T-geometric regression models with applications to zero-inflated count data

Carl Lee[1], Felix Famoye[1] and Alfred Akinsete[2]
[1]: Department of Mathematics, Central Michigan University, MI 49959
[2]: Department of Mathematics and Statistics, Marshall University, WV 25755

Abstract
A method of developing generalized parametric regression models for modeling count data is proposed and studied. The method is based on the framework of the *T*-geometric family of distributions. A *T*-geometric distribution is the discrete analogue of the corresponding continuous distribution. The general methodology is applied to derive several generalized regression models for count data. These regression models can fit count data with under-dispersion or over-dispersion. The extension to model truncated or zero inflated data is addressed. Some new generalized *T*-geometric regression models are applied to real world data sets to illustrate the flexibility of these models.

Key Words: discrete analogue, generalized parametric models, under- and over-dispersion

1. Introduction

Regression model for count data has been and continues to be an important research topic for over half century due to its applications in many different disciplines such as actuarial science, biostatistics, demography, economic, social sciences and many others. Early work started with the regression models for count data that follow standard Poisson distribution. For example, Jorgenson (1961) consider a multiple linear regression analysis for count data following the Poisson distribution with applications to number of purchases over a time period, and number of failures over a time period. Frome et al. (1973) considered the Poisson distribution in the context of non-linear regression analysis for count data where the sample mean and sample variance are about equal. An important milestone in the development of count data models was the generalized linear models by Nelder and Wedderburn (1972), and later elaborated by McCullagh and Nelder (1989). Poisson regression is a special case of the generalized linear model by the logarithm transformation of the Poisson response variable.

The standard Poisson regression models have the limitation of equ-dispersion; the mean and variance of a count data are almost equal. Many count data do not satisfy the equ-dispersion property. Instead, they are either over-dispersion (variance > mean) or under-dispersion (variance < mean). There has been a rich development of flexible models for modeling over-or under-dispersion count data (e.g., Lawless (1987), McCullagh and Nelder (1989), Famoye (1993)). For detailed review on count data modeling, one may refer to Hibe (2011) and Cameron and Trivedi (2013), and the references therein.

In addition to over- or under-dispersion, there are other types of departures from standard Poisson regression models. Cameron and Trivedi (2013) described eight types of departures including failure of equ-dispersion, truncation and censoring, zero-inflation problem, violation of iid assumption among observations, and others. Various existing methods for dealing with the departures are reviewed in Cameron and Trivedi (2013) and the references therein.

2. A framework for developing generalized discrete distributions

The proposed method for developing generalized models for count data relies on the method of developing discrete generalized distributions proposed by Alzaatreh et al. (2013), namely the T-R(W) family. Let $f_T(t)$ be a probability density function (PDF) of a continuous random variable $T \in [a,\ b]$, $-\infty \le a < b \le \infty$. Suppose further that $W(F_R(y))$ is a monotonic and absolutely continuous function of the cumulative distribution function (CDF), $F_R(y)$, of any random variable *R*. The CDF $F_Y(y)$ of a new random variable *Y* is given by

$$F_Y(y) = \int_a^{W(F_R(y))} f_T(t)dt = F_T\left\{W\left(F_R(y)\right)\right\}. \tag{2.1}$$

Many continuous distributions have been defined by using the result in (2.1). By taking $W(F_R(y)) = -\ln(1-F_R(y))$, (2.1) defines the CDF and PDF of a T-R(W) random variable as: $F_Y(y) = F_T\{H_R(y)\}$ and $f_Y(y) = h_R(y) f_T\{H_R(y)\}$, where $H_R(y)$ is the survival function to R and $h_R(y)$ is the hazard function of R. Further by taking R to be the discrete Geometric random variable with $F_R(y) = 1 - p^{(y+1)}, y = 0,1,2,.....$, where p is the probability of failure , Alzaatreh et al. (2012) defined the *T*-geometric family. The *T*-geometric family defines the discrete analogue to the distribution of any continuous non-negative random variable *T*. A *T*-geometric random variable *Y* has the CDF and PDF, respectively:

$$F_Y(y) = F_T(-\ln p^{y+1}) \ , \tag{2.2}$$

$$f_Y(y) = F_T(-\ln p^{y+1}) - F_T(-\ln p^{y}) . \tag{2.3}$$

Alzaatreh et al. (2012) defined and studied in details the exponentiated-exponential geometric distribution (EEGD). Akinsete et al. (2014) defined and studied the Kumaraswamy-geometric distribution in detail.

A generalized regression model for modeling the response *Y* following a member of *T*-geometric distribution, namely the exponentiated-exponential geometric regression model (EEGR) for count data and zero-inflated count data was studied in Famoye and Lee (2016). The purpose of this study is to propose a modeling framework for developing models that are capable of modeling over- and under- dispersed as well as zero-inflated count data. These models can be modified to model truncated and censored data. Kumaraswamy-geometric regression model (KGR) will be studied in detail. The KGR model will be applied to some real world data and compared with the EEGR model and other well-known models including generalized Poisson regression (GPR) and Negative binomial regression (NBR) for modeling over- or under-disposed count data, as well as zero-inflated count data.

3. *T*-geometric regression models

The equation (2.3) defines the probability mass function (PMF) of the discrete analogue random variable *Y* for the corresponding continuous random variable. In the following we define two regression models for modeling the discrete random variable *Y* defined from (2.3).

3.1 Exponentiated-exponential geometric regression model

Taking *T* to be exponeiated-exponental (EE) random variable with CDF: The CDF of the EE is given by (Gupta and Kundu, 2001):

$$F_T(t) = (1-e^{-\lambda t})^c \text{, for } t > 0,\ c > 0 \text{ and } \lambda > 0. \tag{3.1}$$

The probability mass function (PMF) of exponentiated-exponential geometric distribution (EEGD) is (Alzaatreh et al. ,2012)

$$\begin{aligned} f_Y(y) = F_Y(y) - F_Y(y-1) &= \left(1-p^{\lambda(y+1)}\right)^c - \left(1-p^{\lambda y}\right)^c \\ &= \left(1-\theta^{y+1}\right)^c - \left(1-\theta^{y}\right)^c,\ y = 0, 1, 2, \ldots, \end{aligned} \tag{3.2}$$

where $c > 0$ and $0 < p^{\lambda} = \theta < 1$. The EEGD is unimodal and right skewed. The parameter *c* affects the shape of the distribution, which can be over-dispersed, equi-dispersed or over-dispersed. Alzaatreh et al. (2012) discussed the conditions when the distribution is over- or under-dispersed according to the shape parameter c.

Suppose *Y* is a count response variable that follows the EEGD in equation (3.2) and is associated with a set of covariates. Let $x_i = (x_{i0} = 1, x_{i1}, x_{i2}, \ldots, x_{i,k-1})'$ be a (*k* – 1)-dimensional vector of predictor variables. There are two parameters of EEGD (θ, *c*). The parameter θ is associated with moments, which is also depend on parameter c. For example, for *c*=1, mean of *Y* is $\theta(1-\theta)^{-1}$ and variance of *Y* is $\theta(1-\theta)^{-2}$ with variance/mean = $1/(1-\theta)$, an over-dispersed distribution. The parameter *c* is associated with the shape. Famoye and Lee (2016) defines the EEGR model assuming that parameter θ of EEGD is a function of x_i given by $\theta(x_i) = f(x_i, \beta)$, where $0 < f(x_i, \beta) < 1$ is a known function of x_i and a *k*-dimensional vector $\beta = (\beta_0, \beta_1, \beta_2, \ldots, \beta_{k-1})'$ of regression parameters. Since

$0 < \theta < 1$, the logit transformation of $f(x_i,\beta)$, $\theta(x_i) = \theta_i = f(x_i,\beta) = e^{x_i'\beta} / (1+e^{x_i'\beta})$, is modeled, and the EEGR model is defined as

$$P(Y = y_i \mid x_i) = \left(1-[\theta(x_i)]^{y_i+1}\right)^c - \left(1-[\theta(x_i)]^{y_i}\right)^c,\ y_i = 0,1,2,\ldots \tag{3.3}$$

If the shape parameter c is of interest, one can also assume the shape parameter c to be a function of the same covariates with $c(x_i) = \exp(x_i'\gamma)$ or a set of m different covariates, $\{z_i\}$ with $c(z_i) = \exp(z_i'\gamma)$.

3.2 Kumaraswamy-geometric regression model

Akinsete et al. (2014) defined and studied the Kumaraswamy-geometric distribution (KGD) by letting the random variable *T* follow the Kumaraswamy's distribution with CDF and PDF, respectively:

$$F_T(y) = 1-(1-y^a)^b, 0 < y < 1, a > 0, b > 0, \tag{3.4}$$

$$f_T(y) = aby^{a-1}(1-ya)^{b-1}, 0 < y < 1, \tag{3.5}$$

The CDF and PMF of the KGD is defined as

$$F_Y(y) = 1-\{1-(1-p^{y+1})^a\}^b, y = 0,1,2,..... \tag{3.6}$$

$$f_Y(y) = \{1-(1-p^y)^a\}^b - \{1-(1-p^{y+1})^a\}^b, y = 0,1,2,..., a > 0, b > 0 \tag{3.7}$$

The KGD has three parameters (*a* ,*b*, *p*). When $b=1$, the KGD reduces to the EEGD. Parameters (*a*, *b*) are shape parameters and parameter p is associated with the moments. Akinsete et al. (2014) derived expressions for computing moments. Their numerical results indicate about the mean and variance of KGD can be over-, equ- or under-dispersed.

Applying a similar definition as EEGR, we assume the parameter *p* is associated with the covariates with the relation $p(x_i) = f(x_i,\beta)$ and apply the logit transformation of $f(x_i,\beta)$, to define the KGR model as

$$P(Y = y_i \mid x_i) = \{1-[1-(p(x_i))^{y_i}]^a\}^b - \{1-[1-(p(x_i))^{y_i+1}]^a\}^b, \tag{3.8}$$

where $y_i = 0,1,2,..., a > 0, b > 0$. Using similar method described above, one can define the *T*-geometric regression model for count data for any continuous distribution.

3.3 Zero-inflated T-geometric regression models

Excessive zero count data occur in many real world problems. For this situation, a zero-inflated regression model can be applied (e.g., Famoye and Singh ,2006). A zero-inflated *T*-geometric regression model defines the relationship between *Y* and covariates in two pieces:

$$P(Y = y_i \mid x_i, z_i) = \begin{cases} \omega_i + (1-\omega_i) f_Y(y_i), & y_i = 0 \\ (1-\omega_i) f_Y(y_i), & y_i = 1,2,3,\ldots, \end{cases} \tag{3.9}$$

where $f_Y(y_i)$ is a *T*-geometric regression model, e.g., EEGR or KGR, and $0 < \omega_i < 1$. The probability ω_i is taken to be a function of covariates $z_i = (z_{i0} = 1, z_{i1}, z_{i2}, \ldots, z_{i,r-1})'$ and it is defined as logit function $\omega_i = \exp(z_i'\delta)/[1+\exp(z_i'\delta)]$, where δ is an *r*-dimensional vector $\delta = (\delta_0, \delta_1, \delta_2, \ldots, \delta_{r-1})'$ of parameters. The covariates z_i may be a subset of the x_i or be completely different from the x_i. Other types of departures from standard Poisson distribution, such as truncated and censored models can be easily modified from the regular T-geometric regression models based on the type of departure.

3.4 Parameter estimates for Kum-Geometric regression model

The maximum likelihood method is applied to estimate the regression parameters and the nuisance parameters. Let a random sample of size *n* be taken from KGR, with likelihood function expressed as

$$L(x \mid a,b,\beta) = \prod_{i=0}^{n} \left\{ \left\{1-[1-(p(x_i))^{y_i}]^a\right\}^b - \left\{1-[1-(p(x_i))^{y_i+1}]^a\right\}^b \right\}, \tag{3.10}$$

and log-likelihood function

$$\begin{aligned} \ell(a,b,\beta) = \ell &= \sum_{i=1}^{n} \log P(y_i \mid x_i) \\ &= \sum_{i=1}^{n} \log \left\{ \left\{1-[1-(p(x_i))^{y_i}]^a\right\}^b - \left\{1-[1-(p(x_i))^{y_i+1}]^a\right\}^b \right\} \end{aligned} \tag{3.11}$$

By taking the first derivatives with respect to the nuisance parameters (*a*,*b*) and the regression parameters, $\beta = (\beta_0, \beta_1, \beta_2, \ldots, \beta_{k-1})'$, setting the first derivatives to zeros and solving the equations numerically, we obtain the maximum likelihood estimators for (*a*,*b*) and β. The initial estimates of β is the modified linear regression estimate given by $\hat{\beta}_* = (X'X)^{-1}[X'\ln(Y+0.5)]$, where *X* is an $n \times k$ matrix and *Y* is an $n \times 1$ column vector of count response variable. The same approach can be used to find the initial estimate of the zero inflated regression parameters δ. The initial estimate of the shape parameters (a,b) can be taken to be 1 or the final estimates from fitting the KGD to the count response variable *Y* without the covariates. The PROC NLMIXED procedure in SAS is applied to obtain the parameter estimates for KGR and ZIKGR. In a similar fashion, the second partial derivatives of the log-likelihood can be obtained, forming the elements of the score Fisher's information matrix.

3.5 Goodness-of-fit statistics and comparison test

Various goodness test statistics will be applied for selecting models such as Root Mean Square Error, AIC, BIC and others. When a model involves nuisance parameters, a common practice in model selection is to compare full and reduced models based on the different values of the nuisance parameters. Likelihood ratio test is applied to test the significance of the nuisance parameter. For EEGR, the parameter *c* can be treated a nuisance parameter. While, for the KGR, both parameters (*a*,*b*) can be treated as nuisance parameters.

When comparing models derived from different modeling techniques, the measures RMSE, BIC or AIC are computed to compare the model performance. Alternatively, Vuong's (1989) Kullback-Liebler Information Criterion can also be applied to discriminate between two models that are not nested.

4. Applications

The data set to be analyzed is the 1977-1978 Australian Health Survey data with n = 5190 single-person households obtained from the *Journal of Applied Econometrics 1997* Data Archive. The interest is to model the numbers of Doctor's visits in the past 2 weeks on health utilization (DOCVISIT) using (a) social-economic variables: SEX, AGE, AGESEQ, INCOME, (b) health insurance status indicators: LEVYPLUS, FREEPOOR, FREEREPA, (c) recent health status measures: ILLNESS, ACTDAYS, and (d) long-term health status measures: HSCORE, CHCOND1, CHCOND2. Details for twelve predictor variables are given in Cameron and Trivedi (2003). Mullahy (1997) and Cameron and Johnson (1997) used the data to illustrate univariate count regression models. The response variable DOCVISIT is highly over-dispersed and excessively zero-inflated with mean 0.3017 and variance 0.637 and percent of zeros 79.79%.

The strategy of modeling the DOCVISIT is as follows. We first fit the response variable using EEGD, KGD and GPD without covariates and assess the adequacy of EEGD over the sub model geometric distribution (GD), KGD over the sub model EEGD, and GPD over the sub model Poisson distribution (PD). The GPR regression defined in Famoye (1993) has $E(Y_i \mid x_i) = \mu(x_i)$ and $V(Y_i \mid x_i) = \mu(x_i)\left(1 + c\mu(x_i)\right)^2$, where the parameter where parameter *c* is any real number that describes the under-, equ- or over-dispersion of the response with c <0, =0 or > 0, respectively. The three hypotheses are:

(I) Test EEGD over GD: $H_0 : c = 1$ against $H_a : c \neq 1$

(II) Test KGD over EEGD: $H_0 : b = 1$ against $H_a : b \neq 1$

(III) Test GPD over PD: $H_0 : c = 0$ against $H_a : c \neq 0$

The asymptotic Wald test, $z = (\hat{\theta} - \theta_0) / s.e.(\hat{\theta})$, is used to test these hypotheses. The maximum likelihood estimates of the parameters of fitted EEGD and KGD distributions are summarized in Table 1. The parameter estimates indicate that the response DOCVISIT is over-dispersed. The estimates of the shape parameters are used as the initial estimate for the regression model with covariates.

The next in our analysis is to fit the regression models without and with covariates, then, test if the hypothesis that each response variable is zero-inflated over non-zero inflated by testing the following hypothesis:

(IV) H_0 : All δ_i's $= 0$ agaist H_a : Not all δ_i's $= 0,\ i = 1, 2, \cdots, r-1,$, where $\delta = (\delta_0, \delta_1, \delta_2, \ldots, \delta_{r-1})'$ is the parameter vectors for modeling the zero-inflated component. The log-likelihood ratio test is used to test the hypothesis (IV). For this health care application, number of parameters for the zero-inflated component is (*r*-1) = 12, which is the degrees of freedom for the log-likelihood ratio test. The results of the above four hypotheses tests are summarized in Table 2, There are three tests for the Hypothesis (IV). We provide only the comparison between EEGR and ZIEEGR in the summary table, since the results from the other two model comparisons are very similar. Table 2 indicates the following conclusion for the DOCVISIT response variable:

i) EEGD fits better than GD,
ii) KGD fits better than EEGD,
iii) GPD is better than PD model,
iv) The zero-inflated component is statistically significant.

Based on the above conclusion, we fit ZIEEGR, ZIKGR and ZIGPR, and summarize the parameter estimates in Tables 3.

The actual percent of zero is 79.79%. The estimated percent of zero from each of the three models are, respectively, 80.64% from ZIGPR, 80.15% from ZIEEGR and 80.25% from ZIKGR. To compare ZIEEGR and ZIPGR, the Kullback-Liebler Information Criterion proposed by Vuong (1989) is used, since ZIEEGR and ZIEEGR are two independent models. The results are summarized in Table 4. The likelihood ratio test indicates the ZIKGR fits better than ZIEEGR for all three responses, while the performance between ZIEEGR and ZIGPR are not significantly different.

Table 1: The maximum likelihood estimates of fitted distributions of DOCVISIT without covariates

EEGD	$\hat{\theta}(s.e.)$	$\hat{c}(s.e.)$	
	0.3558(0.0143)	0.5092(0.0307)	
KGD	$\hat{p}(s.e.)$	$\hat{a}(s.e.)$	$\hat{b}(s.e.)$
	0.0342(0.0050)	0.2170(0.0397)	0.3272(0.0167)

Table 2: Full and Sub model comparisons: (p-value)

	Wald’s Z-test	p-value
Hypothesis (I)	z=(0.5092-1)/0.0307=-15.99	< 0.0001*
Hypothesis (II)	z=(.3272-1)/0.0167=-40.29	< 0.0001*
Hypothesis (III)	z=(0.5142-0)/0.0461=11.16	< 0.0001*
Hypothesis (IV)	z=(6398.0-6194.8)=203.2	< 0.0001*

Table 3: Parameter estimates (standard errors in parentheses) for modeling DOCVISIT response variable

Variable x_i / z_i	ZIGPR		ZIEEGR		ZIKGR	
	$\hat{\beta}$	$\hat{\delta}$	$\hat{\beta}$	$\hat{\delta}$	$\hat{\beta}$	$\hat{\delta}$
Constant	-1.225(0.300)*	0.647(0.758)	-1.595(0.309)*	0.732(0.678)	1.670(1.978)	0.710(0.753)
Sex	0.018(0.085)	-0.581(0.230)*	-0.001(0.087)	-0.532(0.205)*	0.007(0.054)	-0.553(0.232)*
Age	2.010(1.572)	10.507(4.420)*	2.436(1.609)	10.29(3.923)*	1.146(1.027)	10.229(4.388)*
Agesq	-2.076(1.675)	-13.61(5.040)*	-2.643(1.712)	-13.35(4.457)*	-1.190(1.101)	-13.274(5.006)*
Income	-0.205(0.135)	-0.348(0.349)	-0.248(0.138)	-0.379(0.313)	0.129(0.088)	-0.371(0.353)
Levyplus	-0.107(0.117)	-0.662(0.267)*	-0.083(0.119)	-0.545(0.236)*	-0.080(0.075)	-0.675(0.277)*
Freepoor	-0.492(0.286)	0.094(0.666)	-0.489(0.294)	0.171(0.604)	-0.294(0.190)	0.073(0.732)
Freerepa	-0.194(0.143)	-1.383(0.451)*	-0.215(0.145)	-1.278(0.385)*	-0.129(0.094)	-1.378(0.445)*
Illness	0.050(0.030)	-0.683(0.159)*	0.049(0.031)	-0.544(0.119)*	0.027(0.020)	-0.626(0.160)*
Actdays	0.105(0.008)	-1.792(0.665)*	0.100(0.008)*	-1.494(0.380)*	0.075(0.007)*	-1.546(0.471)*
Hscore	0.024(0.014)	-0.104(0.056)	0.023(0.014)	-0.100(0.047)*	0.016(0.009)	-0.096(0.053)
Chcond1	0.001(0.110)	-0.117(0.282)	0.008(0.112)	-0.096(0.244)	0.019(0.070)	-0.061(0.285)
Chcond2	0.066(0.123)	-0.467(0.416)	0.042(0.125)	-0.513(0.366)	0.076(0.078)	1.251(0.403)
	$\hat{c}$ = 0.2767(0.035)*		$\hat{c}$ = 1.6945(0.143)*		$\hat{a}$ = 1.2513(0.104)*	
					$\hat{b}$ = 13.1000(26.579)	
LogL	–3103.69		–3097.38		-3094.95	
AIC	6261.4		6248.8		6245.9	
BIC	6438.3		6425.7		6429.4	
RMSE	0.7459		0.7406		0.7364	

Table 4: Model comparisons between ZIKGR and ZIEEGR using likelihood ratio test and Vuong Test for comparing ZIEEGR and ZIGPR

ZIKGR Vs. ZIEEGR	ZIEEGR Vs. ZIGPR
$\chi^2 = 4.86$ (p-value = 0.0275)*	z=0.9140 (p-value = 0.3707)

5. Conclusions

A method for developing generalized regression models for count data and zero-inflated count data is proposed using the *T*-geometric framework. Two models based on this method are investigated in detail. One is the EEGR model, which was first studied in Famoye (2016). The other is a generalization of the EEGR model, namely the KGR model. To illustrate the potential of this method, we analyze the 1977-1978 Australian health care data. The response variable, DOCVISIT is modeled using three different count regression models EEGR, KGR and GPR and their modifications of ZIEEGR, ZIKGR and ZIGPR. These three models can be used to model under-dispersed or over-dispersed count data. The models EEGR and KGR are more versatile than the NBR model, which can handle only over-dispersed data. For the application we analyze in this article, the ZIKGR performs the best, while the ZIEEGR performs equally well as the ZIGPR.The method proposed in this paper can be applied to develop other types of generalized regression models for discrete responses, which provide a broad selection of generalized modeling techniques in addition to the current available regression models for discrete data.

References

Alzaatreh, A., Lee, C. and Famoye, F. (2012) On the discrete analogues of continuous distributions. *Statistical Methodology*, 9, 589-603.

Alzaatreh, A., Lee, C. and Famoye, F. (2013) A new method for generating families of continuous distributions. *Metron*, 71(1), 63-79.

Basu, B. and Famoye, F. (2004) Domestic violence against women and their economic dependence: A count data analysis. *Review of Political Economy*, 16(4), 457-472.

Cameron, A.C. and Johansson, P. (1997) Count data regression using series expansion: With applications. *Journal of Applied Econometrics*, 12, 203-223.

Cameron, A.C. and Trivedi, P.K. (2013) *Regression Analysis of Count Data, 2nd Edition*, Cambridge University Press, Cambridge, UK.

Cameron, A.C., Trivedi, P.K., Milne, F. and Piggott, J. (1988) A microeconomic model of the demand for health care and health insurance in Australia. *Review of Econometric Studies*, LV, 85-106.

Consul, P.C. (1989) *Generalized Poisson Distributions: Properties and Applications*, Marcel Dekker, New York.

Famoye, F. (1993) Restricted generalized Poisson regression model. *Communications in Statistics-Theory and Methods*, 22(5), 1335-1354.

Famoye, F. (2010) On the bivariate negative binomial regression model. *Journal of Applied Statistics*, 37(6), 969-981.

Famoye, F. (2015) A multivariate generalized Poisson regression m *Communications in Statistics-Theory and Methods*, 44, 497-511.

Famoye, F. and Singh, K.P. (2006) Zero-inflated generalized Poisson regression model with applications to domestic violence data. *Journal of Data Science*, 4(1), 117-130.

Famoye, F. and Wang, W. (2004) Censored generalized Poisson regression model. *Journal of Computational Statistics and Data Analysis*, 46, 547-560.

Frome, E.L. (1983) The analysis of rates using Poisson regression models. *Biometrics*, 39, 665-674.

Frome, E.L., Kurtner, M.H. and Beauchamp, J.J. (1973) Regression analysis of Poisson-distributed data. *Journal of the American Statistical Association*, 68, 288-298.

Gupta, R.D. and Kundu, D. (2001) Exponentiated-exponential family: An alternative to gamma and Weibull distributions. *Biometrical Journal*, 43, 117-130.

Gurmu, S. and Elder, J. (2000) Generalized bivariate count data regression models. *Economics Letters*, 68, 31-36.

Hinde, J. (1982) Compound Poisson regression models. *GLIM 82: Proceedings International Conference on Generalized Linear Models* (R. Gilchrist ed.), Springer, Berlin, 109-121.

Holford, T.R. (1983) The estimation of age, period and cohort effects of vital rates. *Biometrics*, 39, 311-324.

Jorgenson, D.W. (1961) Multiple regression analysis of a Poisson process. *Journal of the American Statistical Association*, 56, 235-245.

Kendall, M.G. (1961) Natural law in the social sciences. *Journal of the Royal Statistical Society. Series A (General)*, 124, 1-19.

Kokonendji, C.C. (2014) Over- and underdispersion models. In N. Balakrishnan (Ed.), *Methods and Applications of Statistics in Clinical Trials, Volume 2: Planning, Analysis, and Inferential Methods* (pp. 506-526). John Wiley & Sons, Hoboken, NJ, USA.

Lambert, D. (1992) Zero-inflated Poisson regression, with an application to defects in manufacturing. *Technometrics*, 34, 1-14.

Lawless, J.F. (1987) Negative binomial and mixed Poisson regression, *Canadian Journal of Statistics*, 15(3), 209-225.

Lawless, J.F. (2003) *Statistical Models and Methods for Lifetime Data, 2nd edition*, John Wiley & Sons, Inc., Hoboken, NJ.

McCullagh, P. and Nelder, J.A. (1989) *Generalized Linear Models, 2nd edition*, Chapman and Hall, London.

Mullahy, J. (1997) Heterogeneity, excess zeros, and the structure of count data models. *Journal of Applied Econometrics*, 12, 337-350.

Terza, J.V. (1985) A tobit-type estimator for the censored Poisson regression model. *Economic Letters*, 18, 361-365.

Vuong, Q.H. (1989) Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica* 57(2), 307-333.

Xekalaki, E. (2014) On the distribution theory of over-dispersion. *Journal of Statistical Distributions and Applications*, *1*(19), 1-22.