

COPYRIGHT NOTICE:

Christian Gourieroux & Joann Jasiak: The Econometrics of Individual Risk

is published by Princeton University Press and copyrighted, © 2007, by Princeton University Press. All rights reserved. No part of this book may be reproduced in any form by any electronic or mechanical means (including photocopying, recording, or information storage and retrieval) without permission in writing from the publisher, except for reading and browsing via the World Wide Web. Users are not permitted to mount this file on any network servers.

Follow links for Class Use and other Permissions. For more information send email to: permissions@pupress.princeton.edu

1

Introduction

1.1 Market Risk and Individual Risk

People and businesses operate in uncertain environments and bear a variety of risks. As the service sector of the economy grows rapidly, the risk exposure of financial institutions, insurers, and marketers becomes more and more substantial. The risks grow and diversify in parallel with the offering of market and retail financial products, insurances, and marketing techniques. Private businesses adapt to the increasingly risky environment by implementing quite sophisticated, and often costly, systems of risk management and control. Recent corporate history has proven, however, that these are far from flawless and that financial losses can be devastating.

The risk can be viewed from four perspectives. The first concerns the occurrence of a loss event. One can think of it as an answer to the question, Did a loss event occur or not? The answer is either yes or no.¹ The second is about the frequency or count of loss events in a period of time. It answers the question, How many losses were recorded in a year? The answer is zero or any positive integer. The third refers to timing. It is about determining when a loss event has occurred. The answer is an interval of time, usually measured with reference to a fixed point of origin, such as the beginning of a contract, for example. The last is the severity. It tells us how much money is spent to cover the losses caused by a risk event. The answer is measured in currency units, such as dollars.

From all four perspectives, risk is quantifiable. Therefore, it is easy to imagine that risk can be formalized using statistical methodology. The elementary approach consists of determining what type of random variable would match the four aspects of risk listed in the last paragraph. Accordingly, the occurrence can be modeled by a qualitative dichotomous variable, the frequency by a count variable, the timing by a duration variable, and the severity by any continuous, positive variable.

The econometric analysis concerns modeling, estimation, and inference on random variables. In order to proceed to risk assessment we need to first establish the assumptions about the mechanism that generates the risk. The concept that acts as a guideline for this book is the notion that any risk is associated with an individual who is either bearing the risk or is perceived as risky by another individual. At this

¹For clarity, the answer “uncertain” is not admissible.

point, an individual can be a person, a company, an insurance policy, or a credit agreement. It is crucial that it is an entity that can be depicted by some individual characteristics, which, like risk, can be quantified and recorded in a data set. The individual characteristics are an essential part of any model for individual risk assessment. Their statistical summary is called a score.

Among the approaches to risk modeling it is important to distinguish between the parametric and nonparametric methods. The parametric methods consist in choosing a model based on a specific distribution, characterized by a set of parameters to be estimated. The nonparametric approach is to some extent model free and relies on generic parameters, such as means, variances, covariances, and quantiles. The semi-parametric methods bridge the gap and share some features of the pure parametric and nonparametric approaches.

In the remainder of this chapter we elaborate more on the four types of risk variable and the score. In the final part, we discuss the organization of the book.

1.2 Risk Variable

Any loss event, such as a road accident or default on a corporate loan, can be viewed as the outcome of a random phenomenon generating economic uncertainty. An event associated with a random phenomenon that has more than one possible outcome is said to be contingent or random.

Let us consider an individual car owner whose car is insured for the period January–December 2000. After this period, the realization of a random risk variable is known. According to the classification given earlier one can consider the following risk variables.

- (i) **Dichotomous qualitative variable.** The dichotomous qualitative variable indicates if any road accidents were reported to the insurance agency, or equivalently if any claims on the automobile insurance were filed in the given year. To quantify the two possible outcomes, “yes” and “no,” the dummies 1 and 0 are assigned to each of them respectively.
- (ii) **Count variable.** The count variable gives the number of claims filed on the automobile insurance in the year 2000.
- (iii) **Duration variable.** The duration variable can represent the time beginning with the issuing of the insurance policy and ending with the first incidence of a claim. It can also be the time from the incidence of a claim to the time of its report to the insurer, or else the time from the reporting of a claim to its settlement.
- (iv) **Continuous variable.** The continuous positive variable can represent the amount of money paid by the insurer to settle each claim, or the total cost of all claims filed in the year 2000.

Let us consider a series of loss events recorded sequentially in time, along with a characteristic, such as the severity. This sequence of observations associated with

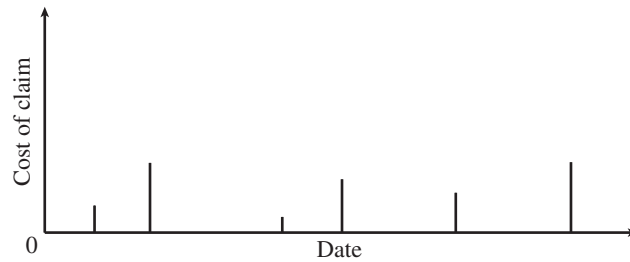


Figure 1.1. A claim history.

specific points in time forms the so-called *marked-point process*. A marked-point process can model, for example, the individual risk history of a car owner, represented by a sequence of timed road accidents. A trajectory of a marked-point process² is shown in Figure 1.1.

Each event is indicated by a vertical bar. The height of each bar is used to distinguish between accidents of greater or lesser severity. The bars are irregularly spaced in time because the time intervals between subsequent accidents are not of equal length.

1.3 Scores

The score is a quantified measure of individual risk based on individual characteristics. The dependence between the probability of default and individual characteristics was established for the first time by Fitzpatrick (1932) for corporate credit, and by Durand (1941) for consumer credit. Nevertheless, it took about 30 years to develop a technique that would allow the quantification of the individual propensity to cause financial losses. In 1964, Smith computed a risk index, defined as the sum of default probabilities associated with various individual characteristics. Even though this measure was strongly biased (since it disregards the fact that individual characteristics may be interrelated), it had the merit of defining risk as a scalar function of covariates that represent various individual characteristics. It was called the *score* and became the first tool that allowed for ranking the individuals in a sample. Scores are currently determined by more sophisticated methods, based on models such as the linear discriminant, or the logit. In particular, scores are used in credit and insurance to distinguish between low-risk (good-risk) and high-risk (bad-risk) individuals. This procedure is called *segmentation*. In marketing, segmentation is used to distinguish the potential buyers of new products or to build mailing lists for advertising by direct mail.

²This a selected trajectory among many possible ones.

1.4 Organization of the Book

The book contains eleven chapters. Chapters 2–6 present models associated with various types of risk variable. The risk models based on (1) a dichotomous qualitative variable appear in Chapter 2, (2) a count variable appear in Chapter 4, and (3) a duration variable appear in Chapter 6. Basic estimators and simple sample-based modeling techniques are given in Chapter 3. Up to Chapter 6 the methodology relies on the assumption of independent and identically distributed (i.i.d.) variables. Chapters 7 and 8 cover departures from the i.i.d. assumption and full observability of variables. Chapter 9 discusses multiple scores. Chapters 10 and 11, on panel data and the “Value-at-Risk” (VaR), respectively, can be seen as smorgasbords of selected topics, as comprehensive coverage of these subjects is beyond the scope of this text.

Chapters 2–7 can be taught to graduate students at either the master’s or doctorate levels. At the master’s level, the sections on technically advanced methods can be left out. The material covered in the first seven chapters can be taught in a course on risk management offered in an MBA program or in an M.A. program in Mathematical Finance, Financial Engineering, Business and Economics, and Economics. The text in its entirety can be used as required reading at the Ph.D. level in a course on topics in advanced econometrics or advanced risk management.³ The text can also be used as suggested reading in a variety of economic and financial courses.

A detailed description of the book follows.

Chapter 2 considers a dichotomous qualitative risk variable. The links between this variable and individual covariates can be examined by comparing the distribution of the characteristics of individuals who defaulted on a loan with the characteristics of those who repaid the debt. Econometric models introduced in this chapter include the discriminant analysis and the logit model.

Chapter 3 presents the maximum likelihood estimation methods, their implementation and related tests.

In practice, a quality of a score may deteriorate over time and regular updating may be required to preserve its quality. Chapter 4 introduces statistical methods that allow for monitoring the score performance.

The models for count variables of risk are introduced in Chapter 5. These include the Poisson and the negative-binomial regression models, the latter accommodating unobserved heterogeneity. This chapter describes their application to automobile insurance for determination and updating of risk premiums.

Chapter 6 examines the timing of default, with the focus on the analysis of durations. We describe the basic exponential model, and study the effect of unobservable

³The first draft of the book was taught in the Advanced Econometrics class at the University of Toronto, attended by Ph.D. and M.A. students in Economics and Finance, including a group of students from the Rotman School of Management. The last five chapters were taught in Ph.D. courses on credit risk in Paris, Geneva, and Lausanne.

heterogeneity. We also discuss semi-parametric models with accelerated and proportional hazards. Applications include the design of pension funds and the pricing of corporate bonds.

Chapter 7 covers the problems related to endogenous selection of samples of individuals for risk modeling. Endogenous selection can result in biased score, wrong segmentation, and unfair pricing. Various examples of endogenous selection and the associated correction techniques are presented.

Chapter 8 introduces the transition models for dynamic analysis of individual risks. These models are used to predict risk on a portfolio of individual contracts with different termination dates.

In the presence of multiple risks, the total risk exposure has to be summarized by several scores (ratings). Examples of the use of multiple scores are given in Chapter 9. In this framework, profit maximization is discussed, and the approach for selecting the minimal number of necessary scores is outlined.

Chapter 10 examines serial dependence in longitudinal data. The Poisson and the compound Poisson models, the nonlinear autoregressive models, and models with time-dependent heterogeneity are presented.

The econometric models for credit quality rating transitions and management of credit portfolios are discussed in Chapter 11. As in Chapter 10, the content is limited to selected topics as comprehensive coverage is beyond the scope of this text.

References

- Durand, D. 1941. *Risk Elements in Consumer Installment Lending*. Studies in Consumer Installment Financing, Volume 8. New York: National Bureau of Economic Research.
- Fabozzi, F. 1992. *The Handbook of Mortgage Backed Securities*. Chicago, IL: Probus.
- Fabozzi, F., and F. Modigliani. 1992. *Mortgage and Mortgage Backed Securities*. Harvard, MA: Harvard Business School Press.
- Fitzpatrick, P. 1932. A comparison of ratios of successful industrial enterprises with those of failed firms. *Certified Public Accountant* 12:598–605, 659–62, 721–31.
- Frachot, A., and C. Gouriéroux. 1995. *Titrisation et Remboursements Anticipés*. Paris: Economica.
- Hassett, M., and D. Stewart. 1999. *Probability for Risk Management*. Winsted, CT: Actex.
- Klugman, S., H. Panjer, and G. Willmot. 2004. *Loss Models: From Data to Decisions*. Wiley.
- Sandhusen, R. 2000. *Marketing*. Hauppauge, NY: Barron's.
- Smith, P. 1964. Measuring risk installment credit. *Management Science* November:327–40.