

Randomized Response, Statistical Disclosure Control and Misclassification: a Review

Ardo van den Hout and Peter G.M. van der Heijden

Utrecht University, Faculty of Social Sciences, Dept. of Methodology and Statistics, Heidelberglaan 2, 3584 CS Utrecht, The Netherlands

Summary

This paper discusses analysis of categorical data which have been misclassified and where misclassification probabilities are known. Fields where this kind of misclassification occurs are randomized response, statistical disclosure control, and classification with known sensitivity and specificity. Estimates of true frequencies are given, and adjustments to the odds ratio are discussed. Moment estimates and maximum likelihood estimates are compared and it is proved that they are the same in the interior of the parameter space. Since moment estimators are regularly outside the parameter space, special attention is paid to the possibility of boundary solutions. An example is given.

Key words: Contingency table; EM algorithm; Misclassification; Odds ratio; Randomized response; Sensitivity; Specificity; Statistical disclosure control.