



Zero to k Inflated Poisson Regression Models with Applications

Hadi Saboori¹ · Mahdi Doostparast¹

Received: 31 May 2023 / Accepted: 15 October 2023 / Published online: 13 November 2023
© The Author(s) 2023

Abstract

In the count data set, the frequency of some points may occur more than expected under the standard data analysis models. Indeed, in many situations, the frequencies of zero and of some other points tend to be higher than those of the Poisson. Adapting existing models for analyzing inflated observations has been studied in the literature. A method for modeling the inflated data is the inflated distribution. In this paper, we extend this inflated distribution. Indeed, if inflations occur in three or more of the support point, then the previous models are not suitable. We propose a model based on zero, one, \dots , and k inflated points with probabilities w_0, w_1, \dots , and w_k , respectively. By choosing the appropriate values for the weights w_0, \dots, w_k , various inflated distributions, such as the zero-inflated, zero–one-inflated, and zero– k -inflated distributions, are derived as special cases of the proposed model in this paper. Various illustrative examples and real data sets are analyzed using the obtained results.

Keywords Regression · Zero to k inflated distribution · Poisson distribution · Generalized linear model · Estimation

Mathematics Subject Classification Primary 62H10 · Secondary 62J12

1 Introduction

One of the important causes of overdispersion in the count data is an inflated number of zeros in excess of the number expected under the distribution. In such cases, one of the appropriate models is zero-inflated Poisson (ZIP) distribution. There are numerous papers in the literature dealing with the ZIP model. The earliest studies on the ZIP model were done by Cohen [6] and Yoneda [37]. Lambert [14] introduced

✉ Mahdi Doostparast
doustparast@um.ac.ir

¹ Department of Statistics, Ferdowsi University of Mashhad, P.O. Box 91775-1159, Mashhad, Iran

and studied the ZIP regression model using the Expectation-Maximization (EM) approach. Vandenbroek [34] gave a score test for testing a standard Poisson regression model. Jansakul and Hinde [11] extended this test to a more general situation where the zero probability depends on covariates. Ridout et al. [27] derived a score test for testing a ZIP regression model against the zero-inflated negative binomial alternatives. Agarwal et al. [1] applied the ZIP regression model for analyzing spatial count data sets. The score test using the normal approximation might underestimate the nominal significance level for small sample cases. Jung et al. [12] proposed a parametric bootstrap method for this problem. They observed that the ZIP regression model for prediction is more robust than the usual Poisson regression model. Long et al. [20] developed a marginalized ZIP regression model approach for independent responses to model the population mean count directly, allowing straightforward inference for overall exposure effects and derived an empirical robust variance estimation for overall incidence density ratios. Zhu et al. [40] have extended zero-inflated count models to account for random effects. Lim et al. [16] proposed the ZIP regression mixture model to account for both excess zeros and overdispersion caused by unobserved heterogeneity. A Bayesian latent factor ZIP model was proposed by Neelon and Chung [25] to analyze the molecular differences among breast cancer patients. Furthermore, ZIP models for censored data were studied by Saffari and Adnan [28] and Yang and Simpson [36]. Research on these models is still in progress. For instance, the ZIP model's parameters have been estimated using the Genetic Algorithm by Dikheel and Jouda [9] and the Shrinkage estimator by Zandi et al. [38]. ZIP models continue to be frequently utilized in many studies, even with the development of more comprehensive models [24, 32]. Empirical evidence shows that inflation may occur at more than one point. For example, Lin and Tsai [17] discussed a model that can be applied to both excessive zeros and ones known and called the zero–one inflated Poisson (ZOIP) model. Zhang et al. [39] initially studied the likelihood-based ZOIP model without covariates. When the covariates are available, it is essential to build a ZOIP regression model to clarify the relationship between the covariates and the response variable. Tang et al. [33], Liu et al. [18], Liu et al. [19], and Arora and Chaganty [5] studied the statistical inference for the ZOIP model. Melkersson and Rooth [23] proposed a zero–two inflated Poisson model, which accounts for a relative excess of both zero–two children in modeling complete female fertility.

In this article, we attempt to provide a generalization for an inflated regression model based on the Poisson distribution. For this purpose, we generalized the inflated points to $0, \dots, k$, for $k = 0, 1, 2, \dots$. This generalization provides various benefits for the modeling of inflated and non-inflated data. For example, this generalization includes all previous inflated regression models based on Poisson distribution. Also, this model provides a wide range of models to the researcher, who can choose the most appropriate ones according to the data. In summary, the originality of this paper lies in the development of the family of inflated Poisson distribution and its application to generalized linear models. This generalization is significant because it gives the researcher access to the entire family of Poisson-based inflated distributions (the most used family in discrete inflated distributions) at a single model. Thus researcher chooses the appropriate model for the data set by choosing

the appropriate value of k according to the inflation at each point or any number of points. Besides, with other choices of k and distribution weights, new models can also be introduced. Selecting the appropriate value for k might be a challenge when using the ZKIP models. An effective approach for selecting the correct value for k in the ZKIP models is described in order to address this issue. Section 7 provides an explanation of this algorithm along with useful examples.

Therefore, the rest of the paper is organized as follows. In Sect. 2, we introduce the zero to k inflated Poisson (ZKIP) distribution and some special cases. Estimation by the maximum likelihood (ML) method and standard errors of the ML estimates are outlined in Sect. 3. Section 4 deals with the regression ZKIP model and the ML estimation, EM algorithm, and hypotheses testing for this model. In Sect. 5, the method of the randomized quantile residual (RQR) for the adequacy of the proposed model is introduced. Simulations are conducted in Sect. 6 to assess the usefulness of this model. In Sect. 7, two real data sets are used to demonstrate the flexibility and superiority of the proposed model against the existing ones. Finally, in Sect. 8, conclusions are given.

2 Proposing the New Model

We can build the zero to k inflated distribution with introducing the function $g(y)$ as

$$g(y) = \begin{cases} w_0 + (1 - \eta(\mathbf{w}))f(0), & y = 0, \\ w_1 + (1 - \eta(\mathbf{w}))f(1), & y = 1, \\ \vdots & \vdots \\ w_k + (1 - \eta(\mathbf{w}))f(k), & y = k, \\ (1 - \eta(\mathbf{w}))f(y), & y > k, \end{cases} \quad (1)$$

where $f(y)$ is a discrete distribution, say Poisson, $0 \leq w_i, i = 0, \dots, k, \eta(\mathbf{w}) = \sum_{i=0}^k w_i$ and $0 \leq \eta(\mathbf{w}) \leq 1$. By substituting the Poisson probability mass function (PMF) $f(y) = \exp(-\lambda)\lambda^y/y!, y = 0, 1, \dots$, into Equation (1), we have

$$g(y; \mathbf{w}, \lambda) = \begin{cases} w_0 + (1 - \eta(\mathbf{w})) \exp(-\lambda), & y = 0, \\ w_1 + (1 - \eta(\mathbf{w})) \lambda \exp(-\lambda), & y = 1, \\ \vdots & \vdots \\ w_k + (1 - \eta(\mathbf{w})) \frac{\lambda^k}{k!} \exp(-\lambda), & y = k, \\ (1 - \eta(\mathbf{w})) \frac{\lambda^y}{y!} \exp(-\lambda), & y > k, \end{cases} \quad (2)$$

where $\mathbf{w} = (w_0, \dots, w_k)$. The discrete random variable Y with the PMF (2) is called to follow the ZKIP distribution and denoted by $Y \sim \text{ZKIP}(\mathbf{w}, \lambda)$. Here, some special cases of the family of ZKIP models are given as follows:

- $k = 0 \rightarrow$ ZIP distribution [9, 24, 38] and [32].
- $k = 1 \rightarrow$ zero–one inflated Poisson distribution [17, 19, 33, 39], and [13].
- $k = 2 \rightarrow$ zero–one–two inflated Poisson distribution [32].

- $k > 1, w_1 = w_2, \dots, w_{k-1} = 0 \rightarrow$ zero and k inflated Poisson distribution [5, 23, 30].

If $Y \sim \text{ZKIP}(y; \mathbf{w}, \lambda)$, then the moment generating function is obtained from Eq. (2) as

$$M_Y(t) = E[\exp(tY)] = \sum_{i=0}^k w_i \exp(it) + (1 - \eta(\mathbf{w})) \exp[\lambda(\exp(t) - 1)]. \quad (3)$$

Equation (3) gives the expectation and variance of the random variable Y as

$$E(Y) = \sum_{i=1}^k iw_i + (1 - \eta(\mathbf{w}))\lambda, \quad (4)$$

$$\text{Var}(Y) = \sum_{i=1}^k i^2w_i - \left(\sum_{i=1}^k iw_i \right)^2 - \lambda(1 - \eta(\mathbf{w})) \left[2 \sum_{i=1}^k iw_i - \lambda\eta(\mathbf{w}) - 1 \right]. \quad (5)$$

Figure 1 shows PMF of the ZKIP distribution for $\lambda = 1$ and some selected values of k and \mathbf{w} .

3 Estimation

Let $\mathbf{Y} = (Y_1, \dots, Y_n)$ be a random sample arising from the ZKIP distribution. In this section, we study the problem of estimating unknown parameter vector $\theta = (\mathbf{w}, \lambda)$ on the basis of \mathbf{Y} . Let

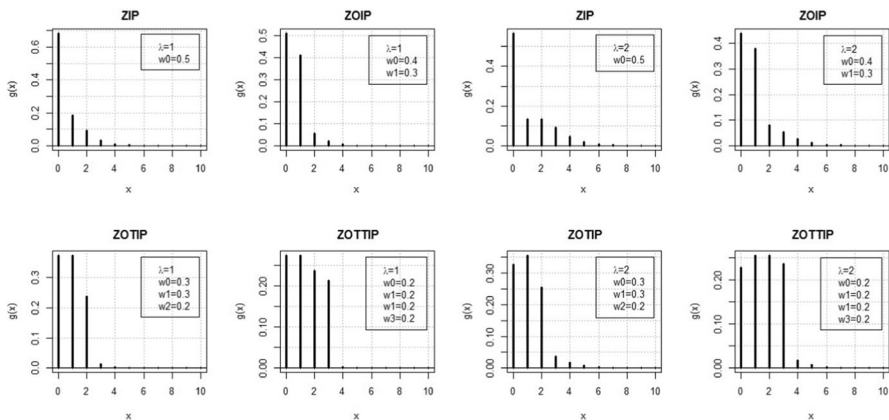


Fig. 1 PMF of the ZIP, ZOIP, ZOTIP, and ZOTTIP distributions, for $\lambda = 1$ and $\lambda = 2$

$$I_i(y) = \begin{cases} 1, & y = i, \\ 0, & y \neq i, \quad i = 0, \dots, k, \end{cases} \quad (6)$$

We can rewrite PMF (2) as

$$g(y; \mathbf{w}, \lambda) = \prod_{i=0}^k \left\{ w_i + (1 - \eta(\mathbf{w})) \frac{\lambda^y}{y!} \exp(-\lambda) \right\}^{I_i(y)} \\ \times \left\{ (1 - \eta(\mathbf{w})) \frac{\lambda^y}{y!} \exp(-\lambda) \right\}^{I(y>k)},$$

where $I(y > k) = 1 - \sum_{i=0}^k I_i(y)$. Thus, the likelihood function (LF) of the observed sample $\mathbf{Y} = \mathbf{y}$ reads

$$L_{obs}(\mathbf{y}; \mathbf{w}, \lambda) = \prod_{j=1}^n \left\{ \prod_{i=0}^k \left(w_i + (1 - \eta(\mathbf{w})) \frac{\lambda^{y_j}}{y_j!} \exp(-\lambda) \right)^{I_i(y_j)} \times \right. \\ \left. \left((1 - \eta(\mathbf{w})) \frac{\lambda^{y_j}}{y_j!} \exp(-\lambda) \right)^{I(y_j>k)} \right\}. \quad (7)$$

Then the logarithm of the LF (LLF) is

$$\ell_{obs}(\mathbf{w}, \lambda; \mathbf{y}) \propto \sum_{j=1}^n \left\{ \sum_{i=0}^k \left[I_i(y_j) \log \left(w_i + (1 - \eta(\mathbf{w})) \frac{\lambda^{y_j}}{y_j!} \exp(-\lambda) \right) \right] \right. \\ \left. + I(y_j > k) (\log(1 - \eta(\mathbf{w})) - \lambda + y_j \log(\lambda)) \right\}. \quad (8)$$

From Eq. (8), the ML estimates are derived by solving the following equations with respect to the parameters:

$$\frac{\partial \ell_{obs}(\mathbf{w}, \lambda; \mathbf{y})}{\partial w_l} = \sum_{j=1}^n \left\{ \sum_{i=0, i \neq l}^k \left[I_i(y_j) \times K_{1i}(y_j, \mathbf{w}, \lambda) \right] + I_l(y_j) \times K_{2l}(y_j, \mathbf{w}, \lambda) \right. \quad (9)$$

$$\left. - I(y_j > k) \left(\frac{1}{1 - \eta(\mathbf{w})} \right) \right\} = 0, \quad l = 0, \dots, k. \quad (10)$$

$$\frac{\partial \ell_{obs}(\mathbf{w}, \lambda; \mathbf{y})}{\partial \lambda} = \sum_{j=1}^n \left\{ \sum_{i=0}^k \left[I_i(y_j) \times K_{3i}(y_j, \mathbf{w}, \lambda) \right] + I(y_j > k) (y_j / \lambda - 1) \right\} = 0, \quad (11)$$

where

$$\begin{aligned}
 K_{1i}(y_j, \mathbf{w}, \lambda) &= \frac{-\frac{\lambda^{y_j}}{y_j!} \exp(-\lambda)}{w_i + (1 - \eta(\mathbf{w})) \frac{\lambda^{y_j}}{y_j!} \exp(-\lambda)}, \\
 K_{2i}(y_j, \mathbf{w}, \lambda) &= \frac{1 - \frac{\lambda^{y_j}}{y_j!} \exp(-\lambda)}{w_i + (1 - \eta(\mathbf{w})) \frac{\lambda^{y_j}}{y_j!} \exp(-\lambda)}, \\
 K_{3i}(y_j, \mathbf{w}, \lambda) &= -\frac{v_{1j}(\mathbf{w})}{w_i + (1 - \eta(\mathbf{w})) \frac{\lambda^{y_j}}{y_j!} \exp(-\lambda)}, \\
 v_{1j}(\mathbf{w}) &= \frac{(1 - \eta(\mathbf{w}))(\lambda - y_j) \exp(-\lambda) \lambda^{y_j-1}}{y_j!}.
 \end{aligned}$$

Since there is no closed-form solution for this system of equations, we may use numerical methods to estimate the parameters. To do this, we need to derive the elements of the observed information, which are given by

$$\begin{aligned}
 \frac{\partial^2 \ell_{obs}(\mathbf{w}, \lambda; \mathbf{y})}{\partial w_l^2} &= \sum_{j=1}^n \left\{ -\sum_{i=0, i \neq l}^k \left[I_i(y_j) \times K_{1i}^2(y_j, \mathbf{w}, \lambda) \right] \right. \\
 &\quad \left. - I_l(y_j) \times K_{2l}^2(y_j, \mathbf{w}, \lambda) - \frac{I(y_j > k)}{(1 - \eta(\mathbf{w}))^2} \right\}, \quad l = 0, \dots, k, \\
 \frac{\partial^2 \ell_{obs}(\mathbf{w}, \lambda; \mathbf{y})}{\partial w_l \partial w_s} &= \sum_{j=1}^n \left\{ \sum_{i=0, i \neq l, s}^k \left[I_i(y_j) \times K_{1i}^2(y_j, \mathbf{w}, \lambda) \right] + I_l(y_j) \right. \\
 &\quad \times \frac{v_{3j}(\mathbf{w}, \lambda)(1 - v_{3j}(\mathbf{w}, \lambda))}{(v_{2ij}(\mathbf{w}, \lambda))^2} \\
 &\quad \left. + I_s(y_j) \times \frac{v_{3j}(\mathbf{w}, \lambda)(1 - v_{3j}(\mathbf{w}, \lambda))}{(v_{2sj}(\mathbf{w}, \lambda))^2} - \frac{I(y_j > k)}{(1 - \eta(\mathbf{w}))^2} \right\}, \quad l, s = 0, \dots, k, \\
 \frac{\partial^2 \ell_{obs}(\mathbf{w}, \lambda; \mathbf{y})}{\partial w_l \partial \lambda} &= \sum_{j=1}^n \sum_{i=0, i \neq l}^k I_i(y_j) \frac{\lambda^{y_j-1} \exp(-\lambda)(\lambda - y_j) w_i}{y_j! (v_{2ij}(\mathbf{w}, \lambda))^2} \\
 &\quad + I_l(y_j) \frac{\lambda^{y_j-1} \exp(-\lambda)(\lambda - y_j)(w_l + 1 - \eta(\mathbf{w}))}{y_j! (v_{2ij}(\mathbf{w}, \lambda))^2} \quad l = 0, \dots, k, \\
 \frac{\partial^2 \ell_{obs}(\mathbf{w}, \lambda; \mathbf{y})}{\partial \lambda^2} &= \sum_{j=1}^n \left\{ \sum_{i=0}^k \left[I_i(y_j) \times \frac{v_{1j}(\mathbf{w}) \exp(-\lambda)(\lambda - x_j)(1 - \eta(\mathbf{w}))}{(v_{2ij}(\mathbf{w}, \lambda))^2 \times y_j!} \right] \right. \\
 &\quad \left. - I(y_j > k)(y_j / \lambda^2) \right\},
 \end{aligned}$$

where

$$v_{2ij}(\mathbf{w}, \lambda) = w_i + (1 - \eta(\mathbf{w})) \frac{\lambda^{y_j}}{y_j!} \exp(-\lambda),$$

and

$$v_{3j}(\mathbf{w}, \lambda) = \frac{\lambda^{y_j}}{y_j!} \exp(-\lambda).$$

For the problems of interval estimation and testing statistical hypotheses, the Fisher information (FI) matrix is useful. The FI matrix $(\mathbf{I}(\Theta) = \mathbf{I}_{ij}, i, j = 0, \dots, k + 1)$, obtained from the observed information matrix by taking the expected values of each entry, is given as follows:

$$\begin{aligned} \mathbf{I}_{ll} &= -E\left(\frac{\partial^2 \ell_{obs}(\mathbf{w}, \lambda; \mathbf{y})}{\partial w_l^2}\right) \\ &= n \sum_{i=0, i \neq l}^k [K_{1i}^2(i, \mathbf{w}, \lambda) \times g(i; \mathbf{w}, \lambda)] + nK_{2l}^2(l, \mathbf{w}, \lambda) \times g(l; \mathbf{w}, \lambda) \\ &\quad + nPr(Y > k)/(1 - \eta(\mathbf{w}))^2, \quad l = 0, \dots, k. \end{aligned}$$

$$\begin{aligned} \mathbf{I}_{ls} &= -E\left(\frac{\partial^2 \ell_{obs}(\mathbf{w}, \lambda; \mathbf{y})}{\partial w_l \partial w_s}\right) \\ &= -n \sum_{i=0, i \neq l, s}^k [K_{1i}^2(i, \mathbf{w}, \lambda) \times g(i; \mathbf{w}, \lambda)] \\ &\quad - \frac{nv_{3l}(\mathbf{w}, \lambda)(1 - v_{3l}(\mathbf{w}, \lambda))}{(v_{2ll}(\mathbf{w}, \lambda))^2} \times g(l; \mathbf{w}, \lambda) \\ &\quad - \frac{nv_{3s}(\mathbf{w}, \lambda)(1 - v_{3s}(\mathbf{w}, \lambda))}{(v_{2ss}(\mathbf{w}, \lambda))^2} \times g(s; \mathbf{w}, \lambda) \\ &\quad + \frac{nPr(Y > k)}{(1 - \eta(\mathbf{w}))^2}, \quad l, s = 0, \dots, k, \text{ and } l \neq s, \end{aligned}$$

$$\begin{aligned} \mathbf{I}_{l(k+1)} &= \mathbf{I}_{(k+1)s} = -E\left(\frac{\partial^2 \ell_{obs}(\mathbf{w}, \lambda; \mathbf{y})}{\partial w_l \partial \lambda}\right) \\ &= -n \sum_{i=0, i \neq l}^k \left[\frac{n\lambda^{i-1} \exp(-\lambda)(\lambda - i)w_i}{i!(v_{2ii}(\mathbf{w}, \lambda))^2} \times g(i; \mathbf{w}, \lambda) \right] \\ &\quad - \frac{n\lambda^{l-1} \exp(-\lambda)(\lambda l)(w_l + 1 - \eta(\mathbf{w}))}{l!(v_{2il}(\mathbf{w}, \lambda))^2} \\ &\quad \times g(l; \mathbf{w}, \lambda), \quad l, s = 0, \dots, k, \end{aligned}$$

$$\begin{aligned} \mathbf{I}_{(k+1)(k+1)} &= -E\left(\frac{\partial^2 \ell_{obs}(\mathbf{w}, \lambda; \mathbf{y})}{\partial \lambda^2}\right) \\ &= -n \sum_{i=0}^k \left[\frac{v_{1i}(\mathbf{w}) \exp(-\lambda)(\lambda - i)(1 - \eta(\mathbf{w}))}{(v_{2ii}(\mathbf{w}, \lambda))^2 i!} \times g(i; \mathbf{w}, \lambda) \right] \\ &\quad + \frac{nPr(Y > k)}{\lambda^2}. \end{aligned}$$

If θ is the arbitrary parameter of model and $\hat{\theta}$ is the ML estimates for this parameter, then from Lehmann et al. [15], we have

$$n^{1/2}(\hat{\theta} - \theta) \xrightarrow{d} N(0, I^{-1}(\theta)). \tag{12}$$

Therefore, we can use asymptotic distribution (12) to construct an asymptotic confidence interval and hypotheses testing for the parameters of the proposed model (see Lehmann et al. [15]) for more details).

4 ZKIP Regression Model

In Sect. 2, we introduce and motivate the ZKIP distribution. Recall that the PMF of the ZKIP distribution is

$$g(y; \mathbf{w}, \lambda) = \begin{cases} w_y + (1 - \eta(\mathbf{w}))P_y(\lambda), & y = 0, \dots, k, \\ (1 - \eta(\mathbf{w}))P_y(\lambda), & y > k, \end{cases} \tag{13}$$

where $P_\lambda(y) = \exp(-\lambda)\lambda^y/y!$ for $y = 0, 1, \dots$. So the ZKIP is a distribution with $k + 2$ parameters $\mathbf{w} = (w_0, \dots, w_k)$ and λ . In fact, it is a mixture of $k + 2$ distributions. The first distribution is degenerate at zero with weight w_0 , and the second is degenerate at one with weight w_1 , and the same goes for up to k . Finally, the $(k + 2)$ th distribution is the Poisson with mean λ with weight $(1 - \eta(\mathbf{w}))$.

Suppose that we have a vector $\mathbf{y} = (y_1, \dots, y_n)$ of n independent count responses from a ZKIP distribution. We assume that, associated with each y_i , the vector of covariates (\mathbf{x}_i^T) has been observed. The layout of the observed data is shown in Table 1.

From (13), the LF of the available data in Table 1 can be rewritten as

$$L_{obs}(\mathbf{w}, \lambda | \mathbf{y}) = \prod_{j=1}^n \left\{ \prod_{i=0}^k \left([w_i + [1 - \eta(\mathbf{w})]P_{\lambda_j}(y_j)]^{I(y_j)} \right) \times \left([1 - \eta(\mathbf{w})]P_{\lambda_j}(y_j) \right)^{I(y_j > k)} \right\}, \tag{14}$$

where $\lambda = (\lambda_1, \dots, \lambda_n)$ and $P_{\lambda_j}(y_j) = \exp(-\lambda_j)\lambda_j^{y_j}/y_j!$, $(y_j \geq 0)$. To connect the parameters with the covariates, we follow the standard generalized linear model (GLM) framework for the multinomial distribution; see Agresti [2] for further reading. The $k + 2$ mixing distributions can be viewed as $k + 2$ nominal categories. Thus, the probabilities of the $k + 2$ (degenerate(0), degenerate(1), ..., degenerate(k), Poisson) categories are $w_0, w_1, \dots, w_k, (1 - \eta(\mathbf{w}))$, respectively. Following the GLM baseline category logit model for the multinomial, let

$$\gamma_0 = \log \left(\frac{w_0}{1 - \eta(\mathbf{w})} \right), \gamma_1 = \log \left(\frac{w_1}{1 - \eta(\mathbf{w})} \right), \dots, \gamma_k = \log \left(\frac{w_k}{1 - \eta(\mathbf{w})} \right). \tag{15}$$

Here, we treat the Poisson distribution as the baseline category, and thus we have $((k + 2) - 1) = k + 1$ equations for the other $k + 1$ categories. As in log-linear models, the ZKIP regression model assumes that the Poisson parameter λ_i is a loglinear function of the covariates; that is,

$$\log(\lambda_j) = \mathbf{x}_j^T \boldsymbol{\beta}, \quad j = 1, \dots, n,$$

where $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^T$ is a p -dimensional unknown regression parameters vector. For the sake of brevity, we assume that the parameters vector $\boldsymbol{\gamma} = (\gamma_0, \dots, \gamma_k)$ is constant. The generalization where these $k + 1$ parameters are functions of the covariates is straightforward. Thus, the parameters of the ZKIP regression model are $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$. In what follows, the problems of estimating parameters and hypotheses testing are discussed in detail.

4.1 Estimation of Regression Parameters

In the following, we study methods for estimating the parameters of the ZKIP regression model. Two popular methods are the ML and EM methods. The ML method involves optimizing the LF (14) or the logarithm of the LF with respect to the unknown parameters $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$. By substituting the reparameterizations (15) into the LF (14), we can rewrite (8) as follows:

$$\begin{aligned} \ell_{obs}(\boldsymbol{\beta}, \boldsymbol{\gamma}) &= \log L_{obs}(\boldsymbol{\beta}, \boldsymbol{\gamma} | \mathbf{y}) \\ &= \sum_{j=1}^n \left\{ \sum_{i=0}^k \left[I_i(y_j) \log \left(\left[\exp(\gamma_i) + P_{y_j}(\lambda_j) \right] \gamma^* \right) \right] \right. \\ &\quad \left. + I(y_j > k) \log \left(P_{y_j}(\lambda_j) \right) - \log \left(1 + \sum_{l=0}^k \exp(\gamma_l) \right) \right\}, \end{aligned} \tag{16}$$

where $\log(\lambda_j) = \mathbf{x}_j^T \boldsymbol{\beta}$ and $\gamma^* = 1 / \{1 + \sum_{l=0}^k \exp(\gamma_l)\}$. The ML estimates can be obtained by maximizing the LLF (16) directly with respect to the parameters. Alternatively, one can take the partial derivatives of the LLF and solve the $(k + 2)$ score equations as follows:

Table 1 Form of data for regression analyses on inflated ZKIP distribution

Observation	Response	Covariates
1	y_1	$x_{11} \dots x_{1p}$
2	y_2	$x_{21} \dots x_{2p}$
\vdots	\vdots	\vdots
i	y_i	$x_{i1} \dots x_{ip}$
\vdots	\vdots	\vdots
n	y_n	$x_{n1} \dots x_{np}$

$$\sum_{j=1}^n \left(I_i(y_j) \times \frac{\exp(\gamma_i)}{\exp(\gamma_i) + P_{y_j}(\lambda_j)} \right) = \frac{n \exp(\gamma_i)}{1 + \sum_{l=0}^k \exp(\gamma_l)}, \quad i = 0, \dots, k,$$

$$\sum_{j=1}^n \left\{ \sum_{i=0}^k \left(I_i(y_j) \times \frac{P_{y_j}(\lambda_j)(y_j - \lambda_j)/\lambda_j}{\exp(\gamma_i) + P_{y_j}(\lambda_j)} \right) + I(y_j > k) \times \frac{P_{y_j}(\lambda_j)(y_j - \lambda_j)/\lambda_j}{P_j(\lambda_j)} \right\} = 0.$$

An alternative and popular method for parameter estimation is the EM approach. The EM approach treats the observed data $\mathbf{y} = (y_1, \dots, y_n)$ as a part of complete data that includes $\mathbf{z} = (\mathbf{z}_1, \dots, \mathbf{z}_n)$, which is regarded as missing. Here each $\mathbf{z}_i = (z_{i0}, \dots, z_{i(k+1)})$ is a $k + 2$ component vector with PMF (17). By the definition of latent variable \mathbf{z} , we have

$$Pr(\mathbf{z} = (z_0, \dots, z_{k+1})) = \begin{cases} w_0, & \mathbf{z} = (1, 0, \dots, 0), \\ w_1, & \mathbf{z} = (0, 1, 0, \dots, 0), \\ \vdots & \vdots \\ w_k, & \mathbf{z} = (0, 0, 0, \dots, 1, 0), \\ w_{k+1}, & \mathbf{z} = (0, 0, 0, \dots, 0, 1), \end{cases} \tag{17}$$

where $w_{k+1} = (1 - \eta(\mathbf{w}))$.

Then, the conditional PMF of y_i given \mathbf{z}_i is calculated. Thus, the joint distribution of the observed and missing data, which is given by the ZKIP distribution, is a mixture of Poisson with $k + 1$ degenerate distributions at zero to k . Consider a latent variable $\mathbf{z} = (z_0, \dots, z_{k+1})$, which is distributed as a multinomial with parameters $(1, w_0, \dots, w_{k+1})$. Note that \mathbf{z} takes values $(0, \dots, 0, 1_i, 0, \dots, 0)$ with probability $w_i, i = 0, \dots, (k + 1)$. That is,

$$Pr(y_j, \mathbf{z}_j) = \begin{cases} w_0, & z_{0j} = 1, y_j = 0, \\ w_1, & z_{1j} = 1, y_j = 1, \\ \vdots & \vdots \\ w_k, & z_{kj} = 1, y_j = k, \\ w_{k+1} \times \frac{\exp(-\gamma_j)\lambda_j^{y_j}}{y_j!}, & z_{(k+1)j} = 1, y_j \geq 0. \end{cases} \tag{18}$$

Thus the conditional distribution of Y given $\mathbf{z} = (z_0, \dots, z_{k+1})$ is

$$Pr(Y = y | \mathbf{z} = (z_0, \dots, z_{k+1})) = \begin{cases} 1, & z_0 = 1, y = 0, \\ 1, & z_1 = 1, y = 1, \\ \vdots & \vdots \\ 1, & z_k = 1, y = k, \\ \frac{\exp(-\lambda)\lambda^y}{y!}, & z_{k+1} = 1, y \geq 0. \end{cases} \tag{19}$$

Finally, the joint PMF of (Y, \mathbf{z}) is obtained from (17) and (19) as

$$\begin{aligned}
 &Pr\left(Y = y, \mathbf{z} = (z_0, \dots, z_{k+1})\right) \\
 &= Pr\left(Y = y | \mathbf{z} = (z_0, \dots, z_{k+1})\right) \times Pr\left(\mathbf{z} = (z_0, \dots, z_{k+1})\right) \\
 &= \begin{cases} w_0, & z_0 = 1, y = 0, \\ w_1, & z_1 = 1, y = 1, \\ \vdots & \vdots \\ w_k, & z_k = 1, y = k, \\ w_{k+1} \times \frac{\exp(-\lambda)^y}{y!}, & z_{k+1} = 1, y \geq 0. \end{cases} \tag{20}
 \end{aligned}$$

Therefore from (20), the complete LF of the ZKIP model is

$$L_{comp}(\mathbf{w}, \lambda; \mathbf{y}, \mathbf{z}) = \prod_{j=1}^n \left\{ \prod_{i=0}^k \left[w_i^{z_{ij} \times I_i(y_j)} \right] \times \left(w_{k+1} P_{y_j}(\lambda_j) \right)^{z_{(k+1)j}} \right\}. \tag{21}$$

By (15), the LLF of the complete data (\mathbf{y}, \mathbf{z}) reads

$$\begin{aligned}
 \ell_{comp}(\mathbf{w}, \lambda; \mathbf{y}, \mathbf{z}) &= \log L_{comp}(\mathbf{w}, \lambda; \mathbf{y}, \mathbf{z}) \\
 &= \sum_{j=1}^n \left\{ \sum_{i=0}^k z_{ij} I_i(y_j) \left[\gamma_i - \log\left(1 + \sum_{l=0}^k e^{\gamma_l}\right) \right] + z_{(k+1)j} \log w_{k+1} \right. \\
 &\quad \left. + z_{(k+1)j} \log P_{y_j}(\lambda_j) \right\}. \tag{22}
 \end{aligned}$$

For $w_1 = \dots = w_k = 0$, the ZKIP is reduced to the ZIP model. From (22), the LLF of the ZIP for the complete data is derived as

$$\begin{aligned}
 \ell_{comp}(\mathbf{w}, \lambda; \mathbf{y}, \mathbf{z}_1) &= \sum_{j=1}^n \left\{ z_{0j} [\gamma_0 - \log(1 + \exp(\gamma_0))] \right. \\
 &\quad \left. + (1 - z_{0j}) \log w_1 + (1 - z_{0j}) \log P_{y_j}(\lambda_j) \right\}.
 \end{aligned}$$

Lambert [14] used Eq. (22) as the LLF of the complete data for the ZIP model to obtain the EM estimates.

We now proceed to describe the EM algorithm (Dempster et al. [8]);, Wu [35] for the ZKIP model. The first step in the EM algorithm involves selecting some initial values for the unknown parameters. The choice of the initial values is important for the convergence of the algorithm. An incorrect choice of the initial values could result in slow convergence or breakdown of the algorithm. We recommend using the proportions of zeros, ..., k 's, respectively, in the observed data as initial values for the parameters w_0, \dots, w_k and use the relations (15) to get initial values $\gamma_{00}, \dots, \gamma_{k0}$, for the parameters $\gamma_0, \dots, \gamma_k$, respectively. The next step involves filling the latent values \mathbf{z}_i by their expectations, which is the E-step. We

will use the conditional expected values of $E(\mathbf{z}|\mathbf{y})$ given in Table 2 to generate \mathbf{z}_i 's.

We use Table 2 to estimate the missing values in the expectation step of the EM algorithm as follows:

$$\hat{z}_{ij} = E(z_{ij}|y_j = i) = \frac{\exp(\gamma_i)}{\exp(\gamma_i) + P_i(\lambda_j)} \quad (23)$$

$$\text{and } \hat{z}_{ij} = E(z_{ij}|y_j \neq i) = 0, \quad i = 0, \dots, k, j = 1, \dots, n. \quad (24)$$

For the maximization step in the EM algorithm, we solve the following score equations instead of maximizing the complete LF directly:

$$\frac{\partial \ell_{comp}}{\partial \boldsymbol{\beta}} = \sum_{j=1}^n \hat{z}_{(k+1)j} (y_j - \exp[\mathbf{x}_j^T \boldsymbol{\beta}]) \mathbf{x}_j^T = 0, \quad (25)$$

$$\frac{\partial \ell_{comp}}{\partial \gamma_i} = \sum_{j=1}^n \hat{z}_{ij} I_i(y_j) - \frac{n \exp(\gamma_i)}{1 + \sum_{l=0}^k \exp(\gamma_l)} = 0, \quad i = 0, \dots, k, \quad (26)$$

where $\hat{z}_{(k+1)j} = 1 - \sum_{i=0}^k \hat{z}_{ij}$. In summary, the EM algorithm to estimate the parameters $(\gamma_0, \dots, \gamma_k)$ and the regression parameter $\boldsymbol{\beta}$ for the ZKIP regression model is as follows.

1. Select initial values $\boldsymbol{\beta}_0, (\gamma_{00}, \dots, \gamma_{k0})$ for the parameters $\boldsymbol{\beta}$ and the vector of parameters $(\gamma_0, \dots, \gamma_k)$, respectively.
2. **E-step:** Estimate \hat{z}_{ij} 's, $i = 0, \dots, k$ using Eqs. (23) and (24).
3. **M-step:** Solve Eqs. (25) and (26) and obtain updated estimates $\boldsymbol{\beta}_1, (\gamma_{01}, \dots, \gamma_{k1})$.
4. Repeat E-step and M-step until the parameter estimates converge.

In the next subsection, we discuss how to obtain the standard errors of the estimates obtained by the EM algorithm.

4.2 Standard Errors for the EM Algorithm

The most commonly used method to get the standard errors in the mixture models is to compute the matrix of partial derivatives of the LLF for the observed data, that is, to calculate the information matrix from the observed data. Lambert [14] used this method for computing the standard errors for the ZIP regression model. Lin and Tsai [17] used the Hessian matrix to get the standard errors for the zero- k inflated Poisson model without actually computing second-order partial derivatives of the LLF. Recall that the Hessian matrix comes out as a byproduct of the nonlinear optimization methods in the common statistical packages.

Table 2 $E(\mathbf{z}|\mathbf{y})$ for the ZKIP regression model

\mathbf{z}	$y = 0$	$y = 1$...	$y = l$...	$y = k$	$y \geq k$
z_0	$\frac{\exp(\gamma_0)}{\exp(\gamma_0)+P_0(\lambda_j)}$	0	...	0	...	0	0
z_1	0	$\frac{\exp(\gamma_1)}{\exp(\gamma_1)+P_1(\lambda_j)}$...	0	...	0	0
\vdots	\vdots	\vdots	...	\vdots	...	\vdots	\vdots
z_l	0	0	...	$\frac{\exp(\gamma_l)}{\exp(\gamma_l)+P_l(\lambda_j)}$...	0	0
\vdots	\vdots	\vdots	...	\vdots	...	\vdots	\vdots
z_k	0	0	...	0	...	$\frac{\exp(\gamma_k)}{\exp(\gamma_k)+P_k(\lambda_j)}$	0
z_{k+1}	$\frac{P_0(\lambda_j)}{\exp(\gamma_0)+P_0(\lambda_j)}$	$\frac{P_1(\lambda_j)}{\exp(\gamma_1)+P_1(\lambda_j)}$...	$\frac{P_l(\lambda_j)}{\exp(\gamma_l)+P_l(\lambda_j)}$...	$\frac{P_k(\lambda_j)}{\exp(\gamma_k)+P_k(\lambda_j)}$	1

To compute the standard errors of the estimates obtained by the EM algorithm, we follow the approach described by Louis [22]. The relation between the likelihood of the complete, observed, and missing data is given by

$$L_{comp}(\mathbf{w}, \lambda|\mathbf{y}, \mathbf{z}) = L_{obs}(\mathbf{w}, \lambda|\mathbf{y})L_{miss}(\mathbf{w}, \lambda|(\mathbf{z}|\mathbf{y})), \tag{27}$$

where \mathbf{y} and \mathbf{z} stand for the observed and missing data, respectively. By taking the logarithm from Equation (27), we get

$$\ell_{obs}(\mathbf{w}, \lambda|\mathbf{y}) = \ell_{comp}(\mathbf{w}, \lambda|\mathbf{y}, \mathbf{z}) - \ell_{miss}(\mathbf{w}, \lambda|(\mathbf{z}|\mathbf{y})), \tag{28}$$

where $\ell_{obs}(\mathbf{w}, \lambda|\mathbf{y}) = \log \{L_{obs}(\mathbf{w}, \lambda|\mathbf{y})\}$, $\ell_{comp}(\mathbf{w}, \lambda|\mathbf{y}, \mathbf{z}) = \log \{L_{comp}(\mathbf{w}, \lambda|\mathbf{y}, \mathbf{z})\}$ and $\ell_{miss}(\mathbf{w}, \lambda|(\mathbf{z}|\mathbf{y})) = \log \{L_{miss}(\mathbf{w}, \lambda|(\mathbf{z}|\mathbf{y}))\}$. Taking second-order partial derivatives from Eq. (28), the information matrices for the complete, observed, and missing data satisfy the following identity:

$$I_{obs} = I_{comp} - I_{miss}, \tag{29}$$

where the matrices I_{comp} and I_{miss} are, respectively, the Hessian matrix of the LLF of complete data and the LLF of missing data. Since the right hand side of Eq. (29) depends on the missing data, Louis [22] suggested to take the expected value of the missing data given the observed. This gives us the identity

$$I_{obs} = E(I_{obs}|\mathbf{y}) = E(I_{comp}|\mathbf{y}) - E(I_{miss}|\mathbf{y}). \tag{30}$$

In other words, Louis [22] estimated of the observed information matrix by

$$\hat{I}_{obs} = E(I_{obs}|\mathbf{y}) = E(I_{comp}|\mathbf{y}) - E(I_{miss}|\mathbf{y}). \tag{31}$$

Note that the LLF of the complete data for the ZKIP regression model is given by Eq. (22), and the corresponding first-order derivatives are shown in Eqs. (25) and (26). The elements of the matrix $E(I_{comp}|\mathbf{y})$ are the expected values of the negative of second-order partial derivatives of the complete LLF (22) given by

$$\begin{aligned}
 E \left[-\frac{\partial^2 \ell_{comp}}{\partial \gamma_i^2} \right] &= n \times \frac{\exp(\gamma_i)(1 + \sum_{l=0}^k \exp(\gamma_l)) - \exp(2\gamma_i)}{(1 + \sum_{l=0}^k \exp(\gamma_l))^2}, \quad i = 0, \dots, k, \\
 E \left[-\frac{\partial^2 \ell_{comp}}{\partial \gamma_l \partial \gamma_s} \right] &= n \times \frac{\exp(\gamma_l)(1 + \sum_{l=0}^k \exp(\gamma_l)) - \exp(\gamma_l + \gamma_s)}{(1 + \sum_{m=0}^k \exp(\gamma_m))^2}, \quad l, s = 0, \dots, k, \\
 E \left[-\frac{\partial^2 \ell_{comp}}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^T} \right] &= \sum_{j=1}^n \frac{[P_0(\lambda_j) \times \dots \times P_k(\lambda_j) - \exp(\gamma_0 + \dots + \gamma_k)] \lambda_j}{[\exp(\gamma_0) + P_0(\lambda_j)] \times \dots \times [\exp(\gamma_k) + P_k(\lambda_j)]} (\mathbf{x}_j \mathbf{x}_j^T).
 \end{aligned}$$

To see this, from Eq. (15), we have

$$1 - \eta(\mathbf{w}) = w_{k+1} = \frac{1}{1 + \sum_{l=0}^k \gamma_l},$$

and then Eq. (22) yields

$$\frac{\partial \ell_{comp}}{\partial \gamma_i} = \sum_{j=1}^n \left\{ z_{ij} I_i(y_j) \left[1 - \frac{e^{\gamma_i}}{1 + \sum_{l=0}^k e^{\gamma_l}} \right] + z_{(k+1)j} \frac{e^{\gamma_i}}{1 + \sum_{l=0}^k e^{\gamma_l}} \right\}. \tag{32}$$

From Eq. (32) we conclude that

$$\begin{aligned}
 \frac{\partial^2 \ell_{comp}}{\partial \gamma_i^2} &= \sum_{j=1}^n \left\{ z_{ij} I_i(y_j) \left[\frac{\exp(\gamma_i)(1 + \sum_{l=0}^k \exp(\gamma_l)) - \exp(2\gamma_i)}{(1 + \sum_{l=0}^k \exp(\gamma_l))^2} \right] \right. \\
 &\quad \left. + z_{(k+1)j} \frac{\exp(\gamma_i)(1 + \sum_{l=0}^k \exp(\gamma_l)) - \exp(2\gamma_i)}{(1 + \sum_{l=0}^k \exp(\gamma_l))^2} \right\}.
 \end{aligned} \tag{33}$$

After some algebraic manipulations and the value of $E(\mathbf{z}|\mathbf{y})$ from Table 2, the following result is concluded

$$E \left[-\frac{\partial^2 \ell_{comp}}{\partial \gamma_i^2} \right] = n \times \frac{\exp(\gamma_i)(1 + \sum_{l=0}^k \exp(\gamma_l)) - \exp(2\gamma_i)}{(1 + \sum_{l=0}^k \exp(\gamma_l))^2}. \tag{34}$$

Similarly, $E[-\partial^2 \ell_{comp} / \partial \gamma_l \partial \gamma_s]$ and $E[-\partial^2 \ell_{comp} / \partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^T]$ are obtained. The LLF of the missing data for the ZKIP regression model is

$$\begin{aligned}
 \ell_{miss}(\boldsymbol{\beta}, \boldsymbol{\gamma}) &= \sum_{j=1}^n \left\{ \sum_{i=0}^k [z_{ij} \gamma_i] + z_{(k+1)j} \log P_{y_j}(\lambda_j) \right\} \\
 &\quad - \sum_{j=1}^n \left\{ \sum_{i=0}^k I_i(y_j) \log \left(\exp(\gamma_i) + P_{y_j}(\lambda_j) \right) \right\} - \sum_{j=1}^n \left(I(y_j > k) \log P_{y_j}(\lambda_j) \right).
 \end{aligned} \tag{35}$$

Finally, the elements of the matrix $E(I_{miss}|\mathbf{y})$ are the negative of the expected value of the second-order derivatives of (35). These are given by

$$\begin{aligned}
 E \left[-\frac{\partial^2 \ell_{miss}}{\partial \gamma_i^2} \right] &= \sum_{j=1}^n \left\{ I_j(y_j) \frac{\exp(\gamma_j) P_{y_j}(\lambda_j)}{(\exp(\gamma_j) + P_{y_j}(\lambda_j))^2} \right\}, \quad i = 0, \dots, k, \\
 E \left[-\frac{\partial^2 \ell_{miss}}{\partial \gamma_l \partial \gamma_s} \right] &= 0, \quad l, s = 0, \dots, k \text{ and } l \neq s, \\
 E \left[-\frac{\partial^2 \ell_{miss}}{\partial \beta \partial \beta^T} \right] &= \sum_{j=1}^n \left\{ I_j(y_j) \frac{\exp(\gamma_j) P_{y_j}(\lambda_j)}{(\exp(\gamma_j) + P_{y_j}(\lambda_j))^2} \right\}, \quad i = 0, \dots, k, \\
 E \left[-\frac{\partial^2 \ell_{miss}}{\partial \beta \partial \beta^T} \right] &= \sum_{j=1}^n \frac{[P_0(\lambda_j) \times \dots \times P_k(\lambda_j) - \exp(\gamma_0 + \dots + \gamma_k)] \lambda_j \mathbf{x}_j \mathbf{x}_j^T}{[\exp(\gamma_0) + P_0(\lambda_j)] \times \dots \times [\exp(\gamma_k) + P_k(\lambda_j)]} \\
 &\quad - \sum_{j=1}^n \left\{ \sum_{i=0}^k I_i(y_j) \frac{[\exp(\gamma_i) P_{y_j}(\lambda_j) (\lambda_j - (y_j - \lambda_j)^2) + P_{y_j}^2(\lambda_j) \lambda_j] \mathbf{x}_j \mathbf{x}_j^T}{(\exp(\gamma_j) + P_{y_j}(\lambda_j))^2} \right\}.
 \end{aligned}$$

4.3 Hypothesis Testing

Testing the impact of the j th covariate on the count response is equivalent to testing $H_0 : \beta_j = 0$ vs. $H_1 : \beta_j \neq 0$. This hypothesis test is straightforward and can be done using the standard Wald statistic, $z = \hat{\beta}_j / SE(\hat{\beta}_j)$, which has asymptotically standard normal distribution under the null hypothesis H_0 . Here, $SE(\hat{\beta}_j)$ minus twice the LLLF stands for the standard error of the estimate $\hat{\beta}_j$. An alternative approach for testing $H_0 : \beta_j = 0$ is to use the generalized likelihood ratio test statistic defined by

$$-2 \log \Lambda = -2 \log \frac{L_{obs}(\tilde{\beta}, \tilde{\gamma}, \beta_j = 0)}{L_{obs}(\hat{\beta}, \hat{\gamma})}, \tag{36}$$

which has asymptotically the chi-Square distribution with one degree of freedom, where $\tilde{\beta}$ and $\tilde{\gamma}$ denote the ML estimates under the hypothesis $\beta_j = 0$ and $\hat{\beta}$ and $\hat{\gamma}$ are the ML estimates under the ZKIP model.

Since $0 \leq w_i \leq 1$, ($i = 0, \dots, k$), the null hypothesis $H_0 : w_{i_1} = \dots = w_{i_r} = 0$, $0 \leq i_r \leq k$, corresponds to testing the parameters to check the necessity of their presence in the model. In this case, the regularity conditions are not met. That is, the standard asymptotic theory for the LRT statistic (36) is not applicable, and in fact, its asymptotic distribution is a mixture of chi-Square distributions.

5 The Model Diagnosis

The model diagnosis is an essential step to ensure that a fitted model is adequate for the observed data set. However, diagnosing counts models is still a challenging research problem. Pearson and deviance residuals are often used in practice for diagnosing counts models, despite wide recognition that these residuals are far from normality when applied to count data. RQRs, proposed by Dunn and Smyth [10] and Park et al. [26], are used to overcome the above-mentioned problems. The key idea of the RQR

is to randomize the lower tail probability into a uniform random number between the discontinuity gap of the cumulative density function (CDF). It can be shown that the RQRs are normally distributed under the true model. To the best of our knowledge, the RQR has not been applied to the residual analyses for zero-inflated or modified mixed effects models.

To do this, we follow the Dunn and Smyth [10] approach. The RQR inverts the fitted distribution function for each response value and finds the equivalent standard normal quantile. Let $G(y; \mathbf{w}, \lambda)$ denote the CDF for random variable y . If the CDF is continuous, then $G(y_i; \mathbf{w}, \lambda)$ is uniformly distributed on the unit interval. RQRs can thus be defined as

$$q_i = \Phi^{-1}[G(y_i; \hat{\mathbf{w}}_i, \hat{\lambda}_i)], \quad (37)$$

where $\Phi^{-1}(\cdot)$ is the quantile function of the standard normal distribution. However, if the CDF is discrete, then randomization is added to modify q_i in Eq. (37). To be more specific, let $g(y_i; \mathbf{w}, \lambda)$ denote the PMF of y . The CDF can be redefined as

$$G^*(y; \mathbf{w}, \lambda) = G(y^-; \mathbf{w}, \lambda) + U \times g(y; \mathbf{w}, \lambda),$$

where U is a uniform random variable on $[0, 1]$, and $G(y^-; \mathbf{w}, \lambda)$ is the lower limit of G in y . When G is discrete, we let $a_i = \lim_{y \rightarrow y_i^-} G(y; \hat{\mathbf{w}}_i, \hat{\lambda}_i)$ and $b_i = G(y_i; \hat{\mathbf{w}}_i, \hat{\lambda}_i)$. Then, the randomized quantile residual is defined by

$$q_i = \Phi^{-1}[G_i^*],$$

where G_i^* is a uniform random variable on the interval (a_i, b_i) , and $q_i \sim N(0, 1)$. Here, $N(0, 1)$ stands for the standard normal distribution. Therefore, the only information that is required for calculating RQRs is the CDF of the response variable. In numerical analysis, we will use RQRs to investigate the adequacy of the used models.

6 Simulation Studies

In this simulated scenario, we have evaluated the efficiency of the ZKIP regression model with the link function $\log(\lambda) = \beta_0 + \beta_1 \mathbf{x}_1 + \beta_2 \mathbf{x}_2$, where $\beta_0 = 0.5$, $\beta_1 = \beta_2 = -0.05$, $w_0 = w_1 = w_2 = w_3 = 0.2$ and \mathbf{x}_1 and \mathbf{x}_2 are, respectively, generated from $N(0, 1)$ and the Bernoulli distribution with parameter 0.5 and $k = 0, 1, 2, 3$. In each step, we select a sample with sizes of $n = 100, 500, 1000$ from ZKIP regression ($k = 0, 1, 2, 3$) models and estimate the parameters of the assumed models using the ML method. We repeat this step 10,000 times. If the model and its related definitions work correctly, the estimates obtained for the parameters of this models should be close to their real values. Also, as the sample size rises, the means of bias and standard error should be decreased. This assertions are supported by Table 3's results.

For the next simulated scenario, we have selected a sample with sizes $n = 100, 500, 1000$, from Z3IP regression model (with the parameters similar to

Table 3). Then, we have fitted the regression ZKIP with $k = 0, 1, 2$ and Poisson models, for the same sample. As anticipated, the proposed model performed admirably results. The regression Z3IP model shows the best outcomes based on the comparative criteria in Table 4.

7 Real Data Analysis

To assess the performance of the proposed ZKIP model and the corresponding regression version, two real data sets are analyzed.

7.1 Example 1 (Alcohol Consumption)

DeHart et al. [7] described a study in which “moderate to heavy drinkers” (at least 12 alcoholic drinks/week for women, 15 for men) were recruited to keep a daily record of each drink that they consumed over a 30-day study period. Participants also completed a variety of rating scales covering daily events in their lives and items related to self-esteem. Among the researchers’ hypotheses, the negative events, particularly those involving romantic relationships, are suspected to be related to the amount of alcohol consumed, especially among those with low self-esteem.

In this example, we consider *numall* (number of alcoholic beverages, or “drinks,” consumed in one day) as a response variable. In Tables 5 and 6, we fit the ZKIP distribution to this data set. In the following, *negevent* (an index for combining the total number and intensity of negative events experienced during the day) and *nrel* (a measure of negative relationship interactions) are considered covariate variables. Tables 7 and 8 show the fitted regression ZKIP models.

The results in Tables 5 and 6 show that the ZKIP regression model with $k = 6$ performs well and dominates the rest models. Tables 7 and 8 also confirm that the regression ZKIP with $k = 6$ is the best model among the considered models. In Table 9, we consider the problem of the hypothesis testing $H_0 : w_6 = 0$ against the alternative $H_1 : w_5 > 0$ and $H_1 : w_7 = 0$, with significance level $\alpha = 0.05$. The asymptotic distribution is $0.5\chi_0^2 + 0.5\chi_1^2$, where $\chi_0^2 \equiv 0$, for the Alcohol consumption data set (for more details, see [29]). We see that $k = 6$ is the best choose for the ZKIP model.

With these explanations, an algorithm can be presented to choose the appropriate value of k . We start fitting the inflated model from $k = 0$ and continuing this fitting until k , so there is no significant difference between the fitting of k and the larger $k + 1$.

7.2 Example 2 (Lung Data Set)

This data set is about survival in those with advanced lung cancer from the “North Central Cancer Treatment Group” (Loprinzi et al. [21]). Performance scores rate how well the patient can perform usual daily activities. The Eastern Cooperative Oncology Group (ECOG) performance score is measured by the physician as

Table 3 Bias of ML estimates and standard errors (in parentheses), and model diagnostics in the simulation study

<i>n</i>	Parameter	<i>k</i> = 0	<i>k</i> = 1	<i>k</i> = 2	<i>k</i> = 3
100	$\hat{\beta}_0$	- 0.0301 (0.0015)	- 0.0635 (0.0023)	- 0.1405 (0.0162)	- 1.2247 (0.0470)
	$\hat{\beta}_1$	- 0.0024 (0.0010)	0.0001 (0.0012)	- 0.0060 (0.0093)	- 0.0676 (0.0409)
	$\hat{\beta}_2$	- 0.0031 (0.0020)	- 0.0024 (0.0024)	- 0.0237 (0.0114)	- 0.5169 (0.0707)
	\hat{w}_0	- 0.0148 (0.0007)	- 0.0373 (0.0009)	- 0.0631 (0.0011)	- 0.0735 (0.0010)
	\hat{w}_1	-	- 0.0224 (0.0009)	- 0.0408 (0.0011)	- 0.0336 (0.0010)
	\hat{w}_2	-	-	- 0.0077 (0.0007)	0.0052 (0.0007)
	\hat{w}_3	-	-	-	0.0019 (0.0005)
500	$\hat{\beta}_0$	- 0.0049 (0.0007)	- 0.0144 (0.0011)	- 0.0415 (0.0019)	- 0.1350 (0.0059)
	$\hat{\beta}_1$	- 0.0003 (0.0004)	0.0001 (0.0005)	0.0011 (0.0007)	- 0.0005 (0.0036)
	$\hat{\beta}_2$	- 0.0009 (0.0009)	0.0000 (0.0011)	0.0005 (0.0014)	0.0087 (0.0064)
	\hat{w}_0	- 0.0024 (0.0003)	- 0.0080 (0.0004)	- 0.0295 (0.0007)	- 0.0676 (0.0010)
	\hat{w}_1	-	0.0057 (0.0004)	- 0.0249 (0.0007)	- 0.0569 (0.0010)
	\hat{w}_2	-	-	- 0.0081 (0.0004)	- 0.0206 (0.0005)
	\hat{w}_3	-	-	-	- 0.0031 (0.0002)
1000	$\hat{\beta}_0$	- 0.0024 (0.0005)	- 0.0075 (0.0007)	- 0.0227 (0.0014)	- 0.0668 (0.0030)
	$\hat{\beta}_1$	- 0.0023 (0.0003)	- 0.0002 (0.0004)	0.0005 (0.0005)	0.0001 (0.0009)
	$\hat{\beta}_2$	- 0.0005 (0.0006)	- 0.0003 (0.0007)	0.0009 (0.0010)	- 0.0034 (0.0020)
	\hat{w}_0	- 0.0016 (0.0002)	- 0.0041 (0.0003)	- 0.0156 (0.0005)	- 0.0508 (0.0009)
	\hat{w}_1	-	- 0.0032 (0.0003)	- 0.0136 (0.0005)	- 0.0445 (0.0009)
	\hat{w}_2	-	-	- 0.0044 (0.0003)	- 0.0169 (0.0005)
	\hat{w}_3	-	-	-	0.0028 (0.0004)

0 = asymptomatic, 1 = symptomatic but completely ambulatory, 2 = in bed < 50% of the day, 3 = in bed > 50% of the day but not bedbound, 4 = bedbound (ph.ecog). We consider the ph.ecog as the response variable and the Age and Sex of patients as covariates. The fitted ZKIP model for the response variable in the Lung data set is shown in Tables 10 and 11 . It is observed that the ZKIP model with *k* = 2 is the best one. Tables 12 and 13 confirm the same results in the regression ZKIP model. Table 14 also introduces *k* = 2 as an appropriate value for the regression ZKIP model for this data set.

Figures 3 and 5 for, respectively, real data sets 1 and 2 show the RQRs of the regression ZKIP model for all observations. For the best *k*-obtained regression ZKIP fitted, these figures approximately show a random scatter around zero for RQR in the three real data sets, which seems to be reasonable. Figures 2 and 4 also show that the RQRs follow approximately the *N*(0, 1) distribution for the obtained *k* (Figs. 2, 3, 4, 5).

Table 4 Frequency comparisons associated with simulated data from the ZKIP regression distribution

n	Count	Observed	$k = 0$	$k = 1$	$k = 2$	$k = 3$	Poisson
100	0	15	14.99	35.58	14.78	14.54	12.36
	1	20	21.02	28.65	21.45	20.48	22.97
	2	27	22.53	17.78	28.26	27.42	23.71
	3	22	17.73	9.36	12.34	21.54	18.05
	4	4	11.46	4.53	9.07	5.04	11.33
	5	3	6.45	2.11	6.00	3.74	6.20
	6	3	3.25	0.99	3.67	2.64	3.05
	7	3	1.50	0.48	2.11	1.79	1.38
	8 <	3	1.01	0.45	2.04	2.28	0.90
		– Log L		187.1926	187.0785	184.3587	179.1421
	ABE		24.1239	59.5964	23.1806	5.8907	27.1555
	χ^2		14.1239	68.2423	12.9983	1.5044	15.4826
500	0	114	118.46	202.91	116.42	116.17	82.02
	1	110	106.51	130.03	111.84	109.15	136.76
	2	111	106.68	81.50	117.47	113.13	124.18
	3	108	78.14	43.89	60.97	107.45	81.53
	4	18	46.80	21.94	41.12	17.41	43.31
	5	14	24.28	10.52	24.76	11.95	19.75
	6	10	11.30	4.93	13.72	8.10	8.01
	7	4	4.82	2.28	7.13	5.46	2.96
	8 <	11	2.98	1.90	6.30	8.89	1.46
		– Log L		884.4426	884.1051	880.0346	840.6413
	ABE		91.3488	225.866	103.1793	13.8187	142.0298
	χ^2		55.7628	198.331	60.2662	1.8025	107.152
1000	0	244	246.34	395.49	242.06	237.75	188.77
	1	232	238.08	275.01	244.24	234.09	287.97
	2	229	216.23	167.21	235.71	230.04	241.64
	3	203	145.28	85.27	121.15	199.13	148.85
	4	32	81.24	40.48	73.72	35.44	75.78
	5	18	40.32	18.78	40.72	22.68	34.04
	6	14	18.49	8.76	21.17	14.36	14.07
	7	9	8.07	4.19	10.63	9.07	5.51
	8 <	19	5.91	4.59	10.38	15.19	3.36
		– Log L		1719.624	1718.899	1712.996	1655.383
	ABE		168.9827	407.7418	184.6011	25.6067	155.3624
	χ^2		96.2978	305.8896	102.2458	2.5274	257.0112

Table 5 Estimates, standard errors (in parentheses), and model diagnostics for the Alcohol consumption data set

Parameter	$k = 0$	$k = 1$	$k = 2$	$k = 3$
$\hat{\lambda}$	3.1200 (0.0837)	3.5303 (0.1024)	4.3945 (0.1458)	5.2217 (0.1956)
\hat{w}_0	0.1922 (0.0176)	0.2079 (0.0171)	0.2221 (0.0162)	0.2262 (0.0153)
\hat{w}_1	–	0.1091 (0.01671)	0.1544 (0.0154)	0.1706 (0.0144)
\hat{w}_2	–	–	0.1561 (0.0172)	0.1877 (0.0157)
\hat{w}_3	–	–	–	0.0881 (0.0140)
– Log L	1392.4570	1371.9350	1335.2620	1317.9980
AIC	2788.9140	2749.8700	2678.5240	2645.9960
Parameter	$k = 4$	$k = 5$	$k = 6$	<i>Poisson</i>
$\hat{\lambda}$	5.8876 (0.2443)	7.2689 (0.3643)	8.7945 (0.5517)	2.5201 (0.0636)
\hat{w}_0	0.2273 (0.0146)	0.2281 (0.0127)	0.2280 (0.0129)	–
\hat{w}_1	0.1758 (0.0137)	0.1790 (0.0120)	0.1796 (0.0122)	–
\hat{w}_2	0.2001 (0.0147)	0.2093 (0.0126)	0.2114 (0.0127)	–
\hat{w}_3	0.1068 (0.0130)	0.1237 (0.0113)	0.1285 (0.0113)	–
\hat{w}_4	0.0446 (0.0114)	0.0674 (0.0100)	0.0755 (0.0098)	–
\hat{w}_5	–	0.0526 (0.0098)	0.0635 (0.0097)	–
\hat{w}_6	–	–	0.0304 (0.0085)	–
– Log L	1311.7030	1298.8080	1289.6930	14543.280
AIC	2635.4060	2611.6150	2595.3850	2970.6570

Table 6 Observed number of illness and the corresponding expected values under the fitted models in Table 5

Count	Observed	$k = 0$	$k = 1$	$k = 2$	$k = 3$	$k = 4$	$k = 5$	$k = 6$	<i>Poisson</i>
0	142	124.01	141.98	142.03	141.98	142.02	142.00	142.02	50.13
1	112	69.32	112.01	11.97	112.00	112.01	112.01	112.00	126.33
2	132	108.48	77.68	131.97	132.04	131.99	132.00	131.99	159.17
3	81	112.48	91.41	50.82	80.99	81.00	80.98	81.01	133.69
4	49	87.74	80.67	55.84	34.11	48.99	49.00	48.99	84.22
5	43	54.75	56.96	49.08	35.62	25.01	43.00	42.99	42.44
6	24	28.47	33.51	35.95	31.00	24.54	12.949	24.00	17.83
7	6	12.69	16.90	22.57	23.12	20.64	12.96	6.33	6.42
8	9	4.95	7.46	12.40	15.09	15.19	11.78	6.96	2.02
9	7	1.72	2.92	6.06	8.76	9.93	9.51	6.80	0.57
10	7	0.54	1.03	2.66	4.57	5.85	6.91	5.97	0.14
11 ≤	11	0.20	0.46	1.64	3.71	5.82	10.32	13.72	0.04
ABE		186.2634	152.9210	89.7472	64.0255	48.6561	24.5705	6.3987	259.6502
χ^2		734.6944	344.4936	97.1405	40.7461	31.5483	15.7199	1.3393	3558.4550

Table 7 Estimates, standard errors (in parentheses), and model diagnostics (log-likelihood, and AIC) for the Alcohol consumption data set

Parameter	$k = 0$	$k = 1$	$k = 2$	$k = 3$
$\hat{\beta}_0$	1.1962 (0.0269)	1.2979 (0.0289)	1.5145 (0.0332)	1.6634 (0.0374)
$\hat{\beta}_1$	- 0.2204 (0.0486)	- 0.1659 (0.0527)	- 0.1405 (0.0591)	- 0.0546 (0.0672)
$\hat{\beta}_2$	0.0736 (0.0223)	0.0597 (0.0239)	0.0437 (0.0252)	0.0176 (0.0261)
\hat{w}_0	0.1880 (0.0177)	0.2052 (0.0172)	0.2214 (0.0163)	0.2260 (0.0154)
\hat{w}_1	-	0.1047 (0.0168)	0.1524 (0.0154)	0.1701 (0.0144)
\hat{w}_2	-	-	0.1549 (0.0173)	0.1872 (0.0157)
\hat{w}_3	-	-	-	0.0872 (0.0140)
- Log L	1387.778	1369.5200	1334.0820	1317.8610
AIC	2783.5507	2749.0390	2680.1640	2649.7210
Parameter	$k = 4$	$k = 5$	$k = 6$	<i>Poisson</i>
$\hat{\beta}_0$	1.7602 (0.0416)	1.9071 (0.0498)	2.0064 (0.0579)	1.0195 (0.0116)
$\hat{\beta}_1$	0.0352 (0.0806)	0.2483 (0.1037)	0.4829 (0.1107)	- 0.3355 (0.0235)
$\hat{\beta}_2$	0.0030 (0.0292)	- 0.0163 (0.0333)	- 0.0522 (0.0330)	0.1239 (0.0152)
\hat{w}_0	0.2273 (0.0146)	0.2279 (0.0126)	0.2279 (0.0129)	-
\hat{w}_1	0.1759 (0.0137)	0.1791 (0.0120)	0.1794 (0.0122)	-
\hat{w}_2	0.2003 (0.0147)	0.2094 (0.0125)	0.2111 (0.0127)	-
\hat{w}_3	0.1072 (0.0130)	0.1245 (0.0112)	0.1285 (0.0113)	-
\hat{w}_4	0.0451 (0.0114)	0.0682 (0.0098)	0.0753 (0.0099)	-
\hat{w}_5	-	0.0533 (0.0097)	0.0629 (0.0097)	-
\hat{w}_6	-	-	0.0298 (0.0084)	-
- Log L	1311.6340	1297.8250	1287.6180	1470.9550
AIC	2639.2680	2613.6500	2595.2360	2947.9100

Table 8 Observed number of illness and the corresponding expected values under the fitted models in Table 7

Count	Observed	$k = 0$	$k = 1$	$k = 2$	$k = 3$	$k = 4$	$k = 5$	$k = 6$	Poisson
0	142	141.96	141.02	141.99	141.98	141.98	141.98	142.02	50.14
1	112	69.35	11.99	111.99	112.01	112.02	11.98	112.00	126.33
2	132	108.17	77.68	132.00	132.00	131.99	132.02	132.01	159.17
3	81	112.49	91.41	50.81	80.99	81.00	81.01	81.00	133.69
4	49	84.74	80.67	55.83	34.12	49.00	49.01	49.01	84.22
5	43	54.74	56.95	49.08	35.63	25.01	42.99	42.99	42.44
6	24	28.46	33.51	35.95	31.01	24.54	12.49	23.98	17.82
7	6	12.69	16.90	22.57	23.13	20.64	12.96	6.33	6.42
8	9	4.95	7.64	12.40	15.10	15.19	11.77	6.95	2.02
9	7	1.71	2.92	6.06	8.76	9.94	9.51	6.79	0.57
10	7	0.54	1.03	2.66	4.57	5.85	6.91	5.97	0.14
11≤	11	0.20	0.46	1.64	3.71	5.83	10.31	13.72	0.04
ABE		186.2441	152.9047	89.6884	64.0094	48.6745	24.6268	6.4005	259.6528
χ^2		735.2491	344.5580	97.1210	40.7388	31.5452	15.7120	1.3421	3559.0930

Table 9 Testing to choose k in Tables 7 and 8

	$k = 6$ vs $k = 5$	$k = 6$ vs $k = 7$
Likelihood ratio	10.207	0.007
P-Value	0.0007	0.4667

Table 10 Estimates, standard errors (in parentheses), and model diagnostics for the Lung data set

Parameter	$k = 0$	$k = 1$	$k = 2$	Poisson
$\hat{\lambda}$	0.9518 (0.0207)	0.939 (0.0279)	0.4348 (0.0306)	0.9519 (0.0166)
\hat{w}_0	0.0000 (0.0079)	0.0000 (0.0073)	0.0000 (0.0072)	–
\hat{w}_1	–	0.2100 (0.0085)	0.3798 (0.0083)	–
\hat{w}_2	–	–	0.1930 (0.0071)	–
– Log L	264.1790	255.8550	241.4580	264.1790
AIC	532.3590	517.7100	490.9170	530.3590

Table 11 Observed number of illness and the corresponding expected values under the fitted models in Table 10

Count	Observed	$k = 0$	$k = 1$	$k = 2$	Poisson
0	63	88.03	70.44	63.04	88.02
1	114	83.78	114.01	113.98	83.78
2	50	39.86	31.05	50.01	39.87
3≤	1	12.65	9.72	0.86	12.65
ABE		77.0387	35.1200	0.2085	77.0203
χ^2		31.3244	20.1757	0.0226	31.3169

Table 12 Estimates, standard errors (in parentheses), and model diagnostics (log-likelihood, and AIC) for Lung data set

Parameter	$k = 0$	$k = 1$	$k = 2$	Poisson
$\hat{\beta}_0$	- 1.1029 (0.0124)	- 1.4226 (0.0136)	- 4.8461 (0.0136)	- 1.1007 (0.0116)
$\hat{\beta}_1$	0.0163 (0.0252)	0.0212 (0.0162)	0.0682 (0.0274)	0.0165 (0.0235)
$\hat{\beta}_2$	0.0198 (0.0162)	0.0065 (0.0177)	- 0.0697 (0.0177)	0.0059 (0.0152)
\hat{w}_0	0.0000 (0.0078)	0.0000 (0.0075)	0.1581 (0.0075)	-
\hat{w}_1	-	0.2190 (0.0086)	0.4426 (0.0088)	-
\hat{w}_2	-	-	0.1944 (0.0072)	-
- Log L	261.8590	252.7920	239.8060	261.8520
AIC	531.7180	515.5830	491.6120	529.7050

Table 13 Observed number of illness and the corresponding expected values under the fitted models in Table 12

Count	Observed	$k = 0$	$k = 1$	$k = 2$	Poisson
0	63	88.02	70.44	63.04	88.01
1	114	83.78	114.02	114.00	83.78
2	50	39.87	31.04	49.99	39.88
$3 \leq$	1	12.65	9.71	0.86	12.65
ABE		77.0289	35.1302	0.1932	77.0074
χ^2		31.3204	20.1811	0.0227	31.3116

Table 14 Testing to choose k in Tables 12 and 13

	$k = 2$ vs $k = 1$	$k = 2$ vs $k = 3$
Likelihood ratio	12.986	1.540
P-Value	0.0002	0.1073

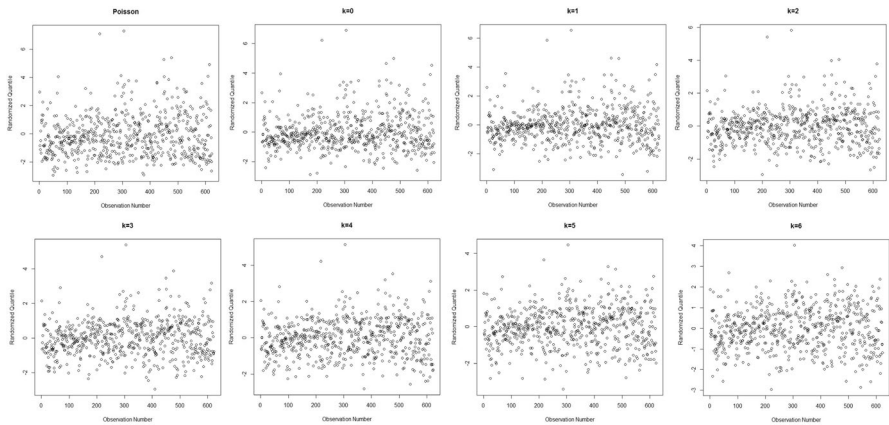


Fig. 2 RQR's plots for the alcohol consumption data set

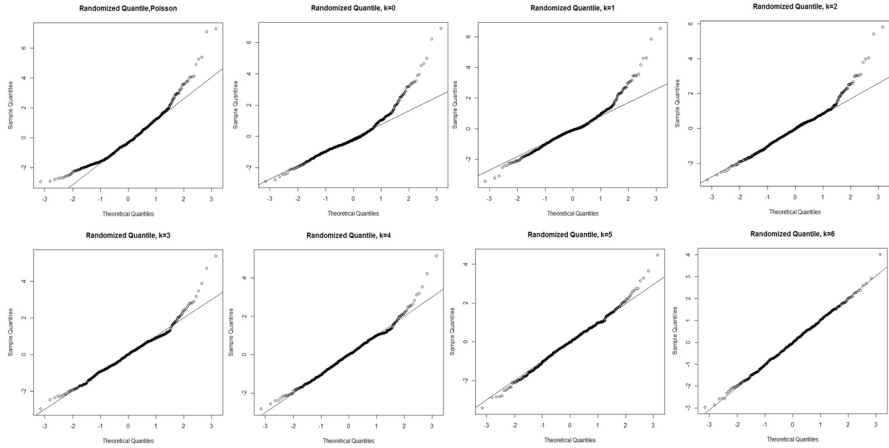


Fig. 3 QQ plots for the RQR's of the alcohol consumption data set

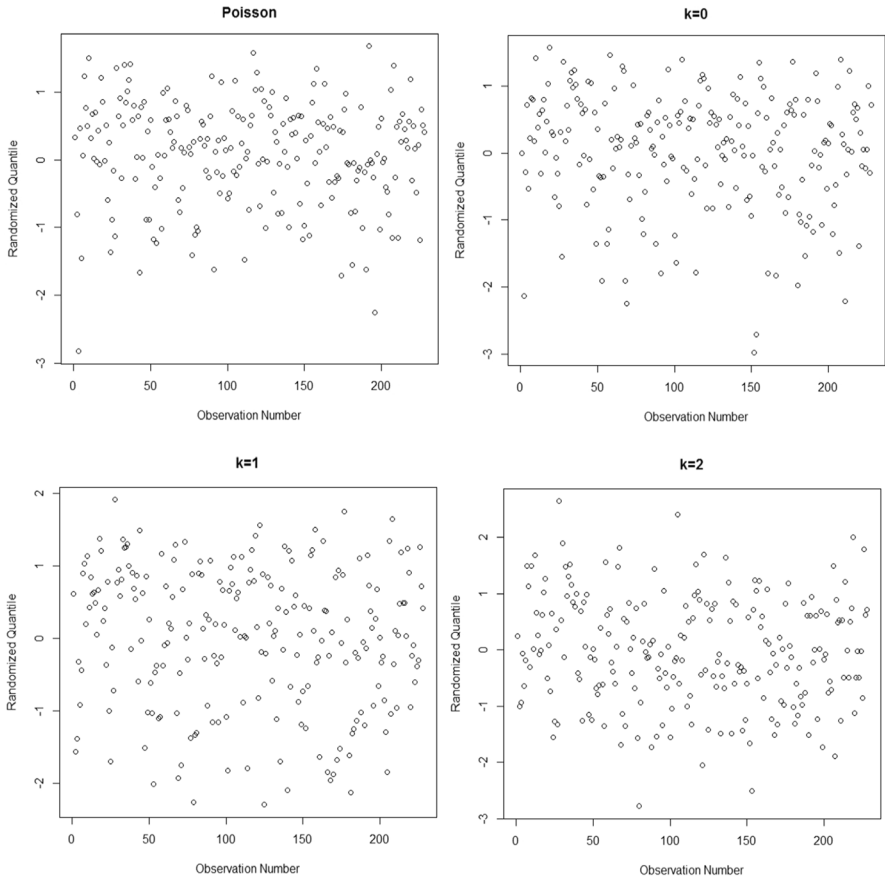


Fig. 4 RQR's plots for the Lung data set

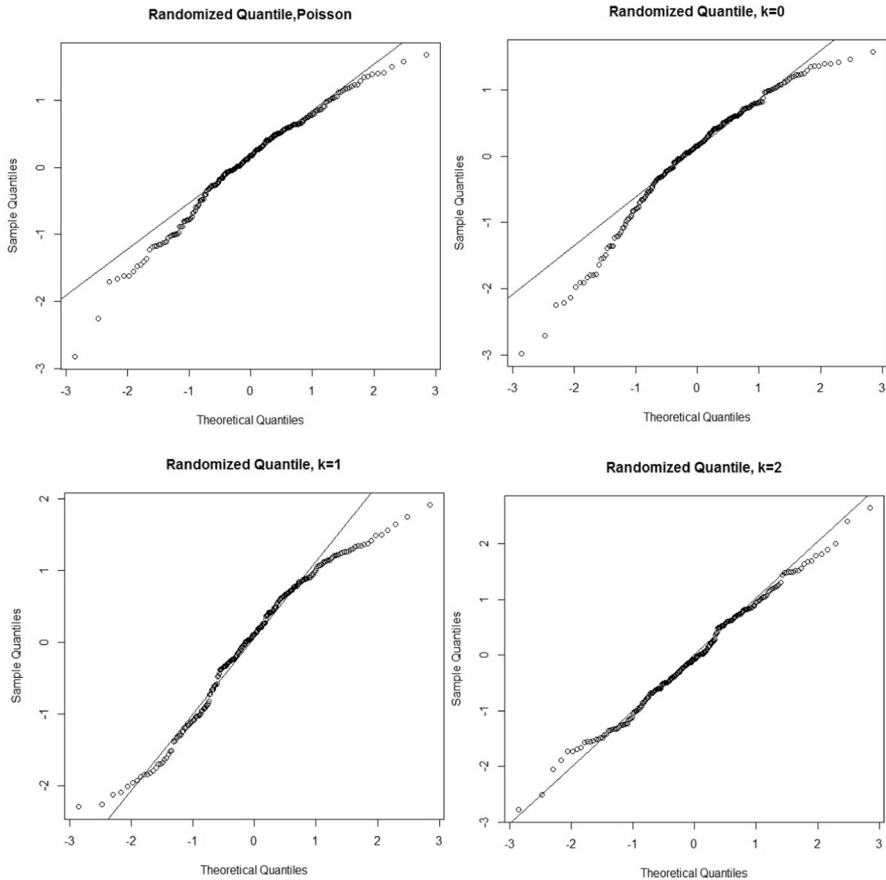


Fig. 5 QQ plots for the RQR's of the Lung data set

8 Conclusions

Previous research about the inflated Poisson distribution has suggested that inflation may occur at one or two points. By introducing the ZKIP distribution, in addition to covering all previous studies, we allowed the inflated points to be three or even more. Also, in this work, we studied the properties of the ZKIP distribution and the ZKIP regression model. The flexibility of the ZKIP distribution (which was perfectly illustrated with real data examples) shows that it may be used as a suitable model for real data set analysis with discrete responses. This flexibility was also observed in the regression version of the ZKIP distribution.

The introduced model can be used in decision trees, random forests and statistical quality control or any model that can be used for inflated discrete data. Also, this model can be developed for the neutrosophic statistics [3, 4, 31].

Acknowledgements The authors are grateful to the editor-in-chief and the anonymous referees for their helpful suggestions.

Author Contributions HS and MD contributed to the design and implementation of the research, to the analysis of the results and to the writing of the manuscript.

Funding This research received no specific grant from any funding agency in the public, commercial, or nonprofit sectors.

Data availability The dataset used in example 1 is available in reference [7] and example 2 is in reference [21].

Declarations

Conflict of Interest The authors declare they have no conflicts of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Agarwal, D.K., Gelfand, A.E., Citron-Pousty, S.: Zero-inflated models with application to spatial count data. *Environ. Ecol Stat.* **9**, 341–355 (2002)
2. Agresti, A.: *Analysis of Ordinal Categorical Data*, 3rd edn. Wiley, Hoboken (2013)
3. AlAita, A., Aslam, M.: Analysis of covariance under neutrosophic statistics. *J. Stat. Comput. Simul.* **93**(3), 397–415 (2023)
4. AlAita, A., Talebi, H., Aslam, M., Al Sultan, K.: Neutrosophic statistical analysis of split-plot designs. *Soft Comput.* **27**, 7801–7811 (2023)
5. Arora, M., Chaganty, N.R.: EM estimation for zero-and k-inflated Poisson regression model. *Computation* **9**(9), 94 (2021)
6. Cohen, A.C.: Estimating the parameters of a modified Poisson distribution. *J. Am. Stat. Assoc.* **55**, 139–143 (1960)
7. DeHart, T., Tennen, H., Armeli, S., Todd, M., Affleck, G.: Drinking to regulate romantic relationship interactions: the moderating role of self-esteem. *J. Exp. Soc. Psychol.* **44**, 527–538 (2008)
8. Dempster, A.P., Laird, N.M., Rubin, D.B.: Maximum likelihood estimation from incomplete data via the EM algorithm (with discussion). *J. R. Stat. Soc. Ser. B (Stat. Methodol.)* **39**, 1–38 (1977)
9. Dikheel, T.R., Jouda, H.A.: Zero inflation Poisson regression estimation by using Genetic Algorithm (GA). *J. Surv. Fish. Sci.* **10**(3S), 5257–5268 (2023)
10. Dunn, P.K., Smyth, G.K.: Randomized quantile residuals. *J. Comput. Graph. Stat.* **5**(3), 236–244 (1996)
11. Jansakul, N., Hinde, J.P.: Score tests for zero-Inflated Poisson models. *Comput. Stat. Data Anal.* **40**, 75–96 (2002)
12. Jung, B.C., Jhun, M., Lee, J.W.: Bootstrap tests for overdispersion in a zero-inflated Poisson regression model. *Biometrics* **61**, 626–629 (2005)
13. Kang, Y., Wang, S., Wang, D., Zhu, F.: Analysis of zero-and-one inflated bounded count time series with applications to climate and crime data. *TEST* **32**(1), 34–73 (2023)
14. Lambert, D.: Zero-inflated Poisson regression, with an application to defects in manufacturing. *Technometrics* **34**, 1–14 (1992)

15. Lehmann, E.L., Romano, J.P., Casella, G.: Testing Statistical Hypotheses. Springer Texts in Statistics, 3rd edn. Springer, New York (2005)
16. Lim, H.K., Li, W.K., Yu, P.L.: Zero-inflated Poisson regression mixture model. *Comput. Stat. Data Anal.* **71**, 151–158 (2014)
17. Lin, T.H., Tsai, M.H.: Modeling health survey data with excessive zero and K responses. *Stat. Med.* **32**(9), 1572–1583 (2013)
18. Liu, W.C., Tang, Y.C., Xu, A.C.: A zero-and-one inflated Poisson model and its application. *Stat. Interface* **11**, 339–351 (2018)
19. Liu, W.C., Tang, Y.C., Xu, A.C.: Zero-and-one-inflated Poisson regression model. *Stat. Pap.* **62**, 915–934 (2021)
20. Long, D.L., Preisser, J.S., Herring, A.H., Golin, C.E.: Zero-inflated Poisson regression with application to defects in manufacturing. *Stat. Med.* **33**, 5151–5165 (2014)
21. Loprinzi, C.L., Laurie, J.A., Wieand, H.S., Krook, J.E., Novotny, P.J., Kugler, J.W., Bartel, J., Law, M., Bateman, M., Klatt, N.E.: Prospective evaluation of prognostic variables from patient-completed questionnaires. North Central Cancer Treatment Group. *J. Clin. Oncol.* **12**(3), 601–607 (1994)
22. Louis, T.A.: Finding the observed information matrix when using the EM algorithm. *J. R. Stat. Soc.* **44**(2), 226–233 (1982)
23. Melkersson, M., Rooth, D.O.: Modeling female fertility using inflated count data models. *J. Popul. Econ.* **13**(2), 189–203 (2000)
24. Motalebi, N., Owlia, M.S., Amiri, A., Fallahnezhad, M.S.: Monitoring social networks based on Zero-inflated Poisson regression model. *Commun. Stat. Theory Methods* **52**(7), 2099–2115 (2023)
25. Neelon, B., Chung, D.J.: The LZIP: a Bayesian latent factor model for correlated zero-inflated counts. *Biometrics* **73**, 185–196 (2017)
26. Park, K., Jung, D., Kim, J.M.: Control charts based on randomized quantile residuals. *Appl. Stoch. Models Bus. Ind.* **36**(4), 716–729 (2020)
27. Ridout, J., Hinde, J., Demetrio, G.B.: A score test for testing a zero-inflated Poisson regression model against zero-inflated negative binomial alternatives. *Biometrics* **57**, 219–223 (2001)
28. Saffari, S.E., Adnan, R.: Zero-inflated Poisson regression models with right censored count data. *Matematika* **27**, 21–29 (2011)
29. Self, S.G., Liang, K.Y.: Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. *J. Am. Stat. Assoc.* **82**(398), 605–610 (1987)
30. Serra, I.J.A., Polestico, D.L.L.: On the zero and k-inflated negative binomial distribution with applications. *Adv. Appl. Stat.* **88**(1), 1–23 (2023)
31. Smarandache, F.: Introduction to neutrosophic statistics (2014). <https://doi.org/10.13140/2.1.2780.1289>
32. Sun, Y., Zhao, S., Tian, G.L., Tang, M.L., Li, T.: Likelihood-based methods for the zero-one-two inflated Poisson model with applications to biomedicine. *J. Stat. Comput. Simul.* **93**(6), 956–982 (2023)
33. Tang, Y.C., Liu, W.C., Xu, A.C.: Statistical inference for zero-and-one-inflated Poisson models. *Stat. Theory Relat. Fields* **1**, 216–226 (2017)
34. Vandebroek, J.: A score test for zero inflation in a Poisson distribution. *Biometrics* **51**, 738–743 (1995)
35. Wu, C.F.J.: On the convergence properties of the EM algorithm. *Ann. Stat.* **11**, 95–103 (1983)
36. Yang, Y., Simpson, D.G.: Conditional decomposition diagnostics for regression analysis of zero-inflated and left-censored data. *Stat. Methods Med. Res.* **21**, 393–408 (2012)
37. Yoneda, K.: Estimations in some modified Poisson distributions. *Yokohama Math. J.* **10**, 73–96 (1962)
38. Zandi, Z., Bevrani, H., Belaghi, R.A.: Shrinkage estimation in the zero-inflated Poisson regression model with right-censored data. *Commun. Stat. Theory Methods* (2023). <https://doi.org/10.1080/03610926.2023.2196751>
39. Zhang, C., Tian, G.L., Ng, K.W.: Properties of the zero-and-one inflated Poisson distribution and likelihood-based inference methods. *Stat. Interface* **9**, 11–32 (2016)
40. Zhu, H., Luo, S., DeSantis, S.M.: Zero-inflated count models for longitudinal measurements with heterogeneous random effects. *Stat. Methods Med. Res.* **26**, 1774–1786 (2017)