Efficient estimation in ZIP models with applications to count data

Nurul N. Mohamad, Ibrahim Mohamed*, Ng Kok-Haur, Mohd S. Yahya
Institute of Mathematical Sciences, Faculty of Science, University of Malaya, 50603 Kuala Lumpur, Malaysia.
* Corresponding author: imohamed@um.edu.my

Abstract

Estimating functions have been used in estimating parameters of many continuous time series models. However, this method has not been applied to models involving count data. In this paper, we use quadratic estimating functions (QEF) to derive estimators for the joint estimation of the conditional mean and variance parameters of count data models, specifically the basic zero-inflated Poisson (ZIP) model, ZIP regression model and integer-valued generalized autoregressive heteroscedastic model with ZIP conditional distribution. Results show that the estimators derived from QEF method, which uses information from combined estimating functions, is more informative than linear estimating functions (LEF) method that only uses information from component estimating functions. Finally, we also fit the real data sets using the ZIP models via QEF, LEF and maximum likelihood methods, and in so doing, demonstrate the superiority of the QEF method in practice.

Keywords: Count data; information matrix; linear estimating functions; quadratic estimating functions; zero-inflated Poisson

1. Introduction

Count data are frequently encountered in many biomedical, epidemiological, industrial and public health applications. In practice, especially in the medical field, many count data sets have a high frequency of zeroes. For example, for diseases with low infection rates, the observed counts typically contains a large number of zeroes, although the counts can also be very large during the outbreak period. For such data set, Lambert (1992) introduced a zero-inflated Poisson (ZIP) regression model. His study showed that the ZIP model is not only easy to interpret, but also leads to more refined data analysis as it can accommodate overdispersion.

Following the findings, many studies and applications of the ZIP model have been conducted. For instance, Baksh et al. (2011) proposed the overdispersion test for the ZIP model, while Zhu (2012) introduced the model inspired by the generalized autoregressive conditional heteroscedastic (GARCH) model. In the GARCH model, the integer-valued case with a conditional distribution has ZIP distribution instead of normal distribution. It is denoted as ZIPINGARCH \((p, q)\), where \(p\) and \(q\) are positive integers.

The maximum likelihood (ML) method is commonly used for estimating the parameters of ZIP models when the distribution of the data is known. However, the method does not always perform well under certain circumstances, see for example, Bahadur (1958), Crowder (1987) and Vinod (1997). Furthermore, as pointed out by Nanjundan and Naika (2012), the ML estimator of ZIP models does not have closed-form expression. Therefore, various estimation methods have been proposed as an alternative to the ML method. Some of these are a recursive technique based on the two-step least squares estimator (Abaza, 1982), Monte Carlo EM method (Chan and Ledolfer, 1995), method of moments (Kharrati-Kopaei and Faghih, 2011) and quasi-likelihood (Staub and Winkelmann, 2012).

The semiparametric approach based on the theory of estimating functions (EF) (Godambe, 1985) has been proposed to estimate the parameters of time series models. The EF approach uses the information based on
the first two conditional moments, which is known as linear estimating functions (LEF). This method has been successfully applied in continuous time series models, see Thavaneswaran and Abraham (1988), Chandra and Taniguchi (2001), Bera et al. (2006), Merkouris (2007), Allen et al. (2013) and Thavaneswaran et al. (2012 & 2015). Meanwhile, the advantages of the LEF method have been highlighted in many papers, Bera et al. (2006) stated that the LEF approach is a sufficiently flexible moment-based estimation method. It is very useful in econometric applications. In addition, Godambe and Heyde (2010) showed that the LEF estimator yields asymptotically the shortest confidence interval. Moreover, Ng and Peiris (2013) found that the LEF method is more computationally efficient and easy to apply in practice than the ML method. They also argued that the LEF method is easier to evaluate, and estimates can be obtained without the requirement of the assumption of the distribution of errors.

Liang et al. (2011) extended the LEF method to the quadratic estimating functions (QEF) method, which involves the first four conditional moments. Their results showed that the QEF method is more informative than the LEF method and gives lower standard errors of the estimated parameters. Furthermore, Thavaneswaran et al. (2015) showed that, this extension leads to an improvement in terms of the efficiency of resulting estimate. It has standard asymptotic properties, such as consistency and asymptotic normality. The QEF method also removes the problem of identifiability. On top of that, the Monte Carlo simulation results presented in Ng et al. (2015) also showed that the QEF estimators outperform the LEF estimators in almost all cases when applied on the autoregressive conditional duration model.

To our knowledge, the QEF method has never been applied to count data models. Therefore, it is in our interest to investigate the performance of the QEF method as an alternative method in the parameter estimation of these models. We have derived the optimal estimating functions of QEF for the three types of ZIP models, namely the basic ZIP, ZIP regression and ZIPINGARCH($p,q$) time series models. We also obtained the information for these ZIP models and then compared them with that from the LEF method. Concurrently, the QEF, LEF and ML methods were applied into real data sets to estimate the model parameters together with their respective standard errors. The Akaike information criterion (AIC) and Bayesian information criterion (BIC) values were calculated to determine the best fitted model.

This paper is organized as follows: Section 2 discusses the theoretical basis for the QEF and LEF methods. In Section 3, we use the QEF method to derive the optimal estimating functions and the information for the three ZIP models. The applicability of the QEF method on empirical examples is presented in Section 4. Finally, concluding remarks are given in Section 5.

2. Parameter estimation methods

This section discusses the LEF and QEF estimation methods.

2.1 Linear estimating functions

Godambe (1985) introduced the estimating functions approach to estimate the parameters of linear and non-linear time series models.

Let $\{y_t\}$ be a discrete time series process depending on a vector parameter $\theta$ that belongs to an open subset $\Theta$ of the $p$-dimensional Euclidean space. Let $\mathcal{F}_t = \sigma(y_1, y_2, \ldots, y_{t-1})$ for $t \geq 1$. Consider a $q$-dimensional vector $h_t = h_t(y_1, y_2, \ldots, y_{t-1}, \theta)$ for $1 \leq t \leq n$ which is a martingale difference and let $a_{t-1} = p \times q$ matrices depending on $\{y_1, y_2, \ldots, y_{t-1}\}$. Let $\mathcal{M}$ be the set of $p$-dimensional estimating functions $g_n(\theta)$

$$\mathcal{M} = \{g_n(\theta) : g_n(\theta) = \sum_{t=1}^{n} a_{t-1} h_t\}. \tag{1}$$

An estimate of $\theta$ can be obtained by solving the equation $g_n(\theta) = 0$. Godambe (1985) assumed that the estimating functions $g_n(\theta)$...
are almost surely differentiable with respect to the components of $\theta$ so that $E\left[ \frac{\partial h_t}{\partial \theta} | \mathcal{Y}_{t-1} \right]$ and $E\left[ g_h(\theta) \frac{\partial g_h(\theta)}{\partial \theta} | \mathcal{Y}_{t-1} \right]$ are nonsingular for all $\theta$ for each $t \geq 1$. Moreover, the $p \times p$ matrix $E\left[ g_h(\theta) g_h(\theta) | \mathcal{Y}_{t-1} \right]$ is assumed to be positive for all $\theta$. Therefore, the optimal $g^*_h(\theta)$ is given by

$$g^*_h(\theta) = \sum_{t=1}^{n} a^*_t \cdot h_t$$

$$= \sum_{t=1}^{n} \left( E\left[ \frac{\partial h_t}{\partial \theta} | \mathcal{Y}_{t-1} \right] \right)' \times \left( E\left[ h_t h'_t | \mathcal{Y}_{t-1} \right] \right)^{-1} h_t,$$

and the corresponding optimal information matrix is

$$I_{g^*_h}(\theta) = \sum_{t=1}^{n} \left( E\left[ \frac{\partial h_t}{\partial \theta} | \mathcal{Y}_{t-1} \right] \right)' \times \left( E\left[ h_t h'_t | \mathcal{Y}_{t-1} \right] \right)^{-1} \times \left( E\left[ \frac{\partial h_t}{\partial \theta} | \mathcal{Y}_{t-1} \right] \right).$$

The optimal estimating functions based on the martingale differences $m_t(\theta)$ and $s_t(\theta)$ are

$$g^*_m(\theta) = -\sum_{t=1}^{n} \frac{\partial \mu_t(\theta)}{\partial \theta} \frac{m_t(\theta)}{\langle m \rangle_t}$$

$$g^*_s(\theta) = -\sum_{t=1}^{n} \frac{\partial \sigma^2_t(\theta)}{\partial \theta} \frac{s_t(\theta)}{\langle s \rangle_t},$$

respectively. The information associated with $g^*_m(\theta)$ and $g^*_s(\theta)$ are given as

$$I_{g^*_m}(\theta) = \sum_{t=1}^{n} \frac{\partial \mu_t(\theta)}{\partial \theta} \frac{\partial \mu_t(\theta)}{\partial \theta'} \frac{1}{\langle m \rangle_t},$$

$$I_{g^*_s}(\theta) = \sum_{t=1}^{n} \frac{\partial \sigma^2_t(\theta)}{\partial \theta} \frac{\partial \sigma^2_t(\theta)}{\partial \theta'} \frac{1}{\langle s \rangle_t},$$

respectively. The optimal QEF and its corresponding information matrix are given by Liang et al. (2011) in Theorem 1.

**Theorem 1:** In the class of all QEF of the form

$$G_Q = \left\{ g_Q(\theta) = \sum_{t=1}^{n} (a_{t-1} m_t(\theta) + b_{t-1} s_t(\theta)) \right\},$$

(a) the optimal estimating function is given by

$$g^*_Q(\theta) = \sum_{t=1}^{n} (a^*_t m_t(\theta) + b^*_t s_t(\theta)),$$

where $a^*_t = R_t q_t$ and $b^*_t = R_t \omega_t$;

(b) the information matrix $I_{g^*_Q}(\theta)$ is

$$I_{g^*_Q}(\theta) = \sum_{t=1}^{n} R_t \left( \frac{N_t}{\langle m \rangle_t} + \frac{V_t}{\langle s \rangle_t} - K_t \right);$$

(c) the gain in information $I_{g^*_Q}(\theta) - I_{g_m}(\theta)$ is

$$I_{g^*_Q}(\theta) - I_{g_m}(\theta) = \sum_{t=1}^{n} R_t \left( \frac{N_t}{\langle m \rangle_t} \frac{\langle s \rangle_t^2}{\langle s \rangle_t} + \frac{V_t}{\langle s \rangle_t} - K_t \right);$$

(d) the gain in information $I_{g^*_Q}(\theta) - I_{g_s}(\theta)$ is

$$I_{g^*_Q}(\theta) - I_{g_s}(\theta) = \sum_{t=1}^{n} R_t \left( \frac{N_t}{\langle m \rangle_t} + V_t R_t \frac{\langle m \rangle_t \langle s \rangle_t}{\langle m \rangle_t \langle s \rangle_t} - K_t \right),$$

where $\gamma_t(\theta) = \frac{1}{\sigma^2_t(\theta)} E[ (y_t - \mu_t(\theta))^4 | \mathcal{Y}_{t-1} ] - 3$.

We intend to estimate the parameter $\theta$ using two classes of martingale differences

$$\{ m_t(\theta) = y_t - \mu_t(\theta), t = 1, 2, \ldots, n \},$$

$$\{ s_t(\theta) = m^2_t(\theta) - \sigma^2_t(\theta), t = 1, 2, \ldots, n \},$$

such that

$$\langle m \rangle_t = E[ m^2_t(\theta) | \mathcal{Y}_{t-1} ]$$

$$= \sigma^2_t(\theta),$$

$$\langle s \rangle_t = E[ s^2_t(\theta) | \mathcal{Y}_{t-1} ]$$

$$= \sigma^4_t(\theta)(\gamma_t(\theta) + 2),$$

$$\langle m, s \rangle_t = E[ m_t(\theta) s_t(\theta) | \mathcal{Y}_{t-1} ]$$

$$= \sigma^3_t(\theta) \gamma_t(\theta).$$
3. Zero-inflated models

There are three types of ZIP models considered: basic ZIP, ZIP regression and ZIPINGARCH\((p, q)\) time series models. Let us define the probability mass function (pmf) of a zero inflated count data model as

\[
f(y) = \begin{cases} 
\omega + (1 - \omega)g(y) & \text{for } y = 0, \\
(1 - \omega)g(y) & \text{for } y = 1, 2, 3, \ldots, \end{cases} \tag{4}
\]

where \(y\) is a count-valued random variable, \(\omega \in [0, 1]\) is a zero-inflation parameter (the probability of a strategic zero), and \(g(\cdot)\) is the probability function of the parent count model.

The mean of the zero-inflated count data model is

\[
E(y) = \sum_{y=0}^{\infty} y f(y) = (1 - \omega)E_g(y),
\]

where \(E_g(y)\) denotes the mean of the parent distribution. A full parametric zero-inflated count data model is obtained once the probability function of the parent count model is specified. For the next three subsections, we will illustrate the three ZIP models.

3.1 Basic ZIP model

The pmf for this model can be obtained from equation (4) by letting

\[
g(y; \lambda) = \frac{\exp(-\lambda)\lambda^y}{y!}, \quad \lambda > 0,
\]

with mean \(E_g(y) = \lambda, \mu_1(\theta) = E \left[ y | \Sigma \right] = (1 - \omega)\lambda\) and \(\mu_2(\theta) = E \left[ y^2 | \Sigma \right] = \lambda(1 - \omega)(\lambda + 1)\). The parameter of interest for this model is \(\theta = (\lambda, \omega)'\). Following Kharrati-Kopaei and Faghih (2011), we define two martingale differences, \(m_t(\theta) = y_t - \mu_1(\theta)\) and \(S_t(\theta) = y_t^2 - \mu_2(\theta)\), respectively. Using the results, the elements of variance-covariance of martingale differences are defined as

\[
\sigma_{11} = Var \left[ y_t | \Sigma_{t-1} \right] = \mu_1(\theta) (1 + \lambda) - \left( \mu_1(\theta) \right)^2,
\]

\[
\sigma_{12} = Cov \left[ y_t y_t | \Sigma_{t-1} \right] = \mu_1(\theta) \left( \lambda^2 + 3\lambda + 1 - \mu_1(\theta) (1 + \lambda) \right),
\]

\[
\sigma_{22} = Var \left[ y_t^2 | \Sigma_{t-1} \right] = \mu_1(\theta) \left( \lambda^3 + 6\lambda^2 + 7\lambda + 1 - \mu_1(\theta) (1 + \lambda)^2 \right).
\]

It is easily shown that \(\langle m_t \rangle = \sigma_{11}, \langle S_t \rangle = \sigma_{22}\) and \(\langle m, S \rangle = \sigma_{12}\). The derivatives of \(\mu_1(\theta)\) and \(\mu_2(\theta)\) with respect to \(\theta\) are

\[
\frac{\partial \mu_1(\theta)}{\partial \theta} = (1 - \omega, -\lambda)',
\]

\[
\frac{\partial \mu_2(\theta)}{\partial \theta} = ((1 - \omega)(1 + 2\lambda), -\lambda - \lambda^2)'.
\]

From Theorem 1, the optimal QEF for each parameter \(\lambda\) and \(\omega\) are given by

\[
g_{Q}(\lambda) = (1 - \omega) \left( \frac{1}{\sigma_{11} \sigma_{22} - \sigma_{12}^2} \right) \sum_{t=1}^{n} J_{1,t},
\]

\[
g_{Q}(\omega) = \lambda \left( \frac{1}{\sigma_{11} \sigma_{22} - \sigma_{12}^2} \right) \sum_{t=1}^{n} J_{2,t},
\]

respectively, where

\[
J_{1,t} = \left( -\sigma_{22} + \sigma_{12} H_{1,t} \right) m_t(\theta) + \left( \sigma_{12} - \sigma_{11} H_{1,t} \right) S_t(\theta),
\]

\[
J_{2,t} = \left( \sigma_{22} - H_{2,t} \sigma_{12} \right) m_t(\theta) + \left( \sigma_{11} H_{2,t} - \sigma_{12} \right) S_t(\theta),
\]

with \(H_{1,t} = 1 + 2\lambda\) and \(H_{2,t} = 1 + \lambda\).

The corresponding information matrix of \(\theta\) is

\[
I_{Q\theta}(\theta) = \begin{bmatrix}
I_{\lambda\lambda}^Q & I_{\lambda\omega}^Q \\
I_{\omega\lambda}^Q & I_{\omega\omega}^Q
\end{bmatrix},
\]

where \(I_{\lambda\lambda}^Q = \frac{n(1 - \omega)^2}{F_t} E_t, I_{\lambda\omega}^Q = \frac{n\lambda^2 L_t}{F_t}, I_{\omega\lambda}^Q = I_{\omega\omega}^Q = \frac{n\lambda^2 I_t}{F_t},\) and symmetrical elements, \(I_{\lambda\lambda}^Q = I_{\omega\omega}^Q = I_{\omega\lambda}^Q = I_{\lambda\omega}^Q = I_{\lambda\lambda}^Q = I_{\omega\omega}^Q\).
\[\frac{n\lambda (1 - \omega) P_t}{F_t}\] such that \(E_t = \sigma_{22} + H_{1,t} (H_{1,t} \sigma_{11} - 2\sigma_{12}), F_t = \sigma_{11} \sigma_{22} - \sigma_{12}^2, L_t = \sigma_{22} + \sigma_{11} H_{2,t}^2 - 2H_{2,t} \sigma_{12}, P_t = -\sigma_{22} - \sigma_{11} (1 + 3\lambda + 2\lambda^2) + \sigma_{12} (2 + 3\lambda)\).

For simplicity, we only compare the information for the parameter \(\lambda\). Hence, the information gain by estimating functions for martingale differences, \(m_t(\theta)\) and \(S_t(\theta)\) are \(I_{\lambda,\lambda}^m = n(1 - \omega)^2 / \sigma_{11}\) and \(I_{\lambda,\lambda}^S = n(1 - \omega)^2 (1 + 2\lambda) / \sigma_{11}\), respectively. It is noted that the value of the denominator of \(I_{\lambda,\lambda}^Q\) is small when the numerator is large, leading to a greater value of information gained compared to \(I_{\lambda,\lambda}^m\) and \(I_{\lambda,\lambda}^S\), i.e., \(I_{\lambda,\lambda}^Q > I_{\lambda,\lambda}^m\) and \(I_{\lambda,\lambda}^Q > I_{\lambda,\lambda}^S\). Hence, one can say that, the combined estimating functions is more informative than the individual elements.

3.2 ZIP regression model

We consider the ZIP regression model. In some cases, we may parameterize both \(\lambda\) and \(\omega\) in terms of exogenous explanatory variables, say \(x\) and \(z\). Following the definition given by Staub and Winkelmann (2012), we assume that \(\lambda = \exp(\lambda_0 + \lambda_1 x)\) and \(\omega = \exp(\delta_0 + \delta_1 z) / [1 + \exp(\delta_0 + \delta_1 z)]\), where \(z\) can be identical to \(x\), overlap with \(x\), or be completely distinct from \(x\).

The parameter is \(\theta = (\lambda_0, \lambda_1, \delta_0, \delta_1)'\) and the conditional expectation function of the corresponding ZIP model is given by

\[E(y \mid x, z) = \frac{\exp(\lambda_0 + \lambda_1 x)}{1 + \exp(\delta_0 + \delta_1 z)}.\]  

Hence, the conditional mean, variance, skewness and kurtosis are defined as

\[\mu_t(\theta) = \frac{\exp(\lambda_0 + \lambda_1 x_t)}{1 + \exp(\delta_0 + \delta_1 z_t)},\]
\[\sigma_t^2(\theta) = \frac{\exp(\lambda_0 + \lambda_1 x_t)T_t}{[1 + \exp(\delta_0 + \delta_1 z_t)]^2},\]
\[\gamma_t(\theta) = \frac{1 + 3\lambda_t \omega_t + \lambda_t^2 \omega_t + 2\lambda_t^2 \omega_t^2}{\mu_t [1 + \lambda_t \omega_t]^{\frac{3}{2}}},\]
\[\kappa_t(\theta) = \frac{\mu_t [1 + \lambda_t \omega_t]^{\frac{3}{2}}}{F_t},\]

respectively, where \(F_t = \omega_t \lambda_t^3 (6\omega_t^2 - 6\omega_t + 1) + 6\omega_t \lambda_t^2 (2\omega_t - 1) + 7\omega_t \lambda_t + 1\) and \(T_t = 1 + \exp(\delta_0 + \delta_1 z_t) + \exp(\delta_0 + \lambda_1 x_t + \delta_0 + \delta_1 z_t)\).

The derivative of \(\mu_t(\theta)\) with respect to each parameter is \(\partial \mu_t(\theta) / \partial \theta = (B_1, B_2, B_3, B_4)'\), where

\[B_{1,t} = \frac{\exp(\lambda_0 + \lambda_1 x_t)}{1 + \exp(\delta_0 + \delta_1 z_t)},\]
\[B_{2,t} = \frac{\exp(\lambda_0 + \lambda_1 x_t) x_t}{1 + \exp(\delta_0 + \delta_1 z_t)},\]
\[B_{3,t} = \frac{\exp(\lambda_0 + \lambda_1 x_t + \delta_0 + \delta_1 z_t)}{[1 + \exp(\delta_0 + \delta_1 z_t)]^2},\]
\[B_{4,t} = \frac{\exp(\lambda_0 + \lambda_1 x_t + \delta_0 + \delta_1 z_t) z_t}{[1 + \exp(\delta_0 + \delta_1 z_t)]^2}.

Now, let \(\Upsilon_t = 1 / (\eta_t \sigma_t^2(\theta)) D_t\), where \(\eta_t = \kappa_t(\theta) + 2 - \gamma_t^2(\theta)\), and \(D_t = -\sigma_t^2(\theta) (\kappa_t(\theta) + 2 - (1 + \lambda_t \omega_t)) + \gamma_t \sigma_t(\theta) (Y_t - \mu_t(\theta)) (1 + \lambda_t \omega_t + (Y_t - \mu_t(\theta))) - \sigma_t^3(\theta) \gamma_t(\theta)\) with \(t = 1, 2, \ldots, n\).

The optimal QEF for \(\lambda_0, \lambda_1, \delta_0\) and \(\delta_1\) respectively

\[g_Q^*(\lambda_0) = \sum_{t=1}^{n} \Upsilon_t B_{1,t}, \quad g_Q^*(\lambda_1) = \sum_{t=1}^{n} \Upsilon_t B_{2,t}, \quad g_Q^*(\delta_0) = \sum_{t=1}^{n} \Upsilon_t B_{3,t}, \quad g_Q^*(\delta_1) = \sum_{t=1}^{n} \Upsilon_t B_{4,t}.
\]

The corresponding information matrix for \(\theta\) based on the QEF method is therefore

\[I_{g_Q^*}(\theta) = \sum_{t=1}^{n} \left( \frac{1}{\sigma_t^2(\theta) \eta_t} \right) \times \begin{pmatrix} \sigma_t(\theta) (\kappa_t(\theta) + 2) W_t^2 \\ -2 \gamma_t(\theta) W_t^2 \end{pmatrix} 3_t, \]

where

\[3_t = \begin{bmatrix} B_{1,t}^2 & B_{1,t} B_{2,t} & B_{1,t} B_{3,t} & B_{1,t} B_{4,t} \\ B_{2,t} B_{1,t} & B_{2,t}^2 & B_{2,t} B_{3,t} & B_{2,t} B_{4,t} \\ B_{3,t} B_{1,t} & B_{3,t} B_{2,t} & B_{3,t}^2 & B_{3,t} B_{4,t} \\ B_{4,t} B_{1,t} & B_{4,t} B_{2,t} & B_{4,t} B_{3,t} & B_{4,t}^2 \end{bmatrix},\]

and \(W_t = \exp(\lambda_0 + \lambda_1 x_t + \delta_0 + \delta_1 z_t).\)
For the purpose of comparison, we focus only on the parameters $\lambda_0$ and $\lambda_1$ with the information gained by the use of estimating functions based on the single element of martingale differences, $m_t(\theta) = y_t - \mu_t(\theta)$ and $s_t(\theta) = m_t^2(\theta) - \sigma_t^2(\theta)$. For $m_t(\theta)$, the information matrices based on the parameter $\lambda$ are

$$I_{\lambda_0,\lambda_0}^m = \sum_{i=1}^n B_{1,t}^2 / \sigma_t^2(\theta),$$

and

$$I_{\lambda_1,\lambda_1}^m = \sum_{i=1}^n B_{2,t}^2 / \sigma_t^2(\theta).$$

While for $s_t(\theta)$, we have

$$I_{\lambda_0,\lambda_0}^s = \sum_{i=1}^n h_t W_t^2 B_{1,t}^2,$$

and

$$I_{\lambda_1,\lambda_1}^s = \sum_{i=1}^n l h_t W_t^2 B_{2,t}^2,$$

where $h_t = [1/\{\sigma_t^4(\theta) (\kappa_t(\theta) + 2)]$. Through the derivation of the information matrix in equation (6), we have the following results, where $I_{\lambda_0,\lambda_0}^Q > I_{\lambda_0,\lambda_0}^m$ and $I_{\lambda_0,\lambda_0}^Q > I_{\lambda_0,\lambda_0}^s$ as well as $I_{\lambda_1,\lambda_1}^Q > I_{\lambda_1,\lambda_1}^m$ and $I_{\lambda_1,\lambda_1}^Q > I_{\lambda_1,\lambda_1}^s$. Therefore, for the ZIP regression model, we can conclude that the information acquired using the QEF method is more informative than that of its linear components.

3.3 ZIPINGARCH($p, q$) time series model

In this last part, we focus on the model analogues to the GARCH($p, q$) model with its conditional distribution following ZIP. The model is denoted as ZIPINGARCH($p, q$). Let $y_t$ denote a count time series with excess zeroes conditional on $\mathcal{Y}_{t-1}$. Let

$$p(y_t|\mathcal{Y}_{t-1}) = \begin{cases} \omega + (1 - \omega) e^{-\lambda_t(\theta)} & \text{for } y_t = 0, \\
(1 - \omega) e^{-\lambda_t(\theta) \frac{y_t}{\lambda_t(\theta)}} & \text{for } y_t > 0. \end{cases}$$

Here, $\lambda_t(\theta)$ is the intensity parameter based on the baseline Poisson distribution and $\omega$ is the zero-inflated parameter with $\lambda_t(\theta)$ defined by Zhu (2012) as

$$\lambda_t(\theta) = \alpha_0 + \sum_{i=1}^p \alpha_i y_{t-i} + \sum_{j=1}^q \beta_j \lambda_{t-j}(\theta).$$

Suppose that the observations $Y = (y_1, y_2, \ldots, y_n)$ are generated from the model. Then, $\theta = (\omega, \alpha_0, \alpha_1, \ldots, \alpha_p, \beta_1, \beta_2, \ldots, \beta_q)'$. The mean, variance, skewness and kurtosis of $y_t$ conditional on $\mathcal{Y}_{t-1}$ are

$$\mu_t(\theta) = (1 - \omega) \lambda_t(\theta),$$

$$\sigma_t^2(\theta) = (1 - \omega) \lambda_t(\theta) (1 + \omega \lambda_t(\theta)),$$

$$\gamma_t(\theta) = \frac{\omega (2\omega + 1) [\lambda_t(\theta)]^2 + 3\omega \lambda_t(\theta) + 1}{[(1 - \omega) \lambda_t(\theta)]^{3/2} [1 + \omega \lambda_t(\theta)]^{3/2}},$$

$$\kappa_t(\theta) = \frac{\mathcal{S}_t}{(1 - \omega) \lambda_t(\theta) [1 + \omega \lambda_t(\theta)]^2},$$

respectively, where $\mathcal{S}_t = \omega (6\omega^2 - 6\omega + 1) [\lambda_t(\theta)]^3 + 6\omega (2\omega - 1) [\lambda_t(\theta)]^2 + 7\omega \lambda_t(\theta) + 1$.

By taking $m_t(\theta) = y_t - \mu_t(\theta)$ and $s_t(\theta) = m_t^2(\theta) - \sigma_t^2(\theta)$, we have $\langle m_t \rangle = \sigma_t^2(\theta)$, $\langle s_t \rangle = \sigma_t^2(\theta) (\kappa_t(\theta) + 2)$ and $\langle m, s \rangle_t = \sigma_t^2(\theta) \gamma_t(\theta)$. For the derivatives of the $\mu_t(\theta)$ and $\sigma_t^2(\theta)$ with respect to $\theta$, we have

$$\frac{\partial \mu_t(\theta)}{\partial \theta} = \begin{pmatrix} -\lambda_t(\theta), (1 - \omega) \frac{\partial \lambda_t(\theta)}{\partial \alpha_0}, \ldots, \\
(1 - \omega) \frac{\partial \lambda_t(\theta)}{\partial \alpha_p}, (1 - \omega) \frac{\partial \lambda_t(\theta)}{\partial \beta_1}, \ldots, \\
(1 - \omega) \frac{\partial \lambda_t(\theta)}{\partial \beta_q} \end{pmatrix},$$

$$\frac{\partial \sigma_t^2(\theta)}{\partial \theta} = \begin{pmatrix} -\lambda_t(\theta) (1 + 2\omega \lambda_t(\theta)), C_t \frac{\partial \lambda_t(\theta)}{\partial \alpha_0}, \\
\ldots, C_t \frac{\partial \lambda_t(\theta)}{\partial \alpha_p}, \\
C_t \frac{\partial \lambda_t(\theta)}{\partial \beta_1}, \ldots, C_t \frac{\partial \lambda_t(\theta)}{\partial \beta_q} \end{pmatrix},$$

where $C_t = (1 - \omega) (1 + 2\omega \lambda_t(\theta))$. Now, let

$$A_{k,t}^m = -A_{k,t}, \quad S_{k,t}^m = \frac{S_{k,t}}{\langle m_t \rangle}, \quad P_{k,t}^v = \frac{A_{k,t}}{\langle m_t \rangle}, \quad \text{and} \quad Q_{k,t}^v = -\frac{S_{k,t}}{\langle s_t \rangle},$$

$$B_t = \left(1 - \frac{\langle m, s \rangle_t}{\langle m_t \rangle} \right)^{-1}.$$
From Theorem 1(a), \( a_{t-1}^* \) and \( b_{t-1}^* \) are given in matrix form by

\[
I_{\omega a_0} = \sum_{t=1}^{n} B_t \left[ \begin{array}{c}
A_{1,t}A_{2,t} + S_{1,t}S_{2,t} \\
\langle m \rangle_t + \langle s \rangle_t \\
- A_{1,t}S_{2,t} + \langle s \rangle_t \\
S_{1,t}A_{2,t}
\end{array} \right],
\]

\[
I_{\omega a_i} = \sum_{t=1}^{n} B_t \left[ \begin{array}{c}
A_{1,t}A_{3,i,t} + S_{1,t}S_{3,i,t} \\
\langle m \rangle_t + \langle s \rangle_t \\
- A_{1,t}S_{3,i,t} + S_{1,t}A_{3,i,t} \\
\langle s \rangle_t
\end{array} \right],
\]

\[
I_{\omega \beta_j} = \sum_{t=1}^{n} B_t \left[ \begin{array}{c}
A_{1,t}A_{4,i,t} + S_{1,t}S_{4,i,t} \\
\langle m \rangle_t + \langle s \rangle_t \\
- A_{1,t}S_{4,i,t} + S_{1,t}A_{4,i,t} \\
\langle s \rangle_t
\end{array} \right],
\]

Thus, the optimal estimating functions for each component of \( \theta \) are

\[
I_{\omega a_i} = \sum_{t=1}^{n} B_t \left[ \begin{array}{c}
A_{1,t}A_{2,t} + S_{1,t}S_{2,t} \\
\langle m \rangle_t + \langle s \rangle_t \\
- A_{1,t}S_{2,t} + \langle s \rangle_t \\
S_{1,t}A_{2,t}
\end{array} \right],
\]

\[
I_{\omega a_i} = \sum_{t=1}^{n} B_t \left[ \begin{array}{c}
A_{1,t}A_{3,i,t} + S_{1,t}S_{3,i,t} \\
\langle m \rangle_t + \langle s \rangle_t \\
- A_{1,t}S_{3,i,t} + S_{1,t}A_{3,i,t} \\
\langle s \rangle_t
\end{array} \right],
\]

\[
I_{\omega \beta_j} = \sum_{t=1}^{n} B_t \left[ \begin{array}{c}
A_{1,t}A_{4,i,t} + S_{1,t}S_{4,i,t} \\
\langle m \rangle_t + \langle s \rangle_t \\
- A_{1,t}S_{4,i,t} + S_{1,t}A_{4,i,t} \\
\langle s \rangle_t
\end{array} \right],
\]

for \( i = 1, 2, \ldots, p \), and

\[
g_{\mathbf{Q}}(\omega) = \sum_{t=1}^{n} \left( \left( A_{m,t} + S_{1,t} \right) m_t(\theta) \right) + (P_{1,t} + Q_{1,t}) s_t(\theta),
\]

\[
g_{\mathbf{Q}}(\alpha_0) = \sum_{t=1}^{n} \left( \left( A_{m,t} + S_{m,t} \right) m_t(\theta) \right) + (P_{2,t} + Q_{2,t}) s_t(\theta),
\]

\[
g_{\mathbf{Q}}(\alpha_i) = \sum_{t=1}^{n} \left( \left( A_{3,i,t} + S_{3,i,t} \right) m_t(\theta) \right) + (P_{3,i,t} + Q_{3,i,t}) s_t(\theta),
\]

\[
g_{\mathbf{Q}}(\beta_j) = \sum_{t=1}^{n} \left( \left( A_{4,j,t} + S_{4,j,t} \right) m_t(\theta) \right) + (P_{4,j,t} + Q_{4,j,t}) s_t(\theta),
\]

for \( j = 1, 2, \ldots, q \).

The corresponding information matrix of the optimal QEF for \( \theta \) is a \((p + q + 2) \times (p + q + 2)\) matrix with the elements are given by

\[
I_{\omega a_0} = \sum_{t=1}^{n} B_t \left[ \begin{array}{c}
A_{1,t}A_{2,t} + S_{1,t}S_{2,t} \\
\langle m \rangle_t + \langle s \rangle_t \\
- A_{1,t}S_{2,t} + \langle s \rangle_t \\
S_{1,t}A_{2,t}
\end{array} \right],
\]

\[
I_{\omega a_i} = \sum_{t=1}^{n} B_t \left[ \begin{array}{c}
A_{1,t}A_{3,i,t} + S_{1,t}S_{3,i,t} \\
\langle m \rangle_t + \langle s \rangle_t \\
- A_{1,t}S_{3,i,t} + S_{1,t}A_{3,i,t} \\
\langle s \rangle_t
\end{array} \right],
\]

\[
I_{\omega \beta_j} = \sum_{t=1}^{n} B_t \left[ \begin{array}{c}
A_{1,t}A_{4,i,t} + S_{1,t}S_{4,i,t} \\
\langle m \rangle_t + \langle s \rangle_t \\
- A_{1,t}S_{4,i,t} + S_{1,t}A_{4,i,t} \\
\langle s \rangle_t
\end{array} \right],
\]

with the symmetrical elements, \( I_{\omega a_0} = I_{\omega a_0}, I_{\omega a_i} = I_{\omega a_i}, I_{\omega \beta_j} = I_{\beta_j \omega}, I_{\alpha_i a_0} = I_{\alpha_i a_0}, I_{\alpha_i a_j} = I_{\beta_j \alpha_i} \) and \( \xi_t = \langle m, s \rangle_t / \langle m, s \rangle_t \).

Again, we compare the information only for parameter \( \lambda_i(\theta) \). Using estimating functions based on \( m_t(\theta) = y_t - \mu_t(\theta) \), the information matrix \( I_{\mathbf{g}_m}(\theta) \) is computed with the elements

\[
I_{\omega a_0} = \sum_{t=1}^{n} A_{3,t}^2 / \omega^2(\theta),
\]

\[
I_{\omega a_i} = \sum_{t=1}^{n} A_{3,t}^2 / \omega^2(\theta),
\]

\[
I_{\beta_j \beta_j} = \sum_{t=1}^{n} A_{4,j}^2 / \omega^2(\theta).
\]

Similarly, the information matrix \( I_{\mathbf{g}_2}(\theta) \) for
component \( s_t(\theta) \) are

\[
I_\alpha^{\theta} = \sum_{t=1}^{n} D_t S_{2,t}^{2},
\]

\[
I_{\alpha,\alpha} = \sum_{t=1}^{n} D_t S_{3,t}^{2},
\]

and

\[
I_{\beta,\beta} = \sum_{t=1}^{n} D_t S_{4,t}^{2},
\]

where \( D_t = [1/(\sigma^4_t(\theta)(\kappa_t(\theta) + 2))] \).

From the information using the QEF method and information via its components \( m_t(\theta) \) and \( s_t(\theta) \), it is clear that

\[
I_Q^{\alpha,\alpha} > I_m^{\alpha,\alpha}, \quad I_Q^{\beta,\beta} > I_m^{\beta,\beta}, \quad I_Q^{\beta,\beta} > I_s^{\beta,\beta}, \quad I_Q^{\alpha,\beta} > I_s^{\alpha,\beta}.
\]

Therefore, we can suggest that the QEF is more informative than the component estimating functions for the ZIPINGARCH\((p, q)\) time series model.

4. Applications

This section discusses the applications of the QEF method on the ZIP models using two real data sets.

4.1 Basic ZIP model

By adopting a count data time series from the Forecasting Principles site at http://www.forecastingprinciples.com, the data represent 144 monthly counts of arson in the 13th police car beat in Pittsburgh, Pennsylvania, USA from January 1990 until December 2001. The data have 54 zeroes, i.e. 37.5\% of the series. The plot of data is given in Figure 1.

![Fig. 1. Monthly counts of arson in the 13th police car beat in Pittsburgh from January 1990 until December 2001](image)

We fitted the basic ZIP model on the data and estimated the parameters using QEF, LEF and ML methods. Using R-CRAN programming language, the ML estimates can be obtained by maximizing the likelihood of the model through a \texttt{nlminb} function, while for QEF and LEF methods, their estimates were obtained by solving their respective simultaneous optimal estimating functions via a \texttt{nleqsv} function.

<table>
<thead>
<tr>
<th>Method</th>
<th>( \hat{\lambda} )</th>
<th>( \hat{\omega} )</th>
<th>AIC</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>QEF</td>
<td>1.156(0.031)</td>
<td>0.075(0.015)</td>
<td>455.82</td>
<td>621.03</td>
</tr>
<tr>
<td>LEF</td>
<td>1.159(0.054)</td>
<td>0.089(0.018)</td>
<td>458.14</td>
<td>627.35</td>
</tr>
<tr>
<td>ML</td>
<td>1.158(0.032)</td>
<td>0.083(0.018)</td>
<td>458.68</td>
<td>624.28</td>
</tr>
</tbody>
</table>
Table 1 reports the estimated parameters, estimated standard errors, AIC and BIC for ZIP models using three different estimation methods. We observe that the standard errors of QEF estimates are lower than those of LEF and ML estimates. In addition, the basic ZIP model with QEF estimates gives the lowest AIC and BIC values. This indicates that a model with QEF estimates gives a better model fit for arson data than the model with LEF and ML estimates.

To assess the performance of these QEF estimates to LEF and ML estimates, a simulation study was carried out. We use a block bootstrap method provided by R-CRAN programming language: tseries package and tsbootstrap command with m=20 blocks to form a new series and estimate the model parameters of the new series using QEF, LEF and ML methods. The procedure is repeated for nb=500 replications and the summary statistics: mean, bias, standard error (SE) and root mean squared error (RMSE) for each parameter are computed. The simulation results are shown in Table 2. We observe that the QEF estimates give lower estimated bias than LEF and ML estimates. In addition, the estimated SEs based on simulation are consistent with the empirical results for all three methods as shown in Table 1.

Table 2. Simulation results based on bootstrap method for basic ZIP model

<table>
<thead>
<tr>
<th>Method</th>
<th>Estimated parameter</th>
<th>Mean</th>
<th>Bias</th>
<th>SE</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>QEF</td>
<td>λ</td>
<td>1.192</td>
<td>0.037</td>
<td>0.031</td>
<td>0.048</td>
</tr>
<tr>
<td></td>
<td>ω</td>
<td>0.079</td>
<td>0.004</td>
<td>0.014</td>
<td>0.015</td>
</tr>
<tr>
<td>LEF</td>
<td>λ</td>
<td>1.227</td>
<td>0.068</td>
<td>0.053</td>
<td>0.086</td>
</tr>
<tr>
<td></td>
<td>ω</td>
<td>0.079</td>
<td>0.010</td>
<td>0.017</td>
<td>0.020</td>
</tr>
<tr>
<td>ML</td>
<td>λ</td>
<td>1.196</td>
<td>0.038</td>
<td>0.033</td>
<td>0.051</td>
</tr>
<tr>
<td></td>
<td>ω</td>
<td>0.077</td>
<td>0.006</td>
<td>0.019</td>
<td>0.020</td>
</tr>
</tbody>
</table>

In order to check the model adequacy for the basic ZIP model based on QEF estimates, the Pearson residual is computed which is defined as $\epsilon_t = \left( y_t \right) / \sqrt{\hat{\lambda}(1-\hat{\omega})(1+\hat{\lambda})}$. We found that the mean and variance of Pearson residuals are 0.083 and 0.956, respectively. These values are relatively close to zero and unity indicates that the data is adequately fitted to the model. The Ljung-Box (LB) test results as given in Table 3 also indicated that there is no significant serial correlation in the residual. This shows that the data fit well using the basic ZIP model with estimated parameters, $\hat{\lambda}$ and $\hat{\omega}$, say ZIP $(\hat{\lambda}, \hat{\omega})$. Furthermore, Figure 2 shows that the cumulative periodogram plot (see Brockwell and Davis (1991)) does not cross the dotted line. Therefore, we can conclude that, the basic ZIP model via QEF estimates gives a better fit for the arson data.

Table 3. Diagnostics for basic ZIP model

<table>
<thead>
<tr>
<th></th>
<th>$LB_{30}(\epsilon_t)$</th>
<th>$LB_{30}(\epsilon_t^2)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\chi^2$</td>
<td>23.6</td>
<td>30.4</td>
</tr>
<tr>
<td>p-value</td>
<td>0.785</td>
<td>0.477</td>
</tr>
</tbody>
</table>

4.2 ZIP regression model

For the ZIP regression model, we consider data from the National Medical Expenditure Survey concerning medical care utilization by the
Efficient estimation in ZIP models with applications to count data

Fig. 2. Cummulative periodogram plot

older American conducted in 1987 and 1988 in the United States. The data set included 4406 people above the age of 66 years that were covered by medicare. The data is available at http://qed.econ.queensu.ca/iae/1997-
v12.3/deb-trivedi/. The number of patients with chronic conditions \(x\) and the number of doctor visits in a hospital \(y\) were the only two variables included from the data (Figure 3).

Fig. 3. Number of doctor visits in hospital versus number of patients (≥ 66 years) with chronic conditions

The estimated parameters together with their standard errors using all three methods are shown in Table 4. It is clear that the estimates using the QEF method give the lowest standard errors when compared to the other two methods. Furthermore, the QEF method produces lower AIC and BIC values than LEF and ML methods. This indicates that the ZIP regression model using QEF estimates gives the best model. For the simulation study, we use the same procedure as described in Section 4.1 (see Table 5). We obtain the same conclusion as discussed in Section 4.1. First, QEF estimates give the smallest bias. Second, the standard errors based on the simulation are consistent with the empirical results in Table 4.
Table 4. Parameter estimates, AIC and BIC for the ZIP regression model. Values in parenthesis are standard errors of parameter estimates

<table>
<thead>
<tr>
<th>Method</th>
<th>$\hat{\lambda}_0$</th>
<th>$\hat{\lambda}_1$</th>
<th>$\hat{\delta}_0$</th>
<th>$\hat{\delta}_1$</th>
<th>AIC</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>QEF</td>
<td>0.531 (0.010)</td>
<td>0.119 (0.015)</td>
<td>0.125 (0.081)</td>
<td>0.189 (0.013)</td>
<td>1258.36</td>
<td>1425.98</td>
</tr>
<tr>
<td>LEF</td>
<td>0.541 (0.014)</td>
<td>0.117 (0.021)</td>
<td>0.129 (0.086)</td>
<td>0.193 (0.015)</td>
<td>1262.23</td>
<td>1428.33</td>
</tr>
<tr>
<td>ML</td>
<td>0.532 (0.013)</td>
<td>0.121 (0.018)</td>
<td>0.127 (0.081)</td>
<td>0.190 (0.013)</td>
<td>1260.11</td>
<td>1426.05</td>
</tr>
</tbody>
</table>

Table 5. Simulation results based on bootstrap method for ZIP regression model

<table>
<thead>
<tr>
<th>Method</th>
<th>Estimated parameter</th>
<th>Mean</th>
<th>Bias</th>
<th>SE</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>QEF</td>
<td>$\hat{\lambda}_0$</td>
<td>0.547</td>
<td>0.016</td>
<td>0.010</td>
<td>0.019</td>
</tr>
<tr>
<td></td>
<td>$\hat{\lambda}_1$</td>
<td>0.122</td>
<td>0.003</td>
<td>0.016</td>
<td>0.017</td>
</tr>
<tr>
<td></td>
<td>$\hat{\delta}_0$</td>
<td>0.183</td>
<td>0.058</td>
<td>0.081</td>
<td>0.100</td>
</tr>
<tr>
<td></td>
<td>$\hat{\delta}_1$</td>
<td>0.188</td>
<td>0.001</td>
<td>0.013</td>
<td>0.014</td>
</tr>
<tr>
<td>LEF</td>
<td>$\hat{\lambda}_0$</td>
<td>0.559</td>
<td>0.018</td>
<td>0.015</td>
<td>0.023</td>
</tr>
<tr>
<td></td>
<td>$\hat{\lambda}_1$</td>
<td>0.127</td>
<td>0.010</td>
<td>0.021</td>
<td>0.024</td>
</tr>
<tr>
<td></td>
<td>$\hat{\delta}_0$</td>
<td>0.237</td>
<td>0.108</td>
<td>0.087</td>
<td>0.174</td>
</tr>
<tr>
<td></td>
<td>$\hat{\delta}_1$</td>
<td>0.186</td>
<td>0.007</td>
<td>0.015</td>
<td>0.016</td>
</tr>
<tr>
<td>ML</td>
<td>$\hat{\lambda}_0$</td>
<td>0.552</td>
<td>0.020</td>
<td>0.013</td>
<td>0.024</td>
</tr>
<tr>
<td></td>
<td>$\hat{\lambda}_1$</td>
<td>0.129</td>
<td>0.008</td>
<td>0.019</td>
<td>0.021</td>
</tr>
<tr>
<td></td>
<td>$\hat{\delta}_0$</td>
<td>0.246</td>
<td>0.119</td>
<td>0.082</td>
<td>0.145</td>
</tr>
<tr>
<td></td>
<td>$\hat{\delta}_1$</td>
<td>0.188</td>
<td>0.002</td>
<td>0.013</td>
<td>0.015</td>
</tr>
</tbody>
</table>

4.3 ZIPINGARCH($p, q$) time series model
For this model, a similar data set as mentioned in Section 4.1 was used. We obtain the parameter estimates, standard errors and AIC and BIC values for ZIPINGARCH (1, 1) model via QEF, LEF and ML estimates (see Table 6). The standard errors of the QEF estimates give comparable results with the ML method, but they are lower than the LEF method. The ZIPINGARCH (1, 1) model using QEF method gives the lowest AIC and BIC values compared to the LEF and ML methods. The same procedure as outlined in Section 4.1 was used. The simulation results can be found in Table 7. The QEF estimates give lower estimated bias than LEF and ML estimates. We also observe that estimated SEs and RMSEs based on QEF estimates are lower than that LEF and ML estimates. The estimated SEs based on the simulation are agreed in agreement with the empirical results for all three methods as shown in Table 6. The mean and variance of Pearson residuals are close to zero and unity which are 0.039 and 0.992, respectively, indicating the adequacy of the model. The results of the LB test suggest that there is no significant serial correlation in the residuals (Table 8).

Table 6. Parameter estimates, AIC and BIC for ZIPINGARCH(1, 1) model. Values in parenthesis are standard errors of parameter estimates

<table>
<thead>
<tr>
<th>Method</th>
<th>$\hat{\alpha}_0$</th>
<th>$\hat{\alpha}_1$</th>
<th>$\hat{\beta}_1$</th>
<th>$\hat{\omega}$</th>
<th>AIC</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>QEF</td>
<td>0.277 (0.007)</td>
<td>0.089 (0.015)</td>
<td>0.217 (0.049)</td>
<td>0.214 (0.017)</td>
<td>394.16</td>
<td>405.36</td>
</tr>
<tr>
<td>LEF</td>
<td>0.281 (0.012)</td>
<td>0.084 (0.023)</td>
<td>0.233 (0.067)</td>
<td>0.208 (0.019)</td>
<td>397.23</td>
<td>408.37</td>
</tr>
<tr>
<td>ML</td>
<td>0.279 (0.007)</td>
<td>0.087 (0.014)</td>
<td>0.213 (0.051)</td>
<td>0.214 (0.017)</td>
<td>396.15</td>
<td>406.63</td>
</tr>
</tbody>
</table>
Table 7. Simulation results based on bootstrap method for ZIPINGARCH(1,1) model

<table>
<thead>
<tr>
<th>Method</th>
<th>Estimated parameter</th>
<th>Mean</th>
<th>Bias</th>
<th>SE</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>QEF</td>
<td>$\hat{\alpha}_0$</td>
<td>0.289</td>
<td>0.012</td>
<td>0.007</td>
<td>0.014</td>
</tr>
<tr>
<td></td>
<td>$\hat{\alpha}_1$</td>
<td>0.094</td>
<td>0.005</td>
<td>0.015</td>
<td>0.016</td>
</tr>
<tr>
<td></td>
<td>$\hat{\beta}_1$</td>
<td>0.226</td>
<td>0.008</td>
<td>0.047</td>
<td>0.048</td>
</tr>
<tr>
<td></td>
<td>$\hat{\omega}$</td>
<td>0.210</td>
<td>0.004</td>
<td>0.017</td>
<td>0.018</td>
</tr>
<tr>
<td>LEF</td>
<td>$\hat{\alpha}_0$</td>
<td>0.296</td>
<td>0.014</td>
<td>0.012</td>
<td>0.019</td>
</tr>
<tr>
<td></td>
<td>$\hat{\alpha}_1$</td>
<td>0.109</td>
<td>0.025</td>
<td>0.024</td>
<td>0.034</td>
</tr>
<tr>
<td></td>
<td>$\hat{\beta}_1$</td>
<td>0.245</td>
<td>0.012</td>
<td>0.067</td>
<td>0.068</td>
</tr>
<tr>
<td></td>
<td>$\hat{\omega}$</td>
<td>0.198</td>
<td>0.010</td>
<td>0.018</td>
<td>0.021</td>
</tr>
<tr>
<td>ML</td>
<td>$\hat{\alpha}_0$</td>
<td>0.290</td>
<td>0.011</td>
<td>0.008</td>
<td>0.013</td>
</tr>
<tr>
<td></td>
<td>$\hat{\alpha}_1$</td>
<td>0.095</td>
<td>0.008</td>
<td>0.015</td>
<td>0.016</td>
</tr>
<tr>
<td></td>
<td>$\hat{\beta}_1$</td>
<td>0.237</td>
<td>0.024</td>
<td>0.049</td>
<td>0.054</td>
</tr>
<tr>
<td></td>
<td>$\hat{\omega}$</td>
<td>0.199</td>
<td>0.015</td>
<td>0.018</td>
<td>0.023</td>
</tr>
</tbody>
</table>

Table 8. Diagnostics for ZIPINGARCH(1,1) model

<table>
<thead>
<tr>
<th></th>
<th>LB$_{30}(\epsilon_t)$</th>
<th>LB$_{30}(\epsilon^2_t)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\chi^2$</td>
<td>25.4</td>
<td>24.22</td>
</tr>
<tr>
<td>p-value</td>
<td>0.705</td>
<td>0.762</td>
</tr>
</tbody>
</table>

This means that the ZIPINGARCH(1,1) model via QEF estimates fit appropriately with the data. The cumulative periodogram plots further support the model’s accuracy (Figure 4).

Fig. 4. Cummulative periodogram plot

5. Conclusions
This paper considers the QEF method for estimating the parameters of ZIP models. We have shown the superiority of the QEF method compared to the LEF method, theoretically. Results also show that the information gain using the QEF method are more informative than that LEF method for the count data in ZIP models. Through the empirical studies, it is found that the ZIP models via QEF esti-
mates provide a better fit than the LEF and ML estimates. Hence, from the findings, the QEF method could serve as an alternative parameter estimation method in estimating the parameters for this class of count data models.

6. Acknowledgements
The authors wish to thank the editors and the referees for their valuable comments and suggestions. This research is financially supported by the Fundamental Research Grant Scheme (No. FP012-2013A).

References


Submitted: 02/11/2016
Revised : 08/02/2017
Accepted : 22/03/2017
التقدير الكفؤ في غازج البواسون
المتضخم عند الصفر (ZIP) وتطبيقاته في بيانات العد

1نورول إن محمد، 2 إبراهيم محمد، 3 إن جي هاور، 4 عمد يحيى
معهد العلوم الرياضية، كلية العلوم، جامعة مالايا، 50603 كوالالمبور، ماليزيا

* imohamed@um.edu.my

الملخص

تم استخدام دوال التقدير لتقدير معلومات العديد من غازج السلاسل الزمنية المستمرة. ومع ذلك، لم يتم تطبيق هذه المنهجية على غازج بيانات العد. في هذا البحث، نستخدم في هذا البحث دوال التقدير من الدرجة الثانية (QEF) للحصول على تقدير مشترك (ZIP) للمعلومات المتوسط والتبان المشروطة لغازج بيانات العد، وخاصة مفهوم بواسون المتضخم عند الصفر (ZIP) ونموذج الانحدار الذاتي غير المتجانس المعمم للأعداد الصحيحة التي يكون توزيعها الشرطي. تشير النتائج إلى أن التقديرات التي تم الحصول عليها باستخدام دوال التقدير من الدرجة الثانية، والتي تستخدم معلومات من دوال تقدير مزودة، هي أكثر معلومانية من دوال التقدير الخطية (LEF)، والتي تستخدم فقط معلومات من دوال التقدير. وأخيراً، نحن بالتطبيق على بيانات حقيقية باستخدام غازج QEF وطرق LEF وQEF وطرق ZIP وطرق الاحتمال الأرجح، وتوضح هذه التطبيقات تقدم طريقة QEF على باقي الطرق عملياً.