The Conway–Maxwell–Poisson Distribution

While the Poisson distribution is a classical statistical model for count data, the distributional model hinges on the constraining property that its mean equals its variance. This text instead introduces the Conway–Maxwell–Poisson distribution and motivates its use in developing flexible statistical methods based on its distributional form.

This two-parameter model not only contains the Poisson distribution as a special case but, in its ability to account for data over- or under-dispersion, encompasses both the geometric and Bernoulli distributions. The resulting statistical methods serve in a multitude of ways, from an exploratory data analysis tool to a flexible modeling impetus for varied statistical methods involving count data.

The first comprehensive reference on the subject, this text contains numerous illustrative examples demonstrating R code and output. It is essential reading for academics in statistics and data science, as well as quantitative researchers and data analysts in economics, biostatistics and other applied disciplines.

KIMBERLY F. SELLERS is Professor in the Department of Mathematics and Statistics at Georgetown University and a researcher at the U.S. Census Bureau. Her work has contributed to count data research and software for the last 15 years. She is a Fellow of the American Statistical Association and an Elected Member of the International Statistical Institute.
IMS Monographs are concise research monographs of high quality on any branch of statistics or probability of sufficient interest to warrant publication as books. Some concern relatively traditional topics in need of up-to-date assessment. Others are on emerging themes. In all cases the objective is to provide a balanced view of the field.

Other Books in the Series

1. Large-Scale Inference, by Bradley Efron
2. Nonparametric Inference on Manifolds, by Abhishek Bhattacharya and Rabi Bhattacharya
3. The Skew-Normal and Related Families, by Adelchi Azzalini
4. Case-Control Studies, by Ruth H. Keogh and D. R. Cox
5. Computer Age Statistical Inference, by Bradley Efron and Trevor Hastie
6. Computer Age Statistical Inference (Student Edition), by Bradley Efron and Trevor Hastie
7. Stable Lévy Process via Lamperti-Type Representations, by Andreas E. Kyprianou and Juan Carlos Pardo
The Conway–Maxwell–Poisson Distribution

KIMBERLY F. SELLERS

Georgetown University
To those that “count” most in my life:

My family, especially my son
Contents

List of Figures  xii
List of Tables  xiv
Preface  xxi
Acknowledgments  xxiii

1 Introduction: Count Data Containing Dispersion  1
   1.1 Poisson Distribution  2
      1.1.1 R Computing  4
   1.2 Data Over-dispersion  5
      1.2.1 R Computing  7
   1.3 Data Under-dispersion  10
      1.3.1 R Computing  11
   1.4 Weighted Poisson Distributions  16
   1.5 Motivation, and Summary of the Book  18

2 The Conway–Maxwell–Poisson (COM–Poisson) Distribution  22
   2.1 The Derivation/Motivation: A Flexible Queueing Model  23
   2.2 The Probability Distribution  25
      2.2.1 R Computing  28
   2.3 Distributional and Statistical Properties  35
      2.3.1 R Computing  39
   2.4 Parameter Estimation and Statistical Inference  40
      2.4.1 Combining COM–Poissonness Plot with Weighted
              Least Squares  40
      2.4.2 Maximum Likelihood Estimation  41
      2.4.3 Bayesian Properties and Estimation  42
      2.4.4 R Computing  43
      2.4.5 Hypothesis Tests for Dispersion  50
## Contents

2.5 Generating Data 52
  2.5.1 Inversion Method 52
  2.5.2 Rejection Sampling 53
  2.5.3 R Computing 56
2.6 Reparametrized Forms 57
2.7 COM–Poisson Is a Weighted Poisson Distribution 64
2.8 Approximating the Normalizing Term, $Z(\lambda, \nu)$ 65
2.9 Summary 69

3 Distributional Extensions and Generalities 71
  3.1 The Conway–Maxwell–Skellam (COM–Skellam or CMS) Distribution 72
  3.2 The Sum-of-COM–Poissons (sCMP) Distribution 75
  3.3 Conway–Maxwell Inspired Generalizations of the Binomial Distribution 77
    3.3.1 The Conway–Maxwell–binomial (CMB) Distribution 78
    3.3.2 The Generalized Conway–Maxwell–Binomial Distribution 81
    3.3.3 The Conway–Maxwell–multinomial (CMM) Distribution 83
    3.3.4 CMB and CMM as Sums of Dependent Bernoulli Random Variables 86
    3.3.5 R Computing 87
  3.4 CMP-Motivated Generalizations of the Negative Binomial Distribution 93
    3.4.1 The Generalized COM–Poisson (GCMP) Distribution 93
    3.4.2 The COM–Negative Binomial (COMNB) Distribution 98
    3.4.3 The COM-type Negative Binomial (COMtNB) Distribution 101
    3.4.4 Extended CMP (ECMP) Distribution 106
  3.5 Conway–Maxwell Katz (COM–Katz) Class of Distributions 112
  3.6 Flexible Series System Life-Length Distributions 113
    3.6.1 The Exponential-CMP (ExpCMP) Distribution 114
    3.6.2 The Weibull–CMP (WCMP) Distribution 117
  3.7 CMP-Motivated Generalizations of the Negative Hypergeometric Distribution 119
    3.7.1 The COM-negative Hypergeometric (COMNH) Distribution, Type I 120
    3.7.2 The COM–Poisson-type Negative Hypergeometric (COMPtNH) Distribution 120
<table>
<thead>
<tr>
<th>Contents</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>3.7.3 The COM-Negative Hypergeometric (CMNH) Distribution, Type II</td>
<td>121</td>
</tr>
<tr>
<td>3.8 Summary</td>
<td>123</td>
</tr>
<tr>
<td>4 Multivariate Forms of the COM–Poisson Distribution</td>
<td>124</td>
</tr>
<tr>
<td>4.1 Trivariate Reduction</td>
<td>125</td>
</tr>
<tr>
<td>4.1.1 Parameter Estimation</td>
<td>127</td>
</tr>
<tr>
<td>4.1.2 Hypothesis Testing</td>
<td>128</td>
</tr>
<tr>
<td>4.1.3 Multivariate Generalization</td>
<td>129</td>
</tr>
<tr>
<td>4.2 Compounding Method</td>
<td>131</td>
</tr>
<tr>
<td>4.2.1 Parameter Estimation</td>
<td>133</td>
</tr>
<tr>
<td>4.2.2 Hypothesis Testing</td>
<td>134</td>
</tr>
<tr>
<td>4.2.3 R Computing</td>
<td>135</td>
</tr>
<tr>
<td>4.2.4 Multivariate Generalization</td>
<td>137</td>
</tr>
<tr>
<td>4.3 The Sarmanov Construction</td>
<td>138</td>
</tr>
<tr>
<td>4.3.1 Parameter Estimation and Hypothesis Testing</td>
<td>140</td>
</tr>
<tr>
<td>4.3.2 Multivariate Generalization</td>
<td>140</td>
</tr>
<tr>
<td>4.4 Construction with Copulas</td>
<td>141</td>
</tr>
<tr>
<td>4.5 Real Data Examples</td>
<td>143</td>
</tr>
<tr>
<td>4.5.1 Over-dispersed Example: Number of Shunter Accidents</td>
<td>143</td>
</tr>
<tr>
<td>4.5.2 Under-dispersed Example: Number of All-Star Basketball Players</td>
<td>147</td>
</tr>
<tr>
<td>4.6 Summary</td>
<td>150</td>
</tr>
<tr>
<td>5 COM–Poisson Regression</td>
<td>153</td>
</tr>
<tr>
<td>5.1 Introduction: Generalized Linear Models</td>
<td>154</td>
</tr>
<tr>
<td>5.1.1 Logistic Regression</td>
<td>154</td>
</tr>
<tr>
<td>5.1.2 Poisson Regression</td>
<td>155</td>
</tr>
<tr>
<td>5.1.3 Addressing Data Over-dispersion: Negative Binomial Regression</td>
<td>156</td>
</tr>
<tr>
<td>5.1.4 Addressing Data Over- or Under-dispersion: Restricted Generalized Poisson Regression</td>
<td>157</td>
</tr>
<tr>
<td>5.2 Conway–Maxwell–Poisson (COM–Poisson) Regression</td>
<td>157</td>
</tr>
<tr>
<td>5.2.1 Model Formulations</td>
<td>158</td>
</tr>
<tr>
<td>5.2.2 Parameter Estimation</td>
<td>160</td>
</tr>
<tr>
<td>5.2.3 Hypothesis Testing</td>
<td>173</td>
</tr>
<tr>
<td>5.2.4 R Computing</td>
<td>173</td>
</tr>
<tr>
<td>5.2.5 Illustrative Examples</td>
<td>178</td>
</tr>
<tr>
<td>5.3 Accounting for Excess Zeroes: Zero-inflated COM–Poisson Regression</td>
<td>190</td>
</tr>
<tr>
<td>5.3.1 Model Formulations</td>
<td>191</td>
</tr>
</tbody>
</table>
### Contents

5.3.2 Parameter Estimation 194
5.3.3 Hypothesis Testing 195
5.3.4 A Word of Caution 196
5.3.5 Alternative Approach: Hurdle Model 197
5.4 Clustered Data Analysis 197
5.5 R Computing for Excess Zeroes and/or Clustered Data 202
5.5.1 Examples 203
5.6 Generalized Additive Model 211
5.7 Computing via Alternative Softwares 213
5.7.1 MATLAB Computing 213
5.7.2 SAS Computing 213
5.8 Summary 217

6 COM–Poisson Control Charts 218
6.1 CMP-Shewhart Charts 219
6.1.1 CMP Control Chart Probability Limits 220
6.1.2 R Computing 221
6.1.3 Example: Nonconformities in Circuit Boards 223
6.1.4 Multivariate CMP-Shewhart Chart 224
6.2 CMP-inspired EWMA Control Charts 226
6.2.1 COM–Poisson EWMA (CMP-EWMA) Chart 227
6.2.2 CMP-EWMA Chart with Multiple Dependent State Sampling 229
6.2.3 CMP-EWMA Chart with Repetitive Sampling 231
6.2.4 Modified CMP-EWMA Chart 233
6.2.5 Double EWMA Chart for CMP Attributes 234
6.2.6 Hybrid EWMA Chart 237
6.3 COM–Poisson Cumulative Sum (CUSUM) Charts 238
6.3.1 CMP-CUSUM charts 239
6.3.2 Mixed EWMA-CUSUM for CMP Attribute Data 243
6.4 Generally Weighted Moving Average 244
6.5 COM–Poisson Chart Via Progressive Mean Statistic 248
6.6 Summary 249

7 COM–Poisson Models for Serially Dependent Count Data 251
7.1 CMP-motivated Stochastic Processes 251
7.1.1 The Homogeneous CMP Process 251
7.1.2 Copula-based CMP Markov Models 257
7.1.3 CMP-Hidden Markov Models 259
7.2 Intensity Parameter Time Series Modeling 263
7.2.1 ACMP-INGARCH 263
### Contents

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.2.2</td>
<td>MCMP1-ARMA</td>
<td>265</td>
</tr>
<tr>
<td>7.3</td>
<td>Thinning-Based Models</td>
<td>267</td>
</tr>
<tr>
<td>7.3.1</td>
<td>Autoregressive Models</td>
<td>268</td>
</tr>
<tr>
<td>7.3.2</td>
<td>Moving Average Models</td>
<td>273</td>
</tr>
<tr>
<td>7.4</td>
<td>CMP Spatio-temporal Models</td>
<td>280</td>
</tr>
<tr>
<td>7.5</td>
<td>Summary</td>
<td>282</td>
</tr>
<tr>
<td>8</td>
<td>COM–Poisson Cure Rate Models</td>
<td>284</td>
</tr>
<tr>
<td>8.1</td>
<td>Model Background and Notation</td>
<td>286</td>
</tr>
<tr>
<td>8.2</td>
<td>Right Censoring</td>
<td>288</td>
</tr>
<tr>
<td>8.2.1</td>
<td>Parameter Estimation Methods</td>
<td>288</td>
</tr>
<tr>
<td>8.2.2</td>
<td>Quantifying Variation</td>
<td>294</td>
</tr>
<tr>
<td>8.2.3</td>
<td>Simulation Studies</td>
<td>294</td>
</tr>
<tr>
<td>8.2.4</td>
<td>Hypothesis Testing and Model Discernment</td>
<td>295</td>
</tr>
<tr>
<td>8.3</td>
<td>Interval Censoring</td>
<td>298</td>
</tr>
<tr>
<td>8.3.1</td>
<td>Parameter Estimation</td>
<td>299</td>
</tr>
<tr>
<td>8.3.2</td>
<td>Variation Quantification</td>
<td>300</td>
</tr>
<tr>
<td>8.3.3</td>
<td>Simulation Studies</td>
<td>301</td>
</tr>
<tr>
<td>8.3.4</td>
<td>Hypothesis Testing and Model Discernment</td>
<td>301</td>
</tr>
<tr>
<td>8.4</td>
<td>Destructive CMP Cure Rate Model</td>
<td>301</td>
</tr>
<tr>
<td>8.4.1</td>
<td>Parameter Estimation</td>
<td>304</td>
</tr>
<tr>
<td>8.4.2</td>
<td>Hypothesis Testing and Model Discernment</td>
<td>306</td>
</tr>
<tr>
<td>8.5</td>
<td>Lifetime Distributions</td>
<td>306</td>
</tr>
<tr>
<td>8.6</td>
<td>Summary</td>
<td>310</td>
</tr>
</tbody>
</table>

**References** 312

**Index** 327
Figures

1.1 Poisson probability mass function illustrations for $\lambda \in \{0.3, 1, 3, 10\}$. page 3

1.2 Negative binomial distribution illustrations for values of $(r, p) \in \{(5, 0.4), (10, 0.7), (15, 0.8), (60, 0.95), (300, 0.99)\}$ and the Poisson($\lambda = 3$) probability mass function. This series of density plots nicely demonstrates the distributional convergence of the negative binomial to the Poisson as $r \to \infty$ and $p \to 1$ such that $r(1 - p) \to \lambda$. 8

1.3 Generalized Poisson probability mass function illustrations for values of $\lambda_1 > 0$, and dispersion parameter $\lambda_2 \in \{-0.5, 0, 0.5\}$. For $\lambda_1 > 0$ and $-1 < \lambda_2 < 1$ such that $\lambda_2 > (\leq) 0$ denotes data over-dispersion (under-dispersion), the generalized Poisson distribution has the mean $E(X) = \frac{\lambda_1}{1 - \lambda_2}$ and variance $V(X) = \frac{\lambda_1}{(1 - \lambda_2)^2}$. 12

1.4 The probability mass function $P(X = x)$ created for $x \in \{0, \ldots, 20\}$ for a generalized Poisson distribution (a) via dgenpois (HMMpa) with $\lambda_1 = 3$, $\lambda_2 = -0.5$; and (b) via dgpois (LaplaceDemon) with $\lambda = 2$, $\omega = -0.5$. The resulting plots should be identical because $\lambda_1 = \lambda(1 - \omega)$ and $\lambda_2 = \omega$. 14

2.1 CMP probability mass function illustrations for the values of $\lambda$ and $\nu$. Respective illustrative plots define the same value for $\lambda$ when $\nu > 0$ for easy distributional comparisons, while for $\nu = 0$, $\lambda$ must be constrained to be less than 1; $\nu < (>) 1$ signifies data over-dispersion (under-dispersion) relative to the Poisson ($\nu = 1$) distribution. 27

2.2 COM–Poissonness plot associated with Macaulay (1923) data. 48

3.1 COM–Skellam($\lambda_1 = \lambda_2 = 5, \nu$) probability mass function illustrations for the values of $\nu \in \{0.25, 0.5, 1, 2, 4, 10\}$. 73
List of Figures xiii

3.2 sCMP(m, \lambda = 1, \nu) probability mass function illustrations for the values of m \in \{2, 3, 4, 5\} and \nu \in \{0.5, 1, 2\}. 76

5.1 Trace and density plots associated with textile fabrics example for (a) \beta (b) \gamma; RL = RollLength. 189

5.2 Credible intervals associated with textile fabrics example for (a) \beta (b) \gamma; RL = RollLength. Inner and outer intervals, respectively, represent 68% and 95% credible intervals. 190

6.1 Control chart associated with nonconformities data analysis via CMPControl package: corresponding R code and output supplied in Code 6.1. Lower CMP-Shewhart bound is Winsorized to 0, while the lower CMP probability bound equals 0. 223
Tables

1.1 Weight functions associated with various examples of weighted Poisson distributions.  
1.2 Levels of model support based on AIC difference values, $\Delta_i = \text{AIC}_i - \text{AIC}_{\text{min}}$, for Model $i$ (Burnham and Anderson, 2002).  
2.1 Special cases of the CMP parametrization distribution.  
2.2 Available R functions for CMP computing.  
2.3 (Cumulative) probability computations via various R packages and their respective functions, illustrated assuming a CMP$(\lambda = 4, \nu)$ random variable $X$ evaluated at the value 2 for $\nu \in \{0.3, 1, 3, 30\}$. Functions produce equal calculations when rounded to three decimal places.  
2.4 Probability $P(X = 2)$ and cumulative probability $P(X \leq 2)$ computations for the CMP$(\lambda = 4, \nu = 1) = \text{Poisson}(\lambda = 4)$ distributed random variable $X$. CMP computations determined using compoisson (Dunn, 2012), CompGLM (Pollock, 2014a), and COMPoissonReg (Sellers et al., 2019); Poisson results obtained using the stats package. All calculations rounded to six decimal places.  
2.5 Probability $P(X = 2)$ and cumulative probability $P(X \leq 2)$ computations for the CMP$(\lambda = 0.25, \nu = 0) = \text{Geom}(p = 0.75)$ distributed random variable $X$. CMP computations determined using compoisson (Dunn, 2012), CompGLM (Pollock, 2014a), and COMPoissonReg (Sellers et al., 2019); geometric results obtained using the stats package. All calculations rounded to six decimal places.  
2.6 Hypothetical frequency table for count data. These data are used for illustrative analyses in Code 2.1.
List of Tables

2.7 Quantile determinations $x$ such that $P(X \leq x) \geq 0.9$ for the CMP$(\lambda = 3, \nu)$ distributed random variable $X$, where $\nu \in \{0.3, 1, 3, 30\}$. Computations conducted via the qcmp function (COMPoissonReg).

2.8 Observed versus CMP estimated frequency of occurrences of the articles “the,” “an,” and “a” in five-word samples from “Essay on Milton” by Macaulay (1923) (Oxford edition). CMP estimated frequencies obtained based on the maximum likelihood estimators $\lambda \approx 1.0995$ and $\nu \approx 3.2864$.

2.9 Observed versus CMP estimated frequency of the number of children born to a random sample of females (Winkelmann, 1995).

2.10 Available R functions based on reparametrized COM–Poisson models.

2.11 Probability computations (to six decimal places) via dcmpois (combayes) and dcomp (mpcmp) for ACM$P(\mu = \lambda^{1/\nu} = 4^{1/\nu}, \nu)$ and MCMP1$P(\mu, \nu) =$ CMP$(\lambda = 4, \nu)$, respectively, and cumulative probability pcmp for the MCMP1 distribution.

3.1 Probability, normalizing constant, expected value, and variance calculations (rounded to three decimal places) for CMB$(m = 10, p = 0.75, \nu)$ distributions where $\nu \in \{0, 1, 2\}$.

3.2 Probability, normalizing constant, expected value, and variance calculations (rounded to three decimal places) for CMM$3(m = 10; p = (0.1, 0.35, 0.55); \nu)$ distributions where $\nu \in \{0, 1, 2\}$.

3.3 Summary of CMP-motivated generalizations of the negative binomial (NB) distribution – the generalized COM–Poisson (GCMP), the Conway-Maxwell-negative binomial (COMNB), the Conway-and-Maxwell-type negative binomial (COMtNB), and the extended Conway–Maxwell–Poisson (ECMP) – and the special cases contained by each of them (noted with $\sqrt{}$), namely any of the following: NB, CMP, Conway-Maxwell-binomial (CMB), COMNB, COMtNB, exponentially weighted Poisson (EWP), and GCMP.

3.4 Distributional properties of the exponential-geometric (EG) and exponential-Poisson (EP) distributions.

4.1 Bivariate copula functions.

4.2 Parameter maximum likelihood estimates (MLEs), log-likelihood (ln $L$), Akaike information criterion (AIC), difference in AIC ($\Delta_i = AIC_i - AIC_{min}$), goodness-of-fit (GOF = $\sum \frac{(O - E)^2}{E}$), where $O$ and $E$ denote the observed and expected cell frequencies,
respectively) measures and associated p-values for various bivariate distributions on the shunters accident dataset: bivariate Poisson (BP); bivariate negative binomial (BNB); bivariate generalized Poisson (BGP); and the BCMP obtained via the trivariate reduction (BCMPtriv), compounding (BCMPcomp), or either Sarmanov family (BCMPsar1 and BCMPsar2, respectively) method. 145

4.3 Observed accident data among 122 shunters along with associated count estimates from various bivariate distributions: bivariate negative binomial (BNB), bivariate Poisson (BP), bivariate generalized Poisson (BGP), bivariate geometric (BG), and BCMP obtained via the compounding (BCMPcomp), trivariate reduction (BCMPtriv), or either Sarmanov family (BCMPsar1 and BCMPsar2) methods. Estimated counts determined from MLEs for respective model parameters reported in Table 4.2. 146

4.4 Observed accident data (continued from Table 4.3) among 122 shunters. 148

4.5 Respective maximum likelihood estimates (MLEs), log-likelihood (ln $L$) values, Akaike information criterion (AIC), and $\Delta_i = AIC_i - AIC_{min}$ values for various bivariate models, namely the bivariate Poisson distribution (BP), bivariate negative binomial (BNB), bivariate generalized Poisson (BGP), and four BCMP models attained via trivariate reduction (BCMPtriv), compounding (BCMPcomp), and two Sarmanov family approaches (BCMPsar1 and BCMPsar2), respectively, on the number of Forward and Center players dataset. 149

4.6 Bivariate CMP development approaches (trivariate reduction; compounding; the Sarmanov families considering the CMP distribution as a weighted Poisson (Sarmanov 1) or based on the CMP probability generating function (Sarmanov 2), respectively; and copulas) and associated qualities. For each of the considered approaches, the correlation range and reported special-case distributions attainable for the bivariate (Biv.) and marginal (Marg.) distributions are supplied. 151

5.1 Structure of various generalized linear models. 154

5.2 Coefficient estimates and standard errors (in parentheses) associated with the number of children from women over 44 years of age and in their first marriage. Respective outputs likewise report the associated log-likelihood and Akaike information criterion (AIC) associated with each model. The glm.cmp
(COMPoissonReg) and glm.comp (CompGLM) functions conduct CMP regression, while the glm.cmp (mpcmp) and glm.CMP (DGLMExtPois) functions conduct MCMP1 regression. NR = not reported.

5.3 Airfreight breakage dataset, where broken denotes the number of broken jars detected following a flight that involved a number of transfers, transfers (Kutner et al., 2003).

5.4 Coefficient estimates and standard errors (in parentheses) associating the number of broken jars detected following a flight that involved a number of transfers. Respective outputs likewise report the associated log-likelihood and Akaike information criterion (AIC) associated with each model. The glm.cmp (COMPoissonReg) and glm.comp (CompGLM) functions conduct CMP regression, while the glm.cmp (mpcmp) and glm.CMP (DGLMExtPois) functions conduct MCMP1 regression. NA = not applicable; NR = not reported.

5.5 Coefficient estimates and standard errors (in parentheses) associating the number of faults in rolls of fabric with the logarithm of the corresponding roll length. Respective outputs likewise report the corresponding Akaike information criterion (AIC) for each model. CMP regression was conducted via the COMPoissonReg package, while MCMP1 regressions were performed via the mpcmp and DGLMExtPois packages, respectively. NR = not reported. Dispersion measures are reported on varying scales (with standard errors rounded to two decimal places) as provided in the respective outputs and do not allow for direct comparison.

5.6 Estimated coefficients and standard errors (in parentheses), negated log-likelihood, and Akaike information criterion (AIC) for various zero-inflated regressions associating the number of unwanted pursuit behavior perpetrations in the context of couple separation trajectories with the levels of education (an indicator function where 1 (0) denotes having at least bachelor’s degree (otherwise)) and anxious attachment (a continuous measure) among 387 participants. Considered models are zero-inflated Poisson (ZIP), negative binomial (ZINB), CMP (ZICMP), Huang (2017) mean-parametrized COM–Poisson (ZIMCMP1), and geometric (ZIG), as well as a hurdle MCMP1 (HMCMP1) model. NR = not reported.
List of Tables

5.7 Estimated coefficients and standard errors (in parentheses), log-likelihood, Akaike information criterion (AIC), and deviance for various epilepsy longitudinal data analyses associating the number of seizures experienced by 59 patients in an eight-week baseline period, followed by four consecutive two-week periods where the patients are treated with progabide. Baseline Poisson and mean-parametrized COM–Poisson (MCMP1) regressions (along with their zero-inflated and hurdle analog models) are considered for constructing generalized linear mixed models, where, for Subject $i$, $T_{ij}$ denotes the length (in weeks) of the time period $j$, $x_{ij1}$ is an indicator function of a period after the baseline (i.e. weeks 8 through 16), $x_{ij2}$ is an indicator function noting whether or not the progabide medication is administered, and $\sigma^2$ is the variance associated with the random intercept. Zero-inflation and hurdle regressions are performed assuming a constant model (i.e. Equation (5.48) reduces to $\text{logit}(\pi^*) = \zeta_0$). The parameter $\nu$ denotes the associated MCMP1 dispersion component under each respective model.

5.8 COUNTREG output from the airfreight breakage example with CMP regression.

5.9 COUNTREG output from the airfreight breakage example with approximate COM–Poisson (ACMP) regression.

6.1 Centerline and Shewhart $k\sigma$ upper/lower control limits for cmpc- and cmpu-charts (Sellers, 2012b).

7.1 R functions provided in the cmpprocess package for CMP process analysis. These functions determine (approximate) MLEs based on the information provided by the analyst.

7.2 Data (presented in sequential order, left to right) regarding the number of alpha particles emitted in successive 7.5-second intervals from a disk coated with polonium via the scintillation method (Rutherford et al., 1910).

7.3 R functions to conduct statistical computing associated with CMP-hidden Markov modeling. Codes available online as supplementary material associated with MacDonald and Bhamani (2020).

7.4 Univariate and multivariate thinning-based time series constructions involving COM–Poisson-motivated distributions. Khan and Jowaheer (2013) and Jowaheer et al. (2018) use a modified CMP notation (namely CMP($\mu$, $\nu$)) that relies on the
approximations for the mean \( \mu = \lambda^{1/\nu} - \frac{\nu-1}{2\nu} \) and variance \( \sigma^2 = \frac{1}{\nu} \lambda^{1/\nu} \). Considered estimation methods are either generalized quasi-likelihood (GQL) or maximum likelihood (ML).

8.1 Special-case CMP cure rate models that have closed forms for \( g(\lambda;\nu) \) and \( \lambda = g^{-1}(1 + \exp(\mathbf{x}'\mathbf{\beta});\nu) \).

8.2 Expectation step functions for fixed \( \nu \) from the CMP cure rate model with right censoring and various associated lifetime distributions.

8.3 Expectation step functions for fixed \( \nu \) from the CMP cure rate model with right censoring and various associated lifetime distributions (continued).
Preface

Welcome to *The Conway–Maxwell–Poisson Distribution* – the first coherent introduction to the Conway–Maxwell–Poisson distribution and its contributions with regard to statistical theory and methods. This two-parameter model not only serves as a flexible distribution containing the Poisson distribution as a special case but also, in its ability to capture either data over- or under-dispersion, it contains (in particular) two other classical distributions. The Conway–Maxwell–Poisson distribution thereby can effectively model a range of count data distributions that contain data over- or under-dispersion, simply through the addition of one parameter. This distribution’s flexibility offers numerous opportunities with regard to statistical methods development. To date, such efforts involve work in univariate and multivariate distributional theory, regression analysis (including spatial and/or temporal models, and cure rate models), control chart theory, and count processes. Accordingly, the statistical methods described in this reference can effectively serve in a multitude of ways, from an exploratory data analysis tool to an appropriate, flexible count data modeling impetus for a variety of statistical methods involving count data.

*The Conway–Maxwell–Poisson Distribution* can benefit a broad statistical audience. This book combines theoretical and applied data developments and discussions regarding the Conway–Maxwell–Poisson distribution and its significant flexibility in modeling count data, where this reference adopts the convention that the counting numbers are the natural numbers including zero, i.e. $\mathbb{N} = \{0, 1, 2, \ldots\}$. Count data modeling research is a topic of interest to the academic audience, ranging from upper-level undergraduates to graduate students and faculty in statistics (and, more broadly, data science). Meanwhile, the compelling nature of this topic and the writing format of the reference
Preface

intend to draw quantitative researchers and data analysts in applied disciplines, including business and economics, medicine and public health, engineering, psychology, and sociology – broadly anyone interested in its supporting computational discussions and examples using R. This reference seeks to assume minimal prerequisite statistics coursework/knowledge (e.g. calculus and a calculus-based introduction to probability and statistics that includes maximum likelihood estimation) throughout the book. More advanced readers, however, will benefit from additional knowledge of other subject areas in some chapters, for example, linear algebra or Bayesian computation.

Along with this reference’s discussion of flexible statistical methods for count data comes an accounting of available computation packages in R to conduct analyses. Accordingly, preliminary R knowledge will also prove handy as this reference brings to light the various packages that exist for modeling count data via the Conway–Maxwell–Poisson distribution through the relevant statistical methods. The Comprehensive R Archive Network (CRAN) regularly updates its system. In the event that any package discussed in this reference is subsequently no longer directly accessible through the CRAN, note that it is archived and thus still accessible for download and use by analysts.
Acknowledgments

This book is not only the culmination of my years of work and research developing this field but also represents the vastness of contributions in statistical methods and computation by the many authors that are cited in this reference. I am particularly thankful to Dr. Galit Shmueli, the first to introduce me to the Conway–Maxwell–Poisson distribution. That initial collaboration sparked my great interest in this topic where I could easily recognize the broad and diverse areas for potential research, some of which has since come to fruition by myself and/or others. Accordingly, I further thank my colleague and student collaborators with whom we have accomplished a great deal. Thanks as well are extended to other researchers in the larger field of count data modeling with whom I have not only established a camaraderie, but some of whom have also provided discussions and/or analysis associated with their respective models that do not provide an R package for direct computation, including Dr. N. Balakrishnan regarding various works on cure-rate models, Dr. Alan Huang regarding the mean-parametrized Conway–Maxwell–Poisson distribution and the mpcmp package, Dr. Felix Famoye regarding the bivariate generalized Poisson, Mr. Tong Li for the trivariate-reduced bivariate COM–Poisson, and Dr. S. H. Ong for both Sarmanov-formulated bivariate COM–Poisson models.

Words cannot fully describe how much I thank and appreciate Sir David Cox for his vision and encouragement for me to author this reference; I am only sorry that he was not able to see it come to fruition. I also am forever thankful and appreciative to Dr. Darcy Steeg Morris for our countless, fruitful conversations as she aided with editing the content in this reference that not only helped to shape its vision, but further ensure its accuracy, understanding, and comprehension.

xxiii
xxiv

Acknowledgments

Completing such a project does not happen in a vacuum. I thank family and friends who, through word and deed, kept me encouraged, motivated, and focused to see this project to and through the finish line. I particularly thank Dr. Kris Marsh for our many hours spent together in our self-made writing group. I am also appreciative to Georgetown University for their financial support through avenues including the Senior Faculty Research Fellowship which aided the preparation of this reference.