Computational Statistics and Data Analysis 87 (2015) 1–17

Contents lists available at ScienceDirect

Computational Statistics and Data Analysis

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

Kappa statistic for clustered physician–patients polytomous data

Zhao Yang [a,*], Ming Zhou [b]

[a] University of Tennessee Health Science Center, 66 N. Pauline St Suite 633, Memphis, TN 38105, USA

[b] Global Biometric Sciences, Bristol-Myers Squibb, Hopewell, NJ 08534, USA

ARTICLE INFO

Article history:
Received 19 February 2014
Received in revised form 8 October 2014
Accepted 14 January 2015
Available online 23 January 2015

Keywords:
Kappa statistic
Clustered polytomous data
Confidence interval
Agreement
Coverage probability
Physician–patients

ABSTRACT

Motivated by the recent advances in the kappa statistic for the clustered physician–patients dichotomous data, we extend the development for the polytomous data. For the clustered physician–patients polytomous data, based on its special correlation and covariance structure, we propose a simple and efficient data generation algorithm, and develop a semi-parametric variance estimator for the kappa statistic. An extensive Monte Carlo simulation study is then conducted to evaluate the performance of empirical coverage probability of the new proposal and alternative methods. The method ignoring dependence within a cluster underestimates the variance, and the variance estimators from new proposal and sampling-based delta method behave reasonably well for at least a moderately large number of clusters (e.g., $K \geq 50$). Moreover, the new proposal has acceptable performance when the number of clusters is small (e.g., $K = 25$), although we need to be cautious about some bias concerns of the kappa statistic when the within-physician correlation is large as $\rho_w > 0.5$. To illustrate the practical application of all the methods, a physician–patients data example from the Enhancing Communication and HIV Outcomes (ECHO) study is analyzed.

1. Introduction

In clinical practice, such as the medical community from general practice and primary health care, the physician–patient relationship is integral to the successful delivery of medical care. The concordance or agreement between the patient and physician on the problem or need and its management may affect the quality of treatment decisions, the quality of patient–physician communication, the satisfaction of both patients and physicians, and then the health outcome of the patients (Vedsted et al., 2002; Beach et al., 2006, 2010; Kerse et al., 2004). Such kind of data is called "physician–patients data" by Zhou and Yang (2014) to emphasize that each physician may have multiple patients. For physician–patients data, patients seen by the same physician form a cluster (we will use the terms of physician and cluster interchangeably), within which patients' responses, and the responses of their physician tend to be more similar. However, till recent works in Kang et al. (2013) and Zhou and Yang (2014), the fact of correlated responses for physician–patients data has been overlooked in the statistical evaluation of agreement using the kappa statistic, a landmark in the development of agreement theory.

As a measure of inter-rater agreement between two observers classifying subjects into mutually exclusive categories, the kappa statistic was introduced by Cohen (1960) to summarize the cross-classification of two nominal variables with identical

* Corresponding author.
E-mail address: tonyyangsxz@gmail.com (Z. Yang).

http://dx.doi.org/10.1016/j.csda.2015.01.007

categories in the matched-pair data. It is a normalized difference between the observed agreement (P_0) and the agreement expected by chance (P_e). Ever since its introduction, the kappa statistic has enjoyed broad applications to different disciplines such as epidemiology (Jakobsson and Westergren, 2005), educational measurement, psychometrics, and diagnostic imaging (Kundel and Polansky, 2003). There are various extensions of the kappa statistic (Zhou and Yang, 2014; Yang and Zhou, 2014, 2015) and reference therein. The most recent contributions by Kang et al. (2013), Zhou and Yang (2014), and Yang and Zhou (2014, 2015) focus on variance estimation of the kappa statistic in the correlated matched-pair data. To evaluate the ratings agreement between physicians and their patients regarding their discussions on the same event, Kang et al. (2013) proposed a simulation-based cluster bootstrap method to analyze the correlated clustered binary data. Motivated by the work of Kang et al. (2013), Zhou and Yang (2014) developed an asymptotic semi-parametric variance estimator for the kappa statistic without further parametric assumptions. However, an apparent limitation of their work is that their method is only applicable for dichotomous data. Hence, based on similar framework, especially the assumption of conditional independence, we develop a new algorithm to generate physician–patients polytomous data, and propose the kappa statistic together with its corresponding semi-parametric variance estimator. The performance of the proposed estimator is studied by comparing it with the non-parametric method from Yang and Zhou (2014, 2015), in terms of empirical coverage probability, root-mean-square error (RMSE), and average width of the 95% confidence interval (CI) for the kappa statistic.

Section 2 presents the proposed kappa statistic and its variance estimator, followed by a description of existing methods in Section 3. The Monte Carlo simulation study for the evaluation of the proposal is summarized in Section 4. In Section 5, we provide a detailed real data analysis example to illustrate the application of the methods. A final conclusion is given in Section 6, and all the technical details are included in Appendices A–F.

2. Kappa statistic for independent data

For a physician and her/his patient, suppose (Y, X) represents a pair of responses from g exhaustive and mutually exclusive categories of a nominal or ordinal scale, there are then g^2 pairs of outcomes. For a random sample of N pairs of responses, such kind of traditional matched-pair data can be summarized in a $g \times g$ contingency table. In particular, let n_{ij} be the count that physician and her/his patient provide the response of categories i and j $(i, j = 1, \ldots, g)$, respectively, n_{i+} be the total count for the response of ith category by the physician, and n_{+j} be similarly defined for the jth category by the patient. Meanwhile, let p_{ij} be the probability of falling into the ith category of physician and the jth category of patient, p_{i+} and p_{+j} be the marginal probabilities such that $p_{i+} = \sum_j p_{ij}$ and $p_{+j} = \sum_i p_{ij}$. Based on the multinomial model, the maximum likelihood estimator of p_{ij} is $\widehat{p}_{ij} = n_{ij}/N$. For dichotomous/binary data, the $g \times g$ table reduces to a 2×2 contingency table.

To define the kappa statistic as an agreement index, Cohen (1960) gave the kappa statistic as the following,

$$\widehat{\mathcal{K}} = \frac{\widehat{P}_o - \widehat{P}_e}{1 - \widehat{P}_e}, \tag{1}$$

where $\widehat{P}_o$ is observed agreement, $\widehat{P}_e$ measures the agreement expected by chance. Their forms are

$$\widehat{P}_o = \sum_{i=1}^{g} \widehat{p}_{ii} = \frac{1}{N}\sum_{i=1}^{g} n_{ii} \quad \text{and} \quad \widehat{P}_e = \sum_{i=1}^{g} \widehat{p}_{i+}\widehat{p}_{+i} = \frac{1}{N^2}\sum_{i=1}^{g} n_{i+}n_{+i},$$

where $\widehat{p}_{i+}$ and $\widehat{p}_{+i}$ are estimators of marginal probabilities p_{i+} and p_{+j} given by

$$\widehat{p}_{i+} = \sum_{j=1}^{g} \widehat{p}_{ij} = \frac{1}{N}\sum_{j=1}^{g} n_{ij} \quad \text{and} \quad \widehat{p}_{+i} = \sum_{j=1}^{g} \widehat{p}_{ji} = \frac{1}{N}\sum_{j=1}^{g} n_{ji}, \quad \text{for } i = 1, \ldots, g.$$

For the kappa statistic, only values between 0 and 1 have useful meaning, a value of 1 indicates a perfect agreement and a value of 0 suggests the agreement is only due to chance. A negative value of kappa statistic can occur when the observed agreement is lower than the agreement expected by chance.

The asymptotic variance of $\widehat{\mathcal{K}}$ in (1) can be estimated (Agresti, 2002, p. 434) as

$$\begin{aligned}\widehat{V}_D \equiv \widehat{\text{Var}}\left(\widehat{\mathcal{K}}\right) = \frac{1}{N\left(1-\widehat{P}_e\right)^2}\Bigg\{&\widehat{P}_o\left(1-\widehat{P}_o\right) + \frac{2\left(1-\widehat{P}_o\right)}{\left(1-\widehat{P}_e\right)}\left[2\widehat{P}_o\widehat{P}_e - \sum_{i=1}^{g}\widehat{p}_{ii}(\widehat{p}_{i+}+\widehat{p}_{+i})\right] \\ &+ \frac{\left(1-\widehat{P}_o\right)^2}{\left(1-\widehat{P}_e\right)^2}\left[\sum_{i=1}^{g}\sum_{j=1}^{g}\widehat{p}_{ij}\left(\widehat{p}_{j+}+\widehat{p}_{+i}\right)^2 - 4\widehat{P}_e^2\right]\Bigg\}.\end{aligned} \tag{2}$$

With $\widehat{\mathcal{K}}$ and its variance estimator $\widehat{V}_D$, an approximate $100(1-\alpha)\%$ CI is $\widehat{\mathcal{K}} \pm z_{1-\alpha/2}\sqrt{\widehat{V}_D}$, where $z_{1-\alpha/2}$ is the upper $100(1-\alpha/2)\%$ percentage point of the standard normal distribution.

3. Kappa statistic for clustered data

Following the setup in Zhou and Yang (2014) for the physician–patients dichotomous data, suppose there are K independent physicians (clusters), and the kth physician has n_k patients $(k = 1, \ldots, K)$, then the total number of patients is