Measurement xxx (2018) xxx–xxx

Contents lists available at ScienceDirect

Measurement

journal homepage: www.elsevier.com/locate/measurement

A law of ordinal random error: The Rasch measurement model and random error distributions of ordinal assessments

David Andrich *, Pender Pedler

The University of Western, Australia

ARTICLE INFO

Article history:
Received 1 May 2018
Received in revised form 22 August 2018
Accepted 27 August 2018
Available online xxxx

Keywords:
Measurement error
Error distributions
Ordered categories
Ordinal counts
Log-concave
Rasch model

ABSTRACT

Assessments in ordered categories are ubiquitous in educational, social and health sciences. These assessments are analogous to measurements in the natural sciences in that an idealised linear continuum is partitioned by successive thresholds into contiguous, ordered *categories*. In advanced analyses, the ordinal assessments are characterised with a probabilistic model as a function of a vector of threshold parameters defining the categories and a scalar parameter for the entity of measurement which is taken to be a measurement on an interval scale with an arbitrary origin and unit. One such model is the Rasch measurement model. If the ordinal assessments fit the model the probability distribution is taken to be a random error distribution of inferred replicated assessments. Therefore, it is analogous to the Gaussian random error distribution of replicated measurements known as *the law of error*. However, the Gaussian distribution is strictly log-concave which makes it unimodal with a smooth transition between probabilities of adjacent measurements. Such a distribution, referred to as *randomly unimodal*, ensures there is no evidence that unknown factors have produced systematic errors, and in turn justifies the mean as an estimate of the measure of the entity. The paper establishes that *random unimodality* arises from the natural ordering of the thresholds in the Rasch measurement model. Then by analogy to the Gaussian *law of error*, a distribution of ordinal assessments that has its thresholds in the natural order and fits the Rasch model may be said to satisfy the *law of ordinal error*. Again by analogy to Gaussian distribution with respect to replicated measurements, the law of ordinal error ensures that no unaccounted-for factors have produced systematic errors and that the estimates of the scalar parameter of the Rasch model can be taken as an estimate of the measure of the entity assessed.

1. Introduction

The concept and successful implementation of measurement has been the basis of the many remarkable advances in the natural sciences. The prototypic measurement involves an instrument which maps the amount of a variable on a real line divided into contiguous intervals of equal size, called the unit, by thresholds sufficiently fine that their own width can be ignored. Then a measurement is a count of the number of thresholds exceeded from a conventional origin in the common unit. Although measurement is a sophisticated concept that has been studied extensively [25], its principles of a count in terms of a conventionally specified unit and origin are understood by elementary school children. We refer to an empirical observation from the *process* of measurement with a measuring instrument as *a measurement*, and to the value that is to be estimated from the process of measurement as *a measure*. We note, without engaging in an epistemological debate on its implications, that a real number line is an idealisation which is approximated to varying degrees in the realisation of measurement which is always discrete in the unit of the instrument.

Generally, a measurement in its context is an important end in itself. For example, the temperature of a person suspected of having a fever or the amount of sugar in grape juice ready for fermenting into wine. However, the *function* of measurement in science is much deeper – it is central to the construction of sophisticated instruments for the measurement of variables, the confirmation of quantitative theories and laws involving these variables, and the disclosure of anomalies in implications of theories [20]. This paper generalises this function to the application of the Rasch model of measurement for discrete, ordinal assessments found in the social sciences. Central to the generalisation is the understanding of the properties of random error distributions of real or inferred replications of ordinal assessments and the disclosure of anomalies in these assessments.

* Corresponding author at: Graduate School of Education, The University of Western Australia, Crawley, Western Australia 6009, Australia.

E-mail address: david.andrich@uwa.edu.au (D. Andrich).

https://doi.org/10.1016/j.measurement.2018.08.062
0263-2241/© 2018 Published by Elsevier Ltd.

1.1. Randomly unimodal – the distribution of random errors of measurement

One of the major advances in measurement in the natural sciences, now generally taken for granted, is the understanding that replicated measurements give a distribution of measurements rather than a single value. As detailed in Section 2, the theoretical distribution of such measurements is the Gaussian distribution known as *the law of error* or the *law of measurement error*. To emphasise the role of randomness of replicated measurements, it may be referred to more expansively as *the law of random measurement error*. The distribution is strictly *unimodal* and its density function is *smooth*. To convey the characteristic of smooth unimodality which arises from random variation, such a distribution is referred to in this paper as *randomly unimodal*. A basic principle of random unimodality, together with the smoothness of the transition between adjacent probabilities, is that the probability of a small error is greater than the probability of a larger error.

The significance of the theoretical *law of measurement error* is that any relevant deviation of ostensibly replicated measurements, real or inferred, renders their mean biased and in some cases unusable as an estimate of the measure of an entity. Equally significantly, the deviation implies that some factor, which needs to be identified empirically, had interfered with the measurements. The paper abstracts the principle and significance of the random unimodality of the Gaussian distribution to discrete distributions of assessments in ordered categories when analysed with the Rasch model for measurement.

1.2. Ordinal assessments – the distribution of random errors of ordinal counts

Where no measuring instrument of the kind used in the natural sciences is available, measurements in the social sciences begin with the ubiquitous assessments in ordered categories. Dawes [9] estimated that approximately 60% of assessed variables had this format. This percentage is likely to have increased since Dawes' estimate.

In psychometric contexts multiple sets of ordered categories are called *items*. For the purpose of emphasis of the analogy with measurement in the physical sciences in this paper, items will be referred to as *instruments* of ordinal assessment. Although the principles developed apply to the assessment of a property of any entity with ordered dichotomous or polytomous sets of categories, we specialise the entities of assessment to persons and take them to be assessed with instruments with more than two ordered categories. The assessment of each person with *multiple instruments* in the social sciences provides the kinds of replications which are elaborated in the paper and which are compared to replications of measurements with the *same instrument* in the physical sciences.

The principles developed in this paper can also be applied to contingency table contexts where the assessment is with one instrument and the persons are classified by background variables such as age or gender. However, in this case it is required to conceptualise the persons in each classification as replications of each other. Because the paper is concerned with the replications of assessments of a single person, the paper emphasises the psychometric context where multiple instruments are used to assess each person.

By analogy to physical measurement, the ordered categories of an instrument are assumed to partition an idealised continuum with successive thresholds into mutually exclusive contiguous intervals called *categories*. We refer to assessments in these categories as *ordinal assessments*. Ordered categories have two main differences from measuring instruments: first the category intervals are not expected to be equal; second the number of categories is finite, often of the order of five or six. However, without loss of generality, the successive categories may be assigned successive integers beginning with 0 and be referred to as *ordinal counts* of the ordinal assessments. In elementary analyses, these ordinal counts are simply treated as measurements. In advanced analyses the ordinal counts are characterised by a discrete, probabilistic model which, given relevant conditions are satisfied, is taken to characterise the distribution of random errors. The paper is concerned with the probabilistic Rasch measurement model (RMM) for modelling responses to instruments and transforming them into measurements with an arbitrary origin and unit.

1.3. Rasch's measurement theory and the model for ordinal assessments

Rasch [21–23] abstracted principles of measurement from his theory of invariant comparisons for deterministic and probabilistic models. His measurement theory requires that, within a specified frame of reference, the comparisons between entities are independent of the instrument, and that the comparisons between instruments are independent of the entities. The resultant models are compatible with those from both representational and classical theories of measurement; moreover it explains them [6]. Implications of Rasch's probabilistic models for measurement in the social sciences have been advanced in Fisher [15], Fisher and Stenner [16], Stone and Stenner [24], Wright [26–28] and others. The paper applies the RMM to ordinal assessments. The RMM characterises the probability of each ordinal assessment as a function of a vector of real-valued parameters for the instrument and a real-valued scalar parameter for the person.

Three related properties of the RMM are relevant to highlight for this paper. First, the ordinal counts of each instrument, beginning with 0 for the first category of each instrument, and the threshold parameters which define the categories, conveniently and naturally appear in the model. These result from the assumed equality of the discriminations at the thresholds, which is again analogous to thresholds of measuring instruments. Second, with replicated assessments with multiple instruments, the relevant statistic that characterises each person is the summed score across the instruments. Third, this total score is sufficient in the sense that no further information regarding the person parameter resides in the profile of responses. As a consequence, the realisation of the separation of the instrument and person parameters, and the facilitation of invariant comparisons, is obtained by a conditional equation which is a function of only the instrument parameters and is independent of the person parameters. Two further relevant features of the RMM are elaborated in Section 3. First, although each person/instrument engagement is unique, the estimation equation is of explicit replications in the sense that every response is a function of exactly the same instrument parameters. Second, in estimating the instrument parameters from this equation no assumptions need to be made regarding the distribution of the person parameters. Then the measurement of each person is estimated individually from the instrument parameters which are taken as known. The measurement procedure is directly analogous to that of measurement in the natural sciences where, given an instrument's operation is validated, entities are measured individually.

The paper shows that under a particular natural condition, the distribution of ordinal counts of ordinal assessments modelled by the RMM satisfies the same mathematical condition of random unimodality satisfied by the Gaussian law of measurement error. Accordingly, when this condition is satisfied it is proposed that the RMM is the *law of ordinal error*. Despite the ubiquity of the modelling of ordinal assessments, at present no such law exists.

Please cite this article in press as: D. Andrich, P. Pedler, A law of ordinal random error: The Rasch measurement model and random error distributions of ordinal assessments, Measurement (2018), https://doi.org/10.1016/j.measurement.2018.08.062