Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Journal of Multivariate Analysis 96 (2005) 172 – 189
Journal of Multivariate Analysis
www.elsevier.com/locate/jmva

Measurement errors in multivariate measurement scales

L. Tarkkonen, K. Vehkalahti*

Department of Mathematics and Statistics, University of Helsinki, P.O. Box 54, FI-00014, Finland

Received 25 October 2003
Available online 11 November 2004

Abstract

Our aim is to construct a general measurement framework for analyzing the effects of measurement errors in multivariate measurement scales. We define a measurement model, which forms the core of the framework. The measurement scales in turn are often produced by methods of multivariate statistical analysis. As a central element of the framework, we introduce a new, general method of estimating the reliability of measurement scales. It is more appropriate than the classical procedures, especially in the context of multivariate analyses. The framework provides methods for various topics related to the quality of measurement, such as assessing the structural validity of the measurement model, estimating the standard errors of measurement, and correcting the predictive validity of a measurement scale for attenuation. A proper estimate of reliability is a requisite in each task. We illustrate the idea of the measurement framework with an example based on real data.

AMS 2000 subject classification: 62H25; 62J99; 62H20; 91C05; 97C40

Keywords: Measurement error; Measurement scale; Factor analysis; Reliability; Validity

1. Introduction

In statistical research, we are often interested in estimating some population parameters based on a random sample. The uncertainty then comes from the sampling. Sometimes, however, our data include all records under study, and no sampling is then needed. It is also

* Corresponding author. Fax: +358 9 191 24 872.
E-mail address: Kimmo.Vehkalahti@helsinki.fi (K. Vehkalahti).

0047-259X/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmva.2004.09.007

possible to work in the individual level, without a need to consider how the results would be generalized to a population level. In some circumstances, it might even be difficult to define the population.

The other source of uncertainty comes from the measurement, which is an essential concept in science. Measurement is needed regardless of the sampling procedure. As the conclusions of empirical studies are based on values measured on research objects, it is crucial to assess the quality of the measurements.

The most important property of measurement is *validity*. Broadly stated, validity is concerned with whether a measuring instrument measures what it is supposed to measure in the context in which it is to be applied. In addition, the measurements should be reliable, in the sense that the researchers can rely on the precision of the measuring instrument. The precision is stated by *reliability*, the ratio of the true variance to the total variance of the measurement. The true variance excludes the variance caused by the random *measurement error*. Reliability defines the resolution of the measurement, and tells us how small differences we can talk about.

Traditional statistical models concentrate on the sampling variation, and often treat the measurement errors with neglect, e.g., by including them in the sampling variation, or simply by assuming that the subjects are measured without error. This is rather vague because sampling and measurement are clearly different procedures. In a given study, it might be useful to find out, which one is the main source of uncertainty. If the measurement is inaccurate, increasing the sample size will not improve it. Instead, the measurement errors should be taken into account by using suitable approaches of modeling.

In this paper, we construct a general framework for analyzing the effects of measurement errors in multivariate measurement scales. The central concepts of the framework are *measurement model* and *measurement scale*. The measurement model relates our framework to the factor analysis model [20] and its generalizations [6,10,18], but it is also a multidimensional generalization of the classical, one-dimensional true score model [22, Chapter 3]. Throughout this paper, we assume that the subjects are random but the items are fixed, i.e. we do not consider random sampling of items (see [21]).

In many application areas it is typical to create one-dimensional scales, e.g., preference scales or predictive regression scales. Nevertheless, the scales are constructed from measurements of several multidimensional attributes. Therefore, we term these scales *multivariate measurement scales*. They connect our measurement framework with various methods of multivariate statistical analysis, such as regression analysis, canonical correlations, or discriminant analysis.

It is essential to ensure that the scales are reliable, i.e. the variation caused by the measurement error is minimized. Therefore, as a central element of the framework, we introduce a new, general method of estimating the reliability of multivariate measurement scales. We establish our method on the classical definition of reliability, and show that the well-known Cronbach's α [12] is a restricted special case of our method.

1.1. Historical background

The concept of reliability is due to Charles Spearman already at the turn of the 20th century. In 1904, Spearman [27] proposed a formula for correcting the effects of