Chapter 7

Survival Analysis and ROC Analysis in Analyzing Credit Risks: Assessing Default Risks Over Time

Nan Hu
University of Utah, USA

Haojie Cheng
University of Utah, USA

ABSTRACT

As the aim of large banks has been changing to select customers of highest benefits, it is important for banks to know not only if but also when a customer will default. Survival analyses have been used to estimate over time risk of default or early payoff, two major risks for banks. The major benefit of this method is that it can easily handle censoring and competing risks. An ROC curve, as a statistical tool, was applied to evaluate credit scoring systems. Traditional ROC analyses allow banks to evaluate if a credit-scoring system can correctly classify customers based on their cross-sectional default status, but will fail when assessing a credit-scoring system at a series of future time points, especially when there are censorings or competing risks. The time-dependent ROC analysis was introduced by Hu and Zhou to evaluate credit-scoring systems in a time-varying fashion and it allows us to assess credit scoring systems for predicting default by any time within study periods.

INTRODUCTION

Risk managements, in general, are the identification, assessment and prioritization of risks followed by application of resources to monitor the probability and impact of unfortunate events (Hubbard 2009). Among different risks, financial risk is the high-priority risk for all businesses. The financial risk can be classified into various types such as market risk, credit risk, liquidity risk, operational risk and legal risk (Ray 2011). In finance, the risk management usually focuses on market risks and credit risks. Mar-DOI: 10.4018/978-1-5225-3932-2.ch007
Survival Analysis and ROC Analysis in Analyzing Credit Risks

Market risks present in the market, and usually arise from market movements. The equity risk, interest risk, currency risk and commodity risk are different types of market risks. The credit risk refers to risk that borrowers will default on their debts by failing to make payments. The credit risk management deals with the probability of non-payment from the debtors and is the fundamental work of financial institutions. With the recent global financial crisis and the following credit crunch, more attentions have been paid to analyzing and managing credit risks. In this chapter, we focus on analyzing credit risks using survival analyses and on evaluating the accuracies of credit scoring systems for predicting default risks using a receiver operating characteristic (ROC) curve. In the Background section, we give the motivation for applying survival models and ROC curve methods for analyzing credit risks. Then, we review the applications of survival analysis in analyzing default risks and illustrate, in detail, how survival analysis methods can be implemented to estimate hazard of default and survival functions (proportion of default-free), to compare survival functions among different groups, and to make statistical inference of covariate effect on the risk of default. In addition, we demonstrate how ROC analysis can be applied to assess accuracies of credit scoring systems for predicting the cross-sectional status of default using a traditional ROC curve. We give mathematical definitions for different classification accuracy parameters, such as sensitivity and specificity, and describe the procedure to establish an ROC curve. Furthermore, we describe the extension from traditional ROC curves to time-dependent ROC curves, which can be used to evaluate the classification accuracies for predicting defaults by a specific future time in a time-varying fashion. Audiences of this book chapter should realize that survival analyses and ROC methods in general are two separate topics. However, the time-dependent ROC method is actually combing traditional ROC curve approach with survival analysis methods. We conclude this chapter by summarizing the application of survival analyses and ROC methods in analyzing credit risks and providing future research directions on these two topics.

BACKGROUND

The primary goals of risk analysis for the credit industry are to monitor the credit risks and to decide whether it should grant credit to an applicant. Traditionally, it is done by estimating the probability that the applicant will eventually default (Stepanova & Thomas 2002). More recently, however, the aim of banks and credit companies has been changing to select customers of highest benefits. This changes implies that it is important not only if but also when a customer will default in a future time (Banasik et al. 1999). It is possible that if the time to default is long, the acquired interest from the customer will compensate or even exceed the losses resulting from default. On the other hand, an early pay-off by the customer can also impact the profitability of financial businesses. Depending on when the actual repayment occurs, the lender will lose a proportion of interest in the loan. It has been shown that survival analysis can be applied to estimate time to default or early repayment (Narian 1992, Banasik et al. 1999).

The survival analysis (also known as the time-to-event analysis), is a branch of statistics dealing with analysis of time until one or more events happens. The major benefit of using the survival analysis lies in the fact that it can handle the censored data when modeling the hazard or risk for the default or the early pay-off. Survival analysis has been intensively applied in medical research and reliability studies. It allows researchers to incorporate information from both censored and uncensored observations in estimating model parameters. In analyzing credit risks for the financial industry, profits realized on loans, credit cards or their related products depend heavily on whether borrowers pay interest regularly,