

Sampling bias in binary random effects models

Peter McCullagh

Department of Statistics, University of Chicago

Abstract

The need to incorporate random effects to handle correlations in binary responses has long been evident. Cox's suggestion in 1958 for matched pairs is to condition on the sum of the responses on the matched sets or blocks. A more flexible approach by Laird and Ware in 1982 to deal with correlations in longitudinal studies inserts the random effects additively on the logistic scale. This method has been widely used without major modification for the past 25 years, and has been the topic of several books. I will describe an alternative approach in which the data (x, y) pairs are generated by a point process with a random intensity function. The response distribution under the point process model depends on the sampling scheme. Under quota sampling with pre-determined x -values, the distribution coincides with the Laird-Ware model. Under the more natural sampling scheme in which events are observed over a fixed temporal interval, the x -values are also random, and the conditional distribution is different from that in the conventional Laird-Ware model. The second sampling scheme also gives a different likelihood.