

Conference abstracts

2nd Channel Network Meeting, Ghent, Belgium, April 6-8, 2009

Model selection via the lasso in conditional logistic regression

Marta Avalos

Biostatistics team of INSERM U897 and University of Bordeaux 2, France
marta.avalos@isped.u-bordeaux2.fr

We propose a model selection procedure in the context of matched case-control studies and, more specifically, for the conditional logistic regression. The method is based on penalized conditional likelihood with an L_1 -type penalty of the regression coefficients, the so-called lasso. This penalty, which shrinks coefficients to improve the accuracy of prediction, is particularly adapted when the number of covariates is large (with respect to the number of events) or in case of collinearity between them. An attractive property of the lasso is that it performs parameter estimation and variable selection simultaneously. The implementation of the method is based on a simple modification of the R-package *penalized* for fitting generalized linear and Cox proportional hazard models with L_1 or L_2 penalties, developed by J. Goeman. The effectiveness of the algorithms allows the use of resampling methods for choosing the regularization term and for evaluating the stability of the selected model. Thus, K-fold likelihood cross-validation and bootstrap are applied, both taking into account the dependent nature of data. The methodology is illustrated with some examples.