The paradox of increasing sample size and decreasing power in testing the difference between two independent binomial proportions

Youyi Fong

University of Washington, Department of Biostatistics
E-mail: youyifong@gmail.com

Abstract: Suppose we are to conduct a trial to compare the probabilities of a binary outcome between two populations. Let the number of subjects to be sampled from the two populations be denoted by m and n, respectively. Denote the number of successes to be observed in the two samples by a and c, respectively. Thus a is a binomial random variable binom(m,p1) and b is a binomial random variable binom(n,p2). We are interested in testing the null hypothesis p1=p2. This hypothesis testing problem has been studied extensively in the literature. It is generally agreed among practitioners that unconditional tests such as Boschloo's test (Boschloo 1970) and Barnard's CSM test (Barnard, 1947) are more powerful than conditional tests like Fisher's exact test. But what to make of this difference in power is not without controversy. For example, Barnard (1989) argued that "the various suggestions that have been made to increase the power of Fisher's test for 2 x 2 tables are shown to give no real increases." In this talk we seek to shed some light on this controversy by considering a case of increasing sample size and decreasing power.