Introduction to
BAYESIAN ESTIMATION
AND COPULA MODELS
OF DEPENDENCE

Arkady Shemyakin
Alexander Kniazev
Introduction to Bayesian Estimation and Copula Models of Dependence
Introduction to Bayesian Estimation and Copula Models of Dependence

Arkady Shemyakin
University of St. Thomas

Alexander Kniazev
Astrakhan State University
To all our families:
special thanks to
Evguenia Savinkina and
Ninel Kostenko
Contents

List of Figures   xiii
List of Tables   xvii
Acknowledgments   xix
Acronyms   xxi
Glossary   xxiii
About the Companion Website   xxv
Introduction   xxvii

Part I   Bayesian Estimation   1

1 Random Variables and Distributions   3
   1.1 Conditional Probability   3
   1.2 Discrete Random Variables   5
   1.3 Continuous Distributions on the Real Line   8
   1.4 Continuous Distributions with Nonnegative Values   13
   1.5 Continuous Distributions on a Bounded Interval   19
   1.6 Joint Distributions   20
   1.7 Time-Dependent Random Variables   27
       References   32

2 Foundations of Bayesian Analysis   33
   2.1 Education and Wages   33
   2.2 Two Envelopes   36
   2.3 Hypothesis Testing   39
       2.3.1 The Likelihood Principle   39
       2.3.2 Review of Classical Procedures   40
       2.3.3 Bayesian Hypotheses Testing   42
   2.4 Parametric Estimation   43
       2.4.1 Review of Classical Procedures   43
2.4.2 Maximum Likelihood Estimation 44
2.4.3 Bayesian Approach to Parametric Estimation 46
2.5 Bayesian and Classical Approaches to Statistics 47
   2.5.1 Classical (Frequentist) Approach 49
   2.5.2 Lady Tasting Tea 50
   2.5.3 Bayes Theorem 53
   2.5.4 Main Principles of the Bayesian Approach 55
2.6 The Choice of the Prior 57
   2.6.1 Subjective Priors 57
   2.6.2 Objective Priors 60
   2.6.3 Empirical Bayes 63
2.7 Conjugate Distributions 66
   2.7.1 Exponential Family 66
   2.7.2 Poisson Likelihood 67
   2.7.3 Table of Conjugate Distributions 68
References 68
Exercises 69

3 Background for Markov Chain Monte Carlo 73
   3.1 Randomization 73
      3.1.1 Rolling Dice 73
      3.1.2 Two Envelopes Revisited 74
   3.2 Random Number Generation 76
      3.2.1 Pseudo-random Numbers 76
      3.2.2 Inverse Transform Method 77
      3.2.3 General Transformation Methods 78
      3.2.4 Accept–Reject Methods 81
   3.3 Monte Carlo Integration 85
      3.3.1 Numerical Integration 86
      3.3.2 Estimating Moments 87
      3.3.3 Estimating Probabilities 88
      3.3.4 Simulating Multiple Futures 90
   3.4 Precision of Monte Carlo Method 91
      3.4.1 Monitoring Mean and Variance 91
      3.4.2 Importance Sampling 94
      3.4.3 Correlated Samples 96
      3.4.4 Variance Reduction Methods 97
   3.5 Markov Chains 101
      3.5.1 Markov Processes 101
      3.5.2 Discrete Time, Discrete State Space 103
      3.5.3 Transition Probability 103
      3.5.4 “Sun City” 104
      3.5.5 Utility Bills 105
<table>
<thead>
<tr>
<th>Contents</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>3.5.6 Classification of States</td>
<td>105</td>
</tr>
<tr>
<td>3.5.7 Stationary Distribution</td>
<td>107</td>
</tr>
<tr>
<td>3.5.8 Reversibility Condition</td>
<td>108</td>
</tr>
<tr>
<td>3.5.9 Markov Chains with Continuous State Spaces</td>
<td>108</td>
</tr>
<tr>
<td>3.6 Simulation of a Markov Chain</td>
<td>109</td>
</tr>
<tr>
<td>3.7 Applications</td>
<td></td>
</tr>
<tr>
<td>3.7.1 Bank Sizes</td>
<td>111</td>
</tr>
<tr>
<td>3.7.2 Related Failures of Car Parts</td>
<td>113</td>
</tr>
<tr>
<td>References</td>
<td>115</td>
</tr>
<tr>
<td>Exercises</td>
<td>117</td>
</tr>
<tr>
<td>4 Markov Chain Monte Carlo Methods</td>
<td>119</td>
</tr>
<tr>
<td>4.1 Markov Chain Simulations for Sun City and Ten Coins</td>
<td>119</td>
</tr>
<tr>
<td>4.2 Metropolis–Hastings Algorithm</td>
<td>126</td>
</tr>
<tr>
<td>4.3 Random Walk MHA</td>
<td>130</td>
</tr>
<tr>
<td>4.4 Gibbs Sampling</td>
<td>134</td>
</tr>
<tr>
<td>4.5 Diagnostics of MCMC</td>
<td>136</td>
</tr>
<tr>
<td>4.5.1 Monitoring Bias and Variance of MCMC</td>
<td>137</td>
</tr>
<tr>
<td>4.5.2 Burn-in and Skip Intervals</td>
<td>140</td>
</tr>
<tr>
<td>4.5.3 Diagnostics of MCMC</td>
<td>142</td>
</tr>
<tr>
<td>4.6 Suppressing Bias and Variance</td>
<td>144</td>
</tr>
<tr>
<td>4.6.1 Perfect Sampling</td>
<td>144</td>
</tr>
<tr>
<td>4.6.2 Adaptive MHA</td>
<td>145</td>
</tr>
<tr>
<td>4.6.3 ABC and Other Methods</td>
<td>145</td>
</tr>
<tr>
<td>4.7 Time-to-Default Analysis of Mortgage Portfolios</td>
<td>146</td>
</tr>
<tr>
<td>4.7.1 Mortgage Defaults</td>
<td>146</td>
</tr>
<tr>
<td>4.7.2 Customer Retention and Infinite Mixture Models</td>
<td>147</td>
</tr>
<tr>
<td>4.7.3 Latent Classes and Finite Mixture Models</td>
<td>149</td>
</tr>
<tr>
<td>4.7.4 Maximum Likelihood Estimation</td>
<td>150</td>
</tr>
<tr>
<td>4.7.5 A Bayesian Model</td>
<td>151</td>
</tr>
<tr>
<td>References</td>
<td>156</td>
</tr>
<tr>
<td>Exercises</td>
<td>158</td>
</tr>
<tr>
<td>Part II Modeling Dependence</td>
<td>159</td>
</tr>
<tr>
<td>5 Statistical Dependence Structures</td>
<td>161</td>
</tr>
<tr>
<td>5.1 Introduction</td>
<td>161</td>
</tr>
<tr>
<td>5.2 Correlation</td>
<td>165</td>
</tr>
<tr>
<td>5.2.1 Pearson’s Linear Correlation</td>
<td>165</td>
</tr>
<tr>
<td>5.2.2 Spearman’s Rank Correlation</td>
<td>167</td>
</tr>
<tr>
<td>5.2.3 Kendall’s Concordance</td>
<td>167</td>
</tr>
<tr>
<td>5.3 Regression Models</td>
<td>170</td>
</tr>
<tr>
<td>5.3.1 Heteroskedasticity</td>
<td>171</td>
</tr>
</tbody>
</table>
5.3.2 Nonlinear Regression 172
5.3.3 Prediction 175
5.4 Bayesian Regression 176
5.5 Survival Analysis 179
5.5.1 Proportional Hazards 180
5.5.2 Shared Frailty 180
5.5.3 Multistage Models of Dependence 182
5.6 Modeling Joint Distributions 182
5.6.1 Bivariate Survival Functions 183
5.6.2 Bivariate Normal 185
5.6.3 Simulation of Bivariate Normal 185
5.7 Statistical Dependence and Financial Risks 186
5.7.1 A Story of Three Loans 186
5.7.2 Independent Defaults 188
5.7.3 Correlated Defaults 189
References 192
Exercises 193

6 Copula Models of Dependence 195
6.1 Introduction 195
6.2 Definitions 196
6.2.1 Quasi-Monotonicity 197
6.2.2 Definition of Copula 198
6.2.3 Sklar’s Theorem 198
6.2.4 Survival Copulas 199
6.3 Simplest Pair Copulas 200
6.3.1 Maximum Copula 200
6.3.2 Minimum Copula 202
6.3.3 FGM Copulas 203
6.4 Elliptical Copulas 204
6.4.1 Elliptical Distributions 204
6.4.2 Method of Inverses 205
6.4.3 Gaussian Copula 205
6.4.4 The t-copula 207
6.5 Archimedean Copulas 209
6.5.1 Definitions 209
6.5.2 One-Parameter Copulas 210
6.5.3 Clayton Copula 212
6.5.4 Frank Copula 213
6.5.5 Gumbel–Hougaard Copula 214
6.5.6 Two-Parameter Copulas 216
6.6 Simulation of Joint Distributions 217
6.6.1 Bivariate Elliptical Distributions 218
6.6.2 Bivariate Archimedean Copulas 219
Contents

6.7 Multidimensional Copulas 222
   References 228
   Exercises 230

7 Statistics of Copulas 233
   7.1 The Formula that Killed Wall Street 233
   7.2 Criteria of Model Comparison 237
      7.2.1 Goodness-of-Fit Tests 237
      7.2.2 Posterior Predictive p-values 239
      7.2.3 Information Criteria 241
      7.2.4 Concordance Measures 243
      7.2.5 Tail Dependence 244
   7.3 Parametric Estimation 245
      7.3.1 Parametric, Semiparametric, or Nonparametric? 246
      7.3.2 Method of Moments 248
      7.3.3 Minimum Distance 248
      7.3.4 MLE and MPLE 249
      7.3.5 Bayesian Estimation 250
   7.4 Model Selection 252
      7.4.1 Hybrid Approach 252
      7.4.2 Information Criteria 254
      7.4.3 Bayesian Model Selection 256
   7.5 Copula Models of Joint Survival 257
   7.6 Related Failures of Vehicle Components 260
      7.6.1 Estimation of Association Parameters 261
      7.6.2 Comparison of Copula Classes 262
      7.6.3 Bayesian Model Selection 265
      7.6.4 Conclusions 267
   References 268
   Exercises 271

8 International Markets 273
   8.1 Introduction 273
   8.2 Selection of Univariate Distribution Models 276
   8.3 Prior Elicitation for Pair Copula Parameter 280
   8.4 Bayesian Estimation of Pair Copula Parameters 286
   8.5 Selection of Pair Copula Model 290
   8.6 Goodness-of-Fit Testing 295
   8.7 Simulation and Forecasting 298
   References 304
   Exercises 307

Index 309
## List of Figures

1.1 Gaussian curve with $\mu = 3, \sigma = 1$ 9  
1.2 Gaussian curves for different values of $\sigma$ 10  
1.3 Australian stock index ASX 10  
1.4 Density of $t$-distribution 12  
1.5 Density of asymmetric $t$-distribution 13  
1.6 Mexican stock index IPC 13  
1.7 Gamma density graphs 15  
1.8 Weibull density graphs 16  
1.9 Densities of $\chi^2$ distribution 18  
1.10 Densities of $F$-distribution 18  
1.11 Beta distribution densities 19  
1.12 Density of bivariate normal distribution, $\rho = 0$ 24  
1.13 Density of bivariate normal distribution, $\rho = -0.8$ 24  
1.14 Density of bivariate normal distribution, $\rho = 0.8$ 25  
1.15 Density of bivariate $t$-distribution, $\rho = 0$ 25  
1.16 Density of bivariate $t$-distribution, $\rho = -0.8$ 26  
1.17 Density of bivariate $t$-distribution, $\rho = 0.8$ 26  
1.18 Scatterplots of daily index returns: NIKKEI versus IPC and DAX versus CAC 27  
1.19 Daily values of SPX 28  
1.20 Daily logarithmic returns on SPX 30  
1.21 Series $z_t$ 31  
2.1 Model function of $E$ and $W$ 34  
2.2 Graphs of p.d.f’s 35  
2.3 Graphs of likelihood functions 36  
2.5 Leonard Jimmie Savage (1917–1971). Courtesy of Gonalo L. Fonseca 52  
2.6 Rev. Thomas Bayes (circa 1702–1761).  
https://commons.wikimedia.org/wiki/
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.7</td>
<td>Sir Harold Jeffreys (1891–1989). Reproduced with permission of the Master and Fellows of St. John's College, Cambridge</td>
</tr>
<tr>
<td>2.8</td>
<td>Arnold Zellner (1927–2010), the first president of ISBA. Reproduced with permission of International Society for Bayesian Analysis</td>
</tr>
<tr>
<td>2.9</td>
<td>Graphs of beta p.d.f.'s</td>
</tr>
<tr>
<td>2.10</td>
<td>Graphs of prior (dotted) and posterior (solid) for subjective prior</td>
</tr>
<tr>
<td>2.11</td>
<td>Graphs of priors and posteriors for objective priors</td>
</tr>
<tr>
<td>2.12</td>
<td>Graphs of prior (dotted) and posterior (solid) for empirical prior</td>
</tr>
<tr>
<td>3.1</td>
<td>Sampling from exponential distribution with ( \lambda = 2 )</td>
</tr>
<tr>
<td>3.2</td>
<td>Uniform sample and Beta(19,15) p.d.f.</td>
</tr>
<tr>
<td>3.3</td>
<td>Accept–reject sampling from beta distribution</td>
</tr>
<tr>
<td>3.4</td>
<td>Accept–reject sampling from gamma distribution</td>
</tr>
<tr>
<td>3.5</td>
<td>Graph of ( h(x) )</td>
</tr>
<tr>
<td>3.6</td>
<td>Cumulative sample means</td>
</tr>
<tr>
<td>3.7</td>
<td>Estimated variance of the estimated means</td>
</tr>
<tr>
<td>3.8</td>
<td>Confidence intervals for means</td>
</tr>
<tr>
<td>3.9</td>
<td>Multiple trajectories for estimated means</td>
</tr>
<tr>
<td>3.10</td>
<td>Antithetic sampling</td>
</tr>
<tr>
<td>3.11</td>
<td>White noise</td>
</tr>
<tr>
<td>3.12</td>
<td>Random walk</td>
</tr>
<tr>
<td>3.13</td>
<td>Sun City simulation</td>
</tr>
<tr>
<td>3.14</td>
<td>Utility bills simulation</td>
</tr>
<tr>
<td>4.1</td>
<td>Simulation with the proposed transition: ( N = 50 )</td>
</tr>
<tr>
<td>4.2</td>
<td>Two-step Markov chain Sun City simulation: ( N = 50 )</td>
</tr>
<tr>
<td>4.3</td>
<td>Ten coins. Proposal sample</td>
</tr>
<tr>
<td>4.4</td>
<td>Ten coins. Histogram of the proposal and target density</td>
</tr>
<tr>
<td>4.5</td>
<td>Ten coins. Markov chain simulation, ( N = 1000 )</td>
</tr>
<tr>
<td>4.6</td>
<td>Ten coins. Markov chain simulation, ( N = 100 )</td>
</tr>
<tr>
<td>4.7</td>
<td>Ten coins. Histogram of the resulting chain and target density</td>
</tr>
<tr>
<td>4.8</td>
<td>Ten coins RWMHA: ( \sigma = 0.2 )</td>
</tr>
<tr>
<td>4.9</td>
<td>Ten coins RWMHA: ( \sigma = 0.1 )</td>
</tr>
<tr>
<td>4.10</td>
<td>Ten coins RWMHA: ( \sigma = 0.01 )</td>
</tr>
<tr>
<td>4.11</td>
<td>Ten coins RWMHA histogram: ( \sigma = 0.1 )</td>
</tr>
<tr>
<td>4.12</td>
<td>Ten coins RWMHA histogram: ( \sigma = 0.01 )</td>
</tr>
<tr>
<td>4.13</td>
<td>Ten coins RWMHA SACF: ( \sigma = 0.1 )</td>
</tr>
<tr>
<td>4.14</td>
<td>Ten coins RWMHA SACF: ( \sigma = 0.01 )</td>
</tr>
<tr>
<td>4.15</td>
<td>Multiple chains for ( \sigma = 0.1 )</td>
</tr>
<tr>
<td>4.16</td>
<td>Multiple chains for ( \sigma = 0.01 )</td>
</tr>
<tr>
<td>4.17</td>
<td>WGM model fit with MLE</td>
</tr>
<tr>
<td>4.18</td>
<td>WGM model fit with Bayes estimates (RWMHA)</td>
</tr>
<tr>
<td>4.19</td>
<td>Trace plots for WGM model</td>
</tr>
</tbody>
</table>
4.20 WS model fit with Bayes estimates (RWMHA) 154
4.21 Trace plots for WS model 155
5.1 Black bears in Pennsylvania 162
5.2 Joint mortality data 163
5.3 Linear regression 171
5.4 Quadratic regression 173
5.5 Logarithmic transform of weight 174
5.6 Histogram of the residuals 176
5.7 Independent defaults 188
5.8 Correlated defaults with $\rho = 0.5$ 190
5.9 Extreme scenario 1 191
5.10 Extreme scenario 2 191
6.1 Rectangular area 197
6.2 Maximum copula, case 1 200
6.3 Maximum copula, case 2 201
6.4 Maximum copula, case 3 201
6.5 Maximum copula, three-dimensional surface, and contour plot 201
6.6 Minimum copula, three-dimensional surface, and contour plot 202
6.7 Product copula, three-dimensional surface, and contour plot 203
6.8 Transformation of variables for Gaussian copula 206
6.9 Density: Gaussian copula, Weibull margins (6.12) and (6.13) 207
6.10 Student $t$-copula, Weibull margins (6.12) and (6.13) 208
6.11 Graphs of $\phi(t) = -\ln(t)$ and its inverse 210
6.12 Graphs of $\phi(t) = 1 - t$ and its pseudo-inverse 210
6.13 Graphs of generator $\phi(t) = \tan(\pi(1 - t)/2)$ and its inverse 211
6.14 Clayton copula with Weibull margins (6.12) and (6.13) 213
6.15 Frank copula with Weibull margins (6.12) and (6.13) 214
6.16 Gumbel–Hougaard copula with Weibull margins (6.12) and (6.13) 215
6.17 Diagonal cross-sections for Clayton, Frank, and Gumbel–Hougaard copulas 215
6.18 Vine structure, centered, $d = 3$ 224
6.19 Vine structure, linear, $d = 3$ 224
6.20 C-vine for $d = 4$ 225
6.21 D-vine for $d = 4$ 226
6.22 Hierarchical Archimedean copulas, $d = 3$ 226
6.23 Hierarchical Archimedean copulas, fully nested 227
6.24 Hierarchical Archimedean copulas, non-nested 227
7.1 Trace plot for two-step Bayesian estimation 252
7.2 Trace plots for one-step Bayesian estimation 253
8.1 Histogram of standardized residuals for SPX and the Gaussian curve 276
8.2 Histogram of $\eta_1$ and $\lambda_1$: transformed parameters of asymmetric $t$-distribution 277
8.3 Histograms of $\eta_1$ and $\lambda_1$ with priors 278
8.4 Scatterplot of MLE for $\eta_1$ and $\lambda_1$ 279
8.5 RWMHA chain for the residuals of FTSU100 (acceptance rate 57.6%) 280
8.6 Scatterplots of $z$-series (with Kendall’s $\hat{\tau}$) 281
8.7 Clayton copula parameter $\alpha_1 = \ln \alpha$ 283
8.8 Gumbel–Hougaard copula parameter 284
8.9 Gauss copula parameter 285
8.10 Frank copula parameter 285
8.11 Student’s $t$-copula parameters 286
8.12 Proposed and accepted chain values. Sigma = 0.2, Acra = 0.49,
A = 2.496 289
8.13 Proposed and accepted chain values. Sigma = 1, Acra = 0.12,
A = 2.485 290
8.14 Scatterplots of relative forecast error for Gaussian and $t$-copulas 301
8.15 Scatterplots of relative forecast error for Clayton, Frank, and Gumbel copulas. Kendall’s function method 302
8.16 Scatterplots of the relative forecast error for Clayton, Frank, and Gumbel copulas. Marshall–Olkin’s method 303
8.17 Forecasting AEX with Frank, Clayton, and Student $t$-copulas 304
List of Tables

3.1 Accept–reject method for Beta(19,15) 84
3.2 Classification of banks 112
3.3 Transition probabilities for bank sizes 112
3.4 Distribution of bank sizes for 2012 113
3.5 Transition probabilities for car repairs 115
4.1 Bayesian parameter estimates for WGM 154
4.2 Parameter estimates for the Bayesian WS model, subprime portfolios 155
5.1 Three samples 168
6.1 Copula results for joint life distribution 216
6.2 Gaussian copula with Weibull margins 218
6.3 Clayton Copula with Weibull Margins 221
6.4 Gumbel–Hougaard copula with Weibull margins 222
7.1 Simulated exponential/Gumbel–Hougaard sample 247
7.2 Parametric and nonparametric estimates of marginals 247
7.3 Ranks, pseudo-observations, and empirical Kendall's function 249
7.4 Comparison of survival Gumbel–Hougaard and Clayton copulas 254
7.5 Parameter estimates (with standard errors) 259
7.6 Estimates of association for six classes of copulas 262
7.7 AIC values for six classes of copulas 263
7.8 BIC values for six classes of copulas 263
7.9 $d_{KS}$ distances for six classes of copulas 263
7.10 Tail dependence induced by the models 264
7.11 Sample $\hat{\tau}$ and model-induced values of $\tau$ 265
7.12 Logs of posterior weights or H1–H6 267
8.1 Sample and model values of $\tau$ for selected pairs 291
8.2 Squared deviations $RSS_{DF}$ for selected pairs 293
8.3 Comparison of copula models 294
8.4 Results of goodness-of-fit testing 298
8.5 Results of the simulation 301
Acknowledgments

We express our sincere appreciation to all of our friends and colleagues who contributed to this book in many different ways. We are grateful for many discussions we had and many comments we received from Paul Alper, Susan Callaway, Oleg Lepekhin, Alex McNeil, Olga Demidova, and many others.

We deeply appreciate the assistance we have received with collecting supplementary materials and preparing files for the companion website from Laura Hanson, Kathryn McKee, Cheryl Heskin, Matthew Galloway, Shannon Currier, Natalie Vandeweghe, and Stephanie Fritz.

We are grateful to our collaborators on the research projects which became the foundation for the book: Heekyung Youn, Natalia Kangina, Alicia Johnson, Vadim Gordeev, Matthew Galloway, Ellen Klingner, Nicole Lenz, Nicole Lopez, Kelsie Sargent, and Katheryn Wifvat.

We want to thank all of those to whom the material in the book was presented over the years and who contributed with their thoughtful comments regarding both the form and the content of the book. Special thanks are due to Sarah Millholland, Valentin Artemiev, Yuri Chepasov, Anastasia Rozhkova, Ekaterina Savinkina, Alexander Zyrianov, Doug Swisher, Laura Fink, Eric Schlicht, and many other students and seminar participants.

We would like to acknowledge the staff of our schools, Astrakhan State University and the University of St. Thomas in Minnesota, for their support of our work and making this project possible.

We also would like to thank all the editorial staff of Wiley: Susanne Steitz-Filler, Sari Friedman, Amy Hendrickson, Divya Narayanan, and others who at various stages helped with the completion of the book.

We are also grateful to Ekaterina Kniazeva and Alexandra Savinkina for proof-reading the text and making suggestions which helped immensely in improving the style of the book and hopefully making it readable.

Finally, our deepest gratitude goes to all our families, without whose support and companionship the project would not be possible.

A.S. and A.K.
### Acronyms

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABS</td>
<td>asset-backed security</td>
</tr>
<tr>
<td>AR</td>
<td>autoregressive</td>
</tr>
<tr>
<td>ARCH</td>
<td>autoregressive conditional heteroskedasticity</td>
</tr>
<tr>
<td>ARIMA</td>
<td>autoregressive integrated moving averages</td>
</tr>
<tr>
<td>BB1</td>
<td>Joe’s BB1</td>
</tr>
<tr>
<td>BG</td>
<td>beta-geometric mixture model</td>
</tr>
<tr>
<td>BUGS</td>
<td>Bayes using Gibbs sampler (software)</td>
</tr>
<tr>
<td>CDO</td>
<td>collateralized debt obligation</td>
</tr>
<tr>
<td>CDS</td>
<td>credit default swap</td>
</tr>
<tr>
<td>CFTP</td>
<td>coupling from the past</td>
</tr>
<tr>
<td>CLT</td>
<td>the Central Limit Theorem</td>
</tr>
<tr>
<td>CMLE</td>
<td>canonical maximum likelihood estimation (estimates)</td>
</tr>
<tr>
<td>c.d.f.</td>
<td>cumulative distribution function</td>
</tr>
<tr>
<td>ENIAC</td>
<td>electronic numerical integrator and computer</td>
</tr>
<tr>
<td>FGM</td>
<td>Farlie–Gumbel–Morgenstern copulas</td>
</tr>
<tr>
<td>FTD</td>
<td>first-to-default (swaps)</td>
</tr>
<tr>
<td>GARCH</td>
<td>generalized autoregressive conditional heteroskedasticity</td>
</tr>
<tr>
<td>HAC</td>
<td>hierarchical Archimedean copula</td>
</tr>
<tr>
<td>HKC</td>
<td>hierarchical Kendall copula</td>
</tr>
<tr>
<td>IFM</td>
<td>inference from margins</td>
</tr>
<tr>
<td>i.i.d.</td>
<td>independent identically distributed</td>
</tr>
<tr>
<td>IMA</td>
<td>independent Metropolis algorithm</td>
</tr>
<tr>
<td>ISBA</td>
<td>International Society for Bayesian Analysis</td>
</tr>
<tr>
<td>JAGS</td>
<td>just another Gibbs sampler (software)</td>
</tr>
<tr>
<td>LLN</td>
<td>the Law of Large Numbers</td>
</tr>
<tr>
<td>MANIAC</td>
<td>mathematical analyzer, numerator, integrator, and computer</td>
</tr>
<tr>
<td>MCMC</td>
<td>Markov chain Monte Carlo</td>
</tr>
<tr>
<td>MHA</td>
<td>Metropolis–Hastings algorithm</td>
</tr>
<tr>
<td>MLE</td>
<td>maximum likelihood estimation (estimates)</td>
</tr>
<tr>
<td>MPLE</td>
<td>maximum pseudo-likelihood estimation (estimates)</td>
</tr>
<tr>
<td>PCC</td>
<td>pair copula construction</td>
</tr>
<tr>
<td>Acronyms</td>
<td>Definition</td>
</tr>
<tr>
<td>----------</td>
<td>------------</td>
</tr>
<tr>
<td>PD</td>
<td>probability of default</td>
</tr>
<tr>
<td>p.d.f</td>
<td>probability density function</td>
</tr>
<tr>
<td>RLUF</td>
<td>Rodriguez Lallena and Ubeda Flores copulas</td>
</tr>
<tr>
<td>RTO</td>
<td>regression through the origin</td>
</tr>
<tr>
<td>RWMHA</td>
<td>random walk Metropolis–Hastings algorithm</td>
</tr>
<tr>
<td>SACF</td>
<td>sample autocorrelation function</td>
</tr>
<tr>
<td>TTD</td>
<td>time-to-default</td>
</tr>
<tr>
<td>TTF</td>
<td>time-to-failure</td>
</tr>
<tr>
<td>WG</td>
<td>Weibull-gamma model</td>
</tr>
<tr>
<td>WGM</td>
<td>Weibull-gamma mixture model</td>
</tr>
<tr>
<td>WS</td>
<td>Weibull segmentation model</td>
</tr>
</tbody>
</table>
Glossary

\[ E(X) \] expected value (mean)
\[ \text{Var}(X) \] variance
\[ \text{Bin}(n, p) \] binomial distribution with \( n \) trials and success probability \( p \)
\[ \text{Poiss}(\lambda) \] Poisson distribution
\[ \text{NB}(p, r) \] negative binomial distribution
\[ N(\mu, \sigma^2) \] normal (Gaussian) distribution with mean \( \mu \) and variance \( \sigma^2 \)
\[ T(\eta) \] (Student) \( t \)-distribution with \( \eta \) degrees of freedom
\[ \text{Gamma}(\alpha, \lambda) \] gamma distribution with shape \( \alpha \) and rate \( \lambda \)
\[ \text{Beta}(\alpha, \beta) \] beta distribution
\[ \text{Cov}(X, Y) \] covariance
\[ \rho(X, Y) \] correlation
\[ MN(\mu, \Sigma) \] multivariate normal distribution
\[ L(x; \theta) \] likelihood function with data \( x \) and parameter \( \theta \)
\[ \Lambda(x) \] likelihood ratio
\[ PG(1, r) \] polygamma function order \( r \)
\[ X \sim F \] random variable \( X \) has distribution \( F \)
\[ \propto \] proportional
\[ \leq \] less or equal in sense of a special natural ordering
About the Companion Website

This book is accompanied by a companion website:

http://www.wiley.com/go/shemyakin/bayesian_estimation

The website includes:

- Solutions to selected exercises
- Excel dataset
- Excel simulation templates
- Appendices for Chapter 8
  - Datasets and results
  - Code in R
Introduction

Why does anyone need another book in Bayesian statistics? It seems that there already exist a lot of resources for those interested in the topic. There are many excellent books covering specific aspects of Bayesian analysis or providing a wide and comprehensive background of the entire field: Berger, Bernardo and Smith, Gamerman and Freitas Lopes, Gelman et al., Robert and Cassella, and many others. Most of these books, though, will assume a certain mathematical and statistical background and would rather fit a reader’s profile of a graduate or advanced graduate level. Out of those aimed at a less sophisticated audiences, we would certainly recommend excellent books of William Bolstad, John Kruschke, and Peter Lee. There also exist some very good books on copulas: comprehensive coverage by Nelsen and Joe, and also more application-related Cherubini et al., Emrechts et al., and some others. However, instead of just referring to these works and returning to our extensive to-do lists, we decided to spend considerable amount of time and effort putting together another book—the book we presently offer to the reader.

The main reason for our endeavor is: we target a very specific audience, which as we believe is not sufficiently serviced yet with Bayesian literature. We communicate with members of this audience routinely in our day-to-day work, and we have not failed to register that just providing them with reading recommendations does not seem to satisfy their needs. Our perceived audience could be loosely divided into two groups. The first includes advanced undergraduate students of Statistics, who in all likelihood have already had some exposure to main probabilistic and statistical principles and concepts (most likely, in classical or “frequentist” setup), and may (as we probably all do) exhibit some traces of Bayesian philosophy as applicable to their everyday lives. But for them these two: statistical methods on one hand and Bayesian thinking on the other, belong to very different spheres and do not easily combine in their decision-making process.

The second group consists of practitioners of statistical methods, working in their everyday lives on specific problems requiring the use of advanced quantitative analysis. They may be aware of a Bayesian alternative to classical
methods and find it vaguely attractive, but are not familiar enough with formal Bayesian analysis in order to put it to work. These practitioners populate analytical departments of banks, insurance companies, and other major businesses. In short, they might be involved in predictive modeling, quantitative forecasting, and statistical reporting which often directly call for Bayesian approach.

In the recent years, we have frequently encountered representatives of both groups described above as our collaborators, be it in undergraduate research or in applied consulting projects, or both at once (such things do happen). We have discovered a frequent need to provide to them a crash course in Bayesian methods: prior and posterior, Bayes estimation, prediction, MCMC, Bayesian regression and time series, Bayesian analysis of statistical dependence. From this environment we get the idea to concisely summarize the methodology we normally share with our collaborators in order to provide the framework for successful joint projects. Later on this idea transformed itself into a phantasy to write this book and hand it to these two segments of the audience as a potentially useful resource. This intention determines the content of the book and dictates the necessity to cover specific topics in specific order, trying also to avoid any unnecessary detail. That is why we do not include a serious introduction to probability and classical statistics (we believe that our audience has at least some formal knowledge of the main principles and facts in these fields). Instead, in Chapter 1 we just offer a review of the concepts we will eventually use. If this review happens to be insufficient to some readers, it will hopefully at least inspire them to hit the introductory books which will provide a more comprehensive coverage.

Chapter 2 deals with the basics of Bayesian statistics: prior information and experimental data, prior and posterior distributions, with emphasis on Bayesian parametric estimation, just barely touching Bayesian hypothesis testing. Some time is spend on addressing subjective versus objective Bayesian paradigms and brief discussion of noninformative priors. We spend just so much time with conjugate priors and analytical derivation of Bayes estimators that will give an idea of the scope and limitations of the analytical approach. It seems likely that most readers in their practical applications will require the use of MCMC—Markov chain Monte Carlo method—the most efficient tool in the hands of modern Bayesians. Therefore, Chapter 3 contains the basic mathematical background on both Markov chains and Monte Carlo integration and simulation. In our opinion, successful use of Markov chain Monte Carlo methods is heavily based on good understanding on these two components. Speaking of Monte Carlo methods, the central idea of variance reduction nicely transitions us to MCMC and its diagnostics. Equally important, both Markov chains and Monte Carlo methods have numerous important applications outside of Bayesian setting, and these applications will be discussed as examples.

Chapter 4 covers MCMC per se. It may look suspicious from traditional point of view that we do not particularly emphasize Gibbs sampling, rather deciding