Computational Statistics and Data Analysis 84 (2015) 85–98

Contents lists available at ScienceDirect

Computational Statistics and Data Analysis

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

A pseudo-likelihood approach for estimating diagnostic accuracy of multiple binary medical tests

Wei Liu [a], Bo Zhang [b,*], Zhiwei Zhang [b], Baojiang Chen [c], Xiao-Hua Zhou [d]

[a] *Department of Mathematics, Harbin Institute of Technology, Harbin 150001, PR China*

[b] *Division of Biostatistics, Office of Surveillance and Biometrics, Center for Devices and Radiological Health, Food and Drug Administration, Silver Spring, MD 20993, USA*

[c] *Department of Biostatistics, College of Public Health, University of Nebraska Medical Center, Omaha, NE 68198, USA*

[d] *Department of Biostatistics, School of Public Health and Community Medicine, University of Washington, HSR&D Center of Excellence, U.S. Department of Veterans Affairs Puget Sound Health Care System, Seattle, WA 98198, USA*

ARTICLE INFO

Article history:
Received 23 March 2014
Received in revised form 10 October 2014
Accepted 11 November 2014
Available online 20 November 2014

Keywords:
Sensitivity and specificity
Random effects
Latent class models
Composite likelihood
Imperfect reference standards

ABSTRACT

Latent class models with crossed subject-specific and test(rater)-specific random effects have been proposed to estimate the diagnostic accuracy (sensitivity and specificity) of a group of binary tests or binary ratings. However, the computation of these models are hindered by their complicated Monte Carlo Expectation–Maximization (MCEM) algorithm. In this article, a class of pseudo-likelihood functions is developed for conducting statistical inference with crossed random-effects latent class models in diagnostic medicine. Theoretically, the maximum pseudo-likelihood estimation is still consistent and has asymptotic normality. Numerically, our results show that not only the pseudo-likelihood approach significantly reduces the computational time, but it has comparable efficiency relative to the MCEM algorithm. In addition, dimension-wise likelihood, one of the proposed pseudo-likelihoods, demonstrates its superior performance in estimating sensitivity and specificity.

1. Introduction

Pharmaceutical and regulatory statisticians who work in the fields of developing state-of-the-art medical devices and radiology diagnostic tests have shown tremendous interests in the strategies for accurately estimating diagnostic accuracy of multiple binary medical tests or raters. It has been widely recognized that sensitivity and specificity are two primary measures that characterize the diagnostic accuracy of binary tests. Statistical methodologies have been proposed to estimate sensitivity and specificity of binary tests (Zhou et al., 2002; Pepe, 2003). When investigators are committed to estimating the average sensitivity and specificity of a group of tests or raters, latent class models (Qu et al., 1996; Hui and Zhou, 1998; Qu and Hadgu, 1998; Albert et al., 2001; Albert and Dodd, 2008), in which the true disease status is considered as a latent variable, have been proved to be an effective and strategic approach. Qu et al. (1996) and Qu and Hadgu (1998) proposed a random-effects latent class model that characterized conditional dependence between tests through random effects. Albert et al. (2001) proposed a latent class model with a finite mixture structure to account for dependence between tests. Albert and Dodd (2008) extended latent class models in Qu and Hadgu (1998) and Albert et al. (2001) to incorporate information from both the verified and nonverified subjects during estimation. More recently, Zhang et al. (2012) introduced a latent

---

[*] Corresponding author.
E-mail address: Bo.Zhang@fda.hhs.gov (B. Zhang).

http://dx.doi.org/10.1016/j.csda.2014.11.006
0167-9473/Published by Elsevier B.V.

class model with crossed random effects (with both subject-specific and test-specific random effects) for estimating sensitivity and specificity of a group of binary tests or raters. Despite the success of these random-effects latent class models in diagnostic statistics, computational complexity has been the major flaw of the latent class models with crossed random effects. This article is devoted to proposing a competitive pseudo-likelihood approach that reduces the computational burden of crossed random effects latent class models in estimating diagnostic accuracy.

Statistical inference in random-effects models (or mixed-effects models) (Laird and Ware, 1982; McCulloch and Searle, 2001) starts from parameter estimation step, in which maximum likelihood estimation (MLE) is achieved via numerical optimization methods, such as Newton–Raphson algorithm and its variants. However, these optimization procedures require numerical approximation of the likelihoods of random-effects models that may involve evaluating high-dimensional integrals. In the latent class models with crossed random effects, as we show in Section 2, this computational issue becomes the major obstacle because the high-dimensional integral in the full likelihood of a crossed random effects latent class model is intractable. Computational strategies have been proposed to overcome the numerical difficulties posed by high-dimensional integrations in random-effects models. Breslow and Clayton (1993) proposed a pseudo quasi-likelihood method that seeks the parameter estimation by maximizing the joint distribution of the observed data and the random effects. Approximate MLE computation using the Laplace approximation also proposed by Steele (1996) and Lee and Nelder (2001). Yet, these methods cannot provide generally consistent estimation (Lin and Breslow, 1996). Another option is to develop simulation-based Monte Carlo algorithms for obtaining the MLE, which include Monte Carlo Markov chain (MCMC) algorithms proposed by Zeger and Karim (1991) and McCulloch (1997) and Monte Carlo Expectation–Maximization (MCEM) algorithm proposed by Booth and Hobert (1999) and Booth et al. (2001). These algorithms have drawbacks that include the computational duration and MCEM convergence assessment, so that they cannot provide instant statistical inference results to practitioners. Statistical inference turns to be even harder in the latent class models with crossed random effects. Unlike the conventional random-effects models with independent between-subject observations, the full likelihood of a crossed random effects latent class model is an integral whose dimension increases with the number of subjects and the number of tests. In a crossed diagnostic design with I subjects and J tests, the full likelihood appears to be a $(I \times J)$-dimensional integral embedded in an I-dimensional summation, which particularly deteriorates the statistical inference procedure when using crossed random effects latent class models.

In this article, we propose a pseudo-likelihood approach as a competitive statistical analysis strategy for the crossed random effects latent class models that estimate the sensitivity and specificity of a group of binary medical tests or raters. A class of pseudo-likelihood functions is created in Section 2, including pairwise likelihood (Bellio and Varin, 2005), triplewise likelihood, hybrid likelihoods, and dimension-wise likelihood. The benefits of using the proposed pseudo-likelihood approach are enormous. The proposed pseudo-likelihood functions contain the integrations with a reduced dimension and possess more succinct integrands than the full likelihood. The implementation of the pseudo-likelihood approach is rather simple with the aid of a numerical optimization package. Parameter estimates obtained from maximizing the proposed pseudo-likelihoods are consistent and asymptotically follow normal distributions (Lindsay, 1988; Molenberghs and Verbeke, 2005; Varin et al., 2011). The parameter variance estimation can be achieved by bootstrapping. Estimation efficiency of the estimators obtained from maximizing dimension-wise likelihoods is comparable to the full likelihood maximized by the MCEM algorithm. Yet, the efficiency of pairwise and hybrid likelihoods may not be satisfactory especially when imperfect reference standards do not exist. In Section 5, we analyze a colon cancer detection data published by Zhou et al. (2002), for the purpose of demonstrating the pseudo-likelihood approach.

2. Methodology

2.1. Crossed random effects models for estimating diagnostic accuracy

Let Y_{ij} denote the binary diagnostic result of a disease ($Y_{ij} = 1$ for having the disease and $Y_{ij} = 0$ for not having the disease) for the ith subject from the jth test (rater), $i = 1, 2, \ldots, I$ and $j = 1, 2, \ldots, J$. We denote D_i, a binary latent variable, as the true disease status of the ith subject. Under the circumstances that there is no gold or imperfect reference standard, we consider the following model with two crossed random effects for Y_{ij}:

$$P(Y_{ij} = 1|D_i = d_i, b_i, c_j) = \hbar^{-1}\left(\beta_{d_i} + \sigma_{d_i} b_i + \tau_{d_i} c_j\right), \quad \sigma_{d_i}, \tau_{d_i} > 0, \tag{1}$$

where $\hbar^{-1}(\cdot)$ denotes the inverse of a general link function (such as a probit link function or a logit link function), b_i is the subject-specific random effect with probability density distribution (p.d.f.) $g_1(x)$, and c_j is the test-specific random effect with p.d.f. $g_2(x)$. The three unobserved latent variables D_i, b_i, and c_j are assumed to be independent of each other. Let $\pi_{d_i} = P(D_i = d_i)$, and π_1 is consequently the prevalence of disease among the subjects. Under the probit link and the normality assumption of b_i and c_j, (1) can be used to estimate the average sensitivity and specificity of the rater (test) population: $S_e = \Phi(\beta_1/\sqrt{1 + \sigma_1^2 + \tau_1^2})$ and $S_p = \Phi(-\beta_0/\sqrt{1 + \sigma_0^2 + \tau_0^2})$, where $\Phi(\cdot)$ denotes the standard normal cumulative density function. Contrast to the latent class models in Qu et al. (1996) and Albert et al. (2001) with a single subject-specific random effect, (1) describes test-specific variation through an additional random effect c_j. Investigators usually assume b_i and