Richard J. Cook and Jerald F. Lawless

The Statistical Analysis of Recurrent Events

SPIN Springer's project number

April 9, 2007

Springer
Berlin Heidelberg New York
Hong Kong London
Milan Paris Tokyo

Contents

Preface V

Glossary IX

1 Introduction 1
- 1.1 The Scope of Recurrent Events 1
- 1.2 Some Preliminary Examples 2
  - 1.2.1 Mammary Tumors in a Carcinogenicity Study 2
  - 1.2.2 Testing and Debugging a Large Software System 5
  - 1.2.3 Pulmonary Exacerbations in Cystic Fibrosis 5
  - 1.2.4 Automobile Warranty Claims 7
- 1.3 Notation and Frameworks 9
  - 1.3.1 Methods Based on Event Counts 11
  - 1.3.2 Methods Based on Waiting or Gap Times 12
  - 1.3.3 More General Models 14
  - 1.3.4 Covariates 14
  - 1.3.5 Factors Influencing Model Choice 15
- 1.4 Selection of Individuals and Observation Schemes 16
  - 1.4.1 The Choice of Time Scale 16
  - 1.4.2 Defining Periods "At Risk" 18
  - 1.4.3 Initial Conditions and Selecting Individuals for Study 19
  - 1.4.4 Intermittent Observation and Interval Censoring 20
- 1.5 Multitype Event Data 20
  - 1.5.1 Multivariate Event Processes 20
  - 1.5.2 Recurrent Events with Termination 20
  - 1.5.3 Recurrent Episodes 21
- 1.6 Some Other Aspects of Analysis and Design 23
- 1.7 Bibliographic Notes 24

2 Models and Frameworks for Analysis of Recurrent Events . 27
2.1 Mathematical Background 27
2.2 Poisson Processes and Models for Event Counts 31
2.2.1 Poisson Processes 31
2.2.2 Covariates in Poisson Processes 34
2.2.3 Random Effects in Poisson Processes 35
2.2.4 Example: Mammary Tumors in Rats 37
2.3 Renewal Processes and Models for Gap Times 39
2.3.1 Models for Gap Times Between Events 39
2.3.2 Example: Bowel Motility Cycles 42
2.4 General Intensity-Based Models 43
2.5 Discrete-Time Models and Time-Varying Covariates 45
2.6 Likelihood for Selection and Observation Schemes 47
2.7 Bibliographic Notes 51
2.8 Problems and Supplements 52

3 Methods Based on Counts and Rate Functions 59
3.1 Introduction 59
3.2 Parametric Maximum Likelihood for Poisson Models 61
3.2.1 Score and Information Functions 61
3.2.2 A General Parametric Rate Function 62
3.2.3 Time Transform Models 63
3.2.4 Using Survival Software 64
3.3 Poisson Models with Piecewise-Constant Rates 65
3.4 Nonparametric and Semiparametric Poisson Models 68
3.4.1 Nonparametric Inference 68
3.4.2 Semiparametric Regression 70
3.4.3 Stratification 73
3.4.4 Additive Models 75
3.5 Poisson Models with Random Effects 76
3.5.1 Formulation 76
3.5.2 Models for Zero-Inflated Data 78
3.5.3 Negative Binomial Models 79
3.6 Robust Methods for Rate and Mean Functions 82
3.6.1 Nonparametric Estimation 82
3.6.2 Parametric Estimation 83
3.6.3 Robust Semiparametric Methods 84
3.6.4 Robust Methods with Semiparametric Variances 86
3.6.5 Methods Based on Multivariate Failure Time Data 86
3.7 Some Useful Tests for Rate Functions 88
3.7.1 Tests for Trend 88
3.7.2 Tests for Multiplicative Covariate Effects 90
3.7.3 Generalized Residuals, Martingales, and Assessment of Fit 92
3.7.4 Tests for Extra-Poisson Variation 97

3.7.5 Two-Sample Test Statistics Based on Rates 97
3.8 Applications and Illustrations 100
3.8.1 Rat Mammary Tumor Data 100
3.8.2 A Trial of Treatment for Herpes Simplex Virus 105
3.8.3 Fitting and Prediction from a Software Debugging Model 107
3.8.4 Comparing Warranty Claim Histories 110
3.9 Bibliographic Notes 112
3.10 Problems and Supplements 114

4 Analysis of Gap Times 121
4.1 Renewal Processes and Related Methods of Analysis 121
4.2 Extensions of Renewal Models 126
4.2.1 Conditional Analysis of Successive Gap Times 126
4.2.2 Models with Random Effects 128
4.2.3 Joint Gap Time Distributions 129
4.2.4 Modulated Renewal Processes 132
4.3 Examples 133
4.3.1 Bowel Motility Cycles 133
4.3.2 Pulmonary Exacerbations and rhDNase Treatment 134
4.4 Estimation of Marginal Gap Time Probabilities 137
4.4.1 Nonparametric Estimation of Marginal Gap Time Distributions 139
4.4.2 Estimation for Marginal Regression Models 143
4.4.3 Pulmonary Exacerbations in Cystic Fibrosis 143
4.5 Left Truncation of First Gap Times and Initial Conditions 146
4.5.1 Initial Conditions and Renewal Processes 147
4.5.2 Initial Conditions and General Gap Time Analysis 150
4.6 Bibliographic Notes 152
4.7 Problems and Supplements 153

5 General Intensity-Based Models 161
5.1 Time Scales and Intensity Modeling 161
5.2 Parametric Analysis for Two Useful Models 163
5.2.1 Log-Linear Intensity Models 163
5.2.2 A Trend-Renewal Model 165
5.2.3 Model Checking 167
5.2.4 An Illustration: Air-Conditioning System Failures 167
5.3 Semiparametric Markov Analysis 171
5.3.1 Models with Dependence on Prior Counts 171
5.3.2 Markov Nonparametric Estimation 173
5.3.3 Models with Covariates 175
5.3.4 Analysis of Outbreaks Due to Herpes Simplex Virus 177
5.4 Semiparametric Modulated Renewal Analysis 183
5.4.1 Analysis Based on the Cox Model 183

5.4.2 Illustration: Cerebrospinal Fluid Shunts . . . 185
5.5 Some Additional Illustrations . . . 189
5.5.1 Pulmonary Exacerbations in the Study of rhDNase . . . 190
5.5.2 Analysis of Asthma Exacerbations . . . 193
5.6 Bibliographic Notes . . . 200
5.7 Problems and Supplements . . . 201

6 Multitype Recurrent Events . . . 205
6.1 Multivariate Event Data . . . 205
6.2 Intensity-Based Methods . . . 206
6.2.1 Notation and Intensity Functions . . . 206
6.2.2 Remarks on Intensity-Based Models . . . 207
6.3 Random Effect Models for Multitype Events . . . 209
6.4 Robust Methods for Multitype Events . . . 212
6.4.1 Methods Based on Working Independence Assumptions 212
6.4.2 Robust Methods with Covariance Functions . . . 214
6.5 Alternating Two-State Processes . . . 216
6.6 Recurrent Events with a Terminal Event . . . 218
6.6.1 Intensity-Based Approaches . . . 219
6.6.2 Random Effects Models . . . 220
6.6.3 Robust Methods for Marginal Features . . . 222
6.6.4 Partially Conditional Methods . . . 224
6.7 Applications and Illustrations . . . 227
6.7.1 Cerebrospinal Fluid Shunt Failures . . . 227
6.7.2 Exacerbations in Patients with Chronic Bronchitis . . . 232
6.7.3 Skeletal Complications in Metastatic Cancer . . . 236
6.7.4 Relationships Between Skeletal Complications . . . 241
6.8 Bibliographic Notes . . . 246
6.9 Problems and Supplements . . . 247

7 Observation Schemes Giving Incomplete or Selective Data . 251
7.1 Intermittent Observation During Followup . . . 251
7.1.1 Methods Based on Poisson Processes . . . 252
7.1.2 Robust Estimation of Rate and Mean Functions . . . 254
7.1.3 Illustration: Superficial Tumors of the Bladder . . . 256
7.1.4 Interval-Count Data for Multiple Events . . . 260
7.1.5 Illustration: Joint Damage in Psoriatic Arthritis . . . 263
7.2 Dependent Censoring or Inspections . . . 264
7.2.1 Dependent Censoring and Weighted Estimating Functions . . . 264
7.2.2 Rate or Mean Function Estimation with Event-Dependent Censoring . . . 268
7.2.3 Intermittent Observation . . . 270
7.3 Event-Dependent Selection . . . 273
7.3.1 Some Examples . . . 273

7.3.2 Supplementary Information on Selection . . . 281
7.4 Bibliographic Notes . . . 284
7.5 Problems and Supplements . . . 286

8 Other Topics . . . 293
8.1 Event Processes with Marks . . . 293
8.2 Models for Cumulative Costs . . . 295
8.2.1 Introduction . . . 295
8.2.2 Estimation for Cost Processes . . . 297
8.2.3 Examples . . . 299
8.3 Prediction . . . 302
8.3.1 Introduction . . . 302
8.3.2 Predictive Probabilities and Calibration . . . 304
8.3.3 Some Examples of Prediction . . . 306
8.4 Recurrent Events in Randomized Trials . . . 311
8.4.1 Specification and Testing of Treatment Effects . . . 311
8.4.2 Trial Design for Mixed Poisson Processes . . . 315
8.4.3 Use of Baseline Count Data . . . 316
8.4.4 Interim Monitoring with Recurrent Events . . . 320
8.5 Clustered Data . . . 324
8.6 Missing Covariate Values . . . 327
8.7 Covariate Measurement Error . . . 329
8.8 Bayesian Methods . . . 331
8.9 Bibliographic Notes . . . 332
8.10 Problems and Supplements . . . 333

A Estimation and Statistical Inference . . . 337
A.1 Maximum Likelihood . . . 337
A.1.1 Introduction . . . 337
A.1.2 Asymptotic Pivotals . . . 338
A.1.3 Confidence Regions or Intervals . . . 340
A.1.4 Tests of Hypotheses . . . 341
A.2 Estimating Functions . . . 342

B Computational Methods . . . 345
B.1 Software for Recurrent Events . . . 345
B.2 Optimization Methods . . . 346
B.3 Simulation and Resampling Methods . . . 346

C Code and Remarks for Selected Examples . . . 349
C.1 Tumorgenicity Data Analysis of Chapter 3 . . . 349
C.1.1 Poisson Analysis with Weibull Baseline Rate . . . 349
C.1.2 Poisson Analysis with Piecewise-Constant Rates . . . 351
C.1.3 Nonparametric and Semiparametric Poisson Analysis . . . 353
C.1.4 Semiparametric Mixed Poisson Analysis . . . 353

C.1.5 Robust Semiparametric Analysis 354
C.2 Code for rhDNase Data Analyses of Chapter 4 355
C.3 Code for Chronic Bronchitis Trial of Chapter 6 358

D Datasets 363
D.1 Bladder Cancer Data 363
D.2 Bowel Motility Data 364
D.3 Pulmonary Exacerbations and rhDNase 365
D.4 Software Debugging Data 367
D.5 Artificial Field Repair Data 367

References 371