More efficient unconditional tests for exchangeable binary data with equal cluster sizes

Guogen Shan

Department of Environmental and Occupational Health, Epidemiology and Biostatistics Program, University of Nevada Las Vegas, Las Vegas, NV 89154, USA

Abstract

We consider exact unconditional procedures for testing independence of exchangeable binary data with equal-sized clusters. The approximate unconditional approach is recommended as a basic procedure. The exact unconditional procedure based on estimation followed by maximization is recommended for the small number of clusters.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

Correlated binary data are frequently encountered in a wide variety of applications, including longitudinal studies, toxicity experiments, and randomized clinical trials. Several methodologies were developed to analyze such data; for example, the generalized estimating equations methods by Liang and Zeger (1986). A likelihood ratio procedure for analyzing binary data under the assumption of exchangeability was proposed by Bowman and George (1995) and George and Bowman (1995). A set of binary random variables \( X_1, X_2, \ldots, X_m \) is defined to be exchangeable if

\[
P(X_{\pi(1)} = x_1, X_{\pi(2)} = x_2, \ldots, X_{\pi(q)} = x_q) = P(X_1 = x_1, X_2 = x_2, \ldots, X_q = x_q)
\]

is true for any \( q (1 \leq q \leq m) \) and any permutation \( \pi(1), \pi(2), \ldots, \pi(q) \) of \( 1, 2, \ldots, q \), where \( x_i = 0, 1, i = 1, 2, \ldots, m \). In other words, any permutation of the variables \( X_1, X_2, \ldots, X_m \) has the same distribution. In comparison to the aforementioned procedures, the exchangeable binary model involves fewer and more realistic assumptions: the number of clusters is finite, and responses within a cluster are exchangeable (George and Bowman, 1995).

Let \( \lambda_k = P(X_1 = 1, X_2 = 1, \ldots, X_k = 1), k = 1, 2, \ldots, m, \) and \( \lambda_0 = 1 \). Let \( X_i = (X_{i1}, X_{i2}, \ldots, X_{im}), i = 1, 2, \ldots, n, \) be \( n \) independent \( m \times 1 \) vectors of exchangeable binary random variables, where \( n \) is the number of clusters and \( m \) is the common cluster size. The null hypothesis of independence within clusters is given by

\[
\lambda_k = \lambda_1^k, \quad k = 1, 2, \ldots, m.
\]

Let \( W_t \) be the number of \( X_i \) such that \( R_t = t \), where \( R_t = \sum_{j=1}^{m} X_{ij} \). Let \( W = (W_0, W_1, \ldots, W_m) \) and \( w \) represent a realization of \( W \). Under the assumption of exchangeability, \( W \) follows a multinomial distribution with parameters

E-mail address: guogen.shan@unlv.edu.

0167-7152/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.spl.2012.11.014