SORT 46 (1) January-June 2022, 39-66 DOI: 10.2436/20.8080.02.117

Unusual-event processes for count data

Wanrudee Skulpakdee[1] and Mongkol Hunkrajok[2]

Abstract

At least one unusual event appears in some count datasets. It will lead to a more concentrated (or dispersed) distribution than the Poisson, gamma, Weibull, Conway-Maxwell-Poisson (CMP), and Faddy (1997) models can accommodate. These well-known count models are based on the monotonic rates of interarrival times between successive events. Under the assumption of non-monotonic rates and independent exponential interarrival times, a new class of parametric models for unusual-event (UE) count data is proposed. These models are applied to two empirical applications, the number of births and the number of bids, and yield considerably better results to the above well-known count models.

MSC: *62J99, 62M05, 62P99.*

Keywords: *Poisson count model, Gamma count model, Weibull count model, Conway-Maxwell-Poisson count model, Faddy count model.*

1. Introduction

Count data regression analysis is a collection of statistical techniques for modeling and investigating the conditional count distributions of count response variables given sets of covariates. The conditional-variance-mean function of these distributions can be classified into two different categories: linear and non-linear.

1. If the distributions are equidispersed (variance $=$ mean), this function is linear.
2. If the distributions are overdispersed (variance $>$ mean), this function is either linear or non-linear.

[1] Graduate School of Applied Statistics, National Institute of Development Administration, Bangkok, Thailand. wanrudee.sku@nida.ac.th

[2] Independent Researcher, Bangkok, Thailand. hunkrajokmongkol@gmail.com

Received: June 2021
Accepted: November 2021

3. If the distributions are underdispersed (variance $<$ mean), this function is either linear or non-linear.

4. If the distributions are over-, under-, and equidispersed, this function is non-linear.

A renewal process is a counting process. Its times between successive events are independent and identically distributed with a non-negative distribution (Ross 2010). The primary assumption of the Poisson model is that the times between events are exponential. It follows that the Poisson model is equidispersed, and the Poisson regression model has a linear conditional-variance-mean function. The exponential distribution replaced by a less restrictive non-negative distribution such as the gamma and Weibull distributions leads to the gamma (Winkelmann 1995) and Weibull (McShane et al. 2008) count models. They allow for both overdispersion and underdispersion. The gamma and Weibull regression models have linear conditional-variance-mean functions when the additional parameter (α) equals 1, that is, the Poisson regression model. Furthermore, they have nearly linear conditional-variance-mean functions shown in Figures 1(a) and 1(b), although α does not approach 1.

The Conway-Maxwell-Poisson (CMP) model was originally introduced by Conway and Maxwell (1962). In contrast to the above models, the CMP model is not derived from an underlying renewal process. The proof can be found in the Supplementary Material. Surprisingly, however, the graphs in Figures 1(a) and 1(c) of the conditional-variance-mean functions for the gamma and the CMP are hardly distinguishable. A plausible explanation for this similarity is the equality of their approximate variance-mean ratios. These ratios are equal to a constant $1/\alpha$ (Winkelmann 1995, p. 470; Sellers and Shmueli 2010, p. 946). Likewise the gamma and Weibull count models, the CMP model consists of the rate and dispersion parameters. Thus, it allows for both over- and underdispersion.

As previously mentioned, the conditional variance and mean of the above well-known regression models are (nearly) linearly related. In some applications, these regression models are either unsatisfactory or inappropriate when the sample relative frequency distribution is created as a mixture of distributions whose relationship between the variance and the mean is non-linear.

The common assumption that the rates of interarrival times are equal may cause a (nearly) linear conditional-variance-mean function. One potential solution to this problem is to allow the unequal rates. Faddy (1997) suggested the generalization of the Poisson process $\lambda_n = \lambda\,(b+n)^{\alpha}$, $n = 0,1,,2,...$, in which the rate at which new events occur depends on the number of events. The rate sequence of the Faddy (1997) process is either non-decreasing or non-increasing. The Faddy (1997) regression model has both (nearly) linear and non-linear conditional-variance-mean functions shown in Figures 1(d), but it displays only one of over-, under-, and equidispersion. Therefore, this regression model is either unsatisfactory or inappropriate when the sample relative frequency distribution is created as a mixture of over-, under-, and equidispersed distributions whose relationship between the variance and the mean must be non-linear. Note that the conditional variances and means in Figure 1 were computed in **R**

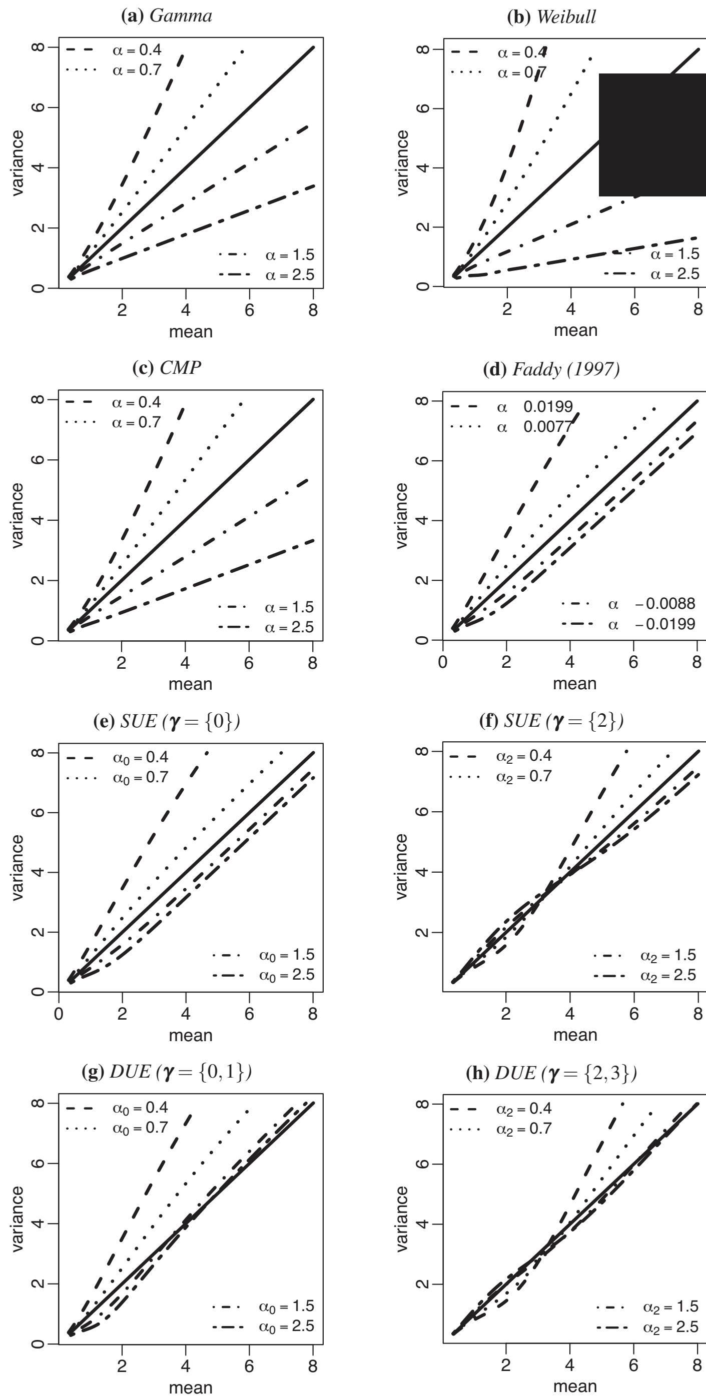


Figure 1. *Graphs showing the linear and non-linear functions of variance and mean. The Faddy (1997), DUE ($\boldsymbol{\gamma} = \{0,1\}$), and DUE ($\boldsymbol{\gamma} = \{2,3\}$) models present the cases in which b, α_1, and α_3 are $1 \times 10^{-20}, 0.687$, and 0.687, respectively.*

(R Core Team 2019) by the `dCount-conv-bi` function in the **Countr** package (Kharrat and Boshnakov 2018) for the gamma and Weibull count models, the `dcmp` function in the **COMPoissonReg** package (Sellers, Lotze and Raim, 2018) for the CMP count model, and the `Faddyprob.general` function in the **CountsEPPM** package (Smith and Faddy 2018) for the Faddy (1997) count model.

The limitation that the above regression models present only one dispersion type may be easily removed by allowing for non-monotonic rate sequences. The two examples are the single-unusual-event (SUE ($\boldsymbol{\gamma} = \{2\}$)) and double-unusual-event (DUE ($\boldsymbol{\gamma} = \{2,3\}$)) models shown in Figures 1(f) and 1(h). Their curves corresponding to $\alpha_2 \neq 1$ always cross the 45-degree (Poisson) lines. Thus, these models can fit a dataset that is a mixture of over-, under-, and equidispersion. The development and exploration of a new class of unusual-event (UE) models is the main objective of the present article. Note that the SUE ($\boldsymbol{\gamma} = \{0\}$) model is a special case of the Faddy (1997) (see Figures 1(d) and 1(e)), as described later.

The rest of this article is organized as follows. Section 2 presents the UE models and their properties, with additional details provided in Appendices A and B at the end of the paper. Section 3 discusses numerical strategies for computing UE probabilities. Section 4 provides and analyses the experimental results from the number of births and the number of bids. Finally, Section 5 concludes the paper.

2. Unusual-event models

Let $X(t)$ be a discrete random variable, representing the total number of events that occur before or at exactly time t. $\{X(t);\, t \geq 0\}$ is a pure birth process with $X(0) = 0$ and birth rates $\lambda_n\,(n \geq 0)$. The probabilities $P_n(t) = P\{X(t) = n \mid X(0) = 0\}$, for $n = 0,1,2,...$, satisfy the Chapman-Kolmogorov forward differential equations (Cox and Miller 1965), namely

$$\begin{aligned} P_0'(t) &= -\lambda_0 P_0(t), \\ P_n'(t) &= -\lambda_n P_n(t) + \lambda_{n-1} P_{n-1}(t), \quad n > 0, \end{aligned} \tag{1}$$

with boundary conditions $P_0(0) = 1$ and $P_n(0) = 0$, $n > 0$.

Different distributions correlate with different birth rate sequence λ_n patterns. The simple Poisson process, which has a constant rate parameter λ, restricts that the variance equals the mean. The birth rate, which depends on the number of events, may allow for overdispersion and underdispersion. Increasing the number of parameters in the process almost always improves the goodness of fit (as assessed by the log-likelihood function), but it may cause overfitting. Thus, the rate λ_n must be a parametric function of n, as stated by Faddy and Smith (2008). Examples of pure birth processes follow below.

1. A sequence of rates

   $$\lambda_n = \lambda, \quad \text{for } n = 0,1,2,..., \quad \lambda > 0,$$

   exhibits the Poisson distribution, which is a one-parameter count model.

2. A sequence of rates

$$\lambda_n = \begin{cases} \lambda & \text{for } n > 0 \\ \lambda_0 & \text{for } n = 0, \quad \lambda \text{ and } \lambda_0 > 0, \end{cases}$$

exhibits the Faddy (1994) distribution, which is a two-parameter count model.

3. A sequence of rates

$$\lambda_n = \begin{cases} \lambda & \text{for } n > 1 \\ \lambda_0 & \text{for } n = 0 \\ \lambda_1 & \text{for } n = 1, \quad \lambda,\ \lambda_0, \text{ and } \lambda_1 > 0, \end{cases}$$

exhibits the extended Faddy (1994) distribution, which is a three-parameter count model.

4. A sequence of rates

$$\lambda_n = \lambda\,(b+n)^{\alpha}, \quad \text{for } n = 0,1,2,..., \quad \lambda > 0,\ b > 0, \text{ and } \alpha \leq 1,$$

exhibits the Faddy (1997) distribution, which is a three-parameter count model.

The Faddy (1994), extended Faddy (1994), and Faddy (1997) models have greater flexibility than the Poisson model at the cost of additional parameters. Covariates can be incorporated into these models by setting λ as a function of the linear predictor $\beta_0 + \beta_1 x_{j1} + ... + \beta_r x_{jr}$, where x_{jk}, $k = 1,...,r$, is the jth observation of the kth covariate, and β_l, $l = 0,...,r$, is the lth unknown parameter to be estimated. The rates $\lambda_n\,(n \geq 0)$ of the Poisson and Faddy (1997) distributions depend on the covariates, but the rates λ_0 and λ_1 of the Faddy (1994) and extended Faddy (1994) do not. One might argue that λ_0 and λ_1 can be written as a function of the linear predictor. However, the approximately doubled (Faddy (1994)) and tripled (extended Faddy (1994)) parameters comparing to the above two distributions may lead to overfitting. Perhaps the rate sequences of the Faddy (1994) and extended Faddy (1994) can be easily modified as follows:

$$\lambda_n = \begin{cases} \lambda & \text{for } n > 0 \\ \alpha_0\lambda & \text{for } n = 0, \quad \alpha_0 \text{ and } \lambda > 0, \end{cases}$$

and

$$\lambda_n = \begin{cases} \lambda & \text{for } n > 1 \\ \alpha_0\lambda & \text{for } n = 0 \\ \alpha_1\lambda & \text{for } n = 1, \quad \alpha_0,\ \alpha_1, \text{ and } \lambda > 0. \end{cases}$$

We call λ the base rate. These modified rate sequences can avoid the risk of overfitting, and the rates λ_n depend on covariates. In other words, these distributions with the fewest numbers of parameters occur when λ is a function of the linear predictor.

We call the pure birth process with this pattern of the rate sequences the unusual-event (UE) process because at least one rate differs from the base rate λ. It is defined as

$$\lambda_n = \begin{cases} \lambda & \text{for } n \notin \boldsymbol{\gamma} = \{\gamma_1, \gamma_2, \ldots, \gamma_m\} \\ \alpha_{\gamma_i}\lambda & \text{for } n \in \boldsymbol{\gamma}, \end{cases} \tag{2}$$

where γ_i is a non-negative integer, and $\alpha_{\gamma_i} > 0$, $i = 1, 2, \ldots, m$. We call α_{γ_i} the shape parameter. The UE process permits a wide range of regression models for count data, including the combinations of distributions with either one or three dispersion types. These possibilities are illustrated using the single-unusual-event (SUE) and double-unusual-event (DUE) processes.

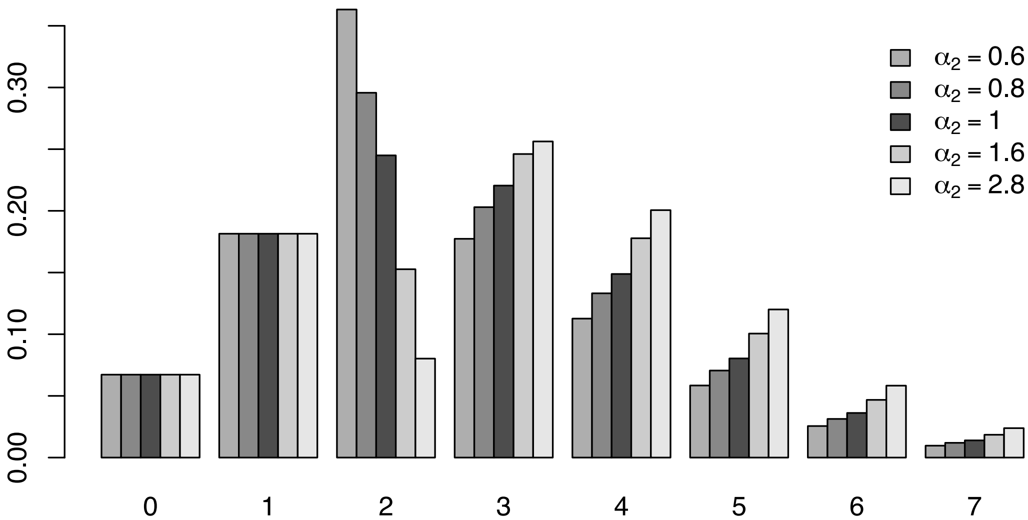


Figure 2. *SUE ($\boldsymbol{\gamma} = \{2\}$ and $\lambda = 2.7$) distributions with unequal means and dispersions.*

2.1. SUE models

Perhaps the simplest example of UE processes is a SUE process with

$$\lambda_n = \begin{cases} \lambda & \text{for } n \notin \boldsymbol{\gamma} = \{\gamma\} \\ \alpha_{\gamma}\lambda & \text{for } n = \gamma. \end{cases} \tag{3}$$

For $\alpha_\gamma = 1$, the SUE process simplifies to the Poisson process. It is noted that the SUE distribution is characterized by independently exponential distributed interarrival times. Figure 2 compares the probability functions of the SUE ($\boldsymbol{\gamma} = \{2\}$ and $\lambda = 2.7$) distribution for five values of α_2. It is more concentrated ($\alpha_2 < 1$) or more dispersed ($\alpha_2 > 1$) than the Poisson distribution ($\alpha_2 = 1$). The overdispersion case $\alpha_2 = 2.8$ shows a probability distribution with two distinct modes (1 and 3) referred to as a bimodal distribution.

The SUE ($\boldsymbol{\gamma} = \{2\}$) probability function is given by

$$P_n(t) = \begin{cases} \dfrac{(\lambda t)^n e^{-\lambda t}}{n!} & \text{for } n < 2 \\ (\lambda t)^n e^{-\lambda t} \displaystyle\sum_{i=0}^{\infty} \frac{((1-\alpha_2)\lambda t)^i}{(i+n)!} & \text{for } n = 2 \\ \alpha_2 (\lambda t)^n e^{-\lambda t} \displaystyle\sum_{i=0}^{\infty} \frac{((1-\alpha_2)\lambda t)^i}{(i+n)!} & \text{for } n > 2. \end{cases} \tag{4}$$

The derivations of SUE probability distributions can be found in Appendix A. For $n < 2$, the SUE ($\boldsymbol{\gamma} = \{2\}$) probability function is Poisson. It is not a function of α_2, and thus the probabilities in Figure 2 are equal at $n = 0$ and 1. For $n = 2$, the SUE ($\boldsymbol{\gamma} = \{2\}$) probability function can be simplified to $\frac{(\lambda t)^2 e^{-\lambda t}}{2!} + (1-\alpha_2)(\lambda t)^3 e^{-\lambda t} \sum_{i=0}^{\infty} \frac{((1-\alpha_2)\lambda t)^i}{(i+3)!}$. The first term is the Poisson probability function for $n = 2$. Since the second term is positive ($\alpha_2 < 1$) and negative ($\alpha_2 > 1$), the SUE ($\boldsymbol{\gamma} = \{2\}$) probability value is greater and smaller than the Poisson ($\alpha_2 = 1$) , respectively.

The Faddy (1994) process is equivalent to the SUE ($\boldsymbol{\gamma} = \{0\}$) process, but only when their regression models are not considered. Therefore, the proof of the Faddy (1994) by direct calculation that for $t > 0$,

$$Var\{X(t)\} > E\{X(t)\} \quad \text{if} \quad \frac{\lambda_0}{\lambda} = \alpha_0 < 1$$

and

$$Var\{X(t)\} < E\{X(t)\} \quad \text{if} \quad \frac{\lambda_0}{\lambda} = \alpha_0 > 1,$$

is still correct for the SUE ($\boldsymbol{\gamma} = \{0\}$) process. Alternatively, $\alpha_0 < 1$ and $\alpha_0 > 1$ result in non-decreasing and non-increasing rate sequences, which provide overdispersed and underdispersed SUE ($\boldsymbol{\gamma} = \{0\}$) distributions, respectively (see Figure 4(a)). These properties were conjectured by Faddy (1994) and proved by Ball (1995). Note that the non-increasing and non-decreasing rate sequences mean $\lambda_{n+1} \leq \lambda_n$ and $\lambda_{n+1} \geq \lambda_n$, respectively.

It is worth mentioning that the SUE ($\boldsymbol{\gamma} = \{0\}$) model is a special case of the Faddy (1997) model (see Figures 1(d) and 1(e)). Let us consider a rate sequence of the Faddy (1997) model in which the parameter b is given in the form $\sigma^{\frac{1}{|\alpha|}}$, where $0 < \sigma \leq 1$. It can be shown that

$$\lim_{\alpha \to 0^-} \lambda_n = \lim_{\alpha \to 0^-} \lambda \left(\sigma^{\frac{1}{|\alpha|}} + n\right)^{\alpha} = \begin{cases} \lambda & \text{for } n > 0 \\ \frac{\lambda}{\sigma} = \alpha_0 \lambda & \text{for } n = 0,\ \alpha_0 = \frac{1}{\sigma} \geq 1, \end{cases}$$

and

$$\lim_{\alpha \to 0^+} \lambda_n = \lim_{\alpha \to 0^+} \lambda \left(\sigma^{\frac{1}{|\alpha|}} + n\right)^{\alpha} = \begin{cases} \lambda & \text{for } n > 0 \\ \sigma \lambda = \alpha_0 \lambda & \text{for } n = 0,\ 0 < \alpha_0 = \sigma \leq 1. \end{cases}$$

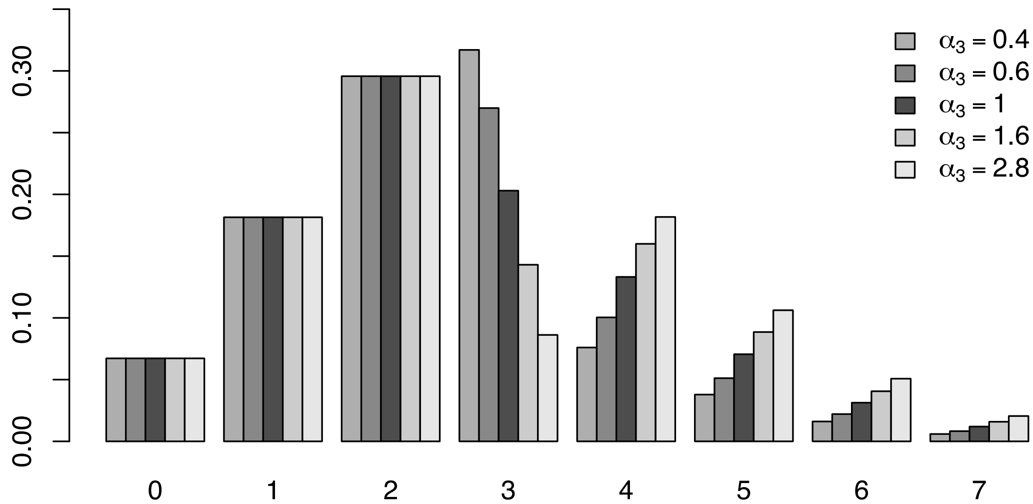


Figure 3. *DUE* $(\boldsymbol{\gamma} = \{2,3\},\ \lambda = 2.7,$ *and* $\alpha_2 = 0.8)$ *distributions with unequal means and dispersions.*

Hence, the SUE ($\boldsymbol{\gamma} = \{0\}$) models with $\alpha_0 \geq 1$ and $0 < \alpha_0 \leq 1$ are the limiting cases of the Faddy (1997) model. They arise when α approaches 0 from the left and the right, respectively.

Figures 4(a)-4(d) show graphs of the variance-mean ratio with $\gamma = 0 - 3$ for various values of λ and α_γ. $\alpha_\gamma \neq 1$, $\gamma > 0$, results in a non-monotonic rate sequence, which provides over-, under-, and equidispersed SUE ($\boldsymbol{\gamma} = \{\gamma > 0\}$) distributions (see Figures 4(b)-4(d)). For fixed α_γ, the three dispersion types of the SUE models are defined by the base rate λ, in contrast to the gamma, Weibull, CMP, and Faddy (1997) models. We conjecture that this property holds for any non-monotonic rate sequence of SUE processes.

2.2. DUE models

Perhaps the simplest example of DUE processes is a pure birth process with

$$\lambda_n = \begin{cases} \lambda & \text{for } n \notin \boldsymbol{\gamma} = \{\gamma,\ \gamma+1\} \\ \alpha_\gamma \lambda & \text{for } n = \gamma \\ \alpha_{\gamma+1}\lambda & \text{for } n = \gamma+1. \end{cases} \tag{5}$$

For $\alpha_\gamma \neq 1$ and $\alpha_{\gamma+1} = 1$, the DUE $(\boldsymbol{\gamma} = \{\gamma, \gamma+1\})$ process simplifies to the SUE $(\boldsymbol{\gamma} = \{\gamma\})$ process. We consider the DUE count model in which the rates, λ_γ and $\lambda_{\gamma+1}$, of two consecutive events are not equal to the base rate. This phenomenon appears to occur in the two empirical applications, the number of births and the number of bids, shown in Tables 4 and 5. Figure 3 compares the probability functions of the DUE $(\boldsymbol{\gamma} = \{2,3\},\ \lambda = 2.7,$ and $\alpha_2 = 0.8)$ distribution for five values of α_3. It is more concentrated $(\alpha_3 < 1)$

or more dispersed ($\alpha_3 > 1$) than the SUE distribution ($\alpha_3 = 1$). The DUE ($\alpha_3 = 2.8$) distribution is a bimodal distribution whose modes are 2 and 4. The bimodal distributions of the SUE ($\boldsymbol{\gamma} = \{2\}$, $\lambda = 2.7$, and $\alpha_2 = 2.8$) and the DUE ($\boldsymbol{\gamma} = \{2,3\}$, $\lambda = 2.7$, $\alpha_2 = 0.8$, and $\alpha_3 = 2.8$) suggest that we should expect a UE distribution to be a multimodal distribution, which is a discrete probability distribution with two or more modes.

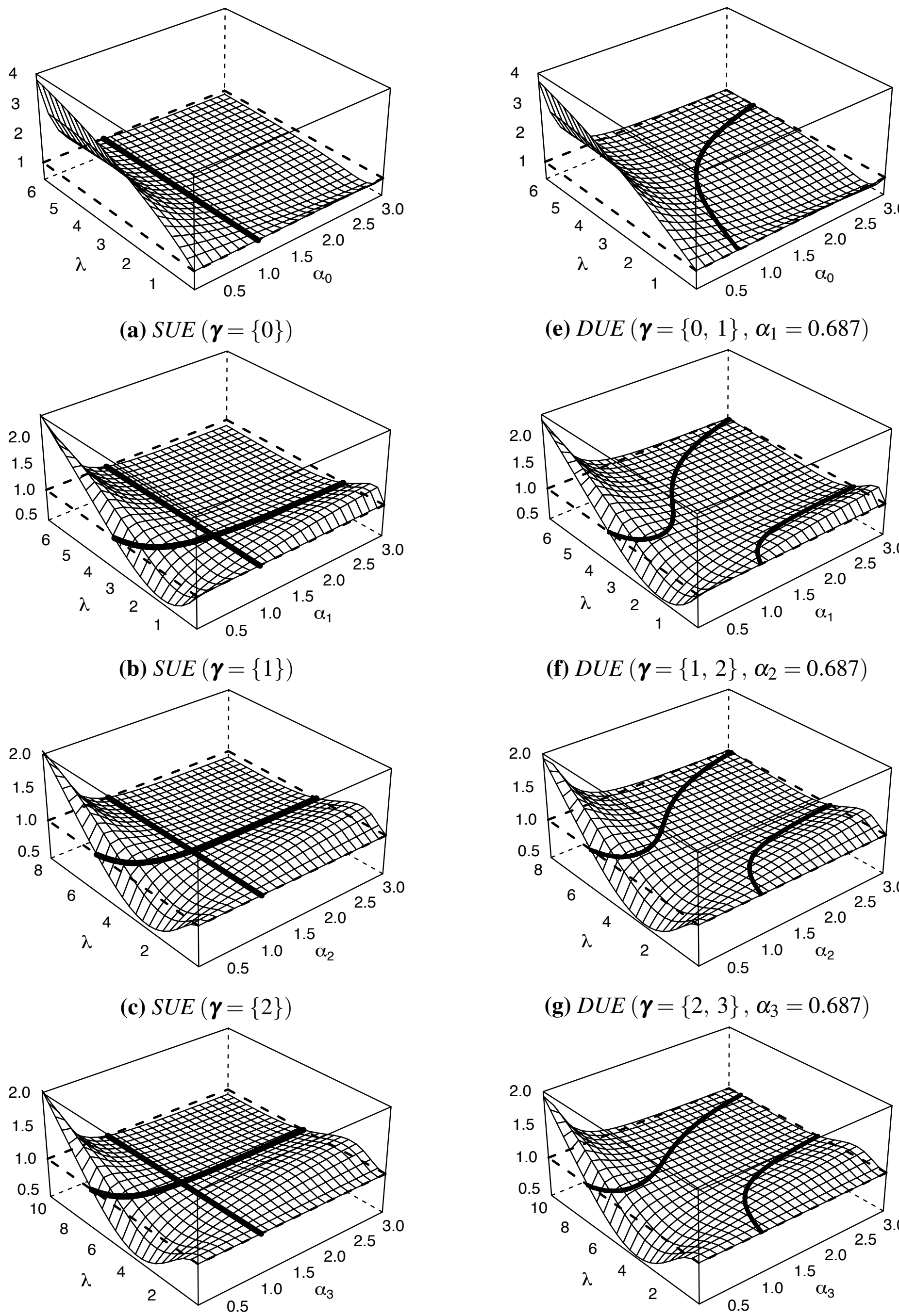
The DUE ($\boldsymbol{\gamma} = \{2,3\}$) probability function is given by

$$P_n(t) = \begin{cases} \dfrac{(\lambda t)^n e^{-\lambda t}}{n!} & \text{for } n < 2 \\ (\lambda t)^n e^{-\lambda t} \sum\limits_{i=0}^{\infty} \dfrac{((1-\alpha_2)\lambda t)^i}{(i+n)!} & \text{for } n = 2 \\ \alpha_2 (\lambda t)^n e^{-\lambda t} \sum\limits_{i=0}^{\infty} \dfrac{c_i (\lambda t)^i}{(i+n)!} & \text{for } n = 3 \\ \alpha_2 \alpha_3 (\lambda t)^n e^{-\lambda t} \sum\limits_{i=0}^{\infty} \dfrac{c_i (\lambda t)^i}{(i+n)!} & \text{for } n > 3, \end{cases} \tag{6}$$

where $c_i = \sum_{k=0}^{i} (1-\alpha_2)^k (1-\alpha_3)^{i-k}$. The derivations of DUE ($\boldsymbol{\gamma} = \{\gamma, \gamma+1\}$) probability distributions can be found in Appendix B. For $n < 3$, the DUE ($\boldsymbol{\gamma} = \{2,3\}$) probability functions are not dependent on α_3, and thus the probabilities in Figure 3 are equal at $n = 0$, 1, and 2. For $n = 3$, the DUE ($\boldsymbol{\gamma} = \{2,3\}$) probability function can be simplified to $\alpha_2 (\lambda t)^3 e^{-\lambda t} \sum_{i=0}^{\infty} \frac{((1-\alpha_2)\lambda t)^i}{(i+3)!} + \alpha_2 (1-\alpha_3)(\lambda t)^4 e^{-\lambda t} \sum_{i=0}^{\infty} \frac{c_i (\lambda t)^i}{(i+4)!}$. The first term is the SUE ($\boldsymbol{\gamma} = \{2\}$) probability function for $n = 3$. Since the second term is positive ($\alpha_3 < 1$) and negative ($\alpha_3 > 1$), the DUE ($\boldsymbol{\gamma} = \{2,3\}$) probability value is greater and smaller than the SUE ($\boldsymbol{\gamma} = \{2\}$), respectively.

Figures 4(e)-4(h) show graphs of the variance-mean ratio with $\gamma = 0-3$ for various values of λ and α_γ. A non-decreasing rate sequence with $\alpha_0 \leq \alpha_1 < 1$ provides only overdispersed DUE ($\boldsymbol{\gamma} = \{0,1\}$) distribution, and a non-monotonic rate sequence with $\alpha_0 > \alpha_1 < 1$ produces over-, under-, and equidispersed DUE ($\boldsymbol{\gamma} = \{0,1\}$) distribution (see Figure 4(e)). $\boldsymbol{\gamma} \neq \{0,1\}$ results in a non-monotonic rate sequence, which provides over-, under-, and equidispersed DUE ($\boldsymbol{\gamma} = \{\gamma, \gamma+1\}$) distributions (see Figures 4(f)-4(h)). For fixed α_γ and $\alpha_{\gamma+1}$, the three dispersion types of the DUE models are defined by the base rate λ, in contrast to the gamma, Weibull, CMP, and Faddy (1997) models. We conjecture that this property holds for any non-monotonic rate sequence of DUE processes and also UE processes.

We conclude that even the simplest generalizations of the Poisson and SUE processes, the SUE and the DUE, are relatively flexible models for count data.



(a) *SUE* ($\boldsymbol{\gamma} = \{0\}$)

(e) *DUE* ($\boldsymbol{\gamma} = \{0,\ 1\}$, $\alpha_1 = 0.687$)

(b) *SUE* ($\boldsymbol{\gamma} = \{1\}$)

(f) *DUE* ($\boldsymbol{\gamma} = \{1,\ 2\}$, $\alpha_2 = 0.687$)

(c) *SUE* ($\boldsymbol{\gamma} = \{2\}$)

(g) *DUE* ($\boldsymbol{\gamma} = \{2,\ 3\}$, $\alpha_3 = 0.687$)

(d) *SUE* ($\boldsymbol{\gamma} = \{3\}$)

(h) *DUE* ($\boldsymbol{\gamma} = \{3,\ 4\}$, $\alpha_4 = 0.687$)

Figure 4. *Variance-mean ratios for SUE and DUE count models with $0 < \alpha_\gamma < 3$. Each SUE surface ($\gamma > 0$) contains a saddle point, which is the intersection of the straight and curved lines.*

3. Computation of UE Probabilities

The solution of the Chapman-Kolmogorov forward differential equation (1) can be written in terms of a matrix-exponential function (Cox and Miller 1965)

$$(P_0(t)\, P_1(t)\, ...\, P_n(t)) = (1\, 0\, ...\, 0)\, exp(\boldsymbol{Q}t), \tag{7}$$

where $\boldsymbol{Q}$ is the matrix of birth rates

$$\boldsymbol{Q} = \begin{pmatrix} -\lambda_0 & \lambda_0 & 0 & \cdots & 0 \\ 0 & -\lambda_1 & \lambda_1 & \cdots & 0 \\ 0 & 0 & -\lambda_2 & \ddots & 0 \\ \vdots & \vdots & \vdots & \ddots & \lambda_{n-1} \\ 0 & 0 & 0 & \cdots & -\lambda_n \end{pmatrix},$$

and an integral function (Bartlett 1978)

$$\begin{aligned} P_0(t) &= e^{-\lambda_0 t}, \\ P_n(t) &= \int_0^t \lambda_{n-1} P_{n-1}(u) e^{-\lambda_n (t-u)} du \quad \text{for } n > 0. \end{aligned} \tag{8}$$

The matrix exponentiation is the most common for computing the probabilities of pure birth processes. Researchers usually rely on this method (e.g. Faddy and Smith 2011 and Smith and Faddy 2016), perhaps because various packages for calculating the matrix exponential have been developed and made the routines available as described by Faddy and Smith (2008). The analytic solution is obtained by the integral function. It is computationally intractable (Faddy 1997; Crawford, Ho and Suchard, 2018) because there is an extremely ill-conditioned problem in the solution. A numerical solution may differ significantly from the exact solution. Therefore, the analytic solution is not appropriate for numerical computation (Podlich et al. 2004). However, this ill-conditioned problem can be solved by a Taylor series expansion. In this research, we will report the computational results from the analytic solution that was previously thought to be infeasible. The fertility and takeover bids datasets are considered in this paper. Their results are obtained by using the matrix-exponential and analytic solution approaches. We confirm here that the results from these two methods are identical.

The probability $P_n(t)$ in Equation (8) can be described as a convolution of two functions. First, the probability density function $\lambda_{n-1} P_{n-1}(u)$ is the n-fold convolution of the exponential density functions of the interarrival times between events. It presents the probability that the nth event occurs at exactly time u. Second, $e^{-\lambda_n (t-u)}$ is the survival function of the interarrival times between the nth and $(n+1)$th event. The survival function denotes the probability that the $(n+1)$th event does not occur after time u and before or at exact time t. Using Equation (8) and letting $\lambda_n = \alpha_n \lambda$, the first few probabilities of the UE count models are obtained:

$$
\begin{aligned}
P_0(t) &= e^{-\alpha_0\lambda t}\\
P_1(t) &= \alpha_0 e^{-\alpha_1\lambda t}\left(\frac{e^{(\alpha_1-\alpha_0)\lambda t}-1}{\alpha_1-\alpha_0}\right)\\
P_2(t) &= \alpha_0\alpha_1 e^{-\alpha_2\lambda t}\left(\frac{(e^{(\alpha_2-\alpha_0)\lambda t}-1)}{(\alpha_1-\alpha_0)(\alpha_2-\alpha_0)}+\frac{(e^{(\alpha_2-\alpha_1)\lambda t}-1)}{(\alpha_0-\alpha_1)(\alpha_2-\alpha_1)}\right)\\
P_3(t) &= \alpha_0\alpha_1\alpha_2 e^{-\alpha_3\lambda t}\left(\frac{(e^{(\alpha_3-\alpha_0)\lambda t}-1)}{(\alpha_1-\alpha_0)(\alpha_2-\alpha_0)(\alpha_3-\alpha_0)}\right.\\
&\qquad\left.+\frac{(e^{(\alpha_3-\alpha_1)\lambda t}-1)}{(\alpha_0-\alpha_1)(\alpha_2-\alpha_1)(\alpha_3-\alpha_1)}+\frac{(e^{(\alpha_3-\alpha_2)\lambda t}-1)}{(\alpha_0-\alpha_2)(\alpha_1-\alpha_2)(\alpha_3-\alpha_2)}\right)
\end{aligned}
$$

From these equations for $P_0(t)$, $P_1(t)$, $P_2(t)$, and $P_3(t)$, one can deduce that the general UE probability function might be of the form

$$
P_n(t) = \begin{cases} e^{-\alpha_0\lambda t} & \text{for } n=0\\ \left(\prod_{i=0}^{n-1}\alpha_i\right) e^{-\alpha_n\lambda t}\sum_{i=0}^{n-1}\frac{(e^{(\alpha_n-\alpha_i)\lambda t}-1)}{\prod_{j=0,j\neq i}^{n}(\alpha_j-\alpha_i)} & \text{for } n>0, \end{cases} \tag{9}
$$

and this expression is similar to Bartlett (1978, eq. (9), p. 55) and Crawford et al. (2018, eq. (55), p. 13). Inserting the Taylor series expansion of $e^{(\alpha_n-\alpha_i)\lambda t}$, the UE probability distribution can be rewritten as follow:

$$
P_n(t) = \begin{cases} e^{-\alpha_0\lambda t} & \text{for } n=0\\ \left(\prod_{i=0}^{n-1}\alpha_i\right) e^{-\lambda t}\sum_{i=0}^{\infty}c_i\frac{(\lambda t)^{i+n}}{(i+n)!} & \text{for } n>0, \end{cases} \tag{10}
$$

where

$$
c_i = \begin{cases} 1 & \text{for } i=0\\ \sum_{k_i=0}^{n}\sum_{k_{i-1}=0}^{k_i}\dots\sum_{k_1=0}^{k_2}\prod_{j=k_1}^{k_i}(1-\alpha_j) & \text{for } i>0. \end{cases}
$$

This expression can also be obtained from Equation (7) by letting $\boldsymbol{Q}=\lambda(\boldsymbol{P}-\boldsymbol{I})$, where

$$
\boldsymbol{P} = \begin{pmatrix} 1-\alpha_0 & \alpha_0 & 0 & \cdots & 0\\ 0 & 1-\alpha_1 & \alpha_1 & \cdots & 0\\ 0 & 0 & 1-\alpha_2 & \ddots & 0\\ \vdots & \vdots & \vdots & \ddots & \alpha_{n-1}\\ 0 & 0 & 0 & \cdots & 1-\alpha_n \end{pmatrix},
$$

and $\boldsymbol{I}$ denotes the identity matrix. If $\alpha_i \leq 1$ for $i = 0, \ldots, n$, this procedure is known as uniformization, originally introduced by Jensen (1953). Another method for obtaining Equation (10) is to use continued fractions (see Parthasarathy and Sudhesh 2006). However, the more details for computing $P_n(t)$ are intricate and cannot be discussed adequately here.

4. Experimental Results

This section shows results from two applications. The fertility data were analysed by Winkelmann (1995) and re-analysed by McShane et al. (2008), Chanialidis et al. (2018), and Kharrat et al. (2019). The takeover bids data were analysed by Jaggia and Thosar (1993) and re-analysed by Cameron and Johansson (1997), Saez-Castillo and Conde-Sanchez (2013), and Smith and Faddy (2016). For more information, the readers are referred to Winkelmann (1995) for the fertility data and Cameron and Johansson (1997) for the takeover bids data.

Experimental results obtained from the Poisson, gamma, Weibull, CMP, and Faddy (mean only) models are computed using the **stats** (R Core Team 2019), **Countr** (gamma and Weibull) (Kharrat and Boshnakov 2018), **COMPoissonReg** (Sellers et al. 2018), and **CountsEPPM** (Smith and Faddy 2018) **R** packages. The fertility and takeover bids datasets are available from the **Countr** and **mpcmp** (Fung et al. 2019) **R** packages, respectively. The UE models are implemented in **R** (R Core Team 2019) and C++. Most of the code is written in C++ via the **Rcpp** (Eddelbuettel et al. 2021) package in order to accelerate computations. The **expm** (Goulet et al. 2020) **R** package enables computation of the matrix exponential for calculating the probabilities of UE processes.

Table 1. *Shape parameter (α_γ), log-likelihood, BIC, and computation time (seconds) values of several SUE models for the fertility and takeover bids data.*

	Fertility data				Takeover bids data			
Model	α_γ	-Log-L	BIC	Time	α_γ	-Log-L	BIC	Time
Poisson	-	2101.8	4282.0	0.02	-	185.0	418.3	0.00
SUE ($\boldsymbol{\gamma} = \{0\}$)	1.46	2078.1	4241.8	0.59	2.96	171.3	395.8	0.05
SUE ($\boldsymbol{\gamma} = \{1\}$)	1.23	2096.0	4277.5	0.57	0.40	174.1	401.4	0.07
SUE ($\boldsymbol{\gamma} = \{2\}$)	0.52	2048.8	4183.1	0.66	1.00	185.0	423.1	0.00
SUE ($\boldsymbol{\gamma} = \{3\}$)	1.00	2101.8	4289.1	0.06	1.00	185.0	423.1	0.00
SUE ($\boldsymbol{\gamma} = \{4\}$)	1.00	2101.8	4289.1	0.05	1.00	185.0	423.1	0.00
SUE ($\boldsymbol{\gamma} = \{5\}$)	1.00	2101.8	4289.1	0.04	1.00	185.0	423.1	0.00

The rate between two consecutive events different from the base rate causes an excess (or a lack) of counts relative to a benchmark model such as the Poisson. This unusual event might be investigated by comparing the histogram of the sample and Poisson

Table 2. *Highest log-likelihood, lowest BIC, and computation time (seconds) values of each k-combination ($1 \leq k \leq 7$) UE regression model.*

Fertility data				Takeover bids data			
$\boldsymbol{\gamma}$	-Log-L	BIC	Time	$\boldsymbol{\gamma}$	-Log-L	BIC	Time
$\{2\}$	2048.8	4183.1	1.6	$\{0\}$	171.3	395.8	0.2
$\{2,3\}$	2040.1	4172.9	10.6	$\{1,2\}$	168.0	394.1	1.2
$\{2,3,4\}$	2037.3	4174.4	31.4	$\{1,2,3\}$	166.9	396.6	3.1
$\{2,3,4,7\}$	2035.8	4178.5	54.6	$\{1,2,3,4\}$	166.2	400.1	4.6
$\{1,2,3,4,7\}$	2034.7	4183.5	55.5	$\{1,2,3,4,5\}$	165.8	404.1	4.3
$\{1,2,3,4,6,7\}$	2034.6	4190.3	31.5	$\{1,2,3,4,5,6\}$	165.8	408.9	2.3
$\{0,1,2,3,4,6,7\}$	2034.3	4196.8	10.2	$\{1,2,3,4,5,6,7\}$	165.8	413.7	0.8

distributions. For example, Figure 5(a) contains an excess of two counts. The "excess two" phenomenon may arise in the situation that is the rate between the second and third events is less than others. In other words, the third event is unusual, and the SUE ($\boldsymbol{\gamma} = \{2\}$ and $\alpha_2 < 1$) model is preferred over other SUE models. The results in Table 1 show that the SUE ($\boldsymbol{\gamma} = \{2\}$) model has a higher log-likelihood and lower BIC values than other models. Therefore, we can conclude that the third event is the unusual event of the fertility data. Similarly, in Figure 5(b), this approach can be applied to the takeover bids data.

Visualizing the histograms can be a method for guessing unusual events, but it is hard to conclude which UE model is the best. Therefore, an exhaustive search is utilized for finding the best UE model because the number of UE models is limited. It is simple and guaranteed to find the best solution. We assume that λ_n's ($n > 7$) are equal to the base rate. For $\boldsymbol{\gamma} = \{0,1,...,7\}$, there are 255 different UE models to choose from using the combinations of all eight unusual events, 8 models (one and seven unusual events), 28 models (two and six unusual events), 56 models (three and five unusual events), and 70 models (four unusual events). Table 2 summarizes the highest log-likelihood, lowest BIC, and computation time values of the k-combination ($1 \leq k \leq 7$) models by fitting the UE regression models to the fertility and takeover bids data. For both datasets, as k increases, the log-likelihood increases monotonically. The BIC attains minimum at $k = 2$, and the DUE model is selected as the best model.

The fertility data, which consists of 10 covariates, are very slightly underdispersed with the variance-mean ratio equalling 2.328/2.384 = 0.977. The Poisson regression model is inappropriate because the mixture of conditional equidispersed distributions is always overdispersed. The gamma, Weibull, CMP, and Faddy (mean only) models display underdispersion. These regression models perhaps provide a good fit for the data because the mixture of conditional underdispersed distributions can be over-, under-, or equidispersion. The SUE ($\boldsymbol{\gamma} = \{2\}$) provides a much better fit to the data than the other

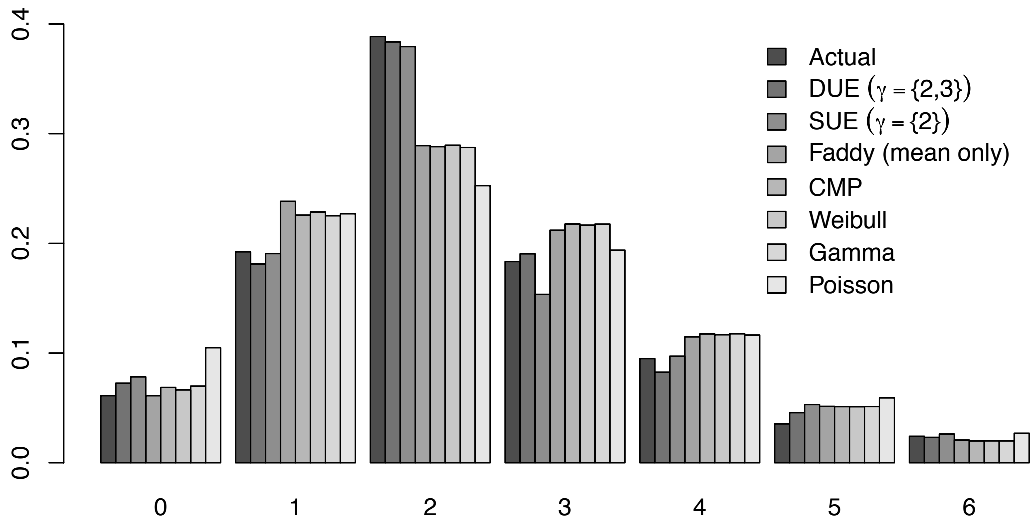
Table 3. *Variance, mean, variance-mean ratio, and BIC values of several regression models for the fertility and takeover bids data. The SUE ($\boldsymbol{\gamma} = \{2\}$) and DUE ($\boldsymbol{\gamma} = \{2,3\}$) models fit to the fertility data. The SUE ($\boldsymbol{\gamma} = \{0\}$) and DUE ($\boldsymbol{\gamma} = \{1,2\}$) models fit to the takeover bids data.*

	Fertility data				Takeover bids data			
Model	Variance	Mean	Ratio	BIC	Variance	Mean	Ratio	BIC
Sample	2.328	2.384	0.977	-	2.035	1.738	1.171	-
Poisson	2.742	2.382	1.151	4281.98	2.227	1.737	1.282	418.26
Gamma	2.175	2.383	0.913	4241.96	1.710	1.736	0.985	413.94
Weibull	2.157	2.383	0.905	4239.55	1.635	1.735	0.943	413.61
CMP	2.166	2.384	0.909	4241.25	1.657	1.738	0.954	413.92
Faddy	2.190	2.387	0.917	4244.23	1.463	1.727	0.847	401.64
SUE	2.512	2.386	1.053	4183.05	1.468	1.727	0.850	395.82
DUE	2.332	2.376	0.981	4172.87	2.142	1.740	1.231	394.12

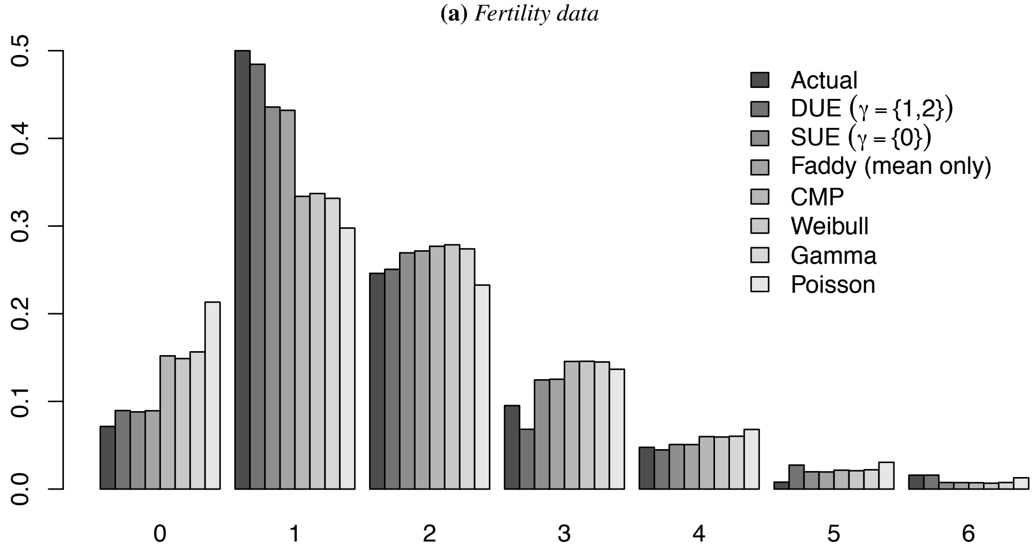
models, excluding the DUE, although its variance-mean ratio disagrees with the actual data (see Table 3). It means that the shape of the fertility data distribution resembles the SUE ($\boldsymbol{\gamma} = \{2\}$) more than the other models (see Figure 5(a)). However, the DUE ($\boldsymbol{\gamma} = \{2,3\}$) provides the best fit in terms of BIC to the data. The log-likelihood value of -2040.12 for this model with 13 parameters is much greater than -2048.77 from the SUE ($\boldsymbol{\gamma} = \{2\}$) model with 12 parameters. Because of the one additional parameter associated with a substantial increase in log-likelihood, the BIC value of 4172.87 is smaller than the SUE ($\boldsymbol{\gamma} = \{2\}$). Note that the fertility data distribution may be the combination of over-, under-, and equidispersed distributions, as described later.

For the takeover bids data, the variance-mean ratio is 2.035/1.738 = 1.171. Therefore, the data present overdispersion. The Poisson provides the worst fit in terms of BIC to the data even though it presents overdispersion as the data do (see Table 3). It interprets that the shape of the takeover bids data distribution resembles the Poisson less than the other models (see Figure 5(b)). The DUE ($\boldsymbol{\gamma} = \{1,2\}$) provides the best fit in terms of BIC to the data, and its variance-mean ratio agrees with the actual data (see Table 3). Note that the takeover bids data distribution may be the combination of over-, under-, and equidispersed distributions, as described later.

Figure 5 presents the sample and predicted probabilities evaluated at individual covariates for the Poisson, gamma, Weibull, CMP, Faddy (mean only), SUE, and DUE models. The fertility and takeover bits datasets contain an excess of two and one outcomes, respectively. It means there are more twos and ones in the two datasets than predicted by the Poisson, the gamma, etc. Figure 5(a) reveals that the models, excluding the SUE ($\boldsymbol{\gamma} = \{2\}$) and DUE ($\boldsymbol{\gamma} = \{2,3\}$) models, greatly underpredict the two outcomes because the third event is unusual. The SUE ($\boldsymbol{\gamma} = \{2\}$) model has the rate between the second and third event differs from others. However, the SUE ($\boldsymbol{\gamma} = \{2\}$) underpredicts the three outcomes because the fourth event is unusual. The DUE ($\boldsymbol{\gamma} = \{2,3\}$) has



(a) *Fertility data*



(b) *Takeover bids data*

Figure 5. *Sample and predicted relative frequency distributions.*

the consecutive rates between the second and fourth event differ from others. Thus, it leads to a considerable improvement of the predicted probabilities in the fertility case. Figure 5(b) shows that the models, excluding the DUE ($\boldsymbol{\gamma} = \{1,2\}$), underpredict the one outcome because the second and third events are unusual. The rate sequences of the SUE ($\boldsymbol{\gamma} = \{0\}$) and DUE ($\boldsymbol{\gamma} = \{1,2\}$) models are $2.962\lambda, \lambda, \lambda, \lambda, \lambda...$ and $\lambda, 0.314\lambda, 0.378\lambda, \lambda, \lambda...$, respectively. The unusual events of the SUE ($\boldsymbol{\gamma} = \{0\}$) and DUE ($\boldsymbol{\gamma} = \{1,2\}$) models disagree, but they have in common the fact that $\lambda_0 > \lambda_1$.

Tables 4 and 5 present the results from regressions for the number of children and number of bids data. The regression results from the gamma and Weibull models are produced by the "nlminb" function and the CMP, Faddy (mean only), SUE, and DUE models by the "optim" function with the "BFGS" method. The six models use the Poisson coefficients in Tables 4 and 5 as starting values of the unknown parameters $\beta_0, \beta_1, ..., \beta_r$, and the initial values of the other parameters are set to zero. The Poisson coefficients are perhaps the best initial guess for these models because the models generalize the Poisson. We note that the estimated parameters for the SUE and DUE regression models reported in Tables 4 and 5 are obtained by using the analytic solution approach. These values are identical to that produced by the matrix-exponential method, thus verifying the accuracy of the Taylor series expansion approach.

Comparing α in Table 4, these values in the gamma, Weibull, and CMP regression models are respectively 1.439, 1.236, and 1.429, which exceed one considerably, so there is an indication of underdispersion. The Faddy (mean only) also displays underdispersion because $\alpha = -0.129$. These four regression models with fixed α exhibit only one of over-, under-, and equidispersion. In other words, the dispersion types of these regression models depend only on α but not on λ. The SUE ($\boldsymbol{\gamma} = \{2\}$ and $\alpha = 0.521$) regression model displays overdispersion ($\lambda > 3.67$), underdispersion ($\lambda < 3.67$), and equidispersion ($\lambda = 3.67$) (see Figure 4(c)). The DUE ($\boldsymbol{\gamma} = \{2,3\}$, $\alpha_2 = 0.503$, and $\alpha_3 = 0.687$) regression model displays overdispersion ($\lambda > 4.31$), underdispersion ($\lambda < 4.31$), and equidispersion ($\lambda = 4.31$) (see Figure 4(g)). The dispersion types of the SUE ($\boldsymbol{\gamma} = \{2\}$) and DUE ($\boldsymbol{\gamma} = \{2,3\}$) regression models depend on α_2, α_3, and λ. It shows the flexibility of the SUE and DUE regression models to allow for over-, under-, and equidispersion, although the shape parameters are fixed. This property does not appear in the gamma, Weibull, CMP, and Faddy (1997) count models.

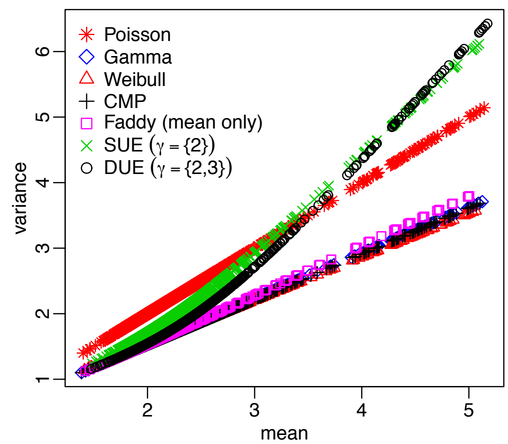
Figure 6 presents scatterplots of the fertility and takeover bids data. The dotted points are an ordered pair of the estimated mean and variance of each response variable produced by the seven models. The estimated mean and variance values of these models have to be determined numerically directly from their probability distributions using a suitable truncation (n). The points below and above the 45-degree (Poisson) line indicate underdispersion and overdispersion, respectively. In Figures 6(a) and 6(b), the SUE ($\boldsymbol{\gamma} = \{2\}$) curved (or the DUE ($\boldsymbol{\gamma} = \{2,3\}$) curved) and Poisson lines cut each other at a point, which is the estimated mean equals the estimated variance. The gamma, Weibull, CMP, and Faddy (mean only) lines are nearly coincident, indicating a similar ability of these four models to handle the fertility data. It is supported by the results in Table 4 that the log-likelihoods of these models are very similar. According to the SUE ($\boldsymbol{\gamma} = \{2\}$) and DUE ($\boldsymbol{\gamma} = \{2,3\}$) regression models, the fertility data are divided into two sets. The first set consists entirely of the underdispersed response variables, and the overdispersed response variables belong to the second. For the SUE ($\boldsymbol{\gamma} = \{2\}$), the first set (1151 members) is about 12.5 times bigger than the second set (92 members). For the DUE ($\boldsymbol{\gamma} = \{2,3\}$), the first set (1175 members) is about 17 times bigger than the second set (68 members). The gamma, Weibull, and CMP models in Figure 6(c) can

Table 4. *Regression model results for fertility data.*

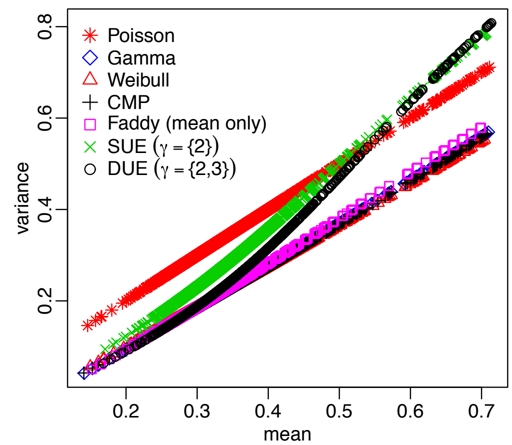
	Model													
	Poisson		Gamma		Weibull		CMP		Faddy (mean only)		SUE $\boldsymbol{\gamma}=\{2\}$		DUE $\boldsymbol{\gamma}=\{2,3\}$	
Variable	Coef	SE	Coef	SE	Coef	SE	Coef	SE	Coef	SE	Coef	SE	Coef	SE
Intercept	1.147	0.302	1.557	0.252	1.397	0.314	1.721	0.357	1.327	0.311	1.335	0.307	1.364	0.308
German	-0.200	0.072	-0.190	0.059	-0.223	0.072	-0.266	0.084	-0.197	0.073	-0.194	0.073	-0.209	0.073
Years of schooling	0.034	0.032	0.032	0.027	0.039	0.033	0.044	0.037	0.029	0.033	0.033	0.033	0.037	0.033
Vocational training	-0.153	0.044	-0.144	0.036	-0.173	0.044	-0.202	0.051	-0.177	0.044	-0.158	0.044	-0.166	0.044
University	-0.155	0.159	-0.146	0.130	-0.181	0.160	-0.207	0.182	-0.158	0.161	-0.136	0.162	-0.148	0.163
Catholic	0.218	0.071	0.206	0.058	0.242	0.070	0.289	0.082	0.241	0.071	0.212	0.071	0.221	0.072
Protestant	0.113	0.076	0.107	0.062	0.123	0.076	0.151	0.088	0.124	0.076	0.097	0.077	0.099	0.077
Muslim	0.548	0.085	0.523	0.070	0.639	0.087	0.742	0.103	0.624	0.087	0.547	0.087	0.572	0.087
Rural	0.059	0.038	0.055	0.031	0.068	0.038	0.078	0.044	0.067	0.038	0.062	0.039	0.070	0.039
Year of birth	0.002	0.002	0.002	0.002	0.002	0.002	0.003	0.003	0.000	0.002	0.001	0.002	0.001	0.002
Age at marriage	-0.030	0.007	-0.029	0.005	-0.034	0.006	-0.040	0.008	-0.035	0.007	-0.030	0.007	-0.031	0.007
ln α			0.364	0.049	0.212	0.027	0.357	0.047						
ln b									-2.317	1.600				
α									-0.129	0.064				
ln α_{γ}											-0.652	0.064	-0.687	0.064
ln $\alpha_{\gamma+1}$													-0.375	0.091
Log likelihood	-2101.80		-2078.23		-2077.02		-2077.88		-2075.80		-2048.77		-2040.12	
AIC (smaller is better)	4225.60		4180.45		4178.04		4179.75		4177.60		4121.54		4106.24	
BIC (smaller is better)	4281.98		4241.96		4239.55		4241.25		4244.23		4183.05		4172.87	
Time (seconds)	0.02		159.78		39.20		4.00		525.72		0.66		0.78	

Table 5. *Regression model results for takeover bids data.*

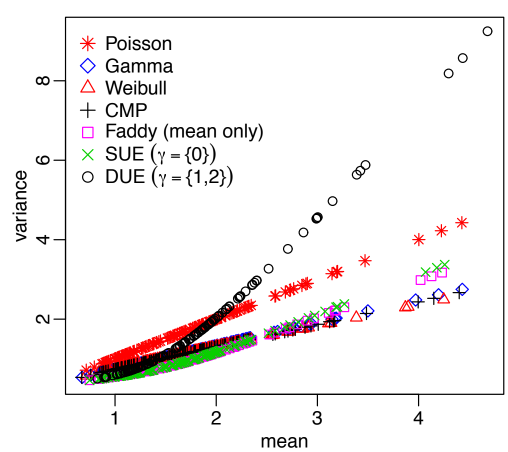
	Model													
	Poisson		Gamma		Weibull		CMP		Faddy (mean only)		SUE $\boldsymbol{\gamma}=\{0\}$		DUE $\boldsymbol{\gamma}=\{1,2\}$	
Variable	Coef	SE	Coef	SE	Coef	SE	Coef	SE	Coef	SE	Coef	SE	Coef	SE
Intercept	0.986	0.534	1.609	0.432	1.330	0.548	1.832	0.720	0.667	0.568	0.653	0.568	1.571	0.582
Leglrest	0.260	0.151	0.234	0.111	0.318	0.153	0.391	0.191	0.346	0.162	0.345	0.162	0.261	0.162
Rearest	-0.196	0.192	-0.168	0.142	-0.244	0.193	-0.307	0.240	-0.380	0.209	-0.385	0.209	-0.379	0.211
Finrest	0.074	0.217	0.072	0.160	0.043	0.217	0.113	0.268	0.015	0.229	0.016	0.229	0.000	0.231
Whtknght	0.481	0.159	0.430	0.117	0.568	0.162	0.710	0.210	0.664	0.175	0.664	0.176	0.572	0.172
Bidprem	-0.678	0.377	-0.616	0.278	-0.791	0.381	-1.009	0.477	-0.877	0.405	-0.876	0.406	-0.721	0.406
Insthold	-0.362	0.424	-0.323	0.313	-0.445	0.426	-0.546	0.521	-0.563	0.460	-0.569	0.461	-0.484	0.462
Size	0.179	0.060	0.164	0.045	0.218	0.062	0.283	0.082	0.251	0.066	0.251	0.065	0.194	0.064
Sizesq	-0.008	0.003	-0.007	0.002	-0.010	0.003	-0.012	0.004	-0.011	0.003	-0.011	0.003	-0.008	0.003
Regulatn	-0.030	0.161	-0.024	0.119	-0.042	0.160	-0.041	0.197	-0.038	0.169	-0.039	0.170	-0.029	0.172
ln α			0.544	0.161	0.331	0.093	0.551	0.152						
ln b									-29.72	42.67				
α									-0.036	0.051				
ln α_γ											1.086	0.218	-1.157	0.213
ln $\alpha_{\gamma+1}$													-0.973	0.287
Log likelihood	-184.95		-180.37		-180.21		-180.36		-171.80		-171.31		-168.04	
AIC (smaller is better)	389.90		382.74		382.41		382.72		367.61		364.62		360.08	
BIC (smaller is better)	418.26		413.94		413.61		413.92		401.64		395.82		394.12	
Time (seconds)	0.00		9.75		2.38		0.33		51.64		0.05		0.09	



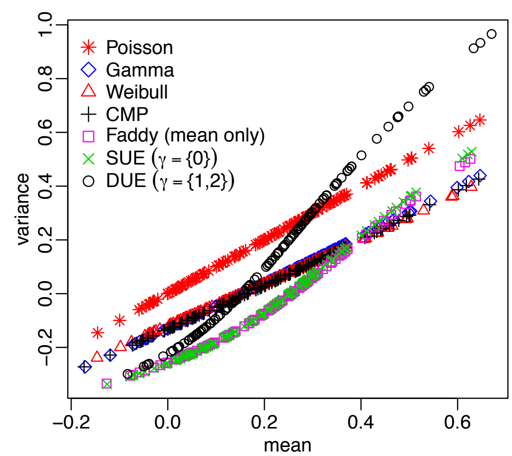
(a) *Linear scale of fertility data*



(b) *Base-10 log scale of fertility data*



(c) *Linear scale of takeover bids data*



(d) *Base-10 log scale of takeover bids data*

Figure 6. *Scatterplots of estimated variances versus estimated means.*

be interpreted similarly to Figure 6(a), but the Faddy (mean only) model is different. The Faddy (mean only) and SUE ($\boldsymbol{\gamma} = \{0\}$) curved lines lie nearly on top of each other because the rate sequence of the SUE ($\boldsymbol{\gamma} = \{0\}$) process is very similar to the Faddy (mean only) process. The rate sequences of the Faddy (mean only) and SUE ($\boldsymbol{\gamma} = \{0\}$) models are $2.92\lambda, 1.00\lambda, 0.98\lambda, 0.96\lambda, \ldots$ and $2.96\lambda, \lambda, \lambda, \lambda, \ldots$, respectively. They are non-increasing, and thus the Faddy (mean only) and SUE ($\boldsymbol{\gamma} = \{0\}$) curved lines do not cross the Poisson line. The DUE ($\boldsymbol{\gamma} = \{1,2\}$) curved line crosses the Poisson line, indicating the takeover bids data distribution is the combination of 90 underdispersed and 36 overdispersed distributions.

5. Conclusion

The Poisson, gamma, Weibull, CMP, and Faddy (1997) count models are well-known, but their underlying assumption of monotonic rate sequence limits their use in many applications. The UE count models, in contrast, are assumed that the rate sequences are non-monotonic, and the distributions of their interarrival times are exponential. One significant advantage of these new count models is the dispersion types defined by the base rate and the shape parameters. Hence, the UE count models can display over-, under-, and equidispersion, although the shape parameters are fixed numbers. In other words, the conditional variance and mean of the UE regression models must not be linearly related, allowing for a mixture of the over-, under-, and equidispersed distributions. The UE regression models are applied to the fertility and takeover bids data, and they offer significant improvements in log-likelihood compared to the above well-known regression models. For fertility data, the results show that the women's intentions to have third and fourth children, unusual events, are considerably less than other children. The behavior of these women cannot be captured by the above well-known count models with monotonic rates. Even though the UE count models offer significant improvements, future studies could improve the models for better results by replacing the exponential distribution with a non-negative distribution such as the gamma, the Weibull, etc.

Acknowledgments

The authors are grateful to the associate editor and anonymous referees for their valuable comments and suggestions.

Appendices

A. Derivations of SUE probability functions

A.1. SUE $(\boldsymbol{\gamma} = \{0\})$

Using Equations (3) and (8), the first few probabilities of the SUE $(\boldsymbol{\gamma} = \{0\})$ count model are obtained:

$$
\begin{aligned}
P_0(t) &= e^{-\alpha_0 \lambda t} \\
P_1(t) &= \int_0^t \alpha_0 \lambda P_0(u) e^{-\lambda(t-u)} du \\
&= \frac{\alpha_0 e^{-\lambda t}}{(1-\alpha_0)} \left(e^{(1-\alpha_0)\lambda t} - 1 \right)
\end{aligned}
$$

$$
\begin{aligned}
P_2(t) &= \int_0^t \lambda P_1(u) e^{-\lambda(t-u)} du \\
&= \frac{\alpha_0 e^{-\lambda t}}{(1-\alpha_0)^2}\left(e^{(1-\alpha_0)\lambda t} - 1 - (1-\alpha_0)\lambda t\right) \\
P_3(t) &= \int_0^t \lambda P_2(u) e^{-\lambda(t-u)} du \\
&= \frac{\alpha_0 e^{-\lambda t}}{(1-\alpha_0)^3}\left(e^{(1-\alpha_0)\lambda t} - 1 - (1-\alpha_0)\lambda t - \frac{((1-\alpha_0)\lambda t)^2}{2!}\right)
\end{aligned}
$$

From these equations for $P_0(t)$, $P_1(t)$, $P_2(t)$, and $P_3(t)$, one can deduce that the general SUE ($\boldsymbol{\gamma} = \{0\}$) probability function might be of the form

$$
P_n(t) = \begin{cases} e^{-\alpha_0 \lambda t} & \text{for } n = 0 \\ \dfrac{\alpha_0 e^{-\lambda t}}{(1-\alpha_0)^n}\left(e^{(1-\alpha_0)\lambda t} - \displaystyle\sum_{i=0}^{n-1} \frac{((1-\alpha_0)\lambda t)^i}{i!}\right) & \text{for } n > 0. \end{cases} \tag{A.1}
$$

Inserting the Taylor series expansion of $e^{(1-\alpha_0)\lambda t}$, the SUE ($\boldsymbol{\gamma} = \{0\}$) probability distribution can be rewritten as

$$
P_n(t) = \begin{cases} e^{-\alpha_0 \lambda t} & \text{for } n = 0 \\ \alpha_0 (\lambda t)^n e^{-\lambda t} \displaystyle\sum_{i=0}^{\infty} \frac{((1-\alpha_0)\lambda t)^i}{(i+n)!} & \text{for } n > 0. \end{cases} \tag{A.2}
$$

A.2. SUE ($\boldsymbol{\gamma} = \{\gamma > 0\}$)

Using Equations (3) and (8), the first few probabilities of the SUE ($\boldsymbol{\gamma} = \{\gamma > 0\}$) count model are obtained:

$$
\begin{aligned}
P_0(t) &= e^{-\lambda t} \\
P_1(t) &= \int_0^t \lambda P_0(u) e^{-\lambda(t-u)} du = \lambda t e^{-\lambda t} \\
&\vdots \\
P_{\gamma-1}(t) &= \int_0^t \lambda P_{\gamma-2}(u) e^{-\lambda(t-u)} du = \frac{(\lambda t)^{\gamma-1} e^{-\lambda t}}{(\gamma-1)!} \\
P_\gamma(t) &= \int_0^t \lambda P_{\gamma-1}(u) e^{-\alpha_\gamma \lambda(t-u)} du \\
&= \frac{e^{-\lambda t}}{(1-\alpha_\gamma)^\gamma}\left(e^{(1-\alpha_\gamma)\lambda t} - \sum_{i=0}^{\gamma-1} \frac{((1-\alpha_\gamma)\lambda t)^i}{i!}\right)
\end{aligned}
$$

$$P_{\gamma+1}(t) = \int_0^t \alpha_\gamma \lambda P_\gamma(u) e^{-\lambda(t-u)} du$$
$$= \frac{\alpha_\gamma e^{-\lambda t}}{(1-\alpha_\gamma)^{\gamma+1}} \left(e^{(1-\alpha_\gamma)\lambda t} - \sum_{i=0}^{\gamma} \frac{\left((1-\alpha_\gamma)\lambda t\right)^i}{i!} \right)$$

From these equations for $P_0(t), P_1(t), \ldots, P_{\gamma-1}(t), P_\gamma(t)$, and $P_{\gamma+1}(t)$, one can deduce that the general SUE ($\boldsymbol{\gamma} = \{\gamma > 0\}$) probability function might be of the form

$$P_n(t) = \begin{cases} \dfrac{(\lambda t)^n e^{-\lambda t}}{n!} & \text{for } n < \gamma \\ \dfrac{e^{-\lambda t}}{(1-\alpha_\gamma)^n} \left(e^{(1-\alpha_\gamma)\lambda t} - \displaystyle\sum_{i=0}^{n-1} \frac{\left((1-\alpha_\gamma)\lambda t\right)^i}{i!} \right) & \text{for } n = \gamma \\ \dfrac{\alpha_\gamma e^{-\lambda t}}{(1-\alpha_\gamma)^n} \left(e^{(1-\alpha_\gamma)\lambda t} - \displaystyle\sum_{i=0}^{n-1} \frac{\left((1-\alpha_\gamma)\lambda t\right)^i}{i!} \right) & \text{for } n > \gamma. \end{cases} \tag{A.3}$$

Inserting the Taylor series expansion of $e^{(1-\alpha_\gamma)\lambda t}$, the SUE ($\boldsymbol{\gamma} = \{\gamma > 0\}$) probability distribution can be rewritten as

$$P_n(t) = \begin{cases} \dfrac{(\lambda t)^n e^{-\lambda t}}{n!} & \text{for } n < \gamma \\ (\lambda t)^n e^{-\lambda t} \displaystyle\sum_{i=0}^{\infty} \frac{\left((1-\alpha_\gamma)\lambda t\right)^i}{(i+n)!} & \text{for } n = \gamma \\ \alpha_\gamma (\lambda t)^n e^{-\lambda t} \displaystyle\sum_{i=0}^{\infty} \frac{\left((1-\alpha_\gamma)\lambda t\right)^i}{(i+n)!} & \text{for } n > \gamma. \end{cases} \tag{A.4}$$

B. Derivations of DUE ($\boldsymbol{\gamma} = \{\gamma, \gamma+1\}$) Probability Functions

B.1. DUE ($\boldsymbol{\gamma} = \{0, 1\}$)

Using Equations (5) and (8), the first few probabilities of the DUE ($\boldsymbol{\gamma} = \{0, 1\}$) count model are obtained:

$$P_0(t) = e^{-\alpha_0 \lambda t}$$
$$P_1(t) = \int_0^t \alpha_0 \lambda P_0(u) e^{-\alpha_1 \lambda (t-u)} du$$
$$= \frac{\alpha_0 e^{-\lambda t}}{(\alpha_1 - \alpha_0)} \left(e^{(1-\alpha_0)\lambda t} - e^{(1-\alpha_1)\lambda t} \right)$$

$$
\begin{aligned}
P_2(t) &= \int_0^t \alpha_1 \lambda P_1(u) e^{-\lambda(t-u)} du \\
&= \frac{\alpha_0 \alpha_1 e^{-\lambda t}}{(\alpha_1 - \alpha_0)} \left(\frac{e^{(1-\alpha_0)\lambda t} - 1}{(1-\alpha_0)} - \frac{e^{(1-\alpha_1)\lambda t} - 1}{(1-\alpha_1)} \right) \\
P_3(t) &= \int_0^t \lambda P_2(u) e^{-\lambda(t-u)} du \\
&= \frac{\alpha_0 \alpha_1 e^{-\lambda t}}{(\alpha_1 - \alpha_0)} \left(\frac{e^{(1-\alpha_0)\lambda t} - 1 - (1-\alpha_0)\lambda t}{(1-\alpha_0)^2} - \frac{e^{(1-\alpha_1)\lambda t} - 1 - (1-\alpha_1)\lambda t}{(1-\alpha_1)^2} \right)
\end{aligned}
$$

From these equations for $P_0(t)$, $P_1(t)$, $P_2(t)$, and $P_3(t)$, one can deduce that the general DUE ($\boldsymbol{\gamma} = \{0,1\}$) probability function might be of the form

$$
P_n(t) = \begin{cases}
e^{-\alpha_0 \lambda t} & \text{for } n = 0 \\
\frac{\alpha_0 e^{-\lambda t}}{(\alpha_1 - \alpha_0)} \left(e^{(1-\alpha_0)\lambda t} - e^{(1-\alpha_1)\lambda t} \right) & \text{for } n = 1 \\
\frac{\alpha_0 \alpha_1 e^{-\lambda t}}{(\alpha_1 - \alpha_0)} \left(\frac{e^{(1-\alpha_0)\lambda t} - \sum_{i=0}^{n-2} \frac{((1-\alpha_0)\lambda t)^i}{i!}}{(1-\alpha_0)^{n-1}} \right. & \\
\qquad \left. - \frac{e^{(1-\alpha_1)\lambda t} - \sum_{i=0}^{n-2} \frac{((1-\alpha_1)\lambda t)^i}{i!}}{(1-\alpha_1)^{n-1}} \right) & \text{for } n > 1.
\end{cases} \tag{B.1}
$$

Inserting the Taylor series expansion of $e^{(1-\alpha_0)\lambda t}$ and $e^{(1-\alpha_1)\lambda t}$, the DUE ($\boldsymbol{\gamma} = \{0,1\}$) probability distribution can be rewritten as

$$
P_n(t) = \begin{cases}
e^{-\alpha_0 \lambda t} & \text{for } n = 0 \\
\alpha_0 (\lambda t) e^{-\lambda t} \sum_{i=0}^{\infty} \frac{c_i (\lambda t)^i}{(i+1)!} & \text{for } n = 1 \\
\alpha_0 \alpha_1 (\lambda t)^n e^{-\lambda t} \sum_{i=0}^{\infty} \frac{c_i (\lambda t)^i}{(i+n)!} & \text{for } n > 1,
\end{cases} \tag{B.2}
$$

where $c_i = \sum_{k=0}^{i} (1-\alpha_0)^k (1-\alpha_1)^{i-k}$.

B.2. DUE ($\boldsymbol{\gamma} = \{\gamma > 0, \gamma + 1\}$)

Using Equations (5) and (8), the first few probabilities of the DUE ($\boldsymbol{\gamma} = \{\gamma, \gamma + 1\}$) count model are obtained:

$$
\begin{aligned}
P_0(t) &= e^{-\lambda t} \\
P_1(t) &= \int_0^t \lambda P_0(u) e^{-\lambda(t-u)} du = \lambda t e^{-\lambda t} \\
P_2(t) &= \int_0^t \lambda P_1(u) e^{-\lambda(t-u)} du = \frac{(\lambda t)^2 e^{-\lambda t}}{2!} \\
&\vdots \\
P_{\gamma-1}(t) &= \int_0^t \lambda P_{\gamma-2}(u) e^{-\lambda(t-u)} du = \frac{(\lambda t)^{\gamma-1} e^{-\lambda t}}{(\gamma-1)!} \\
P_\gamma(t) &= \int_0^t \lambda P_{\gamma-1}(u) e^{-\alpha_\gamma \lambda(t-u)} du \\
&= \frac{e^{-\lambda t}}{(1-\alpha_\gamma)^\gamma} \left(e^{(1-\alpha_\gamma)\lambda t} - \sum_{i=0}^{\gamma-1} \frac{\left(\left(1-\alpha_\gamma\right)\lambda t\right)^i}{i!} \right) \\
P_{\gamma+1}(t) &= \int_0^t \alpha_\gamma \lambda P_\gamma(u) e^{-\alpha_{\gamma+1}\lambda(t-u)} du \\
&= \frac{\alpha_\gamma e^{-\lambda t}}{(\alpha_{\gamma+1}-\alpha_\gamma)} \left(\frac{e^{(1-\alpha_\gamma)\lambda t} - \sum_{i=0}^{\gamma-1} \frac{\left((1-\alpha_\gamma)\lambda t\right)^i}{i!}}{(1-\alpha_\gamma)^\gamma} \right. \\
&\qquad \left. - \frac{e^{\left(1-\alpha_{\gamma+1}\right)\lambda t} - \sum_{i=0}^{\gamma-1} \frac{\left((1-\alpha_{\gamma+1})\lambda t\right)^i}{i!}}{\left(1-\alpha_{\gamma+1}\right)^\gamma} \right) \\
P_{\gamma+2}(t) &= \int_0^t \alpha_{\gamma+1} \lambda P_{\gamma+1}(u) e^{-\lambda(t-u)} du \\
&= \frac{\alpha_\gamma \alpha_{\gamma+1} e^{-\lambda t}}{(\alpha_{\gamma+1}-\alpha_\gamma)} \left(\frac{e^{(1-\alpha_\gamma)\lambda t} - \sum_{i=0}^{\gamma} \frac{\left((1-\alpha_\gamma)\lambda t\right)^i}{i!}}{(1-\alpha_\gamma)^{\gamma+1}} \right. \\
&\qquad \left. - \frac{e^{\left(1-\alpha_{\gamma+1}\right)\lambda t} - \sum_{i=0}^{\gamma} \frac{\left((1-\alpha_{\gamma+1})\lambda t\right)^i}{i!}}{\left(1-\alpha_{\gamma+1}\right)^{\gamma+1}} \right)
\end{aligned}
$$

$$
\begin{aligned}
P_{\gamma+3}(t) &= \int_0^t \lambda P_{\gamma+2}(u) e^{-\lambda(t-u)} du \\
&= \frac{\alpha_\gamma \alpha_{\gamma+1} e^{-\lambda t}}{(\alpha_{\gamma+1} - \alpha_\gamma)} \left(\frac{e^{(1-\alpha_\gamma)\lambda t} - \sum_{i=0}^{\gamma+1} \frac{\left((1-\alpha_\gamma)\lambda t\right)^i}{i!}}{(1-\alpha_\gamma)^{\gamma+2}} \right. \\
&\qquad \left. - \frac{e^{\left(1-\alpha_{\gamma+1}\right)\lambda t} - \sum_{i=0}^{\gamma+1} \frac{\left((1-\alpha_{\gamma+1})\lambda t\right)^i}{i!}}{\left(1-\alpha_{\gamma+1}\right)^{\gamma+2}} \right)
\end{aligned}
$$

From these equations for $P_0(t)$, $P_1(t)$, ..., $P_{\gamma+1}(t)$, $P_{\gamma+2}(t)$, and $P_{\gamma+3}(t)$, one can deduce that the general DUE ($\boldsymbol{\gamma} = \{\gamma, \gamma+1\}$) probability function might be of the form

$$
P_n(t) = \begin{cases}
\dfrac{(\lambda t)^n e^{-\lambda t}}{n!} & \text{for } n < \gamma \\[2ex]
\dfrac{e^{-\lambda t}}{\left(1-\alpha_\gamma\right)^n} \left(e^{\left(1-\alpha_\gamma\right)\lambda t} - \displaystyle\sum_{i=0}^{n-1} \frac{\left(\left(1-\alpha_\gamma\right)\lambda t\right)^i}{i!} \right) & \text{for } n = \gamma \\[2ex]
\dfrac{\alpha_\gamma e^{-\lambda t}}{\left(\alpha_{\gamma+1}-\alpha_\gamma\right)} \left(\dfrac{e^{\left(1-\alpha_\gamma\right)\lambda t} - \sum_{i=0}^{n-2} \frac{\left(\left(1-\alpha_\gamma\right)\lambda t\right)^i}{i!}}{\left(1-\alpha_\gamma\right)^{n-1}} \right. & \\
\qquad \left. - \dfrac{e^{\left(1-\alpha_{\gamma+1}\right)\lambda t} - \sum_{i=0}^{n-2} \frac{\left(\left(1-\alpha_{\gamma+1}\right)\lambda t\right)^i}{i!}}{\left(1-\alpha_{\gamma+1}\right)^{n-1}} \right) & \text{for } n = \gamma+1 \\[2ex]
\dfrac{\alpha_\gamma \alpha_{\gamma+1} e^{-\lambda t}}{\left(\alpha_{\gamma+1}-\alpha_\gamma\right)} \left(\dfrac{e^{\left(1-\alpha_\gamma\right)\lambda t} - \sum_{i=0}^{n-2} \frac{\left(\left(1-\alpha_\gamma\right)\lambda t\right)^i}{i!}}{\left(1-\alpha_\gamma\right)^{n-1}} \right. & \\
\qquad \left. - \dfrac{e^{\left(1-\alpha_{\gamma+1}\right)\lambda t} - \sum_{i=0}^{n-2} \frac{\left(\left(1-\alpha_{\gamma+1}\right)\lambda t\right)^i}{i!}}{\left(1-\alpha_{\gamma+1}\right)^{n-1}} \right) & \text{for } n > \gamma+1.
\end{cases} \tag{B.3}
$$

Inserting the Taylor series expansion of $e^{(1-\alpha_\gamma)\lambda t}$ and $e^{(1-\alpha_{\gamma+1})\lambda t}$, the DUE ($\boldsymbol{\gamma} = \{\gamma, \gamma+1\}$) probability distribution can be rewritten as

$$P_n(t) = \begin{cases} \dfrac{(\lambda t)^n e^{-\lambda t}}{n!} & \text{for } n < \gamma \\ (\lambda t)^n e^{-\lambda t} \sum\limits_{i=0}^{\infty} \dfrac{\left(\left(1-\alpha_\gamma\right)\lambda t\right)^i}{(i+n)!} & \text{for } n = \gamma \\ \alpha_\gamma (\lambda t)^n e^{-\lambda t} \sum\limits_{i=0}^{\infty} \dfrac{c_i (\lambda t)^i}{(i+n)!} & \text{for } n = \gamma+1 \\ \alpha_\gamma \alpha_{\gamma+1} (\lambda t)^n e^{-\lambda t} \sum\limits_{i=0}^{\infty} \dfrac{c_i (\lambda t)^i}{(i+n)!} & \text{for } n > \gamma+1, \end{cases} \tag{B.4}$$

where $c_i = \sum\limits_{k=0}^{i} \left(1-\alpha_\gamma\right)^k \left(1-\alpha_{\gamma+1}\right)^{i-k}$.

References

Ball, F. (1995). A note on variation in birth processes. *The Mathematical Scientist* 20, 50–55.

Bartlett, M. S. (1978). *An Introduction to Stochastic Processes*. Cambridge University Press, Cambridge, UK.

Cameron, A. C. and Johansson, P. (1997). Count data regression using series expansions: with applications. *Journal of Applied Econometrics* 12(3), 203–233.

Chanialidis, C., Evers, L., Neocleous, T., and Nobile, A. (2018). Efficient Bayesian inference for COM-Poisson regression models. *Statistics and Computing* 28, 595–608.

Conway, R. W. and Maxwell, W. L. (1962). A queuing model with state dependent service rates. *Journal of Industrial Engineering* 12, 132–136.

Cox, D. R. and Miller, H.D. (1965). *The Theory of Stochastic Processes*. Chapman and Hall, London, UK.

Crawford, F. W., Ho, L. S. T., and Suchard, M. A. (2018). Computational methods for birth-death processes. *WIREs Computational Statistics* 10(2), 1–22.

Eddelbuettel, D., Francois, R., Allaire, J., Ushey, K., Kou, Q., Russell, N., Bates, D., and Chambers, J. (2021). Rcpp: Seamless R and C++ integration. R package version 1.0.6.

Faddy, M. J. (1994). On variation in Poisson processes. *The Mathematical Scientist* 19, 47–51.

Faddy, M. J. (1997). Extended Poisson process modelling and analysis of count data. *Biometrical Journal* 39(4), 431–440.

Faddy, M. J. and Smith, D. M. (2008). Extended Poisson process modelling of dilution series data. *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 57(4), 461–471.

Faddy, M. J. and Smith, D. M. (2011). Analysis of count data with covariate dependence in both mean and variance. *Journal of Applied Statistics* 38(12), 2683–2694.

Fung, T., Alwan, A., Wishart, J., and Huang, A. (2019). mpcmp: Mean-parametrized Conway-Maxwell Poisson (COM-Poisson) regression. R package version 0.1.3.

Goulet, V., Dutang, C., Maechler, M., Firth, D., Shapira, M., and Stadelmann, M. (2020). expm: Matrix exponential, log, 'etc'. R package version 0.999-5.

Jaggia, S. and Thosar, S. (1993). Multiple bids as a consequence of target management resistance: a count data approach. *Review of Quantitative Finance & Accounting* 3(4), 447–457.

Jensen, A. (1953). Markoff chains as an aid in the study of Markoff processes. *Scandinavian Actuarial Journal* sup1, 87–91.

Kharrat, T. and Boshnakov, G. N. (2018). Countr: Flexible univariate count models based on renewal processes. R package version 3.5.2.

Kharrat, T., Boshnakov, G. N., McHale, I., and Baker, R. (2019). Flexible regression models for count data based on renewal processes. *Journal of Statistical Software* 90(13), 1–35.

McShane, B., Adrian, M., Bradlow, E. T., and Fader, P. S. (2008). Count models based on Weibull interarrival times. *Journal of Business & Economic Statistics* 26(3), 369–378.

Parthasarathy, P. R. and Sudhesh, R. (2006). Exact transient solution of a state-dependent birth-death process. *Journal of Applied Mathematics and Stochastic Analysis* 82(6), 1–16.

Podlich, H. M., Faddy, M. J., and Smyth, G. K.(2004). Semi-parametric extended Poisson process models for count data. *Statistics and Computing* 14, 311–321.

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Ross, S. M. (2010). *Introduction to Probability Models*, 10th Ed. Academic Press, San Diego, CA, USA.

Saez-Castillo, A. J. and Conde-Sanchez, A. (2013). A hyper-Poisson regression model for overdispersed and underdispersed count data. *Journal of Computational Statistics & Data Analysis* 61, 148–157.

Sellers, K. F., Lotze, T., and Raim, A. M. (2018). COMPoissonReg: Conway-Maxwell Poisson (COM-Poisson) regression. R package version 0.6.1.

Sellers, K. F. and Shmueli, G. (2010). A flexible regression model for count data. *The Annals of Applied Statistics* 4(2), 943–961.

Smith, D. M. and Faddy, M. J. (2016). Mean and variance modelling of under- and overdispersed count data. *Journal of Statistical Software* 69(6), 1–23.

Smith, D. M. and Faddy, M. J. (2018). CountsEPPM: Mean and variance modeling of count data. R package version 3.0.

Winkelmann, R. (1995). Duration dependence and dispersion in count-data models. *Journal of Business & Economic Statistics* 13(4), 467–474.