Quantifying measurement error

Stephen Lewis

Abstract

It is important for workers to have some estimate of the degree of error evident when measuring objects. Although many use their own "rule-of-thumb" to give them the personal satisfaction that they are working accurately, measures of error, or conversely reliability, are rarely given in the literature. Some simple, useful equations are given that may be used privately or when reporting metrical work.

Introduction

Reports containing measured parameters rarely include any indication of the error or, alternatively, the reliability of those measurements. Estimates of error are sometimes used by workers privately as something of a "rule-of-thumb" by which to work. For example, sometimes an average difference divided by the mean value of the sample may be calculated to give a crude percentage error. Mueller and Martorell (1988) and Frisancho (1990) have proposed two estimates - the technical error of measurement (TEM) and reliability (R) - which they have applied to anthropometric measurement techniques. Although Ulijaszek and Lourie (1994) have challenged some of the more applied findings by these workers, they support the validity of using these estimates. It may be useful to describe these estimates here so that they may be considered for wider use in different contexts.

Technical Error of Measurement

When a measurement is taken on the same object on more than one occasion, the value obtained will not always be the same. This produces what is referred to as the technical error of measurement (TEM). In its simplest form, this may be used to determine intra-observer error. However, it may also be used to determine inter-observer error as occurs when two workers are independently measuring the same things. TEM is calculated using the equation:

$$TEM = \sqrt{\frac{\sum D^2}{2N}}$$

Eq. 1

where:
D is the difference between measurements made on a given object on two occasions (or by two workers),
and

N is the total number of measurements made on those two occasions i.e. if 10 given objects were measured each time, then N=20.
(More elaborate forms of this equation are available which determine the TEM where more than two measurers are working. (See, for example, Ulijaszek and Lourie, 1994).)

Example 1: Involved in a project to measure a series of bone lengths, a worker wishes to determine his (intra-observer) TEM. That worker takes a series of bones at random and makes the same measurement on each bone on two separate occasions - on different days, for example. For each of these objects, the measurement made on the first occasion is compared with that taken on the second and the difference recorded. Using Eq. 1, the worker determines his TEM and calculates the square root of the sum of the squares of these differences divided by twice the total number of measurements made.
Example 2: Two workers are collaborating on a project which entails taking a series of bone lengths independently and then pooling their results. These workers wish to determine their (inter-observer) TEM. Taking a series of bones at random, each worker measures the same bones once. The difference between the measurements for each bone is recorded. Once again, the TEM is calculated by taking the square root of the sum of the squares of these differences divided by twice the total number of measurements made.

The TEM is expressed in the same units as those used to make the original measurements and is appropriate only for that particular measurement. It is not a generalised indication of a worker's ability to take measurements accurately but rather an indication of the amount of error he demonstrates (or a pair of collaborators demonstrate) when making that measurement.

The coefficient of reliability
The coefficient of reliability (R) is an estimate independent of the units of measurement used. It ranges from 0 to 1 or may be thought of or expressed as a percentage. It is calculated using the equation:

$$R = 1 - \left(\frac{(TEM)^2}{(SD)^2} \right)$$

Eq. 2

where:
TEM is as calculated in Eq. 1, and
SD_2 is the total inter-subject variance (standard deviation squared), irrespective of measurement error, of ALL the measurements taken when determining the TEM.

R, therefore, expresses the proportion of the between-subject variance that is free from measurement error. For example, if R=0.95 then 95% of the variance is due to factors unrelated to measurement error. As with TEM, R is not a generalised indication of a worker's reliability but applies to a specific measured parameter.

R has not been used widely and so values to which workers should strive are largely unknown even for parameters where TEM has been reported. It has been suggested by Himes (1989) that workers should set up their own reliability studies and determine the levels of R necessary for their own particular purposes.

Corollary — Evaluating SEM given R

By re-arranging Eq. 2, it is possible to determine those TEM values necessary to produce a given R value. The re-arranged equation being:

$$TEM = \sqrt{\left(\left(1 - R\right) \times SD^2\right)}$$

Eq. 3

Conclusion

That workers have been reluctant to include in their findings estimates of error is quite understandable. To do so would appear to undermine the efforts they have exerted in obtaining their results. Despite the recommendations of some (Mueller and Martorell, 1988; Himes, 1989; Frisancho, 1990; Ulijaszek and Lourie, 1994), it remains to be seen whether publication of estimates of error becomes common practice. However, the equations provided here will provide workers with a means of measuring their technical error of measurement and reliability - in private at least.

References

Frisancho, A.R. 1990 *Anthropometric Standards for Assessment of Growth and Nutritional Status*. Ann Arbor, University of Michigan Press.
Himes, J.H. 1989 'Reliability of anthropometric methods and replicate measurements', *American Journal of Physical Anthropology* 79: 77-80.
Mueller, W.H. and Martorell, R. 1988 'Reliability and accuracy of measurement', in Lohman, T. G., Roche, A. F. and Martorell, R. (eds). *Anthropometric Standardisation Reference Manual*, pp 83-86. Human Kinetics Books Champaign, Illinois.
Ulijaszek, S.J. and Lourie, J.A. 1994 'Intra- and inter-observer error on anthropometric measurement', in Ulijaszek, S.J. and Mascie-Taylor, C.G.N. (eds), *Anthropometry: The individual and the population*, pp30-55. Cambridge University Press, Cambridge.