
Contents

Preface	V
1 Introduction	1
1.1 Poisson Regression Model	1
1.2 Examples	2
1.3 Organization of the Book	4
2 Probability Models for Count Data	7
2.1 Introduction	7
2.2 Poisson Distribution	7
2.2.1 Definitions and Properties	7
2.2.2 Genesis of the Poisson Distribution	10
2.2.3 Poisson Process	11
2.2.4 Generalizations of the Poisson Process	14
2.2.5 Poisson Distribution as a Binomial Limit	15
2.2.6 Exponential Interarrival Times	16
2.2.7 Non-Poissonness	17
2.3 Further Distributions for Count Data	20
2.3.1 Negative Binomial Distribution	20
2.3.2 Binomial Distribution	25
2.3.3 Logarithmic Distribution	27
2.3.4 Summary	28
2.4 Modified Count Data Distributions	30
2.4.1 Truncation	30
2.4.2 Censoring and Grouping	31
2.4.3 Altered Distributions	32
2.5 Generalizations	33
2.5.1 Mixture Distributions	33
2.5.2 Compound Distributions	36
2.5.3 Birth Process Generalizations	39
2.5.4 Katz Family of Distributions	40

2.5.5	Additive Log-Differenced Probability Models	41
2.5.6	Linear Exponential Families	42
2.5.7	Summary	44
2.6	Distributions for Over- and Underdispersion	45
2.6.1	Generalized Event Count Model	45
2.6.2	Generalized Poisson Distribution	46
2.6.3	Poisson Polynomial Distribution	47
2.6.4	Double Poisson Distribution	49
2.6.5	Summary	49
2.7	Duration Analysis and Count Data	50
2.7.1	Distributions for Interarrival Times	52
2.7.2	Renewal Processes	54
2.7.3	Gamma Count Distribution	56
2.7.4	Duration Mixture Models	59
3	Poisson Regression	63
3.1	Specification	63
3.1.1	Introduction	63
3.1.2	Assumptions of the Poisson Regression Model	63
3.1.3	Ordinary Least Squares and Other Alternatives	65
3.1.4	Interpretation of Parameters	70
3.1.5	Period at Risk	74
3.2	Maximum Likelihood Estimation	77
3.2.1	Introduction	77
3.2.2	Likelihood Function and Maximization	77
3.2.3	Newton-Raphson Algorithm	78
3.2.4	Properties of the Maximum Likelihood Estimator	80
3.2.5	Estimation of the Variance Matrix	82
3.2.6	Approximate Distribution of the Poisson Regression Coefficients	83
3.2.7	Bias Reduction Techniques	84
3.3	Pseudo-Maximum Likelihood	87
3.3.1	Linear Exponential Families	89
3.3.2	Biased Poisson Maximum Likelihood Inference	90
3.3.3	Robust Poisson Regression	91
3.3.4	Non-Parametric Variance Estimation	95
3.3.5	Poisson Regression and Log-Linear Models	97
3.3.6	Generalized Method of Moments	98
3.4	Sources of Misspecification	102
3.4.1	Mean Function	102
3.4.2	Unobserved Heterogeneity	103
3.4.3	Measurement Error	105
3.4.4	Dependent Process	107
3.4.5	Selectivity	107
3.4.6	Simultaneity and Endogeneity	108

- 3.4.7 Underreporting 109
- 3.4.8 Excess Zeros 109
- 3.4.9 Variance Function 110
- 3.5 Testing for Misspecification 112
 - 3.5.1 Classical Specification Tests 112
 - 3.5.2 Regression Based Tests 118
 - 3.5.3 Goodness-of-Fit Tests 118
 - 3.5.4 Tests for Non-Nested Models 120
- 3.6 Outlook 125

- 4 Unobserved Heterogeneity 127**
 - 4.1 Introduction 127
 - 4.1.1 Conditional Mean Function 127
 - 4.1.2 Partial Effects with Unobserved Heterogeneity 128
 - 4.1.3 Unobserved Heterogeneity in the Poisson Model 129
 - 4.1.4 Parametric and Semi-Parametric Models 130
 - 4.2 Parametric Mixture Models 130
 - 4.2.1 Gamma Mixture 131
 - 4.2.2 Inverse Gaussian Mixture 131
 - 4.2.3 Log-Normal Mixture 132
 - 4.3 Negative Binomial Models 134
 - 4.3.1 Negbin II Model 135
 - 4.3.2 Negbin I Model 136
 - 4.3.3 Negbin_k Model 136
 - 4.3.4 Negbin_X Model 137
 - 4.4 Semiparametric Mixture Models 138
 - 4.4.1 Series Expansions 138
 - 4.4.2 Finite Mixture Models 139

- 5 Sample Selection and Endogeneity 143**
 - 5.1 Censoring and Truncation 143
 - 5.1.1 Truncated Count Data Models 144
 - 5.1.2 Endogenous Sampling 144
 - 5.1.3 Censored Count Data Models 146
 - 5.1.4 Grouped Poisson Regression Model 147
 - 5.2 Incidental Censoring and Truncation 148
 - 5.2.1 Outcome and Selection Model 148
 - 5.2.2 Models of Non-Random Selection 149
 - 5.2.3 Bivariate Normal Error Distribution 150
 - 5.2.4 Outcome Distribution 152
 - 5.2.5 Incidental Censoring 153
 - 5.2.6 Incidental Truncation 154
 - 5.3 Endogeneity in Count Data Models 156
 - 5.3.1 Introduction and Examples 156
 - 5.3.2 Parameter Ancillarity 157

5.3.3	Endogeneity and Mean Function	159
5.3.4	A Two-Equation Framework	161
5.3.5	Instrumental Variable Estimation	162
5.3.6	Estimation in Stages	165
5.4	Switching Regression	167
5.4.1	Full Information Maximum Likelihood Estimation	168
5.4.2	Moment-Based Estimation	170
5.4.3	Non-Normality	171
5.5	Mixed Discrete-Continuous Models	171
6	Zeros in Count Data Models	173
6.1	Introduction	173
6.2	Zeros in the Poisson Model	174
6.2.1	Excess Zeros and Overdispersion	174
6.2.2	Two-Crossings Theorem	175
6.2.3	Effects at the Extensive Margin	176
6.2.4	Multi-Index Models	177
6.2.5	A General Decomposition Result	177
6.3	Hurdle Count Data Models	178
6.3.1	Hurdle Poisson Model	181
6.3.2	Marginal Effects	182
6.3.3	Hurdle Negative Binomial Model	183
6.3.4	Non-nested Hurdle Models	183
6.3.5	Unobserved Heterogeneity in Hurdle Models	185
6.3.6	Finite Mixture Versus Hurdle Models	186
6.3.7	Correlated Hurdle Models	187
6.4	Zero-Inflated Count Data Models	188
6.4.1	Introduction	188
6.4.2	Zero-Inflated Poisson Model	189
6.4.3	Zero-Inflated Negative Binomial Model	191
6.4.4	Marginal Effects	191
6.5	Compound Count Data Models	192
6.5.1	Multi-Episode Models	193
6.5.2	Underreporting	193
6.5.3	Count Amount Model	196
6.5.4	Endogenous Underreporting	197
6.6	Quantile Regression for Count Data	199
7	Correlated Count Data	203
7.1	Multivariate Count Data	203
7.1.1	Multivariate Poisson Distribution	205
7.1.2	Multivariate Negative Binomial Model	210
7.1.3	Multivariate Poisson-Gamma Mixture Model	212
7.1.4	Multivariate Poisson-Log-Normal Model	213
7.1.5	Latent Poisson-Normal Model	216

7.1.6	Moment-Based Methods	217
7.1.7	Copula Functions	219
7.2	Panel Data Models	220
7.2.1	Fixed Effects Poisson Model	222
7.2.2	Moment-based Estimation of the Fixed Effects Model	225
7.2.3	Fixed Effects Negative Binomial Model	227
7.2.4	Random Effects Count Data Models	228
7.2.5	Dynamic Panel Count Data Models	230
7.3	Time-Series Count Data Models	232
8	Bayesian Analysis of Count Data	241
8.1	Bayesian Analysis of the Poisson Model	242
8.2	A Poisson Model with Underreporting	245
8.3	Estimation of the Multivariate Poisson-Log-Normal Model by MCMC	247
8.4	Estimation of a Random Coefficients Model by MCMC	248
9	Applications	251
9.1	Accidents	251
9.2	Crime	252
9.3	Trip Frequency	252
9.4	Health Economics	254
9.5	Demography	257
9.6	Marketing and Management	260
9.7	Labor Mobility	261
9.7.1	Economics Models of Labor Mobility	262
9.7.2	Previous Literature	263
9.7.3	Data and Descriptive Statistics	265
9.7.4	Regression Results	269
9.7.5	Model Performance	272
9.7.6	Marginal Probability Effects	274
9.7.7	Structural Inferences	278
A	Probability Generating Functions	281
B	Gauss-Hermite Quadrature	285
C	Software	289
D	Tables	291
	References	299
	Author's Index	321
	Subject Index	327