores. *Psychometrika, 32,* 381-401.
McDonald, R.P., & Mulaik, S. A. (1979). Determinacy of common factors: A nontechnical review. *Psychological Bulletin*, *86(2),* 297-306.
Meehl, P. (1965). Detecting latent clinical taxa by fallible quantitative indicators lacking an accepted criterion (Report No. PR-65-2). Minneapolis: University of Minnesota, Research Laboratories of the Department of Psychiatry.
Meehl, P. (1973). MAXCOV-HITMAX: A taxonomic search method for loose genetic syndromes. In #*97 Psychodiagnosis: Selected papers.* Minneapolis: University of Minnesota Press. pp. 200-224.
Meehl, P. (1977). Specific etiology and other forms of strong influence: Some quantitative meanings. *Journal of Medicine and Philosophy, 2*, 33-53.
Meehl, P. (1986). Diagnostic taxa as open concepts: Metatheoretical and statistical quesions about reliability and construct validity in the grand strategy of nosological revision. In T. Millon & G. Klerman (Eds.*), Contemporary directions in psychopathology.* New York: Guilford.
Meehl, P. (1992). Factors and taxa, traits and types, differences of degree and differences in kind. *Journal of Personality, 60,* 117-174.
Meehl, P., & Golden, R.R. (1982). Taxometric methods. In P.C. Kedall & J.N. Butcher (Eds.), *Handbook of research methods in clinical psychology* (pp. 127-181*).* New York: John Wiley & Sons.
Meredith, W. (1965). Some results based on a general stochastic model for mental tests.*Psychometrika*, *30(4),* 419-440.

"meson" *A Dictionary of Physics*. Ed. Alan Isaacs. Oxford University Press, 2000. *Oxford Reference Online*. Oxford University Press.
Messick, S. (1981). Constructs and their vicissitudes in educational and psychological measurement. *Psychological Bulletin, 89(3),* 575-588.
Mill, J.S. (1843). *System of Logic.* London: Longman's.
Molenaar, P. C., & von Eye, A. (1994). On the arbitrary nature of latent variables. In A. von Eye & C. C. Clogg (Eds*.), Latent variable analysis: Applications for Developmental Research* (pp. 226-242). Thousand Oakes, CA: Sage Publications.
Mueller, G.E. (1904). *Die Geischtpunkte und die Tatsachen der psychophysischen Methodik*. Wiesbaden: Bergmann.
Mulaik, S. (1976). Comments on "The measurement of factorial indeterminacy". *Psychometrika, 41(2),* 249 – 262.
Mulaik, S. (1986). Factor analysis and Psychometrika: Major developments. *Psychometrika, 51(1),* 23-33.
Mulaik, S. (1990). Blurring the distinctions between component analysis and common factor analysis. *Multivariate Behavioral Research, 25 (1),* 53-59.
Mulaik, S. (1993). The critique of pure statistics: Artifact and objectivity in multivariate statistics. In B. Thompson (Ed.), *Advances in Social Science Methodology, Vol.3.* Greenwich, CT: JAI Press.
Mulaik, S. (1994). Kant, Wittgenstein, objectivity, and structural equation models. In C.R. Reynolds (Ed.), *Cognitive Assessment: A Multidisciplinary Perspective*. New York: Plenum.
Mulaik, S., & McDonald, R. (1978). The effect of additional variables on factor indeterminacy in models with a single common factor. *Psychometrika, 43*, 177-192.
Muthen, B. (1978). Contributions to factor analysis of dichotomous variables. *Psychometrika, 43*, 551-560.
Nelder, J. (1984). The role of models in official statistics. In *Recent Developments in the Analysis of Large Scale Data Sets*. Luxembourg: Proceedings of a Seminar.
Norris, S. (1983). The inconsistencies at the foundation of construct validation theory. In E. House (ed.), *Philosophy of Education.* San Francisco: Jossey-Bass.
Overall, J. E. (1964). Note on the scientific status of factors. *Psychological Bulletin*, *61(4)*, 270-276.
Pawlik, K. (1968). *Dimensionen des Verhaltens*. Bern: Huber.
Pearson, K. (1920). Notes of the history of correlation. *Biometrika, 13,* 25-46.
Petrianov-Sokolov, I. (1985). *Elementary order: Mendeleev's periodic system*. (translated from the Russian by Weinstein, J.). Moscow:Mir.
Piaggio, H. (1931). The general factor in Spearman's theory of intelligence. *Nature, 127*, 56-57.
Piaggio, H. (1933). Three sets of conditions necessary for the existence of a *g* that is real and unique except in sign. *British Journal of Psychology, 24,* 88 – 105.
Piaggio, H. (1935). Approximate general and specific factors without indeterminate parts. *British Journal of Psychology, 25,* 485-489.
Richardson, M. W. (1936). The relationship between difficulty and the differential validity of a test. *Psychometrika, 1,* 33-49.

Reiersol (1950). On the identifiability of parameters in thurstone's multiple factor analysis. *Psychometrika, 15* 121-149.
Royce, J. R (1958). The development of factor analysis. *Journal of General Psychology, 58*,. pp. 139-164.
Royce, J. R.(1963). Factors as theoretical constructs. *American Psychologist, 18(7*), 522-527.
Rozeboom, W. (1966). *Foundations of the Theory of Prediction.* Illinois: The Dorsey Press.
Rozeboom, W. (1984). Dispositions do explain. Picking up the pieces after hurricane Walter. In J. Royce & L. Mos (Eds.), *Annals of Theoretical Psychology, Vol.I.* New York: Plenum.
Rozeboom, W. (1988). Factor indeterminacy: The saga continues. *British Journal of Mathematical and Statistical Psychology, 41,* 209-226.
Rozeboom, W. W. (1990). Whatever happened to broad perspective? *Multivariate Behavioral Research*, *25(1*), 61-65.
Rozeboom, W. W. (1996). What might common factors be? *Multivariate Behavioral Research*, *31(4),*. 555-570.
Rummel, R.J. (1970). *Applied Factor Analysis.* Evanston: Northwestern University Press.
Schneeweiss, H. (1989, June). *Models with latent variables.* Paper presented at Joint Summer Research Conference: Statistical Analysis of Measurement Error Models and Applications, Arcata, CA.
Schonemann, P. (1990). Facts, fictions and common sense about factors and components. *Multivariate Behavioral Research, 25 (1),* 47-51.
Schonemann, P. (1978). Some nontrivial implications of factor indeterminacy. Presentation given at the *Society for Multivariate Experimental Psychology,* Cincinnatti, November, 1978.
Schonemann, P. (1971). The minimum average correlation between equivalent sets of uncorrelated factors. *Psychometrika, 36,* 21 – 149.
Schonemann, P., & Haagen, K. (1987). On the use of factor scores for prediction. *Biometrical Journal , 7,* 835-847.
Schonemann, P., & Steiger, J.H. (1976). Regression Component Analysis. *British Journal of Mathematical and Statistical Psychology, 29*, 174-189.
Schonemann, P., & Steiger, J.H. (1978). On the validity of indeterminate factor scores. *Bulletin of the Psychonomic Society, 12(4),* 287 – 290.
Schonemann, P., & Wang M. (1972). Some new results on factor indeterminacy. *Psychometrika, 37*, 61-91.
Seber, G.A.F. (1984). *Multivariate Observations.* New York: Wiley.
"self-esteem". Def. *Oxford English Dictionary.* 2nd Edition, 1989.
Sharp, S. (1997). "Much more at home with 3,999 pupils than with four": The contributions to psychometrics of Sir Godfrey Thomson. *British Journal of Mathematical and Statistical Psychology, 50,* 163-174.
Simon, H. (1954). Spurious correlations: A causal interpretation. *Journal of the American Statistical Association*, *49*, 469-492.
Sobel, Michael E. (1997) Measurement, causation and local independence. in M. Berkane (ed.), *Latent Variable Modeling and Applications to Causality.* New York:

Springer-Verlag.
Slaney, K. (2001). *Categories, Types and Taxa: A Critical Analysis of Paul Meehl's Program for Explicating Taxonicity.* Unpublished master's thesis, Simon Fraser University, Burnaby, British Columbia, Canada.
Spearman, C. (1904). General intelligence, objectively determined and measured. *American Journal of Psychology, 15,* 201-293.
Spearman, C. (1904). The proof and measurement of association between two things. *American Journal of Psychology, 15,* 72-101.
Spearman, C. (1907). Demonstration of formulae for true measurement of correlation. *American Journal of Psychology, 18,* 161-169.
Spearman, C. (1910). Correlation calculated from faulty data. *British Journal of Psychology, 3,* 271-295.
Spearman, C. (1927). *Abilities of Man.* New York: Macmillan.
Spearman, C. (1929). The uniqueness of "g". *Journal of Educational Psychology, 20,* 212 – 216.
Spearman, C. (1931). Response to "The general factor in Spearman's theory of intelligence. *Nature, 127,* 57.
Spearman, C. (1933). The uniqueness and exactness of g. *British Journal of Psychology*, *24*, 106-108.
Steiger, J. (1979). The relationship between external variables and common factors. *Psychometrika, 44(1),* 93-97.
Steiger, J. (1990). Some additional thoughts on components, factors, and factor indeterminacy. *Multivariate Behavioral Research 25 (1),* 41-45.
Steiger, J. (1996). Dispelling some myths about factor indeterminacy. *Multivariate Behavioral Research, 31(4),* 539-550.
Steiger, J. (2001). Driving fast in reverse. *Journal of the American Statistical Association, 96,* 331 – 338.
Steiger, J., & Schonemann, P. (1978). A history of factor indeterminacy. In S. Shye (Ed.), *Theory construction and data analysis in the social sciences*. San Francisco: Jossey Bass.
Steinberg, L., & Thissen, D. (1995). Item response theory in personality research. In P.E. Shrout & S.T. Fiske (Eds.) *Personality research*, *methods*, *and theory*: *A festschrift honoring Donald W. Fiske.* NJ, England: Lawrence Erlbaum Associates, 161-181.
Stephenson, W. (1936). The foundations of psychometry. *Psychometrika, 1,* 195 – 210.
Strauss, B. (1999). Latent trait and latent class models. *International Journal of Sport Psychology*, *30(1)*, 17-40.
Ter Hark, M. (1990). *Beyond the Inner and the Outer: Wittgenstein's Philosophy of Psychology.* London: Kluwer.
Thomson, G. (1946). *Factorial Analysis of Human Ability.* London: University of London Press.
Thomson, G. (1939). *The Factorial Analysis of Human Ability.* London: University of London Press.

Thomson, G. (1935). The definition and measurement of “g” (general intelligence). *The Journal of Educational Psychology, 26(4),* 241-262.
Thomson, G. (1934). The meaning of ‘*i*’ in the estimate of ‘g’. *British Journal of Psychology, 25, 92-99.*
Thorndike, E.L., Lay, W., & Dean, P.R. (1909). The relation of accuracy in sensory discrimination to general intelligence. *American Journal of Psychology, 20,* 364 – 369.
Thurstone, L.L. (1925). A method of scaling psychological and educational tests. *Journal of Educational Psychology, 16,* 433-451.
Thurstone, L. L. (1927). A law of comparative judgment. *Psychological Review, 34,* pp.273–286.
Thurstone, L. L. (1931). Multiple factor analysis. *Psychological Review, 38,* 406-427.
Thurstone, L.L. (1947). *Multiple Factor Analysis.* Chicago: The University of Chicago Press.
Thurstone, L. L (1950). Primary mental abilities.*Psychometric Monographs*, No. 1, 1938. pp. ix + 121.
Thurston, L.L. and Chave, E.J. (1929) *The Measurement of Attitude.* Chicago, University of Chicago Press.
Topping, J. (1955). *Errors of observation and their treatment.* London: Institute of Physics.
Torgerson, W.S. (1958). *Theory and Methods of Scaling*. New York: John Wiley and Sons.
Tucker, L. R (1946). Maximum validity of a test with equivalent items. *Psychometrika*, *11*, 1946, 1-13.
Tucker, L. (1955). The objective definition of simple structure in linear factor analysis. *Psychometrika, 20,* 209-225.
Tucker, L.R. (1971). Relations of factor score estimates to their use. *Psychometrika*, *36(4),* 427-436.
Tuomela, R. (1973). *Theoretical concepts.* New York: Springer-Verlag.
Tyrrell, D. & Fielder, M. (2002). *Cold Wars.* Oxford: Oxford University Press.
Urban, F.M. (1910). Die psychophysichen mass methoden als grundlagen empirischer messung (continued). *Archive Fur Die Gesampte Psychologie, 16,* 168 – 227.
van der Leeden, R. (1990). *Reduced rank regression with structured residuals.* Leiden: DSWO Press.
Velicer, W F., & Jackson, D. N. (1990a). Component analysis versus common factor analysis: Some issues in selecting an appropriate procedure. *Multivariate Behavioral Research, Vol 25(1),* 1-28.
Velicer, W. F., & Jackson, D. N. (1990b). Component analysis versus common factor analysis: Some further observations. *Multivariate Behavioral Research, 25(1),* 97-114.
Velleman, P., & Hoaglin, D. (1981). *Applications, Basics, and Computing of Exploratory Data Analysis.* Boston: Duxbury Press.
Vincent, D. F. (1953). The origin and development of factor analysis. *Applied Statistics, 2,* 107 – 117.

Vittadini, G. (1989). Indeterminacy problems in the LISREL model. *Multivariate Behavioral Research, 24(4),* 397-414.
Waismann, F. (1965). *Principles of linguistic philosophy*. Edited by R. Harré. London: Macmillan.
Waller, N.G., & Meehl, P. (1998). *Multivariate taxometric procedures: Distinguishing types from continua.* Thousand Oaks: Sage.
Wiley, D.E.(1973)."The Identification Problem for Structural Equation Models with Unmeasured Variables", in A.S.Goldberger and O.D. Duncan (Eds.) *Structural Equation Models in the Social Sciences*. New York: Academic Press.
Williams, J. (1978). A definition for the common factor analysis model and the elimination of problems of factor score indeterminacy. *Psychometrika,43*, 293-306.
Wilson, E.B. (1928a). On hierarchical correlation systems. *Proceedings of the National Academy of Science, 14,* 283-291.
Wilson, E. B. (1928b) Review of 'The abilities of man, their nature and measurement' by C. Spearman. *Science, 67*, 244-248.
Wilson, E. B. (1929a). Review of 'Crossroads in the mind of man: A study of differentiable mental abilities' by T. L. Kelley. *Journal of General Psychology, 2,* 153-169.
Wilson, E. B. (1929b). Comment on Professor Spearman's note. *Journal of Educational Psychology*, *20*, 217-223.
Wilson, E. B.. & Worcester, J. (1939). Note on factor analysis. *Psychometrika, 4,* 133 – 148.
Wilson, D., Wood, R., & Gibbons, R. (1991). *TESTFACT. Test scoring, item statistics, and item factor analysis.* Chicago: Scientific Software International.
Wittgenstein, L. (1953). *Philosophical Investigations*. Oxford: Basil Blackwell Ltd.
Wolfle, D. (1940). *Factor Analysis to 1940. Psychometric Monographs*, *3*. Chicago: University of Chicago Press.
Worrall, J. (1982). Scientific realism and scientific change. *The Philosophical Quarterly, 32,* 201-231.
Wright, S. (1918). On the nature of size factors. *Genetics 3,* 367 - 374.
Wright, S. (1921). Correlation and causation. *Journal of Agricultural Research, 20*, 557–585.