Assessment and Development of new Statistical Methods for the Comparability of Scores

Inés Varas [1,2]

[1]Pontificia Universidad Católica de Chile
[2] Laboratorio Interdisciplinario de Estadística Social

Advisor: Jorge González
Co-Advisor: Fernando Quintana

Laboratorio
Interdisciplinario
de Estadística Social

Outline

Laboratorio Interdisciplinario de Estadística Social

Importance of Equating

- It is common for measurement programs to produce different forms of a test that are intended to measure the same attribute.

Laboratorio Interdisciplinario de Estadística Social

Importance of Equating

- It is common for measurement programs to produce different forms of a test that are intended to measure the same attribute.

- Because test scores are used to make important decisions in various settings, it is needed to report scores in a fair and precise way.

Laboratorio Interdisciplinario de Estadística Social

Importance of Equating

- It is common for measurement programs to produce different forms of a test that are intended to measure the same attribute.

- Because test scores are used to make important decisions in various settings, it is needed to report scores in a fair and precise way.

- The main idea behind Equating is *to treat scores as if they come from the same test*.

Laboratorio Interdisciplinario de Estadística Social

Equating: statistical point of view

- Let consider two test forms X and Y. The scores X and Y are assumed to be random variables.

Laboratorio Interdisciplinario de Estadística Social

Equating: statistical point of view

- Let consider two test forms X and Y. The scores X and Y are assumed to be random variables.
- The scores of these two test forms are defined on score sample spaces $\mathscr{X}$ and $\mathscr{Y}$ respectively.

Laboratorio Interdisciplinario de Estadística Social

Equating: statistical point of view

- Let consider two test forms X and Y. The scores X and Y are assumed to be random variables.

- The scores of these two test forms are defined on score sample spaces $\mathscr{X}$ and $\mathscr{Y}$ respectively.

- In the context of educational measurement these spaces represent the scale of these scores.

Laboratorio Interdisciplinario de Estadística Social

Equating: statistical point of view

- Let consider two test forms X and Y. The scores X and Y are assumed to be random variables.
- The scores of these two test forms are defined on score sample spaces $\mathscr{X}$ and $\mathscr{Y}$ respectively.
- In the context of educational measurement these spaces represent the scale of these scores.
- Let $X_1, \ldots, X_{n_X}$ and $Y_1, \ldots, Y_{n_Y}$ be the scores obtained on the test forms X and Y by n_X and n_Y examinees, respectively.

Equating: statistical point of view

- The statistical problem in equating is to establish the *equating transformation*.

Equating: statistical point of view

- The statistical problem in equating is to establish the *equating transformation*.
- González and Wiberg (2017) give a formal definition of the equating transformation: a mapping between these two score sample spaces i.e.

$$\varphi : \mathscr{X} \rightarrow \mathscr{Y}$$

Laboratorio Interdisciplinario de Estadística Social

Equating: statistical point of view

- The statistical problem in equating is to establish the *equating transformation*.

- González and Wiberg (2017) give a formal definition of the equating transformation: a mapping between these two score sample spaces i.e.

$$\varphi : \mathscr{X} \rightarrow \mathscr{Y}$$

- Thus, the equating transformation maps the scores on the scale of one test form into the scale of the other.

Some equating transformations

Equipercentile Equating

- The most popular equating transformation was defined by Braun and Holland (1982). It is called the *equipercentile function* defined by:

$$\varphi(x) \quad = \quad F_Y^{-1}\left(F_X(x)\right)$$

where F_Y and F_X are the CDF functions of Y and X, respectively.

Laboratorio Interdisciplinario de Estadística Social

Some equating transformations

Equipercentile Equating

- It is important to highlight here that we are interested on equate scores which are usually integer i.e. X and Y are subsets of integer numbers.

Laboratorio Interdisciplinario de Estadística Social

Some equating transformations

Equipercentile Equating

- It is important to highlight here that we are interested on equate scores which are usually integer i.e. X and Y are subsets of integer numbers.

- This represent and important drawback of this method: the inverse function F_Y^{-1} is not well defined.

Laboratorio Interdisciplinario de Estadística Social

Some equating transformations

Equipercentile Equating

- It is important to highlight here that we are interested on equate scores which are usually integer i.e. X and Y are subsets of integer numbers.
- This represent and important drawback of this method: the inverse function F_Y^{-1} is not well defined.
- The continuization step is proposed in order to avoid this problem: linear interpolation, polynomial log-linear models and kernel equating.

Laboratorio Interdisciplinario de Estadística Social

Some equating transformations

Linear Equating

- In this method, the equating transformation is defined by:

$$\varphi(x) \quad = \quad \mu_Y + \frac{\sigma_Y}{\sigma_X}(x - \mu_X)$$

where μ_W and σ_W are the expected value and the standard deviation of W.

Laboratorio Interdisciplinario de Estadística Social

Some equating transformations

Linear Equating

- In this method, the equating transformation is defined by:

$$\varphi(x) \quad = \quad \mu_Y + \frac{\sigma_Y}{\sigma_X}(x - \mu_X)$$

where μ_W and σ_W are the expected value and the standard deviation of W.

- In practice, the parameters involved in this function are estimated as the mean and the standard deviation of each sample.

Laboratorio Interdisciplinario de Estadística Social

Goals of this work

- To extend and propose new methods of test equating based on discrete distributions.

Laboratorio Interdisciplinario de Estadística Social

Goals of this work

- To extend and propose new methods of test equating based on discrete distributions.

- Propose improved methods based on continuization tools.

Laboratorio Interdisciplinario de Estadística Social

Goals of this work

- To extend and propose new methods of test equating based on discrete distributions.
- Propose improved methods based on continuization tools.
- To compare both approaches (discrete-based equating and continuization-based equating).

Laboratorio Interdisciplinario de Estadística Social

Discrete-based equating methods

- Since the linear equating transformation does not need continuization, we propose to estimate it using both parametric and nonparametric models for the discrete score distributions.

Laboratorio Interdisciplinario de Estadística Social

Discrete-based equating methods

- Since the linear equating transformation does not need continuization, we propose to estimate it using both parametric and nonparametric models for the discrete score distributions.

- We assume

$$X \sim F(x \mid \theta_X) \quad (1)$$
$$Y \sim F(y \mid \theta_Y) \quad (2)$$

Laboratorio Interdisciplinario de Estadística Social

Discrete-based equating methods

(a) Classical Parametric approach

- For $F(x \mid \theta_X)$ and $F(y \mid \theta_Y)$, assume some parametric distributions that have been used to model discrete scores:

Laboratorio Interdisciplinario de Estadística Social

Discrete-based equating methods

(a) Classical Parametric approach

- For $F(x \mid \theta_X)$ and $F(y \mid \theta_Y)$, assume some parametric distributions that have been used to model discrete scores:

  1. Poisson-Binomial
  2. Negative Binomial

Laboratorio Interdisciplinario de Estadística Social

Discrete-based equating methods

(a) Classical Parametric approach

- For $F(x \mid \theta_X)$ and $F(y \mid \theta_Y)$, assume some parametric distributions that have been used to model discrete scores:

    1. Poisson-Binomial
    2. Negative Binomial

- For these models the mean and standard deviation are functions of θ_X and θ_Y.

Laboratorio Interdisciplinario de Estadística Social

Discrete-based equating methods

(a) Classical Parametric approach

- For $F(x \mid \theta_X)$ and $F(y \mid \theta_Y)$, assume some parametric distributions that have been used to model discrete scores:

  1. Poisson-Binomial
  2. Negative Binomial

- For these models the mean and standard deviation are functions of θ_X and θ_Y.

- By getting estimations(MLE) of θ_X and θ_Y, we can plug them in the linear equating transformation.

Laboratorio Interdisciplinario de Estadística Social

Discrete-based equating methods

(b) Bayesian Parametric approach

- Under the same model (1)-(2), assume some of the parametric distribution for discrete scores mentioned before.

Laboratorio Interdisciplinario de Estadística Social

Discrete-based equating methods

(b) Bayesian Parametric approach

- Under the same model (1)-(2), assume some of the parametric distribution for discrete scores mentioned before.
- Let consider a prior distribution for the parameters θ_X and θ_Y as follows:

$$\begin{aligned} \theta_X &\sim p(\theta_X) \\ \theta_Y &\sim p(\theta_Y) \end{aligned}$$

Discrete-based equating methods

(b) Bayesian Parametric approach

- Under the same model (1)-(2), assume some of the parametric distribution for discrete scores mentioned before.
- Let consider a prior distribution for the parameters θ_X and θ_Y as follows:

$$\begin{aligned} \theta_X &\sim p(\theta_X) \\ \theta_Y &\sim p(\theta_Y) \end{aligned}$$

- Perform Bayesian inference to obtain estimates of the parameters and then compute the linear equating transformation.

Discrete-based equating methods

(c) Bayesian Nonparametric approach

- Assume the following model

Discrete-based equating methods

(c) Bayesian Nonparametric approach

- Assume the following model

$$
\begin{aligned}
X \mid F_X &\sim F_X \\
F_X &\sim \mathscr{P} \\
Y \mid F_Y &\sim F_Y \\
F_Y &\sim \mathscr{P}
\end{aligned}
$$

where $\mathscr{P}$ is a random probability measure.

Discrete-based equating methods

(c) Bayesian Nonparametric approach

- Assume the following model

$$
\begin{aligned}
X \mid F_X &\sim F_X \\
F_X &\sim \mathscr{P} \\
Y \mid F_Y &\sim F_Y \\
F_Y &\sim \mathscr{P}
\end{aligned}
$$

  where $\mathscr{P}$ is a random probability measure.

- Use Bayesian nonparametric inference to estimate the score distributions. Using those estimations, obtain means and standard deviations to be plugged in the linear equating transformation.

Improved continuization methods

- In order to preserve as much as possible the discrete nature of the score data, we propose some statistical tools that take advantages of some results on the continuous framework and relate them with the main characteristic of score data: discreteness.

Improved continuization methods

(a) Latent variable approach

Improved continuization methods

(a) Latent variable approach

- We assume that the scores are realizations of an underlying continuous latent process.

Laboratorio Interdisciplinario de Estadística Social

Improved continuization methods

(a) Latent variable approach

- We assume that the scores are realizations of an underlying continuous latent process.
- Based on a bayesian nonparametric model for the latent process, we can estimate the score distributions and thus the equipercentile equating transformation.

Laboratorio
Interdisciplinario
de Estadística Social

Improved continuization methods

(a) Latent variable approach

- We assume that the scores are realizations of an underlying continuous latent process.
- Based on a bayesian nonparametric model for the latent process, we can estimate the score distributions and thus the equipercentile equating transformation.
- Since there is a one-to-one relation between the discrete score values and the continuous latent process, we can obtain an equating in the discrete setting.

Laboratorio Interdisciplinario de Estadística Social

Improved continuization methods

(b) Testing continuization

Improved continuization methods

(b) Testing continuization

- We can consider the following result

$$\varphi(x) = Lin(x) + R(x)$$

where $Lin(x)$ does not need continuization, whereas $R(x)$ does.

Laboratorio Interdisciplinario de Estadística Social

Improved continuization methods

(b) Testing continuization

- We can consider the following result

$$\varphi(x) = Lin(x) + R(x)$$

where $Lin(x)$ does not need continuization, whereas $R(x)$ does.

- The idea is then look for ways to minimize the use of $R(x)$ and/or to develop measures to decide to what extend is this function important in the estimation of φ.

Laboratorio Interdisciplinario de Estadística Social

Ordinal random variables and latent variables

- Let W be an ordinal random variable with C categories named $w_1, \dots, w_C$.

Ordinal random variables and latent variables

- Let W be an ordinal random variable with C categories named $w_1, \ldots, w_C$.

- One approach to model this kind of random variables is the latent modelling formulation.

Laboratorio Interdisciplinario de Estadística Social

Ordinal random variables and latent variables

- Let W be an ordinal random variable with C categories named $w_1, \ldots, w_C$.
- One approach to model this kind of random variables is the latent modelling formulation.
- Let Z be a continuous latent variable with distribution F_Z.

Laboratorio Interdisciplinario de Estadística Social

Ordinal random variables and latent variables

- The relation between these variables is the following one:

Ordinal random variables and latent variables

- The relation between these variables is the following one:

$$W = \begin{cases} w_1 & \textit{if} & \gamma_0 < Z \leq \gamma_1 \\ w_2 & \textit{if} & \gamma_1 < Z \leq \gamma_2 \\ \vdots & \vdots & \vdots \\ w_C & \textit{if} & \gamma_{C-1} < Z \leq \gamma_C \end{cases} \tag{3}$$

where $\gamma_1, \ldots, \gamma_C$ are called *cutoffs*.

Laboratorio Interdisciplinario de Estadística Social

Ordinal random variables and latent variables

- From the previous relation it follows that for $k = 1, \ldots, C$:

Ordinal random variables and latent variables

- From the previous relation it follows that for $k = 1, \ldots, C$:

$$W = w_k \quad \Leftrightarrow \quad W = w_k \mathbb{1}\{Z \in (\gamma_{k-1}, \gamma_k])\} \tag{4}$$

Laboratorio Interdisciplinario de Estadística Social

Ordinal random variables and latent variables

- From the previous relation it follows that for $k = 1, \ldots, C$:

$$W = w_k \quad \Leftrightarrow \quad W = w_k \mathbb{1}\{Z \in (\gamma_{k-1}, \gamma_k])\} \tag{4}$$

- Then, the probability of W is specified by the probability of Z, for each $k = 1 \ldots, C$:

$$P(W = w_k) \quad = \quad P(Z \leq \gamma_k) - P(Z \leq \gamma_{k-1}) \tag{5}$$

Ordinal random variables and latent variables

- Then, in terms of Z, the distribution of W can be written as a multinomial distribution:

Ordinal random variables and latent variables

- Then, in terms of Z, the distribution of W can be written as a multinomial distribution:

$$W \sim Mult(1, C, p_{Z,\gamma}) \tag{6}$$

where

$$p_{Z,\gamma} = (p_{Z,\gamma_1}, \ldots, p_{Z,\gamma_C}) \tag{7}$$

$$p_{Z,\gamma_k} = P(Z \leq \gamma_k) - P(Z \leq \gamma_{k-1}) \qquad k = 1, \ldots, C \tag{8}$$

Laboratorio Interdisciplinario de Estadística Social

Bayesian nonparametric model approach

- Let $W_1, \ldots, W_n$ be a random sample from W.

Bayesian nonparametric model approach

- Let $W_1, \ldots, W_n$ be a random sample from W.
- For $i = 1, \ldots, n$, let Z_i be the latent variable associated to W_i, from the relation (3)

Laboratorio Interdisciplinario de Estadística Social

Bayesian nonparametric model approach

- Let $W_1, \ldots, W_n$ be a random sample from W.
- For $i = 1, \ldots, n$, let Z_i be the latent variable associated to W_i, from the relation (3)
- Based on the proposal of Kottas et. al(2005), we propose a Bayesian nonparametric model for the latent model F_Z defined by:

Laboratorio Interdisciplinario de Estadística Social

Bayesian nonparametric model approach

- Let $W_1, \ldots, W_n$ be a random sample from W.
- For $i = 1, \ldots, n$, let Z_i be the latent variable associated to W_i, from the relation (3)
- Based on the proposal of Kottas et. al(2005), we propose a Bayesian nonparametric model for the latent model F_Z defined by:

$$Z_1 \ldots, Z_n \mid F_Z \overset{iid}{\sim} F_Z \tag{9}$$

$$F_Z \sim G \tag{10}$$

where G is a random probability measure that leads on a continuous density: a DPM model.

Laboratorio Interdisciplinario de Estadística Social

Equating: latent modelling approach

- Let's consider the score X be an ordinal random variable with J categories defined by each value in the score sample space.

Laboratorio Interdisciplinario de Estadística Social

Equating: latent modelling approach

- Let's consider the score X be an ordinal random variable with J categories defined by each value in the score sample space.
- Let $X_1, \ldots, X_{n_X}$ a random sample of X. Using the latent modelling approach for ordinal variables, we will assume the following model:

Laboratorio Interdisciplinario de Estadística Social

Equating: latent modelling approach

- Let's consider the score X be an ordinal random variable with J categories defined by each value in the score sample space.
- Let $X_1, \ldots, X_{n_X}$ a random sample of X. Using the latent modelling approach for ordinal variables, we will assume the following model:

$$
\begin{aligned}
X_1, \ldots, X_n \mid Z_1, \ldots, Z_n &\overset{ind}{\sim} Mult(1, J, p_{z,\gamma}) && (11)\\
Z_1 \ldots, Z_n \mid \theta_1, \ldots, \theta_n &\overset{ind}{\sim} N(z_i \mid \mu_i, \sigma_i^2) && (12)\\
\theta_1, \ldots, \theta_n \mid G &\overset{iid}{\sim} G && (13)\\
G \mid \phi &\sim PD(a, b, G_0(\phi)) && (14)\\
\phi &\sim p(\phi) && (15)
\end{aligned}
$$

Laboratorio Interdisciplinario de Estadística Social

References

- Holland, P., & Rubin, D. (1982). Test equating. Academic Press, New York
- Gonzalez, J., & Wiberg, M. (2017). Applying Test Equating Methods Using R. Springer.
- Kottas, A., Müller, P., & Quintana, F. (2005). Nonparametric Bayesian modeling for multivariate ordinal data. Journal of Computational and Graphical Statistics, 14, 610–625.