Sampling and validation methods
for hazard estimation

Madorno F., Pafundi P., Artaria A.
Università degli Studi Milano Bicocca
f.madorno@campus.unimib.it — +39 02 6448-5889

Introduction

In this poster we compare in terms of predictive performance sampling and validation methods for hazard estimation in credit risk assessment. More precisely, we compare proportional hazard models based on the Cox model (Cox, 1972) with non parametric survival approaches based on random survival forest (Ishwaran et al., 2008) under different validation settings. On the basis of predictive performance measures (i.e. Brier Score and Prediction Error), we compare validation techniques based on the Bootstrap: Bootstrap without replacement (B=200), BOOT 632 (see e.g. Davison and Hinkley, 1997 and BOOT 632 PLUS (Elton and Tibshirani, 1997). Empirical evidence are given on a real data set provided by a credit rating agency composed of 742 small and medium enterprises, 9 financial ratios, a binary dependent variable which expresses a solvency indicator and a duration indicator.

Main Objectives

1. Explore how validation techniques affect model assessment in predictive models with a special focus on time-dependent survival models.
2. Compare in terms of predictive performance measures (i.e. Prediction Error and Brier score) semi-parametric proportional hazard models with respect to non parametric techniques based on random survival forest.

Models and methods

Models

• Semi-parametric model based on the proportional hazard model (Cox, 1972): the Cox model defines the cumulative hazard function dependent on the vector of predictor variables:

\[ \Lambda(t) = \Lambda_0(t) \exp(\beta X) \]

where \( \Lambda_0(t) \) is the baseline hazard function.

• Non parametric model based on the Random Survival Forests (Ishwaran et al., 2008): is a non parametric predictive dynamic method for the analysis of right-censored survival data that recognize the interaction between pairs of variables.

Selected validation methods (see e.g. Gerds and van de Wiel, 2011, and the references therein)

• Bootstrap cross validation without replacement (Elton and Tibshirani, 1993)
• Bootstrap 632 (Davison and Hinkley, 1997)
• Bootstrap 632a (Elton and Tibshirani, 1997)

Performance Measures

• Brier score (see e.g. Gneiting and Raftery, 2007)
• Prediction error (see e.g. Mogensen et al., 2012)

Data and flow analysis

We analyzed real data coming from a credit risk dataset composed of 742 observations for a binary target variable with its duration (Solvency,Duration) and 9 covariates, i.e. financial ratios:

- Supplier target days
- Capital to asset ratio
- Cost income ratio
- Credit ratio
- Capital to debt ratio
- Trade payable ratio
- Liquidity ratio
- Outstanding capital structure
- Equity ratio

We observe that the classic Bootstrap performs worse than the other techniques. It would seem that the results affirm that the two generalizations of Bootstrap (Bootstrap 632, Bootstrap 632a) perform better than the classical bootstrap.

In the simulation study we controlled the Monte Carlo error through the number of Monte Carlo runs (5000 Monte Carlo runs). So, we control the variation of the estimated value. For all methods this error is lower than 5%.

Conclusions

1. In this poster we have compared in terms of different sampling techniques for validation assessment the performance of Cox model and random survival forest.
2. In terms of the data, at the hand, the best model in terms of prediction performance is the random survival forest and the lowest prediction errors are achieved using the Bootstrap 632 as validation sampling technique. In terms of Brier score the results show that the lowest Brier scores are obtained by Bootstrap 632 validation technique. We note that in both cases the selected models perform better than the reference model.

Forthcoming Research

Improve the comparison considering several techniques of validation as done in (Madorno et al., 2013).

References