ASSESSING PRECISION OF LASSO CONDITIONAL LOGISTIC ESTIMATES IN THE CASE-CROSSOVER DESIGN

Avalos, Marta

INSERM U897 and University Bordeaux Segalen, Bordeaux, France

The case-crossover design introduced by Maclure (1991) is an observational epidemiological study for analyzing the effects of transient exposures on the risk of acute-onset events. This design is based on exclusively selecting subjects that have experienced the event under investigation. The aim is to identify the short-term, transient triggers of event, rather than to identify who is at highest risk of disease. The association between event onset and risk factors is then estimated by comparing exposure during the period of time just prior to the event onset (case period) to the same subject's exposure during one or more control periods. The standard tool for this analysis is conditional logistic regression (Breslow and Day, 1980; Maclure, 1991).

Until present, variable selection issues have not received much attention when modeling the risk of an event of interest under the case-crossover design. On the one hand, time-invariant potential confounders are controlled by design. On the other hand, studies in the literature using the case-crossover design are often concerned by few main risk factors. Nevertheless, advances in technology for data collection lead to complex situations such as those involving many predictors. This may also be the case in studies based on the case-crossover design.

The lasso procedure (Tibshirani, 1996), which relies on an L1 penalty, has been widely used for accommodating high-dimensional predictors. This approach has also been adapted to conditional logistic regression in the context of matched case-control studies (Avalos, 2009). However, a drawback of LASSO regression is that no measure of uncertainty is provided.

In this work we study by simulations the behavior of nonparametric bootstrap percentile confidence intervals in assessing significance of lasso point estimates. We also illustrate the use of this method with some examples under the case-crossover design.

1. Avalos M. (2009) Model selection via the lasso in conditional logistic regression. In: Conference abstracts. Second International Biometric Society Channel Network Conference, Belgium (2009).
2. Breslow, N. E. and Day, N. E. (1980). Statistical methods in cancer research. Volume 1: The analysis of case-control studies. IARC Scientific Publications, Lyon.
3. Maclure, M. (1991). The case-crossover design: a method for studying transient effects on the risk of acute event. Am J Epidemiol. 133, 144–153.
4. Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. J. Royal. Statist. Soc B. 58, 267–288.