## Contents

*Contributors* page xi

*Preface* xv

*Acknowledgments* xix

1 Pure Premium Modeling Using Generalized Linear Models 1

Ernesto Schirmacher

1.1 Introduction 1

1.2 Data Characteristics 3

1.3 Exploratory Data Analysis 6

1.4 Frequency Modeling 12

1.5 Severity Modeling 23

1.6 Pure Premium 30

1.7 Validation 34

1.8 Conclusions 37

References 38

2 Applying Generalized Linear Models to Insurance Data: Frequency/Severity versus Pure Premium Modeling 39

Dan Tevet

2.1 Introduction 39

2.2 Comparing Model Forms 40

2.3 The Dataset and Model Forms 44

2.4 Results 47

Appendix 2.A Proof of Equivalence between Pure Premium Model Forms 55

Conclusion 57

Appendix 2.B The Gini Index 57

References 58
3 Generalized Linear Models as Predictive Claim Models
   Greg Taylor and James Sullivan
   3.1 Review of Loss Reserving
   3.2 Additional Notation
   3.3 GLM Background
   3.4 Advantages of GLMs
   3.5 Diagnostics
   3.6 Example
   3.7 Conclusion
   References

4 Frameworks for General Insurance Ratemaking: Beyond the Generalized Linear Model
   Peng Shi and James Guszcza
   4.1 Introduction
   4.2 Data
   4.3 Univariate Ratemaking Framework
   4.4 Multivariate Ratemaking Frameworks
   4.5 Model Comparisons
   4.6 Conclusion
   References

5 Using Multilevel Modeling for Group Health Insurance Ratemaking: A Case Study from the Egyptian Market
   Mona S. A. Hammad and Galal A. H. Harby
   5.1 Motivation and Background
   5.2 Data
   5.3 Methods and Models
   5.4 Results
   5.5 Conclusions
   Acknowledgments
   Appendix
   References

6 Clustering in General Insurance Pricing
   Ji Yao
   6.1 Introduction
   6.2 Overview of Clustering
   6.3 Dataset for Case Study
   6.4 Clustering Methods
   6.5 Exposure-Adjusted Hybrid (EAH) Clustering Method

References
Contents

6.6 Results of Case Study 171
6.7 Other Considerations 177
6.8 Conclusions 178
References 179

7 Application of Two Unsupervised Learning Techniques to Questionable Claims: PRIDIT and Random Forest 180
Louise A. Francis
7.1 Introduction 180
7.2 Unsupervised Learning 181
7.3 Simulated Automobile PIP Questionable Claims Data and the Fraud Issue 182
7.4 The Questionable Claims Dependent Variable Problem 185
7.5 The PRIDIT Method 185
7.6 Processing the Questionable Claims Data for PRIDIT Analysis 187
7.7 Computing RIDITS and PRIDITS 187
7.8 PRIDIT Results for Simulated PIP Questionable Claims Data 188
7.9 How Good Is the PRIDIT Score? 189
7.10 Trees and Random Forests 192
7.11 Unsupervised Learning with Random Forest 194
7.12 Software for Random Forest Computation 195
7.13 Some Findings from the Brockett et al. Study 201
7.14 Random Forest Visualization via Multidimensional Scaling 202
7.15 Kohonen Neural Networks 204
7.16 Summary 205
References 206

8 The Predictive Distribution of Loss Reserve Estimates over a Finite Time Horizon 208
Glenn Meyers
8.1 Introduction 208
8.2 The CAS Loss Reserve Database 210
8.3 The Correlated Chain Ladder Model 212
8.4 The Predictive Distribution of Future Estimates 213
8.5 The Implications for Capital Management 216
8.6 Summary and Conclusions 223
References 223
## Contents

### 9 Finite Mixture Model and Workers’ Compensation Large-Loss Regression Analysis

Luyang Fu and Xianfang Liu

* 9.1 Introduction 224
* 9.2 DGLM and FMM 230
* 9.3 Data 232
* 9.4 Traditional Distribution Analysis 234
* 9.5 Univariate and Correlation Analyses 239
* 9.6 Regression Analysis 246
* 9.7 Conclusions 257

References 258

### 10 A Framework for Managing Claim Escalation Using Predictive Modeling

Mohamad A. Hindawi and Claudine H. Modlin

* 10.1 Introduction 261
* 10.2 Loss Development Models 262
* 10.3 Additional Data for Triage Models 267
* 10.4 Factor Selection 271
* 10.5 Modeling Method 274
* 10.6 Conclusions 277
* 10.7 Further Research Opportunities 277

Appendix: Penalized Regression 280

References 289

### 11 Predictive Modeling for Usage-Based Auto Insurance

Udi Makov and Jim Weiss

* 11.1 Introduction to Usage-Based Auto Insurance 290
* 11.2 Poisson Model for Usage-Based Auto Insurance 294
* 11.3 Classification Trees 301
* 11.4 Implementing UBI Models with a Traditional Rating Plan 305
* 11.5 Summary and Areas for Future Research 306

Acknowledgments 307

References 308

Index 309