Computational Statistics and Data Analysis 55 (2011) 2354–2362

Contents lists available at ScienceDirect

Computational Statistics and Data Analysis

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

Generalized estimating equations with model selection for comparing dependent categorical agreement data

Miao-Yu Tsai *, Jung-Feng Wang, Jia-Ling Wu

Institute of Statistics and Information Science, College of Science, National Changhua University of Education, Changhua, Taiwan

ARTICLE INFO

Article history:
Received 14 June 2010
Received in revised form 14 November 2010
Accepted 1 February 2011
Available online 23 February 2011

Keywords:
Agreement
Generalized estimating equation
Kappa
QIC
Weighted least square

ABSTRACT

Many studies in biomedical fields are carried out using diagnoses reported by different raters to evaluate the agreement of multiple ratings. The most popular indices of agreement are kappa measures including Cohen's kappa and weighted kappa for binary and ordinal outcomes, respectively. However, when raters assess the same observation on two or more occasions, these ratings are dependent and so the correlation between kappa estimates must be considered when making inferences. In this paper, we focus on testing the equality of correlated kappa coefficients using the generalized estimating equation (GEE) approach and applying *quasi-likelihood under the independence model criterion* (QIC) measures for model selection. Simulation studies are conducted to compare the performance between GEE with and without QIC measures, weighted least squares (WLS) and independence approaches for binary and ordinal data. Two applications are illustrated: a comparison of two methods for assessing cervical ectopy, and similarity in myopic status for monozygous twins and dizygous twins. We conclude that when performing the QIC model-selection procedure in GEE models and taking into account the correlation between kappa measures, it leads to nominal type I errors and larger powers.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

In medical research, disease can be measured with categorical data including a dichotomous, polychotomous or ordinal score for conditions such as the presence or absence of dysplasia or the severity of wheeze. Furthermore, the same observed unit can be assessed by different medical investigators. The agreement between the raters using different methods and the reproducibility of one rater are often essential questions. The most popular indices of agreement for a categorical response are the kappa measures, Cohen's kappa and weighted kappa for binary and ordinal outcomes, respectively (Cohen, 1960, 1968). However, when raters assess one observation more than once, these ratings are dependent, and therefore, one must take into account the correlation between the kappa estimates for making inference.

Generalized estimating equation (GEE) is a common approach for estimating correlated kappa coefficients adjusting for specific covariates and allowing the dependency between replicated samples. Klar et al. (2000) proposed an estimating equations approach using unconditional and conditional residuals to model kappa for binary ratings. Williamson et al. (2000) presented the GEE approach to assess dependent agreement measurements for ordinal data. Williamson and Manatunga (1997) used a latent variable model proposed by Qu et al. (1995) to analyze inter-rater agreement studies with different measurement techniques for clustered ordinal data. In addition, procedures for testing the equality of two dependent kappa statistics have developed. For the case of two raters and a dichotomous outcome variable, a goodness-of-fit procedure

* Corresponding address: Institute of Statistics and Information Science, National Changhua University of Education, No.1, Jin-De Road, Chang-Hua 500, Taiwan.

E-mail address: mytsai@cc.ncue.edu.tw (M.-Y. Tsai).

0167-9473/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.csda.2011.02.002

accounting for the dependency between ratings under a generalized common correlation model has been considered by Donner et al. (2000). Barnhart and Williamson (2002) used a weighted least squares (WLS) approach to test for the equality of dependent kappa measures such as Cohen's kappa, intraclass kappa, and weighted kappa.

As we know, under the GEE approach, a marginal model including all potential explanatory variables may lead to the complexity in interpretation and less precision of parameter estimates. Therefore, when covariates are involved in estimating kappa measures, the variable selection is essential. There is a great volume of literature discussing the model selection for classic linear regressions with independent data (e.g. Miller (1990) and references therein). Akaike's information criterion (AIC) is a well-established goodness-of-fit statistic for likelihood-based model selection with the considerations of model complexity. However, GEE cannot apply AIC to perform model selection directly since no distribution is assumed in GEE. Pan (2001) proposed a modification to AIC based on a quasi-likelihood and a proper adjustment made for the penalty term, called the *quasi-likelihood under the independence model criterion* (QIC), in GEE models for correlated data.

In this paper, we focus on testing the equality of the several dependent kappa statistics using the GEE approach for correlated categorical data with and without choosing the best subset of covariates via the QIC measures. In addition, it will be compared with the results from the testing procedure using the WLS approach with the consideration for the correlated kappa statistics and the independence approach (Independence) using SAS procedure FREQ assuming independent kappa statistics. The rest of the paper is organized as follows. Section 2 introduces the GEE approach for dependent categorical data and the model-selection criterion QIC in GEE. In addition, the inference of the kappa statistic under the WLS approach is also presented in Section 2. In Section 3, simulation studies are conducted to compare the performance of the GEE approach with and without the model-selection procedure via QIC, the WLS, and the independence approaches. Two applications are illustrated in Section 4. The first example is to assess cervical ectopy. The objective of the study is to compare the reliability of a computerized planimetry method for measuring cervical ectopy with direct visual assessment (Gilmour et al., 1997). The second application is a myopia twin study that contained ophthomologic measurements taken on both eyes from monozygous (MZ) twins or dizygous (DZ) twins. The main objective of this study is to measure the similarity of twin pairs with respect to zygosity. In other words, our research question is whether the agreement in high myopia between a MZ twin-pair is higher than that of a DZ twin-pair (Tsai et al., 2009). Finally, in Section 5 we conclude with discussions and remarks.

2. Methods

Among the approaches used for analysis, we consider two versions of kappa measures of agreement between two ratings, Cohen's and weighted kappa for binary and ordinal data, respectively. The two kappa coefficients for categorical outcomes with J categories are defined as follows (Cohen, 1960, 1968):

$$\text{Cohen's kappa } \kappa_C = \frac{\sum_{j=1}^{J}\pi_{jj} - \sum_{j=1}^{J}\pi_{j.}\pi_{.j}}{1-\sum_{j=1}^{J}\pi_{j.}\pi_{.j}} \quad \text{and}$$

$$\text{Weighted kappa } \kappa_W = \frac{\sum_{i=1}^{J}\sum_{j=1}^{J} w_{ij}\pi_{ij} - \sum_{i=1}^{J}\sum_{j=1}^{J} w_{ij}\pi_{i.}\pi_{.j}}{1-\sum_{i=1}^{J}\sum_{j=1}^{J} w_{ij}\pi_{i.}\pi_{.j}},$$

where π_{ij} denotes the (i,j)th cell joint probability of two ratings in a $J \times J$ contingency table, $\pi_{i.} = \sum_{j=1}^{J}\pi_{ij}$, $\pi_{.j} = \sum_{i=1}^{J}\pi_{ij}$, and w_{ij} is the weight for the (i,j)th cell satisfying $w_{jj} = 1$ and $0 \le w_{ij} \le 1$. Two popular settings of weights are quadratic weight $w_{ij} = 1 - \frac{(i-j)^2}{(J-1)^2}$ (Fleiss and Cohen, 1973) and absolute weight $w_{ij} = 1 - \frac{|i-j|}{(J-1)^2}$ (Cicchetti and Allison, 1971). In this paper, we use the GEE and WLS approaches for inference including estimation and tests of equality of agreement, and the two approaches are briefly reviewed in the following subsections.

2.1. Generalized estimating equations

We describe the GEE approach for analyzing dependent agreement data with categorical responses (Gonin et al., 2000; Williamson et al., 2000). In an agreement study, each of the I subjects is assessed R_i times (e.g. R_i raters or R_i methods) and R_i does not have to be the same for all subjects. Let y_{ir} denote the rth rating which can take J categories for the ith subject, where $i = 1, 2, \ldots, I$ and $r = 1, 2, \ldots, R_i$. In other words, $\mathbf{y}_i = (y_{i1}, y_{i2}, \ldots, y_{iR_i})^t$ is a $R_i \times 1$ categorical rating vector for the ith subject. We redefine the response y_{ir} in terms of a $(J-1) \times 1$ binary vector $\mathbf{Y}_{ir} = (Y_{ir,1}, \ldots, Y_{ir,J-1})^t$, where $Y_{ir,j} = 1$ if subject i is assessed to be in level j by rater r and $Y_{ir,j} = 0$ for otherwise, and $j = 1, \ldots, J$. Therefore, the binary response for the ith subject is denoted by the $R_i(J-1) \times 1$ vector $\mathbf{Y}_i = (\mathbf{Y}_{i1}^t, \ldots, \mathbf{Y}_{iR_i}^t)^t$. Let $\mathbf{X}_i = (\mathbf{x}_{i1}, \ldots, \mathbf{x}_{iR_i})^t$ represent the $R_i \times p$ covariate matrix, which may include both subject-specific and rater-specific covariates for the ith subject. The probability

Download English Version:

https://daneshyari.com/en/article/415081

Download Persian Version:

https://daneshyari.com/article/415081

Daneshyari.com