A nonparametrical goodness-of-fit test.

Elke Moons, Marc Aerts, Geert Wets

The assessment of the fit of a model is a very important component in any modeling procedure. Goodness-of-fit tests try to evaluate how well model-based predicted outcomes coincide with the observed data. In transport modeling, logit models have become a standard. However, just in these logistic regression models, investigating goodness-of-fit often is problematic when continuous covariates are modeled, since the approximate chi-squared null distributions for the Pearson test statistic is no longer valid. Categorization might provide a solution for this problem, but it is often not clear how the categories should be defined.

Hosmer and Lemeshow (1980) were the first to propose a goodness-of-fit test that can be used for logistic regression models with continuous predictors. They suggested using a Pearson-like chi-square statistic, but the groups are formed according to deciles of risk. In this way they solve the problem of categorizing, though it is well-known that this is at the cost of power. Many other methods and approaches, see e.g. Azzalini, Bowman and Haerdle (1989), le Cessie and van Houwelingen (1991, 1995), Aerts, Claeskens and Hart (1999, 2000), among others, were examined and most of them were based on nonparametric concepts. They have been shown to have good power characteristics, however, when many explanatory variables are modeled, most of these methods are faced with the curse of dimensionality.

The test statistic that we propose here is similar in approach to the Hosmer and Lemeshow lack-of-fit test statistic in that the observations are classified into distinct groups. However, in our proposed test the grouping is not according to the fitted probabilities under the null model. We let a recursive partitioning algorithm, as used in the classification trees of Breiman et al. (1984), divide the sample space into different groups. This will affect the power characteristics of the test statistic. Classification trees are nonparametric in nature and they can also deal with large and complex datasets with many explanatory variables. Therefore, they are frequently used in data mining applications.

The distribution of this test statistic has been studied through simulations and we will compare its performance to this of the Hosmer and Lemeshow test. An analysis on a large and complex real-world transportation data set will be provided to exemplify the procedure.

key words: Goodness-of-fit test; Logistic regression; Classification trees.

References

Aerts, M., Claeskens, G. and Hart, J. (1999). Testing the fit of a parametric function. *Journal of the American Statistical Association*, **94**, 869--879.

Aerts, M., Claeskens, G. and Hart, J. (2000). Testing lack of fit in multiple regression. *Biometrika*, **87**, 405--424.

Azzalini, A., Bowman, A.W. and Haerdle, W. (1989). On the use of nonparametric regression for model checking. *Biometrika*, **76**, 1--11.

Breiman, L., Friedman, J.H., Olshen, R.A. and Stone, C.J. (1984). *Classification and Regression Trees*. Wadsworth Statistics/Probability Series.

Cressie, N. and Read, T. (1984). Multinomial goodness of fit tests. *Journal of the Royal Statistics Society, Series B*, **46**, 440--464.

Hosmer, D.W., Hosmer, T., Lemeshow, S., and le Cessie, S. (1996). A comparison of goodness-of-fit tests for the logistic regression model. *Statistics in Medicine*, **16**, 965--980.

Hosmer, D.W. and Lemeshow, S. (1980). A goodness-of-fit test for the multiple regression model. *Communications in Statistics*, **A10**, 1043--1069

Hosmer, D.W. and Lemeshow, S. (1989). *Applied Logistic Regression*. Wiley, New York.

Le Cessie, S. and van Houwelingen, J.A. (1991) A goodness-of-fit test for binary data based on smoothing residuals. *Biometrics*, **47**, 1267--1282

Le Cessie, S. and van Houwelingen, J.A. (1993) Building logistic models by means of a non parametric goodness of fit test: a case study. *Statistica Neerlandica*, **47**, 97--109