Biometrika (2004), **91**, 3, *pp.* 738–742

Printed in Great Britain

Measurement exchangeability and normal one-factor models

By HENK KELDERMAN

Department of Work and Organizational Psychology, Vrije Universiteit Amsterdam, van der Boechorststraat 1, 1081 BT Amsterdam, The Netherlands

h.kelderman@psy.vu.nl

Summary

The one-factor model restricts the covariance structure of the observed variables on the basis of assumptions about their relationship with an unobserved variable. It is hard to justify these assumptions on substantive or empirical grounds. In this paper, alternative measurement models are proposed that are based on exchangeability of variables after admissible scale transformations. They provide an alternative interpretation of the model and do not involve unobserved variables. They also yield a new one-factor model for sum scales.

Some key words: Difference scale; Exchangeability; Factor analysis; Factor indeterminancy; Interchangeability; Interval scale; Latent trait model; Measurement model; Permutation invariance; Symmetry model.

1. Introduction

Huynh (1978), and more recently Schuster (2001), based a model for raters on de Finetti's (1937) notion of exchangeability to determine whether or not 'it is indifferent as to which of the raters is used'. The idea that measurements for the same attribute should be exchangeable dates back to Gulliksen (1968). He proposed that measurements be interchangeable in the sense that 'it is indifferent as to which of the measures is used'.

Huynh's idea of exchangeable raters is as follows. Let the random variable X_i denote the rating of rater $i \in M$, with realisation x_i. Let $X = (X_i; i \in M)$ denote the vector of ratings of $|M|$ raters. If it is completely immaterial which of the raters is used, their ratings should have perfectly identical statistical properties. In that case, the joint distribution $f_X(x)$ should be exchangeable, that is permutation invariant in the x_i's:

$$f_X(x) = f_X(x^*), \tag{1·1}$$

where $x^* = \text{perm}(x)$. Obviously, the idea is also applicable to other instruments that purport to measure the same attribute.

By symmetry, (1·1) implies that the measures' associations are all equal and that the marginal distributions are all the same; that is,

$$f_{X_c}(x_c) = f_{X_{c'}}(x_{c'}), \tag{1·2}$$

for $c \cap c' = \varnothing$, $|c| = |c'|$ and $x_c = x_{c'}$. However, because raters may respond on different scales and with different precision, marginal distributions of ratings need not be identical. Let s_i be a monotone increasing transformation of the original measure U_i into rescaled measure X_i and lift the restrictions that the marginal distributions be the same. Measurement exchangeability can then be defined as follows.

Definition 1. *Measurement exchangeability of U obtains if, for all $c \subset M$ and monotone increasing transformations $s_i : U_i \to X_i$,*

$$f_{X_d|X_c}(x_d|x_c) = f_{X_d|X_c}(x_d|x_c^*), \tag{1·3}$$

where d is defined as $M \backslash c$ throughout this paper.

It is easily shown that (1·3) together with (1·2) is equivalent to (1·1).

A specific measurement exchangeable model is obtained by specifying the type of distribution f and an appropriate class of scale transformations. In most applications it suffices to choose the distribution from the exponential family (Barakin & Maitra, 1963; Brown, 1986, p. 22). The conditional distribution of $X_d|X_c$ belongs to an exponential family if the density has the form

$$f_{X_d|X_c}(x_d|x_c) = c^{-1}(\tau_{x_c}) \exp\left\{ \sum_{r=0}^{k} g_r(x_d)\phi_r(\tau_{x_c}) \right\} \tag{1·4}$$

(Andersen, 1980, p. 20), where $\phi_0(\tau_{x_c}) = 1$ and τ_{x_c} are the model parameters. If the functions $g_0(x_d), \ldots, g_k(x_d)$ are linearly independent $\zeta_{x_c} = (\phi_1(\tau_{x_c}), \ldots, \phi_k(\tau_{x_c}))$ are the canonical parameters. Clearly, (1·4) is permutation invariant in the elements of x_c if and only if $\zeta_{x_c} = \zeta_{x_c^*}$. Furthermore, (1·4) is permutation invariant in the elements of x_c if and only if $\tau_{x_c} = \tau_{x_c^*}$ and $\phi = (\phi_1, \ldots, \phi_k)$ is a one-to-one transformation.

Now consider the case of the normal distribution.

2. Measurement exchangeability in the normal distribution

The normal distribution remains normal after linear transformations of the variables. Therefore, admissible transformations are linear, $X_i = a_i + b_i U_i$, additive, $X_i = a_i + U_i$, multiplicative, $X_i = b_i U_i$, and identity, $X_i = U_i$, for interval, difference, ratio and absolute scales respectively (Stevens, 1946). Note that, since the transformations must be monotone increasing, $b_i > 0$. First, consider the identity transformation.

In conditional normal distributions, the parameters $\tau_{u_c} = (\mu_{d|u_c}, \Sigma_{dd|c})$, are one-to-one transformations of the canonical parameters, where $\mu_{d|u_c}$ is the conditional mean vector given u_c and $\Sigma_{dd|c}$ is the conditional covariance matrix. These conditional parameters are related to the unconditional parameters $\tau = (\mu_d, \mu_c, \Sigma_{dd}, \Sigma_{dc}, \Sigma_{cc})$ by

$$\mu_{d|u_c} = \mu_d + B_{d\cdot c}(u_c - \mu_c), \quad \Sigma_{dd|c} = \Sigma_{dd} - B_{d\cdot c}\Sigma_{dc}^{\mathrm{T}}$$

(Morrison, 1990, p. 92), where $B_{d\cdot c} = \Sigma_{dc}\Sigma_{cc}^{-1}$, is the matrix of regression coefficients; for simplicity, assume that Σ_{MM} is nonsingular.

Additive transformations of the U_i's do not affect $\tau(u_c) = (\mu_{d|u_c}, \Sigma_{dd|c})$, because they affect neither covariances nor $u_c - \mu_c$. Consequently, normal measurement exchangeable models that admit identity and additive transformations are observationally equivalent.

The following theorem about the (co)variance parameters of the joint distribution $f_U(u)$ can now be proven; see the Appendix for the proof.

Theorem 1. *If U has a multivariate normal distribution then*

$$\sigma_{ij} = v_i v_j, \tag{2·1a}$$

$$\sigma_i^2 = v_i^2 + v_i\omega, \tag{2·1b}$$

where $i \neq j \in M$, for some $v_i, v_j > 0$ and $\omega > 0$, is equivalent to measurement exchangeability allowing additive transformations $s_i : U_i \to X_i$.

It is easily seen that (2·1) yields correlations

$$\rho_{ij} = \left\{ \frac{v_i v_j}{(v_i + \omega)(v_j + \omega)} \right\}^{\frac{1}{2}}. \quad (2\cdot2)$$

Thus, the sizes of the correlations of a particular measure depend positively on its v_i and the sizes of all correlations negatively on ω.

Next, consider the case of linear transformations. Since (co)variances are bilinear in their arguments, one has, for $X_i = a_i + b_i U_i$,

$$\text{cov}(U_i, U_j) = b_i^{-1} \text{cov}(X_i, X_j) b_j^{-1}, \quad \text{var}(U_i) = b_i^{-2} \text{var}(X_i).$$

If $\text{cov}(X_i, X_j)$ and $\text{var}(X_i)$ satisfy the additive measurement exchangeable model, one has the following theorem for $\sigma_{ij} = \text{cov}(U_i, U_j)$ and $\sigma_i^2 = \text{var}(U_i)$.

THEOREM 2. *If U has a multivariate normal distribution then*

$$\sigma_{ij} = \lambda_i \lambda_j, \quad (2\cdot3a)$$

$$\sigma_i^2 = \lambda_i^2 + \psi_i, \quad (2\cdot3b)$$

where $i \neq j \in M$, $\lambda_i = b_i^{-1} v_i > 0$ *and* $\psi_i = b_i^{-2} \omega v_i > 0$, *is equivalent to measurement exchangeability allowing linear transformations* $s_i : U_i \to X_i$.

From Theorem 2, it is seen that the linear measurement exchangeable model is formally identical to the standard one-factor model with positive loadings (Spearman, 1904; Lawley & Maxwell, 1971, p. 7). In the development of the standard one-factor model, λ_i is defined as the factor loading, i.e. a parameter that describes the dependence of the observed measure U_i on an unobserved common factor θ, and ψ_i is defined as the variance of the residual score $\mu_i + \lambda_i \theta - U_i$. The model for additive measurement exchangeability is a special case of that for linear measurement exchangeability. It is formally identical to, say, an additive one-factor model, of which the distinctive feature is that the residual variances are assumed to be proportional to the loadings, that is $\psi_i = \omega v_i$.

If the orientation, high or low, of the measurements is unknown, one may drop the assumption that the scale transformation function be increasing. In that case, the scale transformation is only assumed to be monotone and b_i may be negative so that λ_i may be negative. The linear measurement exchangeable model then becomes formally identical to the standard one-factor model and the additive measurement exchangeable model then becomes equivalent to an additive one-factor model with $b_i \in \{-1, 1\}$.

All models can be estimated and tested using the standard theory of structural equation models (Jöreskog, 1970). The Mx program (Neale et al., 2002) is pre-eminently suited to computing the estimates of the v and ω, as well as suitably constrained scale-transformation parameters. One may also compare the fit of linear and additive measurement exchangeable models.

3. DISCUSSION

In this paper a new type of measurement model is proposed that is based on relaxations of exchangeability of observed variables. Normal measurement exchangeable models are formally identical to the corresponding one-factor models. However, the assumptions from which both models are derived are quite different. Measurement exchangeable models require that transformed measures of the same attribute should be exchangeable. On the other hand, one-factor models require that they should uniquely depend on a common factor. Bartholomew (1984; 1987, Ch. 1), Goldstein (1980) and Ramsay (1996) have noticed methodological problems with assumptions based on unobserved variables. The upshot of their observation is that, because a factor is completely unobserved, assumptions about its distribution and its relationships with observed measures are essentially arbitrary and hard to justify on substantive or empirical grounds. Researchers may find assumptions about exchangeability, scale types and distributions easier to justify.

Appendix

Proof of Theorem 1

For all $c \in M$ and $u_c \in \Re^{|c|}$, one has the equivalences

$$f_{U_d|U_c}(u_d|u_c) = f_{U_d|U_c}(u_d|u_c^*) \Leftrightarrow \mu_d + B_{d\cdot c}(u_c - \mu_c) = \mu_d + B_{d\cdot c}(u_c^* - \mu_c) \Leftrightarrow$$
$$B_{d\cdot c}u_c = B_{d\cdot c}u_c^* \Leftrightarrow B_{d\cdot c}u_c = B^*_{d\cdot c}u_c \Leftrightarrow B_{d\cdot c} = B^*_{d\cdot c},$$

where B^* denotes the permutation of the columns of B corresponding to the permutation u^* of the elements of u. From the last equation, one has $B_{d\cdot c} = \beta_d^d 1_c^{\mathrm{T}}$, where 1_c is a vector of ones of length $|c|$, and where in β_q^p $(p \geqslant q)$ the subscript denotes the elements of the vector and the superscript the equation in which they occur. From this and $B_{d\cdot c} = \Sigma_{dc}\Sigma_{cc}^{-1}$ one has $\Sigma_{dc} = \beta_d^d \beta_c^{c\mathrm{T}}$, where

$$\beta_c^c = \Sigma_{cc} 1_c. \tag{A·1}$$

The values of the elements of β_d^d and β_c^c vary across different equations $c \subset M$, but, because $\sigma_{ji} = \beta_j^d \beta_i^c = \beta_j^{d'} \beta_i^{c'}$ for $i \in c \cap c' \neq \varnothing$, $j \in d \cap d' \neq \varnothing$ and $c, c' \subset M$, there exist v_c, v_d and scalar e^c such that

$$v_c = \beta_c^c (e^c)^{-1}, \quad v_d = \beta_d^d e^c, \tag{A·2}$$

so that $\Sigma_{dc} = v_d v_c^{\mathrm{T}}$. This proves the necessity of (2·1a).

To prove the necessity of (2·1b), first define $\eta_i = \sigma_i^2 - v_i^2$ and denote the vector by $\eta_c = (\eta_i;\, i \in c)$. Post-multiplying $\Sigma_{cc} = v_c v_c^{\mathrm{T}} + \mathrm{diag}(\eta_c)$ by 1_c, one obtains from (A·1) and (A·2) that

$$\Sigma_{cc} 1_c = \beta_c^c = v_c e^c = v_c(1_c^{\mathrm{T}} v_c) + \eta_c.$$

Solving the last equation for η_c yields

$$\eta_c = \omega^c v_c, \tag{A·3}$$

with $\omega^c = (e^c - 1_c^{\mathrm{T}} v_c)$. From (A·3) one has that ω^c is constant, so that $\omega^c = \omega$. Since values of v_c, η_c and ω do not depend on the equation c, the result holds for all $c \subset M$, and so for $c = M$. Thus the necessity of (2·1b) for measurement exchangeability follows.

To prove the sufficiency of (2·1a) and (2·1b) for measurement exchangeability note that

$$\begin{aligned} B_{d\cdot c} &= \Sigma_{dc}\Sigma_{cc}^{-1} = v_d v_c^{\mathrm{T}} \{v_c v_c^{\mathrm{T}} + \omega\, \mathrm{diag}(v_c)\}^{-1} = v_d 1_c^{\mathrm{T}}\, \mathrm{diag}(v_c)\{v_c 1_c^{\mathrm{T}}\, \mathrm{diag}(v_c) + \omega\, \mathrm{diag}(v_c)\}^{-1} \\ &= v_d 1_c^{\mathrm{T}} (v_c 1_c^{\mathrm{T}} + \omega I_c)^{-1} = v_d 1_c^{\mathrm{T}} \{-v_c 1_c^{\mathrm{T}} \omega^{-1} (\omega + v_c^{\mathrm{T}} 1_c)^{-1} + I_c \omega^{-1}\} \\ &= v_d 1_c^{\mathrm{T}} \omega^{-1} \{-v_c 1_c^{\mathrm{T}} (\omega + v_c^{\mathrm{T}} 1_c)^{-1} + I_c\} = v_d \omega^{-1} \{1_c^{\mathrm{T}} - 1_c^{\mathrm{T}} v_c 1_c^{\mathrm{T}} (e^c)^{-1}\} \\ &= v_d 1_c^{\mathrm{T}} \omega^{-1} \{1 - 1_c^{\mathrm{T}} v_c (e^c)^{-1}\} = v_d 1_c^{\mathrm{T}} \omega^{-1} (e^c)^{-1} (e^c - 1_c^{\mathrm{T}} v_c) = v_d 1_c^{\mathrm{T}} (e^c)^{-1} = \beta_d^d 1_c^{\mathrm{T}}, \end{aligned}$$

where I_c is a $|c| \times |c|$ identity matrix. This proves that the matrix of regression weights has identical columns. Thus (2·1a) and (2·1b) are sufficient for measurement exchangeability.

To prove that $v_i > 0$ and $\omega > 0$, note first that, under the constraints (1·1), by symmetry, all covariances have the same sign. Note further that, from (1·3), all possible conditional distributions can be obtained by integrating out zero or more u_j's $(j \in d)$. By repeated application of the product rule of conditional probabilities, the joint distribution in (1·1) can be written as products of conditional distributions from (1·3) and a one-variable marginal distribution, so that, under (1·1), all one-variable marginal distributions must be the same. Therefore, the difference between (1·1) and (1·3) is the restriction on the one-variable marginal distributions. Since one-variable marginal distributions do not influence the sign of the covariances, the covariances under (1·1) have the same sign as those under (1·3). Furthermore, because under (1·1) the covariances all have the same sign, the covariances under measurement exchangeability (1·3) all have the same sign. The only way to obtain this from the additive measurement exchangeable model (2·1) is to have $v_i > 0$. Furthermore, from (2·2) with $\omega < 0$ and $v_i > 0$, one has $\rho_{ij} > 1$, which is impossible, so that $\omega \geqslant 0$. For $f(U)$ to satisfy the usual regularity conditions, the parameter space should be open so that $\omega > 0$. This completes the proof.

REFERENCES

ANDERSEN, E. B. (1980). *Discrete Statistical Models with Social Science Applications.* Amsterdam: North Holland.

BARAKIN, E. W. & MAITRA, E. P. (1963). Generalisation of the Fisher-Darmois-Koopman-Pitman theorem on sufficient statistics. *Sankhyā* A **25**, 217–44.

BARTHOLOMEW, D. J. (1984). The foundations of factor analysis. *Biometrika* **71**, 221–32.

BARTHOLOMEW, D. J. (1987). *Latent Variable Models and Factor Analysis.* London: Griffin.

BROWN, L. D. (1986). *Fundamentals of Statistical Exponential Families with Applications in Statistical Decision Theory*. Hayward, CA: Institute of Mathematical Statistics.

DE FINETTI, B. (1937). La prévision: ses lois logiques, ses resources subjectives. *Ann. Inst. Henri Poincaré* **7**, 1–68.

GOLDSTEIN, H. G. (1980). Dimensionality, bias, independence and measurement scale problems in latent trait test score models. *Br. J. Math. Statist. Psychol.* **33**, 234–46.

GULLIKSEN, H. (1968). Methods for determining equivalence of measures. *Psychol. Bull.* **70**, 534–44.

HUYNH, H. (1978). Reliability of multiple classifications. *Psychometrika* **43**, 317–25.

JÖRESKOG, K. G. (1970). A general method for the analysis of covariance structures. *Biometrika* **57**, 239–51.

LAWLEY, D. N. & MAXWELL, A. E. (1971). *Factor Analysis as a Statistical Method.* London: Butterworth.

MORRISON, D. F. (1990). *Multivariate Statistical Methods.* New York: McGraw-Hill.

NEALE, M. C., BOKER, S. M., XIE, G. & MAES, H. H. (2002). *Mx*: *Statistical Modeling*, 6th ed. Richmond, VA: Dept Psychiatry, Virginia Commonwealth University.

RAMSAY, J. O. (1996). A geometrical approach to item response theory. *Behaviormetrika* **23**, 3–17.

SCHUSTER, C. (2001). Kappa as a parameter of a symmetry model for rater agreement. *J. Educ. Behav. Statist.* **26**, 331–42.

SPEARMAN, C. (1904), "General intelligence", objectively determined and measured. *Am. J. Psychol.* **5**, 201–92.

STEVENS, S. (1946). On the theory of scales of measurement. *Science* **103**, 677–80.

[*Received May* 2001. *Revised October* 2003]