Medical Teacher

ISSN: 0142-159X (Print) 1466-187X (Online) Journal homepage: http://www.tandfonline.com/loi/imte20

The reliability of assessments: The Bayesian Cronbach's alpha

Mohsen Tavakol

To cite this article: Mohsen Tavakol (2017) The reliability of assessments: The Bayesian Cronbach's alpha, Medical Teacher, 39:5, 561-561, DOI: 10.1080/0142159X.2017.1296121

To link to this article: https://doi.org/10.1080/0142159X.2017.1296121

Published online: 01 Mar 2017.

Submit your article to this journal

Article views: 435

View Crossmark data

Full Terms & Conditions of access and use can be found at
http://www.tandfonline.com/action/journalInformation?journalCode=imte20

MEDICAL TEACHER, 2017
VOL. 39, NO. 5, 561–563

MEDICAL TEACHER
Taylor & Francis
Taylor & Francis Group

LETTERS TO THE EDITOR

The reliability of assessments: The Bayesian Cronbach's alpha

Dear Sir

To ensure the precision of assessment questions as a proxy for measuring student performance, reliability estimates of test scores are required. In the psychometrics literature, an assortment of approaches of reliability estimates has been reported, but Cronbach's alpha (expressed as alpha) is used ubiquitously for providing evidence of the precision of test scores. The reasons for the popularity of alpha lie in the fact that it can be estimated by a single administration of assessment questions or it is easy to compute using statistical applications.

The use of alpha poses several problems. (1) *Test items*. Alpha assumes that all assessment questions are tau-equivalent, meaning that the items of a test measure a single underlying construct (Tavakol & Dennick 2011). If test items are tau-equivalent, the estimated reliability is equivalent to a coefficient alpha. If this assumption is not reached, the estimated alpha is a lower bound of the reliability of test scores. In practice, the tau-equivalent assumption is usually violated, and hence alpha provides a below average measure for the reliability population. (2) *Samples*. Alpha is also sensitive to outliers, skewness and sample size. A deviation from normality and small samples (e.g. sub-groups) can both influence the estimated alpha, resulting in inaccurate estimates of reliability (Oh et al. 2009). (3) *Testing variation*. Last, variations occurring amongst administrations of the exam could lead to variations in the estimated alpha, leaving us uncertain about the reliability of test scores.

The primary goal of Bayesian methods is to identify probability distributions when we are uncertain about the parameter of interest (e.g. alpha). Therefore, to provide stable interpretations of a coefficient alpha, I used a Bayesian method for a synthetic OSCE with a different number of stations and examinees. Using the Markov Chain Monte Carlo (MCMC) with Metropolis-Hastings, I generated 10,000 samples from the distribution of the posterior of alpha. Finally, I derived a Bayesian estimation of Cronbach's alpha (with other statistics, e.g. confidence intervals), assuming a "flat" prior for Cronbach's alpha. Simulations show that the Bayesian alpha provides a better inference of the reliability of test scores compared to the traditional alpha when sample sizes are small. I am currently modeling the precision of assessment questions with a prior knowledge of the alpha when the assumptions of alpha were violated.

Disclosure statement

The author has no declarations of interest.

References

Oh H, Guo H, Walker M. 2009. Improved reliability estimates for small samples using empirical Bayes techniques. ETS Research Report. New Jersey: ETS.

Tavakol M, Dennick R. 2011. Making sense of Cronbach's alpha. Int J Med Educ. 2:53–55.

Mohsen Tavakol
School of Medicine, The University of Nottingham, Nottingham, UK
mohsen.tavakol@nottingham.ac.uk

© 2017 Informa UK Limited, trading as Taylor & Francis Group
10.1080/0142159X.2017.1296121

Almost everything is the realm of the rich!

Dear Sir

What Walsh et al. (2016) stated regarding the number of medical education studies in rich and poor countries did not really surprise me because this problem does not concern only medical education, but all branches of knowledge. According to SCImago Journal & Country Rank portal (www.scimagojr.com), 87% of documents indexed in Scopus database (more than 35 out of 40.5 million records) during 1996–2015 were published by 10% of countries (25 countries), while only 13% of documents were published by the remaining 90% (214 countries). The end of the list 100 countries (mostly poor) only published 42,618 documents (0.1%). Hence, medical education is not alone in this catastrophe! There is also another study that concluded that the majority of medical education articles had been published by only a few countries (Tutarel 2002).

My next remark concerns PubMed (GoPubMed uses PubMed records). Until 2014, PubMed only included the first author's affiliation. Therefore, information in this section is highly imperfect, and number of articles by countries cannot be accurately determined. Perhaps, using Scopus that covers the entire MEDLINE would have been the better choice. However, even if Scopus had been used, it would have made no difference to the author's final conclusion.

Science, culture, research and development, and democracy are usually the first to be neglected in a country experiencing poverty, because the needs in lower strata of Maslow's pyramid are more immediate and must be satisfied first.