Nonlife Actuarial Models

Actuaries must pass exams, but more than that: they must put knowledge into practice. This coherent book gives complete syllabus coverage for Exam C of the Society of Actuaries (SOA) while emphasizing the concepts and practical application of nonlife actuarial models. Ideal for those approaching their professional exams, it is also a class-tested textbook for undergraduate university courses in actuarial science.

All the topics that students need to prepare for Exam C are here, including modeling of losses, risk, and ruin theories, credibility theory and applications, and empirical implementation of loss models. The book also covers more recent topics, such as risk measures and bootstrapping. Readers are assumed to have studied statistical inference and probability at the introductory undergraduate level.

Numerous examples and exercises are provided, with many exercises adapted from past Exam C questions. Computational notes on the use of Excel are included. Teaching slides are available for download.

International Series on Actuarial Science

Christopher Daykin, Independent Consultant and Actuary
Angus Macdonald, Heriot-Watt University

The International Series on Actuarial Science, published by Cambridge University Press in conjunction with the Institute of Actuaries and the Faculty of Actuaries, contains textbooks for students taking courses in or related to actuarial science, as well as more advanced works designed for continuing professional development or for describing and synthesizing research. The series is a vehicle for publishing books that reflect changes and developments in the curriculum, that encourage the introduction of courses on actuarial science in universities, and that show how actuarial science can be used in all areas where there is long-term financial risk.
NONLIFE ACTUARIAL MODELS
Theory, Methods and Evaluation

YIU-KUEN TSE
Singapore Management University
To Vicky
# Contents

<table>
<thead>
<tr>
<th>Preface</th>
<th>page xiii</th>
</tr>
</thead>
<tbody>
<tr>
<td>Notation and convention</td>
<td>xv</td>
</tr>
</tbody>
</table>

## Part I Loss models

1. **Claim-frequency distribution** 3
   1.1 Claim frequency, claim severity, and aggregate claim 4
   1.2 Review of statistics 4
   1.3 Some discrete distributions for claim frequency 6
      1.3.1 Binomial distribution 7
      1.3.2 Geometric distribution 8
      1.3.3 Negative binomial distribution 9
      1.3.4 Poisson distribution 11
   1.4 The \((a, b, 0)\) class of distributions 15
   1.5 Some methods for creating new distributions 20
      1.5.1 Compound distribution 21
      1.5.2 Mixture distribution 31
   1.6 Excel computation notes 34
   1.7 Summary and conclusions 34
   Exercises 36

2. **Claim-severity distribution** 41
   2.1 Review of statistics 42
      2.1.1 Survival function and hazard function 42
      2.1.2 Mixed distribution 44
      2.1.3 Expected value of function of random variable 45
      2.1.4 Distribution of function of random variable 46
   2.2 Some continuous distributions for claim severity 49
      2.2.1 Exponential distribution 49
## Contents

2.2.2 Gamma distribution 50
2.2.3 Weibull distribution 51
2.2.4 Pareto distribution 51
2.3 Some methods for creating new distributions 52
  2.3.1 Transformation of random variable 53
  2.3.2 Mixture distribution 54
  2.3.3 Splicing 58
2.4 Tail properties of claim severity 59
2.5 Effects of coverage modifications 66
  2.5.1 Deductible 66
  2.5.2 Policy limit 72
  2.5.3 Coinsurance 73
  2.5.4 Effects of inflation 76
  2.5.5 Effects of deductible on claim frequency 77
2.6 Excel computation notes 79
2.7 Summary and conclusions 81
Exercises 82

3 Aggregate-loss models 86
  3.1 Individual risk and collective risk models 87
  3.2 Individual risk model 88
    3.2.1 Exact distribution using convolution 89
    3.2.2 Exact distribution using the De Pril recursion 92
    3.2.3 Approximations of the individual risk model 94
  3.3 Collective risk model 96
    3.3.1 Properties of compound distributions 96
    3.3.2 Panjer recursion 98
    3.3.3 Approximations of the collective risk model 100
    3.3.4 Compound Poisson distribution and individual risk model 102
  3.4 Coverage modifications and stop-loss reinsurance 103
  3.5 Summary and conclusions 108
Exercises 108

Part II Risk and ruin 113

4 Risk measures 115
  4.1 Uses of risk measures 116
  4.2 Some premium-based risk measures 117
  4.3 Axioms of coherent risk measures 118
  4.4 Some capital-based risk measures 120
    4.4.1 Value-at-Risk (VaR) 120
Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.4.2</td>
<td>Conditional tail expectation and related measures</td>
<td>123</td>
</tr>
<tr>
<td>4.5</td>
<td>More premium-based risk measures</td>
<td>129</td>
</tr>
<tr>
<td>4.5.1</td>
<td>Proportional hazard transform and risk-adjusted premium</td>
<td>129</td>
</tr>
<tr>
<td>4.5.2</td>
<td>Esscher transform and risk-adjusted premium</td>
<td>132</td>
</tr>
<tr>
<td>4.6</td>
<td>Distortion-function approach</td>
<td>133</td>
</tr>
<tr>
<td>4.7</td>
<td>Wang transform</td>
<td>136</td>
</tr>
<tr>
<td>4.8</td>
<td>Summary and conclusions</td>
<td>138</td>
</tr>
</tbody>
</table>

| Exercises | 139 |

5 Ruin theory | 143

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1</td>
<td>Discrete-time surplus and events of ruin</td>
<td>144</td>
</tr>
<tr>
<td>5.2</td>
<td>Discrete-time ruin theory</td>
<td>145</td>
</tr>
<tr>
<td>5.2.1</td>
<td>Ultimate ruin in discrete time</td>
<td>146</td>
</tr>
<tr>
<td>5.2.2</td>
<td>Finite-time ruin in discrete time</td>
<td>150</td>
</tr>
<tr>
<td>5.2.3</td>
<td>Lundberg’s inequality in discrete time</td>
<td>152</td>
</tr>
<tr>
<td>5.3</td>
<td>Continuous-time surplus function</td>
<td>157</td>
</tr>
<tr>
<td>5.4</td>
<td>Continuous-time ruin theory</td>
<td>159</td>
</tr>
<tr>
<td>5.4.1</td>
<td>Lundberg’s inequality in continuous time</td>
<td>159</td>
</tr>
<tr>
<td>5.4.2</td>
<td>Distribution of deficit</td>
<td>163</td>
</tr>
<tr>
<td>5.5</td>
<td>Summary and conclusions</td>
<td>165</td>
</tr>
</tbody>
</table>

| Exercises | 165 |

Part III Credibility | 169

6 Classical credibility | 171

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1</td>
<td>Framework and notations</td>
<td>172</td>
</tr>
<tr>
<td>6.2</td>
<td>Full credibility</td>
<td>173</td>
</tr>
<tr>
<td>6.2.1</td>
<td>Full credibility for claim frequency</td>
<td>173</td>
</tr>
<tr>
<td>6.2.2</td>
<td>Full credibility for claim severity</td>
<td>177</td>
</tr>
<tr>
<td>6.2.3</td>
<td>Full credibility for aggregate loss</td>
<td>179</td>
</tr>
<tr>
<td>6.2.4</td>
<td>Full credibility for pure premium</td>
<td>181</td>
</tr>
<tr>
<td>6.3</td>
<td>Partial credibility</td>
<td>182</td>
</tr>
<tr>
<td>6.4</td>
<td>Variation of assumptions</td>
<td>184</td>
</tr>
<tr>
<td>6.5</td>
<td>Summary and discussions</td>
<td>185</td>
</tr>
</tbody>
</table>

| Exercises | 186 |

7 Bühlmann credibility | 190

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.1</td>
<td>Framework and notations</td>
<td>191</td>
</tr>
<tr>
<td>7.2</td>
<td>Variance components</td>
<td>192</td>
</tr>
<tr>
<td>7.3</td>
<td>Bühlmann credibility</td>
<td>201</td>
</tr>
<tr>
<td>7.4</td>
<td>Bühlmann–Straub credibility</td>
<td>208</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
<td>Pages</td>
</tr>
<tr>
<td>---------</td>
<td>-------</td>
<td>-------</td>
</tr>
<tr>
<td>7.5</td>
<td>Summary and discussions</td>
<td>216</td>
</tr>
<tr>
<td></td>
<td>Exercises</td>
<td>217</td>
</tr>
<tr>
<td>8</td>
<td>Bayesian approach</td>
<td>223</td>
</tr>
<tr>
<td>8.1</td>
<td>Bayesian inference and estimation</td>
<td>224</td>
</tr>
<tr>
<td></td>
<td>8.1.1 Posterior distribution of parameter</td>
<td>225</td>
</tr>
<tr>
<td></td>
<td>8.1.2 Loss function and Bayesian estimation</td>
<td>228</td>
</tr>
<tr>
<td></td>
<td>8.1.3 Some examples of Bayesian credibility</td>
<td>230</td>
</tr>
<tr>
<td>8.2</td>
<td>Conjugate distributions</td>
<td>234</td>
</tr>
<tr>
<td></td>
<td>8.2.1 The gamma–Poisson conjugate distribution</td>
<td>235</td>
</tr>
<tr>
<td></td>
<td>8.2.2 The beta–geometric conjugate distribution</td>
<td>235</td>
</tr>
<tr>
<td></td>
<td>8.2.3 The gamma–exponential conjugate distribution</td>
<td>235</td>
</tr>
<tr>
<td>8.3</td>
<td>Bayesian versus Bühlmann credibility</td>
<td>235</td>
</tr>
<tr>
<td>8.4</td>
<td>Linear exponential family and exact credibility</td>
<td>242</td>
</tr>
<tr>
<td>8.5</td>
<td>Summary and discussions</td>
<td>248</td>
</tr>
<tr>
<td></td>
<td>Exercises</td>
<td>248</td>
</tr>
<tr>
<td>9</td>
<td>Empirical implementation of credibility</td>
<td>253</td>
</tr>
<tr>
<td>9.1</td>
<td>Empirical Bayes method</td>
<td>254</td>
</tr>
<tr>
<td>9.2</td>
<td>Nonparametric estimation</td>
<td>255</td>
</tr>
<tr>
<td>9.3</td>
<td>Semiparametric estimation</td>
<td>270</td>
</tr>
<tr>
<td>9.4</td>
<td>Parametric estimation</td>
<td>271</td>
</tr>
<tr>
<td>9.5</td>
<td>Summary and discussions</td>
<td>273</td>
</tr>
<tr>
<td></td>
<td>Exercises</td>
<td>274</td>
</tr>
</tbody>
</table>

Part IV Model construction and evaluation

10 Model estimation and types of data

10.1 Estimation

| 10.1.1 | Parametric and nonparametric estimation | 282 |
| 10.1.2 | Point and interval estimation | 282 |
| 10.1.3 | Properties of estimators | 283 |

10.2 Types of data

| 10.2.1 | Duration data and loss data | 286 |
| 10.2.2 | Complete individual data | 287 |
| 10.2.3 | Incomplete individual data | 289 |
| 10.2.4 | Grouped data | 294 |

10.3 Summary and discussions

| Exercises | 297 |

11 Nonparametric model estimation

11.1 Estimation with complete individual data

| 11.1.1 | Empirical distribution | 302 |
| 11.1.2 | Kernel estimation of probability density function | 306 |
## Contents

11.2 Estimation with incomplete individual data 311

11.2.1 Kaplan–Meier (product-limit) estimator 311

11.2.2 Nelson–Aalen estimator 319

11.3 Estimation with grouped data 323

11.4 Excel computation notes 326

11.5 Summary and discussions 326

Exercises 327

12 Parametric model estimation 335

12.1 Methods of moments and percentile matching 336

12.1.1 Method of moments 336

12.1.2 Method of percentile matching 341

12.2 Bayesian estimation method 343

12.3 Maximum likelihood estimation method 344

12.3.1 Complete individual data 347

12.3.2 Grouped and incomplete data 351

12.4 Models with covariates 358

12.4.1 Proportional hazards model 358

12.4.2 Generalized linear model 364

12.4.3 Accelerated failure-time model 365

12.5 Modeling joint distribution using copula 366

12.6 Excel computation notes 369

12.7 Summary and discussions 371

Exercises 372

13 Model evaluation and selection 380

13.1 Graphical methods 381

13.2 Misspecification tests and diagnostic checks 385

13.2.1 Kolmogorov–Smirnov test 386

13.2.2 Anderson–Darling test 388

13.2.3 Chi-square goodness-of-fit test 389

13.2.4 Likelihood ratio test 391

13.3 Information criteria for model selection 393

13.4 Summary and discussions 394

Exercises 395

14 Basic Monte Carlo methods 400

14.1 Monte Carlo simulation 401

14.2 Uniform random number generators 402

14.3 General random number generators 405

14.3.1 Inversion method 406

14.3.2 Acceptance–rejection method 408

14.3.3 Generation of correlated random variables 411
## Contents

14.4 Specific random number generators 414  
14.4.1 Some continuous distributions 414  
14.4.2 Some discrete distributions 417  
14.5 Accuracy and Monte Carlo sample size 418  
14.6 Variance reduction techniques 421  
14.6.1 Antithetic variable 422  
14.6.2 Control variable 423  
14.6.3 Importance sampling 425  
14.7 Excel computation notes 426  
14.8 Summary and discussions 428  
Exercises 428  

15 Applications of Monte Carlo methods 435  
15.1 Monte Carlo simulation for hypothesis test 436  
15.1.1 Kolmogorov–Smirnov test 436  
15.1.2 Chi-square goodness-of-fit test 438  
15.2 Bootstrap estimation of $p$-value 439  
15.3 Bootstrap estimation of bias and mean squared error 442  
15.4 A general framework of bootstrap 445  
15.5 Monte Carlo simulation of asset prices 447  
15.5.1 Wiener process and generalized Wiener process 447  
15.5.2 Diffusion process and lognormal distribution 448  
15.5.3 Jump–diffusion process 453  
15.6 Summary and discussions 455  
Exercises 456  

Appendix: Review of statistics 458  
Answers to exercises 498  
References 518  
Index 521
This book is on the theory, methods, and empirical implementation of nonlife actuarial models. It is intended for use as a textbook for senior undergraduates. Users are assumed to have done one or two one-semester courses on probability theory and statistical inference, including estimation and hypothesis testing. The coverage of this book includes all the topics found in Exam C of the Society of Actuaries (Exam 4 of the Casualty Actuarial Society) as per the 2007 Basic Education Catalog. In addition, it covers some topics (such as risk measures and ruin theory) beyond what is required by these exams, and may be used by actuarial students in general.

This book is divided into four parts: loss models, risk and ruin, credibility, and model construction and evaluation. An appendix on the review of statistics is provided for the benefit of students who require a quick summary. Students may read the appendix prior to the main text if they desire, or they may use the appendix as a reference when required. In order to be self contained, the appendix covers some of the topics developed in the main text.

Some features of this book should be mentioned. First, the concepts and theories introduced are illustrated by many practical examples. Some of these examples explain the theory through numerical applications, while others develop new results. Second, several chapters of the book include a section on numerical computation using Excel. Students are encouraged to use Excel to solve some of the numerical exercises. Third, each chapter includes some exercises for practice. Many of these exercises are adapted from past exam questions of the Society of Actuaries.

I would like to thank Tao Yang for painstakingly going through the manuscript and for providing many useful comments and suggestions. Diana Gillooly has professionally guided me through the publication process with admirable patience and efficiency. Clare Dennison has performed a superb job of
Preface

coordinating the copy editing. I am also grateful to the Society of Actuaries for allowing me to use its past exam questions.

Resources are available at: www.mysmu.edu/faculty/yktse/NAM/NAMbase.htm. Slides in pdf format can be downloaded from this site, which will facilitate classroom teaching by instructors adopting this book. An errata file will be provided, and the solution manual for instructors is obtainable from the author on request.

Yiu-Kuen Tse, Ph.D. FSA
Singapore Management University
yktse@smu.edu.sg
Notation and convention

1 Abbreviations are used in this book without periods. For example, “probability density function” is referred to as pdf (not p.d.f.) and “moment generating function” is referred to as mgf (not m.g.f.).

2 We do not make distinctions between a random variable and the distribution that describes the random variable. Thus, from time to time we make statements such as: “X denotes the binomial distribution”.

3 We use calligraphic fonts to denote commonly used distributions. Discrete distributions are denoted with two alphabets and continuous distributions are denoted with one alphabet. For example, \( \mathcal{P} \) stands for Poisson, \( \mathcal{B} \) stands for binomial, \( \mathcal{N} \) stands for normal, and \( \mathcal{L} \) stands for lognormal.

4 The following conventions are generally used:
   (a) Slanted upper case for random variables, e.g. \( X \).
   (b) Slanted lower case for fixed numbers, e.g. \( x \).
   (c) Slanted bold-faced upper case for vectors of random variables, e.g. \( \mathbf{X} \).
   (d) Slanted bold-faced lower case for vectors of fixed numbers (observations), e.g. \( \mathbf{x} \).
   (e) Upright bold-faced upper case for matrices of fixed numbers (observations), e.g. \( \mathbf{X} \).

5 Natural logarithm is denoted by log, not ln.

Computation notes

1 In some chapters we include a section of Excel computation notes to discuss the use of Excel functions to facilitate computation. These functions require the Excel add-ins Analysis ToolPak and Solver Add-in.

2 Other computer languages for more advanced statistical analysis include R, C++, Splus, Gauss, and Matlab. All graphs in this book were produced using Matlab, and many of the computations were performed using Gauss.