

New modelling method applied to continuous variables in credit risk area

Abstract

This paper presents a new modelling method applied to continuous variables in credit risk area. The method is illustrated on the loss given default prediction but can be applied to any other data and variables, also considering binary variables. The proposed approach is based on data augmentation and construction a model in two phases. In the first phase a comparison between all pairs of observations in the modelling data set is made and an appropriate regression model is constructed to predict the differences between observations. This technique increases the number of observations and focuses on developing a pattern of differences between the original observations rather than looking for a profile for the original observations. Given a certain observation, a wide set of possible estimators can be used for estimation because the Phase I modelling results are a set of estimators based on a comparison of the selected observation with many other comparative ones. The results from the phase I are combined with the additional model built in phase II, which predicts the value of the residual error from the model in phase I. This approach allows to significantly improve the quality of predictions compared to classic regression analysis or others estimators based only on the phase I model. We present the detailed model estimators as well as many potential choices for the final estimators. Additionally a good property for variation coefficient of the proposed method is proven.

Authors & Affiliations

Dr Paweł Kopciuszewski¹, Dr hab. Aneta Ptak-Chmielewska²

¹ING Hubs Poland, Warsaw, Poland; ²Warsaw School of Economics, Warsaw, Poland