Journal of Statistical Planning and Inference 172 (2016) 23–35

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

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

A general approach to categorizing a continuous scale according to an ordinal outcome

Limin Peng [a,*], Amita Manatunga [a], Ming Wang [b], Ying Guo [a], AKM Fazlur Rahman [a]

[a] Department of Biostatistics and Bioinformatics, Emory University, 1518 Clifton Road NE., Atlanta, GA 30322, USA
[b] Department of Public Health Sciences, Penn State University, 500 University Drive, Hershey, PA 17033, USA

ARTICLE INFO

Article history:
Received 19 January 2015
Received in revised form 26 August 2015
Accepted 16 December 2015
Available online 29 December 2015

Keywords:
Agreement
Association
Empirical process
M-estimation
Non-smooth objective function
Subsampling

ABSTRACT

In practice, disease outcomes are often measured in a continuous scale, and classification of subjects into meaningful disease categories is of substantive interest. To address this problem, we propose a general analytic framework for determining cut-points of the continuous scale. We develop a unified approach to assessing optimal cut-points based on various criteria, including common agreement and association measures. We study the nonparametric estimation of optimal cut-points. Our investigation reveals that the proposed estimator, though it has been ad-hocly used in practice, pertains to nonstandard asymptotic theory and warrants modifications to traditional inferential procedures. The techniques developed in this work are generally adaptable to study other estimators that are maximizers of nonsmooth objective functions while not belonging to the paradigm of M-estimation. We conduct extensive simulations to evaluate the proposed method and confirm the derived theoretical results. The new method is illustrated by an application to a mental health study.

1. Introduction

In many biomedical and behavioral studies, to identify a certain disease in human body, different instruments or rating scales are utilized. Typically, measurements are made on a continuous scale; however, researchers are often interested in dividing a continuous scale into ordered categories for reasons such as clinical interpretations of results and simplification of the instrument (O'Brien, 2004). For example, depression is a common problem in medically-ill patients with diabetes and other chronic diseases (Moussavi et al., 2007). The psychiatric diagnostic interview instruments, such as the Mini International Neuropsychiatric Interview (MINI) diagnostic interview (Sheehan et al., 1998), in general provide accurate psychiatric diagnoses in the medically healthy individuals. However, the MINI interview instrument is too time-consuming for sick patients and requires trained psychiatric interviewers, which are not always affordable or available. On the other hand, dimensional psychometric instruments designed to measure the same state of the disease, such as the 20-item Zung Depression rating scale (Zung, 1965), require less time and could be self-administered by patients. The established total observer-rated MINI score has been interpreted with well-accepted graded severity of depression such as no depression, mild depression, and markedly severe depression (Sheehan et al., 1998). While the self-reported Zung scale has many advantages, there are no routine, reliable cut-points of Zung provided to reflect the degree of the severity of depression.

* Corresponding author. Tel.: +1 404 727 7701.
E-mail address: lpeng@emory.edu (L. Peng).

http://dx.doi.org/10.1016/j.jspi.2015.12.006
0378-3758/

Establishment of such cut-points would enhance the utility of the convenient Zung rating scale in the screening or diagnosis of depression, particularly in large medically-ill patient populations.

Analytic methods for determining cut-points of a continuous scale based on validated categorical measurements, despite their practical importance, have not been well studied. For example, Youden index (Youden, 1950) and its variants based on receiver operating curve (ROC) (Pepe, 2003) were studied for identifying the "best" cut-point to dichotomize a continuous scale (Kraemer, 1988; Schisterman et al., 2005; Perkins and Schisterman, 2006). However, this type of approach can only deal with a single cut-point, and moreover, lack formal inference procedures about the estimated cut-point. For the general cases possibly involving two or more cut-points, existing approaches are mostly based on ad hoc arguments and generally lack statistical rigor. For example, the following methods were used in the literature: (a) considering arbitrary cut-points or a certain sample quantile like the median or a cut-point that corresponds to the highest proportions of correct classification with a gold standard (Altman et al., 1994; Mazumdar and Glassman, 2000); (b) finding out cut-points that result in disease rates consistent with a known population disease prevalence (Altman, 1991); and (c) depending on clinicians' experience (Altman et al., 1994). The fundamental deficiency of all these methods is that the criterion levels or cut-points are generally decided subjectively on the basis of the probabilities of Type I/II misclassification, and "judgment" or "experience". Several authors (James, 1978; Brownie and Habicht, 1984) proposed to minimize the variance estimator of the prevalence of a disease by assuming a mixture of normal distributions; however, this criterion is restricted by the distribution and is valid only for the cases with a single cut-point. Baughman et al. (2006) also used a mixture model but the maximum likelihood estimators may not be accurate when the distributions are not well separated, the sample size is small, or the mixture model is misspecified.

In this work, we seek to develop a general and objective analytic framework for addressing the problem of determining cut-points in a continuous scale according to an established ordinal scale. To this end, a fundamental question is, what are the most desirable or optimal cut-points? A common viewpoint in practice is that meaningful cut-points in a continuous scale should produce high agreement or association between the newly categorized continuous scale and established categories. Therefore, we propose to evaluate each set of cut-points by some criterion that reflects a desirable relationship between two ordinal scales (e.g. high agreement or association). From doing so, we obtain a function of cut-points, which we shall refer to as a criterion function. We then define the optimal set of cut-points as the one that optimizes the criterion function.

We propose a general formulation of criterion functions, which are expressed as a smooth function of cell probabilities of the contingency table formed from cross-tabulating an established categorical scale and the newly categorized continuous scale based on a given set of cut-points. Our definition encompasses many important special cases, including those where cut-points are evaluated by weighted kappa (Cohen, 1960; Agresti, 1990), Kendall's τ_b, correct classification rate (Kendall, 1938; Agresti, 1990), or Youden's index in ROC. The general specification of criterion functions forms the foundation of the proposed unified framework for investigating cut-points based on various criteria.

We study the estimation of optimal cut-points without imposing any parametric assumptions on the distributions of the data. We consider a natural approach, which is to optimize a nonparametric estimator of the criterion function, called empirical criterion function hereafter. For instance, one may adopt weighted kappa statistic when the criterion for optimal cut-points is weighted kappa coefficient. While the basic idea is conceptually intuitive and has been adopted in practice in an ad-hoc way, our detailed investigation indicates that such a method is subject to nonstandard theory and requires special attention to its inference procedures. For example, the resulting estimator has a slower convergence rate lower than the usual root n, and may not possess asymptotic normality. The main issue is that an empirical criterion function is usually not smooth, and more specifically, involves cut-points through indicator functions. The challenge resembles the difficulty in M-estimation with non-smooth objective function (Chernoff, 1964, for example). Nevertheless, the proposed estimator is yet not a M-estimator. As a result, existing methods that deal with non-smooth M-estimation, for example, Kim and Pollard (1990), are not directly applicable.

In this work, we employ the technique of empirical processes and conduct rigorous asymptotic studies for the proposed nonparametric cut-point estimator. It is important to point out that our theoretical framework is quite general and may be adapted to many other estimation settings that involve maximization of non-smooth objective functions. Provided the nonstandard asymptotic properties of the proposed estimator, bootstrapping fails to work properly (Kosorok, 2008). We propose to use subsampling (Politis et al., 1999) as a well justified device for inferences including variance estimation and confidence intervals.

We elaborate in Section 2 the proposed method for determining cut-points in a continuous scale. We present the general problem formulation, the proposed nonparametric estimation, and the corresponding asymptotic results and inference. Extensive simulation studies reported in Section 3 demonstrate satisfactory finite-sample performance of our proposals, and also help confirm some of our theoretical results. In Section 4, we illustrate our method via an application to a mental health study. Our analysis suggests a refinement of current empirical rules for categorizing depression among diabetic subjects based on the Zung rating scale. A few concluding remarks are provided in Section 5.

2. The proposed method

2.1. A general formulation of optimal cut-points

Let Y denote an ordinal measurement that takes ordinal values, $1 < \cdots < L$. Let X denote a continuous measurement bounded between x_L and x_U. All possible cut-points for X according to the L categories of Y form a compact parameter space,