Mathematical Biosciences and Engineering

http://www.aimspress.com/journal/MBE

MBE, 18(3): 2303–2330.
DOI: 10.3934/mbe.2021117
Received: 14 December 2020
Accepted: 24 February 2021
Published: 08 March 2021

Research article

Modeling of daily confirmed Saudi COVID-19 cases using inverted exponential regression

Sarah R. Al-Dawsari and Khalaf S. Sultan*

Department of Statistics and Operations Research, College of Science, King Saud University, P.O.Box 2455, Riyadh 11451, Saudi Arabia

* **Correspondence:** Email: ksultan@ksu.edu.sa; Tel: +966-1-467-6263; Fax: +966-1-467-6274.

Abstract: The coronavirus disease 2019 (COVID-19) pandemic caused by the coronavirus strain has had massive global impact, and has interrupted economic and social activity. The daily confirmed COVID-19 cases in Saudi Arabia are shown to be affected by some explanatory variables that are recorded daily: recovered COVID-19 cases, critical cases, daily active cases, tests per million, curfew hours, maximal temperatures, maximal relative humidity, maximal wind speed, and maximal pressure. Restrictions applied by the Saudi Arabia government due to the COVID-19 outbreak, from the suspension of Umrah and flights, and the lockdown of some cities with a curfew are based on information about COVID-19. The aim of the paper is to propose some predictive regression models similar to generalized linear models (GLMs) for fitting COVID-19 data in Saudi Arabia to analyze, forecast, and extract meaningful information that helps decision makers. In this direction, we propose some regression models on the basis of inverted exponential distribution (IE-Reg), Bayesian (BReg) and empirical Bayesian regression (EBReg) models for use in conjunction with inverted exponential distribution (IE-BReg and IE-EBReg). In all approaches, we use the logarithm (log) link function, gamma prior and two loss functions in the Bayesian approach, namely, the zero-one and LINEX loss functions. To deal with the outliers in the proposed models, we apply Huber and Tukey's bisquare (biweight) functions. In addition, we use the iteratively reweighted least squares (IRLS) algorithm to estimate Bayesian regression coefficients. Further, we compare IE-Reg, IE-BReg, and IE-EBReg using some criteria, such as Akaike's information criterion (AIC), Bayesian information criterion (BIC), deviance (D), and mean squared error (MSE). Finally, we apply the collected data of the daily confirmed from March 23 - June 21, 2020 with the corresponding explanatory variables to the theoretical findings. IE-EBReg shows good model for the COVID-19 cases in Saudi Arabia compared with the other models

Keywords: Bayesian generalized linear models; empirical Bayesian; Huber's function; LINEX loss function; log link function; Tukey's bisquare function; zero-one loss function

1. Introduction

Since the beginning of 2020, the world has been facing a coronavirus pandemic (COVID-19) at a rapid and alarming rate, representing a significant challenge for humanity and a serious threat to life. Although many countries have taken multiple and sometimes harsh measures to limit the spread of the virus and reduce its spread, the eyes of the public are now turning to scientists, doctors, and researchers of all scientific disciplines in the hope of finding a quick and successful treatment for this virus. Many authors have investigated the association of environmental and meteorological factors on the spread of COVID-19, see for example, Tello-Leal, et al. [1], Kodera, et al. [2], Meo, et al. [3], Casado-Aranda, et al. [4], Dogan, et al. [5], Fu, et al. [6] and Yuan, et al. [7].

McCullagh and Nelder [8] published a book on GLMs that led to their widespread use and appreciations. They extended the scoring method to maximum-likelihood estimation (MLE) in exponential families. Nelder and Pregibon [9] described methods of jointly estimating parameters of both link and variance functions. The iteratively reweighted least squares (IRLS) algorithm is amenable to statistics and measures that are common to all GLMs. Nelder and Wedderburn [10] used the Newton–Raphson process for regression coefficient estimates. Yuan and Bentler [11] reported that the convergence properties of the Fisher scoring algorithm are affected by many factors. One among the observed variables is multicollinearity. If the sample or model implied covariance matrix is close to singular, the Fisher scoring algorithm may have difficulty reaching a set of converged solutions. Nelder and Wedderburn [10] reported that the Newton–Raphson process with expected second derivatives is equivalent to Fisher's scoring technique. Additionally, de Jong and Heller [12] reported that the Newton–Raphson iteration equation leads to a sequence that often rapidly converges. This includes the D statistic along with specific residuals and influence measures. Liao [13] introduced a systematic way of interpreting commonly used probability models: logit, probit, and other GLMs. For recent works of using these models to the field of epidemiology and to the healthcare, see for example, Richardson and Hartman [14], Song, et al. [15], Mohamadou, et al. [16] and Trunfio, et al. [17].

The inverted exponential (IE) distribution that was introduced by Keller and Kamath [18]. The role of IE distributions is indispensable in many applications of reliability theory for its memoryless property and its constant failure rate. Dey [20] considered IE distribution as life distribution (see Abdel-Aty et al. [19], and Dey [20]). Singh et al. [21] obtained Bayes estimates for parameters of IE distribution by using informative and noninformative priors. They also compared the classical method with the Bayesian through the simulation study.

The Bayesian approach of the statistical modeling provides an alternative to standard GLMs. Posterior-mode estimation is an alternative to full posterior analysis or posterior mean estimation, which avoids numerical integrations or simulation methods. It was proposed by many authors (for more details, see Fahrmeir and Tutz [22] and Cepeda and Gamerman [23]). Dey et al. [24] described how to conceptualize, perform, and critique traditional GLMs from a Bayesian perspective, and how to use modern computational methods to summarize inferences using simulations. Olsson [25] gave an overview of GLMs and presented practical examples. The exponential family of distributions are discussed with maximum-likelihood estimation and ways of assessing the fit of the model. For the Bayesian estimation in this context, a useful asymmetric loss function known as the LINEX loss function was introduced by Varian (1975) and has been widely used by several authors. A highly used one is the zero–one loss function (for more details, see Sano et al. [26]). Robbins [27] has provided a more robust estimate and for estimating the parameters of prior distribution (hyperparameters), studies on the empirical Bayes (EB) method, Wei [28] has proposed the EB test of the regression coefficient, and working out the EB test decision rule by using kernel estimation of multivariate density function and its first-order partial derivatives, Singh [29] has proposed an EB approach in a multiple linear-regression model, Houston and Woodruff [30] have derived EB estimates using an m-group regression model to regress within-group estimates toward common values. More studies by many authors discussed EB test problems for parameters in a class of linear-regression model and other topics, e.g., Wind [31], Huang [32], Karunamuni [33], Yuan [34], Chen [35], Efron [36], Shao [37], Kim and Nembhard [38], and Jampachaisri et al. [39].

In order to reduce the influence of outliers on the estimate, some robust measures were proposed in the literature. The common robust estimation method can be divided into several categories: M, MM, median, L1, Msplit, R, S, least-trimmed squares, and sign-constraint robust least squares estimation. Among these, Huber's M estimation has become one of the main robust estimation methods by virtue of its simple calculation and convenience to implement (see Li et al. [40]). The key aspect is the involvement of a loss function that is applied to data errors that was selected to less rapidly increase than the square loss function that is used in least-squares or maximum-likelihood procedures. There exist several well-known families of loss functions, such as Huber, Hampel, and Tukey's biweight (or bisquare) that can be used for the computation of M estimators (see Sinova and Aelst [41]).

A major contribution of this paper is to propose similarity to GLMs, except that the distribution of the response is not a member of the exponential family using the Bayesian approach. We interest in IE distribution of which the flexible distribution can describe different lifetimes from medicine, reliability, ecology, biological studies, and other areas. We propose the Bayesian and non Bayesian inverted exponential regression models to model and analyze Covid-19-related data with the aim of explaining the relationships between Covid-19 cases and environmental-related variables. The paper is organized as follows: In Section 2, we present an overview of GLMs and propose the IE-Reg model under a log link function. In Section 3, we perform IE-BReg and IE-EBReg under a gamma prior, log link, and two loss functions. We propose Huber's and Tukey's bisquare (biweight) function to improve Bayesian models. We also adopted the iteratively reweighted least squares (IRLS) algorithm to estimate the Bayesian regression coefficients. In Section 4, we apply IE-Reg, IE-BReg, and IE-EBReg models, including an estimation, and use criteria such as AIC, BIC, MSE, and D to the Saudi COVID-19 dataset collected from March 23 to June 21, 2020. Finally, Section 5 draws a succinct conclusion to the findings.

2. Classical approach

Nelder and Wedderburn [10] introduced the class of GLMs, defined according to the assumption that $y_1, y_2, ...y_n$ are observations of the response variable, with density function y_i as follows:

$$f(y_i;\theta_i) = e^{\theta_i y_i - \psi(\theta_i) + c(y_i)}, \quad i = 1, 2, ..., n, \tag{2.1}$$

where $\psi(\cdot)$, $c(\cdot)$ are known functions, with θ_i being the canonical parameter. Link function $g(.)$, related to the regression coefficients, is given by

$$g(\mu_i) = \eta_i = x_i'\beta, \quad i = 1, 2, ..., n, \tag{2.2}$$

where $g(\mu_i) = \theta_i, \beta = (\beta_1, ..., \beta_p)'$ is a vector of p unknown regression parameters, $x_i' = (x_{i1}, x_{i2}, ..., x_{ip})$ is a vector of explanatory variables, and η_i is a linear predictor of vectors x_i' and β. Here, $g(.) : (0, \infty) \longrightarrow R$ is a link function, which is a monotonic differentiable invertible function. The model given by (2.1) and (2.2) is called the GLM. The GLM class includes, as special cases, linear-regression and analysis-of-variance models, logit and probit models for quantal responses, log–linear models, and multinomial response models for counts (for more details, see McCaullagh and Nelder [42]).

Consider that the probability density function of IE distribution is as follows (see Abdel-Aty et al. [19]):

$$f(y;\gamma) = \frac{\gamma}{y^2} e^{-\frac{\gamma}{y}}; \quad y > 0, \quad \gamma > 0, \tag{2.3}$$

which has no mean and γ is a scale parameter. The median value of the response variable is given by

$$\tilde{\mu} = \frac{\gamma}{\log(2)}. \tag{2.4}$$

Since mean does not exist, we use the median $\tilde{\mu}$ instead of it in the link function (see Das and Dey [43]). The cumulative function of IE distribution is given by

$$F(y;\gamma) = e^{-\frac{\gamma}{y}}; \quad y > 0.$$

Let y_i be a random sample from IE, and $\gamma_i = \tilde{\mu}_i \log(2)$, the log-likelihood function based on y_i, is given by

$$l_i = l(\tilde{\mu}_i | y_i) = \log(\tilde{\mu}_i) + \log\left(\frac{\log(2)}{y_i^2}\right) - \frac{\log(2)\tilde{\mu}_i}{y_i}, \quad i = 1, 2, ..., n. \tag{2.5}$$

Regression coefficients are estimated using Fisher's scoring technique (for more details, see Nelder and Wedderburn [10], and McCaullagh and Nelder [42]). In order to develop the GLMs for our models, IE-Reg is similar to GLMs, except that the distribution of the response variable is not a member of the exponential family (Ferrari and Cribari-Neto [44]). We also suggest the logarithm (log) link functions of $g(.)$, in view of (2.2) as in the following lemma.

Lemma 2.1:

Let the response variable Y have an IE distribution, $i = 1, 2, ..., n$, and let the link function of the form be

$$g(\tilde{\mu}_i) = \log(\tilde{\mu}_i) = \eta_i = x_i'\beta, \quad i = 1, 2, ..., n. \tag{2.6}$$

Thus, estimated coefficients $\hat{\beta}' = (\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_p)$ using Fisher's scoring technique at the sth iteration based on the IRLS process are given by

$$\hat{\beta}^{(s)} = (X'X)^{-1} X'Z, \quad s = 1, 2, 3, ..., \tag{2.7}$$

where X is a covariates matrix, $\hat{\beta}_j^{(0)}$ is an initial vector, $Z' = (z_1, z_2, \ldots, z_n)$, and

$$z_i = \sum_{j=1}^{p} x_{ij} \hat{\beta}_j^{(s-1)} + \left(1 - \frac{\tilde{\mu}_i^{(s-1)} \log(2)}{y_i}\right). \tag{2.8}$$

The procedure in (2.7) can be repeated until $|\hat{\beta}^{(s)} - \hat{\beta}^{(s-1)}| \leq \varepsilon$. IE-Reg model in this case is given by

$$\hat{\tilde{\mu}}_i^{(s)} = e^{\hat{\beta}_0^{(s)} + \hat{\beta}_1^{(s)} x_{i1} + ... + \hat{\beta}_p^{(s)} x_{ip}}. \tag{2.9}$$

Proof: See the appendices

3. Bayesian approach

Diaconis and Ylvisker [43] introduced conjugate prior distribution for the exponential family, which, as in (2.1), can be shown as

$$\pi(\theta_i) = k_1 e^{m\mu_0\theta_i - m\psi(\theta_i)}, \quad i = 1, 2, ..., n, \tag{3.1}$$

where k_1 is a normalization constant, and m, μ_0 are natural parameters. θ_i values are connected to the regression coefficients by link function $\eta_i = x_i'\beta$ as

$$g^*(\eta_i) = \theta_i. \tag{3.2}$$

Posterior distribution of θ_i is given by

$$\pi(\theta_i|y_i) = k_2 e^{(y_i+m\mu_0)\theta_i-(1+m)\psi(\theta_i)}. \tag{3.3}$$

Das and Dey [43] suggested a Jacobian of transformation and rewrote (3.3) with term η_i, as

$$\pi(\eta_i|y_i) = k_2 e^{(y_i+m\mu_0)g^*(\eta_i)-(1+m)\psi(g^*(\eta_i))}\frac{\partial g^*(\eta_i)}{\partial \eta_i}, \tag{3.4}$$

where k_2 is a normalization constant, and $\frac{\partial g^*(\eta_i)}{\partial \eta_i} \neq 0$. They used a zero–one loss function to attain the posterior mode of (3.4) as $\hat{\eta}_i = h(y_i)$; hence, estimated coefficients $\hat{\beta}^* = (\hat{\beta}_0^*, \hat{\beta}_1^*, ..., \hat{\beta}_p^*)'$ are given by

$$\hat{\beta}^* = (X'X)^{-1}X'\hat{\eta}, \tag{3.5}$$

where $\hat{\beta}^*$ is the least-squares estimates, and $\hat{\eta}' = (\hat{\eta}_1, \hat{\eta}_2, ..., \hat{\eta}_n)$ (for more details, see Das and Dey [43], and Das and Dey [45]).

3.1. IE-BReg and IE-EBReg models

In order to develop a Bayesian approach, we suggest Bayesian and empirical Bayesian regression models (IE-BReg and IE-EBReg) that are similar to Bayesian GLMs, except that the distribution of the response variable is not a member of the exponential family. We used the general form of the posterior in (3.4), and since $g(\cdot)$ is a monotonic differentiable function, we then attain posterior Bayes estimates. In addition, we use log link function with zero–one and LINEX loss functions to be appropriate of Bayes estimates. IE-BReg and IE-EBReg estimators correspond to the log link function using different loss functions, as in the following lemmas.

Lemma 3.1: (IE-BReg and IE-EBReg models based on zero–one loss function)

Let the response variable Y have an IE distribution, and the link function of the form be as in (2.6). Consider that $\tilde{\mu}$ has a gamma prior $G(\alpha, \lambda)$ with the following density function

$$\pi(\tilde{\mu}) = \frac{\lambda^\alpha}{\Gamma(\alpha)} e^{-\lambda\tilde{\mu}}\tilde{\mu}^{\alpha-1}, \quad \tilde{\mu} > 0, \lambda, \alpha > 0. \tag{3.6}$$

Thus, the posterior mode of η_i by using zero–one loss function can be derived by solving the following equation:

$$\hat{\eta}_i = \log\left(\frac{\alpha+1}{\lambda + \frac{\log(2)}{y_i}}\right), \quad i = 1, 2, ..., n. \tag{3.7}$$

Estimated coefficients $\hat{\beta}^*$ are given as in (3.5). In this case, IE-BReg and IE-EBReg models are given by

$$\hat{\tilde{\mu}}_i^* = e^{\hat{\beta}_0^*+\hat{\beta}_1^* x_{i1}+....+\hat{\beta}_p^* x_{ip}}. \quad (3.8)$$

In the case of IE-EBReg, empirical Bayes estimates for unknown prior distribution parameter λ are given by MLE from the data. Therefore, IE-EBReg estimates are found by placing these estimated prior distribution parameter into Equation (3.7) by $\hat{\lambda}$.

Proof: See the appendices

Lemma 3.2

Let the response variable Y have an IE distribution, and the link function of the form be as in (2.6). Consider that $\tilde{\mu}$ has a gamma prior $G(\alpha, \lambda)$ with the density function as in (3.6). Using Jacobian transformation from $\tilde{\mu}_i$ to η_i, the posterior function of η_i can be written as in the following equation

$$\pi(\eta_i|y_i) \propto e^{(1+\alpha)\eta_i} e^{-e^{\eta_i}\left(\lambda+\frac{\log(2)}{y_i}\right)}.$$

Thus, the scale parameter of the prior $G(\alpha, \lambda)$ is less than or equal one ($\lambda \leq 1$) and its variance (σ^2) is greater than or equal the mean (μ).

Proof: See the appendices

Lemma 3.3: (IE-BReg and IE-EBReg models based on LINEX loss function)

Let the response variable Y have an IE distribution, and let the link function of the form be as given in (2.6). Consider that $\tilde{\mu}$ has a gamma prior with density function as given in (3.6). As a result, the posterior Bayes estimates of η_i, by using the LINEX loss function, can be derived as follows

$$\hat{\eta}_i = \frac{-1}{\alpha} \log\left(\frac{\lambda^\alpha \log(2)}{y_i^2 \Gamma(\alpha)\left(\lambda + \frac{\log(2)}{y_i}\right)}\right), \quad i = 1, 2, ..., n. \quad (3.9)$$

The estimated coefficients $\hat{\beta}^*$ and the IE-BReg and IE-EBReg models and their estimates in this case are given as in (3.5) and (3.8), respectively. In the caes of IE-EBReg model, the regression coefficients estimates is found by placing these estimated prior distribution parameter into Equation (3.9) by $\hat{\lambda}$.

Proof: See the appendices

Lemma 3.4:

Let the response variable Y have an IE distribution, and the link function of the form be as in (2.6). Consider that $\tilde{\mu}$ has a gamma prior $G(\alpha, \lambda)$ with the density function as in (3.6). Using Jacobian transformation from $\tilde{\mu}_i$ to η_i, the posterior function of η_i can be written as in the following equation

$$\pi(\eta_i|y_i) \propto [g^*(\eta_i)]^\alpha e^{-g^*(\eta_i)\left(\lambda+\frac{\log(2)}{y_i}\right)} \frac{\partial g^*(\eta_i)}{\partial \eta_i}. \quad (3.10)$$

Thus, the scale parameter of the prior $G(\alpha, \lambda)$ is greater than or equal one ($\lambda \geq 1$) and its variance (σ^2) is less than or equal the mean (μ).

Proof: See the appendices

3.2. Robust IE-BReg and IE-EBReg models

M-estimation is considered to be the most common method of robust regression. It was proposed by Huber [46] in the presence of outliers, and it is more efficient than ordinary least squares (OLS) (Rousseeuw and Leroy [47], and Chang, et al. [48]). The Huber's function takes the following form (Huber [46, 49]):

$$\rho(r) = \begin{cases} \frac{r^2}{2}, & |r| \leq k, \\ k(|r| - \frac{k}{2}), & |r| > k, \end{cases} \tag{3.11}$$

where k is the tuning constant, r is the residual corresponding to the observation in OLS, and $\rho(\cdot)$ is the objective function that satisfies certain properties. Often, $\rho(\cdot)$ can be formed by using a linear combination of the residuals. Defining function $\frac{\partial}{\partial r}\rho(r)$ and the corresponding weight function in this case is as follows:

$$\frac{\psi(r)}{r} = w(r) = \begin{cases} 1, & |r| \leq k, \\ \frac{k}{|r|}, & |r| > k. \end{cases} \tag{3.12}$$

Another M-estimation function is the Tukey bisquare's (biweight) function. This is based on Tukey's function, taking the form of that in Sinova and Van Aelst [41]

$$\rho(r) = \begin{cases} \frac{k^2}{6}(1 - [1 - (\frac{r}{k})^2]^3), & |r| \leq k, \\ \frac{k^2}{6}, & |r| > k, \end{cases} \tag{3.13}$$

where k is the tuning constant and r is the residual corresponding to the observation in OLS. Defining function $\frac{\partial}{\partial r}\rho(r) = \psi(r)$ and the corresponding weight function in this case is given as follows:

$$\frac{\psi(r)}{r} = w(r) = \begin{cases} [1 - (\frac{r}{k})^2]^2, & |r| \leq k, \\ 0, & |r| > k. \end{cases} \tag{3.14}$$

To make the IE-BReg and IE-EBReg models are robust, we suggest the Huber's and biweight functions for these models based on an adopted IRLS algorithm. There are also many other versions of the M-estimation function that could be used here.

Lemma 3.5: (IE-BReg and IE-EBReg models based on M-estimation functions)

Let the response variable Y have an IE distribution, and let the link function of the form be as given in (2.6). Consider that $\tilde{\mu}$ has a gamma prior with density function is given as in (3.6). Using the Jacobian transformation from $\tilde{\mu}_i$ to η_i and using the log link function, we have the posterior distribution of η_i is given as in (3.10). Thus, the estimated coefficients $\hat{\beta}^* = (\hat{\beta}_0^*, \hat{\beta}_1^*, ..., \hat{\beta}_p^*)'$ are given as

$$\hat{\beta}^{*(s)} = \left(X'W\left(\hat{\beta}^{*(s-1)}\right)X\right)^{-1} X'W\left(\hat{\beta}^{*(s-1)}\right)\hat{\eta}, \quad s = 1, 2, 3, \ldots, \tag{3.15}$$

where $\hat{\eta}_i = h(y)$ and $\hat{\eta}' = (\hat{\eta}_1, \hat{\eta}_2, ..., \hat{\eta}_n)$ are the posterior Bayes estimates of η_i using the zero–one or LINEX loss functions, and $W = diag(w_1, w_2, \ldots, w_n)$, w_i are the selected weights depending on M-estimation functions. In this case, coefficients are estimated using the adopted IRLS Algorithm.

An adopted algorithm based on IRLS and M-estimation is employed as follows:

Equation (3.15) is solved using an adopted algorithm on the basis of the standard IRLS algorithm (for more details, see Maronna et al. [50], Wen and Liu [51], and Kikuchi et al. [52]). This algorithm is proposed for the solution of the IE-BReg and IE-EBReg estimates, that is employed in the following steps:

(i) Setting the iteration counter at $q = 0$, finding an initial estimates of regression coefficients $\hat{\beta}_j^{*(q)}$, $j = 0, 1, 2, ..., p-1$ using IE-Reg estimates.

(ii) The initial residuals $r_{(i)}^{*(q)} = Y_i - e^{(X_i' \hat{\beta}_j^{*(q)})}$ are based on the log link function that is given as in (2.6), and calculate an initial scale estimate $s^{*(q)} = 1.4826(median|r_i^{*(q)}|)$.

(iii) An initial standardized residuals $u_i^{*(q)} = \frac{r_i^{*(q)}}{s^{*(q)}}$ are calculated and used to calculate initial estimates for the weight function. Preliminary weights are $w_i^{*(q)} = w(u_i^{*(q)})$.

(iv) Calculate $\hat{\lambda}$ as the prior distribution $\pi(\tilde{\mu})$ is $G(\alpha, \lambda)$ in the case of IE-EBReg using MLE estimates, or the scale parameter λ is known in the case of IE-BReg model.

(v) Calculate Bayes estimates $\hat{\eta}_i = h(y); i = 1, ..., n$ using the prior $G(\alpha, \lambda)$ and zero–one or LINEX loss functions.

(vi) Using weights from Steps i–iii and Steps iv and v to find estimators in (3.15).

(vii) Set $q = q + 1$; then, go to Step ii. Steps (ii) to (vii) are repeated until the estimate of $\hat{\beta}^{*(q)}$ is stabilized from the previous iteration, which means: $|\hat{\beta}^{*(q+1)} - \hat{\beta}^{*(q)}| \leq \varepsilon$.

Under regularity conditions, estimator $\hat{\beta}^*$ has asymptotically normal distribution $\hat{\beta}^* \equiv N\left(\beta^*, (X'WX)^{-1}\right)$ (see Houston and Woodruff [30], Tellinghuisen [53]).

4. Data analysis

In this section, we apply the IE-Reg, IE-BReg, and IE-EBReg models for the daily confirmed COVID-19 cases in Saudi Arabia. The relevant dataset in this application is COVID-19 data from Saudi Arabia in 2020. These data contain 91 observations from March 23–June 21, 2020 in which the response variable Y is the daily confirmed COVID-19 cases in Saudi Arabia. Explanatory variables are: X_1; daily recovered COVID-19 cases, X_2, daily critical COVID-19 cases; X_3, daily active COVID-19 cases; X_4, tests per million (PCR tests); X_5, curfew hours per day; X_6, maximal temperature in Celsius per day; X_7, maximal relative humidity (%); X_8, maximal wind speed in miles per hour (mph); and X_9, maximal pressure in hectopascal (hPa). In the case of variables $X_i, i = 6, ..., 9$ are the average for the cities of Riyadh, Jeddah, and Dammam that have had the highest number of confirmed and death cases. This dataset was taken from the Ministry of Health of Saudi Arabia (COVID-19 dashboard) and the Ministry of Health's Twitter account.

Lemmas in Sections 2 and 3 were applied to these data. IE-Reg, IE-BReg, and IE-EBReg models based on log link and loss functions were used. Bayes coefficients were obtained using a gamma prior $G(\alpha, \lambda)$ to a known shape α and an unknown scale λ parameter. We used generated data from a gamma prior distribution to estimate λ in the case of IE-BReg model and use our data in the case of IE-EBReg model. We also compared the performance of all these models. Different plots, such as the quantile–quantile (Q–Q) plot, the empirical cumulative distribution function (ECDF), and box plot, were proposed to aid in distributional assessment and identify outliers. In addition, Huber's and biweight functions are suggested in the case of Lemma (3.5) to avoid such distortions due to outliers (for more details, see Sinova and Van Aelst [41]). The backward-selection method is used in the IE-EBReg model to remove the input variable (see Table 3). Modeling performance is measured in terms of some criteria, such as AIC, BIC, D, D/df (divided by its degrees of freedom), and mean square error (MSE) (de Jong and Heller [12]). We also used Thiel's inequality coefficient (TIC) to compare the prediction accuracy of the selected models (Leuthold [54], and Niu et al. [55]). To check the adequacy for the selected models, we consider the deviance residuals (see McCaullagh and Nelder [42]). The predictive results of these models and other numerical results are shown in Tables 1–8. R software was used to carry out calculations. In order to compare with known distributions, the glm() function in "stats" was used to fit the GLMs (Faraway [56]). Functions qqPlot(), ecdf(), boxplot, and ks.test() in R package "stats" were used for the assessment distributions (Fox and Weisberg [57]).

K-S test is used to recommend the IE distribution as a good fit for the daily confirmed COVID-19 cases in Saudi Arabia comparing with some other distributions as given below:

K-S test of daily confirmed COVID-19 cases in Saudi Arabia using different distributions.

Distribution	Gaussian	Exponential	Gamma	Inverted Exponential	Inverse Gaussian
p-value	0.0000	0.0000	0.141	0.169	0.016

Based on the results obtained from the above table, the large p-value for the test indicated that IE distribution fit the response variable in the given data quite well. Figure 1 provides the Q–Q plot, ECDF, and fitted functions of the selected models, and it is clear that IE distribution fits these data well.

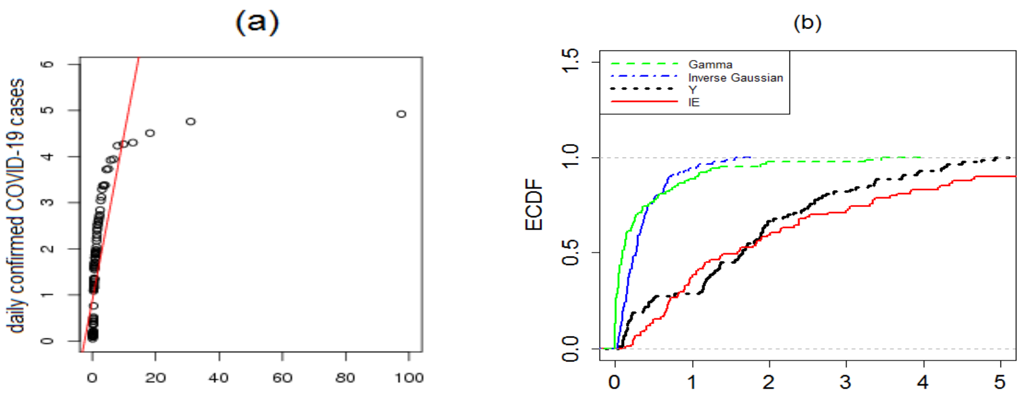

Figure 1. (a) Q-Q plots of the daily confirmed COVID-19 cases in Saudi Arabia based on IE (b) ECDF plot based on different distributions.

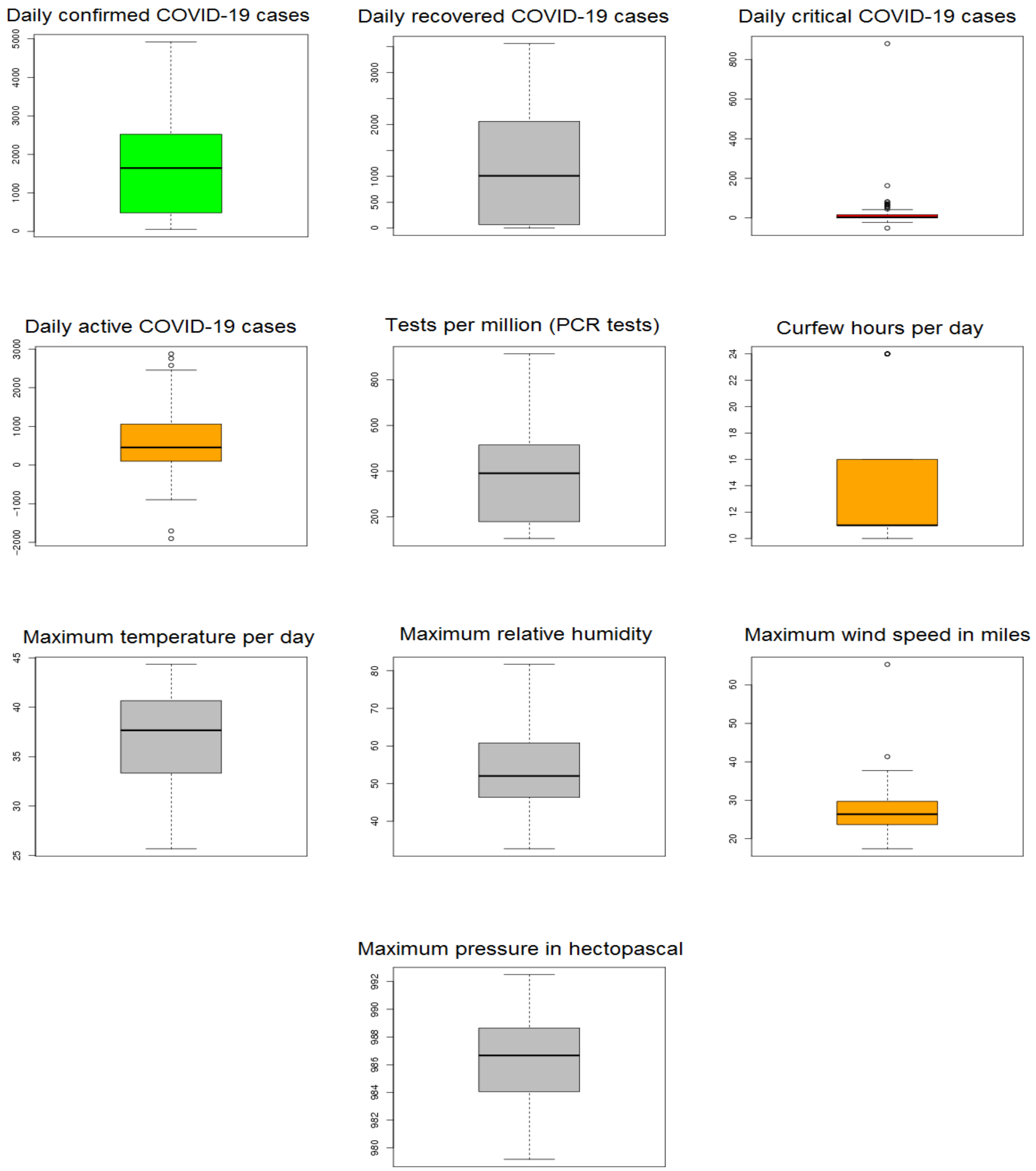

Figure 2. Box plots of the variables of Saudi Arabia COVID-19 dataset.

Figure 2 presents box plots corresponding to each of the Saudi Arabia COVID-19 dataset variables, and the chart also maps outliers that exceeded the values of fences. The plot also displays the maximum, minimum, and median of the data, along with the first and third quantile. The outliers could also be identified as in Figure 2 (shown as unfilled circles) in explanatory variable X_2, and we can observe a big difference between the maximal value and the rest of the observation, which is beyond the outer fence. This case could be considered a very extreme outlier. The chart also shows more outliers on variables X_3, X_5, and X_8. For the X_2 and X_3 box plots, variables having outliers were much higher than the third quartile or much lower than the first quartile of the box was. Possible reasons for this deviation may be when daily active or critical cases exceeded the recovered cases, or daily recovered cases exceeded active or critical cases. There is no outlier in the daily-confirmed-cases variable in the context of box-plot analysis, and this variable was slightly asymmetrically distributed with a relatively heavy tail.

Table 1 presents that the MSE value of the IE-Reg, gamma models was large, while the MSE values of the inverse Gaussian was very large. Table 2 shows that the MSE value of the IE-Reg model was large, while the MSE values of the IE-BReg and IE-EBReg models were small. In the case of IE-Reg model, $\hat{\beta}^{(s)}$ was stabilized when Fisher's scoring procedure converged at $s = 15$ because of $|\hat{\beta}^{(15)} - \hat{\beta}^{(14)}| < 0$. In the case of the IE-BReg and IE-EBReg models, the fitting results based on Huber's function had the largest MSE values compare to the MSE values based on biweight function. The results in this table also show that the MSE values based on biweight function and without use any weight function were almost similar, but fitting results based on biweight function were best in terms of AIC, and D statistics. Comparing between Bayesian fitting results, we observed that the results based on LINEX loss function were better than those of the zero–one loss function. Gamma prior variance was larger than the mean in the case of used the zero–one loss function, while gamma prior variance was less than the mean in the case of used the LINEX loss function.

Table 1. Efficiency of Gamma, inverse Gaussian, Inverted exponential(IE-Reg), IE-Bayesian regression (IE-BReg), and IE-empirical Bayesian regression (IE-EBReg) models.

The Model	IE-Reg	IE-BReg	IE-EBReg	Inverse Gaussian	Gamma
Prior		$G(2.610, 1.389)$			
Loss and Weight function		LINEX and Huber	LINEX and biweight		
AIC	39.339	64.867	65.837	181.256	147.107
D	4.666	30.194	31.163	25.093	18.392
D/df	0.058	0.373	0.385	0.302	0.222
MSE	3.880	0.201	0.188	5935513	1.303

Table 2 also shows that the IE-EBReg model based on gamma prior $G(2.610, 1.420)$ and biweight function was the best on our dataset, since D and D/df ($D/df = 0.380$), $df = 83$ were acceptable with a low MSE compared to that of the other models. Tables 1 and 2 show that the D/df of this model was very close to 1, indicating that the fitting degree was very good. The value of z statistic was large, so there was also a significant relationship among the variables.

Through this comparison, we can conclude that the predictive model was IE-EBReg based on biweight and LINEX functions, and it is given as follows:

$$\hat{Y} = e^{-27.535+0.457x_1+0.415x_3+0.675x_4+21.012x_5+33.815x_6+13.687x_7+24.888x_9}. \tag{4.1}$$

In the side of the adopted IRLS procedure based on M-estimation, $\hat{\beta}^{*(s)}$ is stabilized and it converged at $s = 15$ because of $|\hat{\beta}^{*(15)} - \hat{\beta}^{*(14)}| < 4.2e - 06$. The performance of the selected model was indicative that the adopted algorithm works well. The actual and predicted for the IE-EBReg model compared to IE-Reg are represented graphically in Figure 3. Figure 4 show that the deviance residuals against the indices of the observations suggest that the residuals for the IE-EBReg model are randomly scattered around zero. From Table 2 and 3, we can conclude that, for the selected model among possible IE-EBReg models based on the smallest AIC = 54.151 and lowest MSE = 0.230 with eight variables, this fitting result was the best when using level of significance $\alpha = 0.05$.

Table 4 presents the fitting results for the IE-EBReg model based on biweight function that is the best based on relative errors. For the prediction results, Table 5 shows that the IE-EBReg model based on Huber's function compared to that based on biweight function had the worst prediction accuracy because the TIC value was closer to 1, while the TIC value for the IE-EBReg model based on biweight function was closer to 0.

According to the residuals results of the IE-BReg and IE-EBReg models, it can be found that the errors of the first data are too large, which seriously affects of results, see Table 6 (Case 1). To make a precision comparison, we removed one observation form data $i = 1$ that have large error and we reestimated the model after removing from the data (Case 2). Furthermore, we replaced y_1 with the mean of observation from $i = 2$ to $i = 11$ and we reestimated the model after modifying the data (Case 3). We also replaced $y_i, i = 1, 4, 5, 6, 7, 9$ with the mean of observation from $i = 2$ to $i = 11$ and we reestimated the model after modifying the data (Case 4). However, we removed the observations $i = 1, 4, 5, 6, 7, 9$ and we reestimated the model after removing observations (Case 5). The relative changes in the parameter estimates are presented in Table 6. The the deviance residuals against the indices of the observations suggest that the residuals are randomly scattered around zero at level of significance $\alpha = 0.001$ (Cox Stuart test) for the IE-BReg and IE-EBReg models based on the priors $G(2.3537, 1.410)$ and $G(2.3537, 1.493)$.

Table 6 and 7 show that, the IE-EBReg model (Case 4) is the best for our data, and it is given as follows:

$$\hat{Y} = e^{-27.657+0.510x_1+0.466x_3+0.655x_4+26.546x_5+32.713x_6+14.248x_7+24.847x_9}. \tag{4.2}$$

In the side of the adopted IRLS procedure based on M-estimation, $\hat{\beta}^{*(s)}$ is stabilized and it converged at $s = 15$ because of $|\hat{\beta}^{*(15)} - \hat{\beta}^{*(14)}| < 0$. The actual and predicted for the IE-EBReg model compared to IE-Reg model is represented graphically in Figure 5. From Table 8, we can conclude that, this model has smallest AIC = 50.6634 and a low MSE = 0.2014, and there was also a significant relationship among the variables when using level of significance $\alpha = 0.05$. Figure 6 show that the deviance residuals against the indices of the observations suggest that the residuals for the IE-EBReg model (Case 4) are randomly scattered around zero. In comparison between IE-EBReg models that were shown in (4.1) and (4.2), and based on the results on Table 2 and 6, we can conclude that the IE-EBReg model in (4.2) is the best for our data.

Table 2. AIC, BIC, D and MSE of IE-Reg, IE-BReg, and IE-EBReg models.

Step	Model	Prior	Loss Function	AIC	BIC	D	MSE
Step 1	IE-Reg			39.339	64.448	4.666	3.880
	IE-BReg-Huber	$G(0.893, 0.641)$	zero–one	65.700	90.808	31.026	0.560
	IE-BReg-biweight			66.054	91.163	31.381	0.486
	IE-BReg			66.059	91.167	31.385	0.486
	IE-BReg-Huber	$G(2.610, 1.393)$	LINEX	67.559	89.917	30.135	0.201
	IE-BReg-biweight			65.773	90.882	31.100	0.188
	IE-BReg			65.781	90.890	31.108	0.188
	IE-EBReg-Huber	$G(0.893, 0.642)$	zero–one	65.736	90.845	31.063	0.561
	IE-EBReg-biweight			66.087	91.195	31.413	0.487
	IE-EBReg			66.092	91.200	31.418	0.487
	IE-EBReg-Huber	$G(2.610, 1.389)$	LINEX	64.867	89.976	30.194	0.201
	IE-EBReg-biweight			65.837	90.945	31.163	0.188
	IE-EBReg			65.845	90.954	31.171	0.188
Step 2	IE-Reg			37.381	59.979	4.707	3.887
	IE-BReg-Huber	$G(0.844, 0.616)$	zero–one	62.673	85.271	29.999	0.554
	IE-BReg-biweight			63.082	85.680	30.408	0.473
	IE-BReg			63.093	85.691	30.419	0.472
	IE-BReg-Huber	$G(2.350, 1.672)$	LINEX	54.783	79.891	20.109	0.262
	IE-BReg-biweight			54.128	76.725	21.454	0.237
	IE-BReg			54.155	76.752	21.481	0.237
	IE-EBReg-Huber	$G(0.844, 0.614)$	zero–one	62.600	85.198	29.926	0.554
	IE-EBReg-biwight			63.027	85.625	30.353	0.471
	IE-EBReg			63.038	85.636	30.364	0.469
	IE-EBReg-Huber	$G(2.350, 1.621)$	LINEX	52.522	75.120	19.848	0.261
	IE-EBReg-biweight			54.151	76.749	21.477	0.230
	IE-EBReg			54.186	76.784	21.512	0.230
Step 3	IE-Reg			35.621	55.708	4.947	4.015
	IE-BReg-Huber	$G(0.820, 0.593)$	zero–one	59.890	79.977	29.217	0.539
	IE-BReg-biweight			60.450	80.537	29.777	0.445
	IE-BReg			60.453	80.540	29.780	0.445
	IE-BReg-Huber	$G(2.610, 1.451)$	LINEX	60.320	80.407	29.647	0.205
	IE-BReg-biweight			60.970	81.057	30.296	0.196
	IE-BReg			60.973	81.060	30.299	0.196
	IE-EBReg-Huber	$G(0.820, 0.607)$	zero–one	60.561	80.648	29.887	0.537
	IE-EBReg-biweight			60.970	81.057	30.297	0.465
	IE-EBReg			60.974	81.061	30.300	0.465
	IE-EBReg-Huber	$G(2.610, 1.420)$	LINEX	60.493	80.580	29.819	0.200
	IE-EBReg-biweight			61.472	81.559	30.799	0.192
	IE-EBReg			61.476	81.563	30.803	0.192

Table 3. AIC, BIC, D and MSE of IE-Reg, IE-BReg, and IE-EBReg models.

Step	Variables	$\hat{\beta}$ Huber	z-Statistics	P-value	$\hat{\beta}$ biweight	z-Statistics	P-value
Step 1	Intercept	-31.068	-3.790	0.0002	-27.730	-3.457	0.0005
	X_1	0.460	16.211	0.0000	0.461	16.673	0.0000
	X_2	0.258	2.107	0.0351	0.292	2.401	0.0163
	X_3	0.419	14.949	0.0000	0.420	15.404	0.0000
	X_4	0.757	4.329	0.0000	0.648	3.876	0.0001
	X_5	23.821	5.981	0.0000	23.409	5.973	0.0000
	X_6	32.640	5.503	0.0000	32.170	5.513	0.0000
	X_7	14.379	10.368	0.0000	13.551	9.962	0.0000
	X_8	0.211	0.109	0.9132	-0.668	-0.349	0.7271
	X_9	28.421	3.486	0.0004	25.170	3.154	0.0016
Step 2	Intercept	-36.839	-4.597	0.0000	-32.435	-4.136	0.0000
	X_1	0.542	18.750	0.0000	0.523	19.327	0.0000
	X_2	0.263	2.147	0.0318	0.336	2.763	0.006
	X_3	0.495	17.643	0.0000	0.477	18.048	0.0000
	X_4	0.827	4.687	0.0000	0.729	4.430	0.0000
	X_5	25.401	6.370	0.0000	26.728	6.819	0.0000
	X_6	36.826	6.307	0.0000	37.064	6.467	0.0000
	X_7	16.828	12.124	0.0000	15.534	11.456	0.0000
	X_9	33.688	4.219	0.0000	29.331	3.752	0.0002
Step 3	Intercept	-30.825	-3.867	0.0001	-27.535	-3.518	0.0004
	X_1	0.456	16.450	0.0000	0.457	16.997	0.0000
	X_3	0.415	15.304	0.0000	0.415	15.846	0.0000
	X_4	0.790	4.599	0.0000	0.675	4.132	0.0000
	X_5	21.823	5.600	0.0000	22.012	5.692	0.0000
	X_6	33.856	5.845	0.0000	33.815	5.925	0.0000
	X_7	14.402	10.453	0.0000	13.687	10.102	0.0000
	X_9	28.142	3.543	0.0004	24.888	3.190	0.0014

Table 4. Fitting results of IE-EBReg $\hat{y}_i$ based on LINEX, Huber's, and biweight functions (during fitting interval.

Date	y_i	$\hat{y}_i$	$\lvert y_i - \hat{y}_i \rvert$	Relative error %	$\hat{y}_i$	$\lvert y_i - \hat{y}_i \rvert$	Relative error %
		Huber			biweight		
4/26/2020	1223	1348	125	10.25	1351	128	10.439
4/27/2020	1289	1172	117	9.050	1177	112	8.697
4/28/2020	1266	1.032	234	18.46	1026	240	18.959
4/29/2020	1325	1257	68	5.11	1249	76	5.724
4/30/2020	1351	1.223	128	9.470	1222	129	9.542
5/1/2020	1344	1357	13	0.994	1353	9	0.652
5/2/2020	1362	1514	152	11.15	1491	129	9.490
5/3/2020	1552	1582	30	1.94	1566	14	0.890
5/5/2020	1645	1364	281	17.07	1357	288	17.495

Table 5. Predicted results of IE-EBReg $\hat{y}_i$ based on LINEX, Huber's, and biweight functions with seven variables (out of fitting interval).

Date	y_i	$\hat{y}_i$	$\lvert y_i - \hat{y}_i \rvert$	R.E. %	TIC	$\hat{y}_i$	$\lvert y_i - \hat{y}_i \rvert$	R.E %	TIC
		Huber				biweight			
6/22/2020	3393	3340	53	1.55	0.0612	3165	228	6.731	0.0529
6/23/2020	3139	2789	350	11.16		2693	446	14.205	
6/24/2020	3123	2627	496	15.87		2519	604	19.355	
6/25/2020	3372	4000	628	18.63		3765	393	11.658	
6/26/2020	3938	4272	334	8.48		4012	73.772	1.873	
6/27/2020	3927	4448	521	13.28		4107	180.124	4.587	

R.E.: Relative Error

Table 6. AIC, D, MSE and Cox Stuart test for the deviance residuals of IE-EBReg model based on biweight function.

Model	Cases	Prior	AIC	D	MSE	Cox Stuart test p-value
IE-BReg	1	$G(2.6352, 1.3613)$	63.4138	32.7401	0.1870	0.0001
	2		64.4422	27.8167	0.1727	0.0001
	3		61.1159	28.6864	0.1727	0.0001
	4		58.9767	24.6503	0.1700	0.0001
	5		75.1276	15.6250	0.1617	0.0003
IE-EBReg	1	$G(2.6352, 1.3574)$	63.5012	32.8276	0.1869	0.0001
	2		64.5095	27.8839	0.1726	0.0001
	3		61.1903	28.7608	0.1725	0.0001
	4		59.0385	24.7120	0.1698	0.0001
	5		75.1385	15.6359	0.1612	0.0003
IE-BReg	1	$G(2.3537, 1.4910)$	52.99269	22.3190	0.2354	0.0161
	2		55.6927	19.0672	0.2090	0.0008
	3		51.8461	19.4166	0.2086	0.0008
	4		50.6692	16.3428	0.2015	0.0025
	5		70.3465	10.8439	0.1787	0.0079
IE-EBReg	1	$G(2.3537, 1.4928)$	52.9776	22.3039	0.2352	0.0161
	2		55.6846	19.0591	0.2089	0.0161
	3		51.8356	19.4061	0.2085	0.0066
	4		50.6634	16.3369	0.2014	0.0025
	5		70.3563	10.8536	0.1787	0.0079

Case 1: using the original data; n=91.
Case 2: using the data after removing one observation (i=1); n=90.
Case 3: using the data after replaced one observation by the mean ; n=91.
Case 4: using the data after replaced the observation (i=1,4,5,6,7,9) by the mean; n=91.
Case 5: using the data after removing the observations (i=1,4,5,6,7,9); n=85.

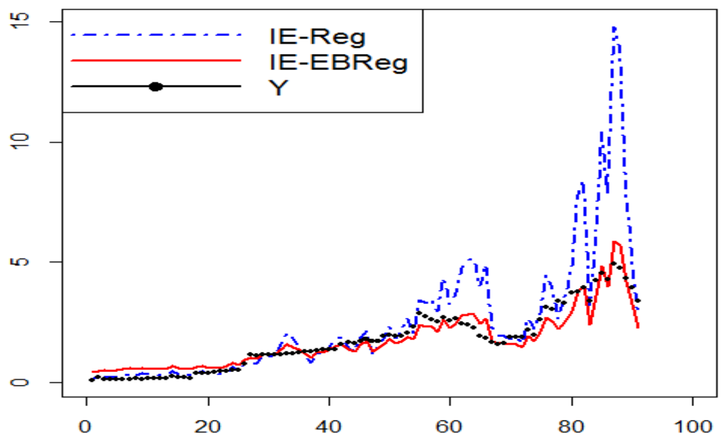

Figure 3. Fitting of IE-Reg and IE-EBReg with 7 variables as in Table 2.

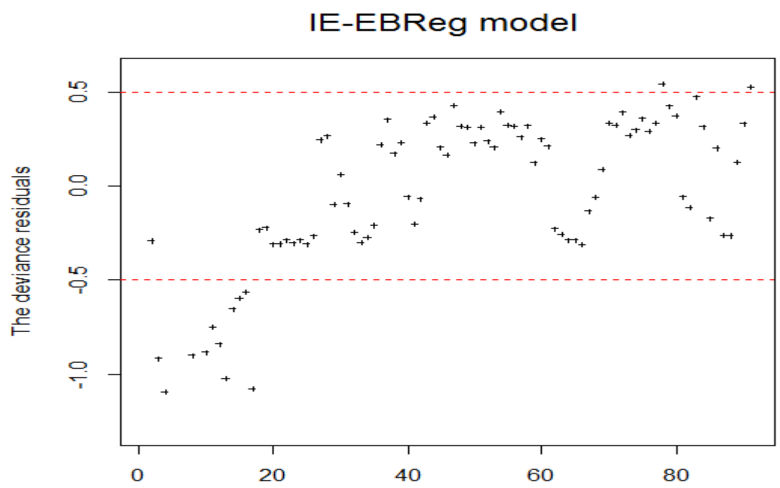

Figure 4. Plot of deviance residuals for the IE-EBReg based on biweight and LINEX as in Table 2.

Figures 3 shows that the IE-EBReg model based on biweight and LINEX loss function generally fits the dataset better than other models. This also clear from Figure 4 as the plot of the deviance residuals against the indices of the observations suggest that the residuals are randomly scattered around zero at level of significant 0.001.

Table 7. The Relative errors % of predicted and fitting results of the IE-EBReg based on prior $G(2.3537, 1.4928)$.

	Date	y_i	Case 1	Case 2	Case 3	Case 4	Case 5
Fitting results	4/26/2020	1.223	6.1574	4.0055	4.3964	5.5198	9.0944
	4/27/2020	1289	13.6654	13.5397	13.6369	12.5813	7.6865
	4/28/2020	1266	24.5773	22.1173	22.6232	22.1945	18.4163
	4/29/2020	1325	10.1971	9.4581	9.5203	9.2789	7.7580
	4/30/2020	1351	14.0620	13.5364	13.7209	12.7740	8.0207
	5/1/2020	1344	3.2130	3.7411	3.7135	2.7615	1.3526
	5/2/2020	1362	6.5542	6.0133	6.1913	6.3622	6.9241
	5/3/2020	1552	1.2706	1.6513	1.5618	1.1365	0.5708
	5/4/2020	1645	20.6544	19.1781	19.3377	19.3725	18.7971
	5/5/2020	1595	23.7866	22.1910	22.2850	22.4925	22.7538
Predicted results	6/22/2020	3.393	0.8852	8.1235	8.1629	8.3268	7.4896
	6/23/2020	3139	10.5945	16.4515	16.4037	16.2531	14.3989
	6/24/2020	3123	16.6312	23.2819	23.0829	23.2841	23.6959
	6/25/2020	3372	21.0604	8.9519	9.4537	8.5929	5.4076
	6/26/2020	3938	11.3371	2.3426	2.5350	1.4769	2.2651
	6/27/2020	3927	14.6806	4.6155	4.7346	3.3727	1.3583
TIC statistics			0.0900	0.0594	0.0594	0.0586	0.0557

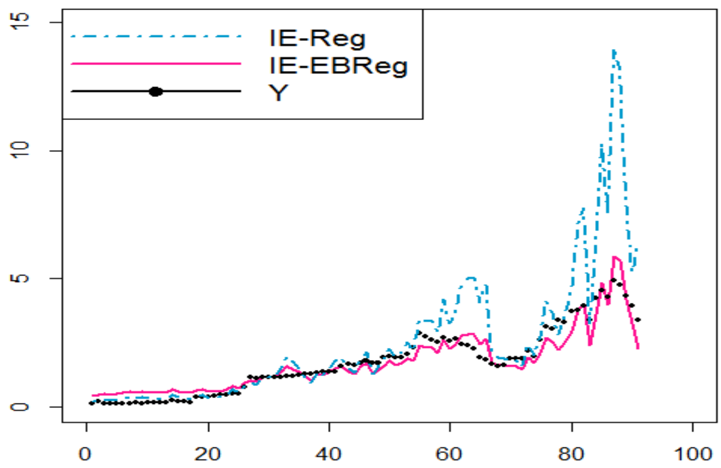

Figure 5. Fitting of IE-Reg and IE-EBReg models with 7 variables as in Table 6.

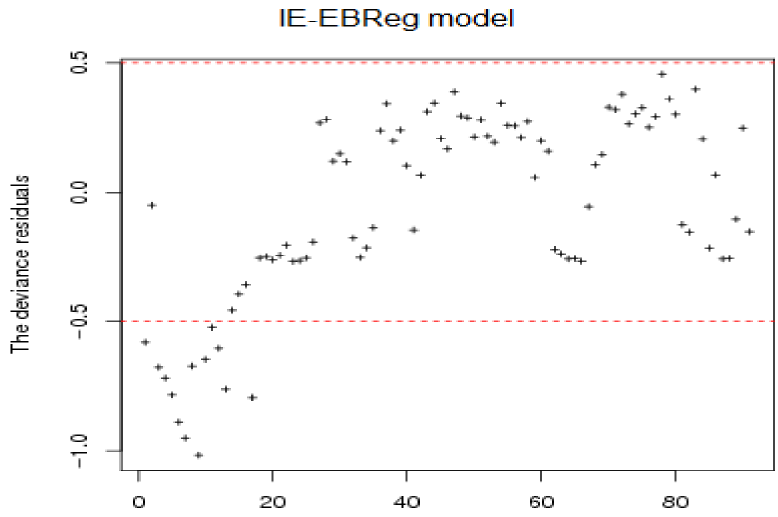

Figure 6. Plot of deviance residuals for the IE-EBReg model (Case 4) as in Table 6.

Figure 5 shows that the IE-EBReg model (Case 4) which based on biweight and LINEX function fits the dataset better. Also, Figure 6 shows the deviance residuals against the indices of the observations suggest that the residuals are randomly scattered around zero at level of significant 0.001.

Table 8. IE-EBReg model based on LINEX, biweight functions (Case 4).

Variables	$\hat{\beta}$	Z-statistics	Standard Error SE	p-value	AIC	D	MSE
Intercept	-27.6574	-3.5193	7.8588	0.0004	50.6634	16.3369	0.2014
X_1	0.5103	20.0674	0.0254	0.0000			
X_3	0.4660	19.3073	0.0241	0.0000			
X_4	0.6548	4.6898	0.1396	0.0000			
X_5	26.5458	6.9168	3.8379	0.0000			
X_6	32.7134	5.9260	5.5203	0.0000			
X_7	14.2477	10.5226	1.3540	0.0000			
X_9	24.8469	3.1705	7.8369	0.0015			

5. Conclusions

In this paper the regression models IE-Reg, IE-BReg and IE-EBReg for modeling the daily confirmed Saudi COVID-19 cases with some envionmental-related variables (covariates) are considered. Zero-one and LINEX loss functions were used to attain the Bayesian and empirical Bayesian estimates based on the log-link function. In a non-Bayesian approach, parameter estimation is done by the Fisher scoring technique, and closed-form expressions are provided for the score function, and for Fisher's information matrix and its inverse.In the Bayesian approach, parameter estimation is performed using a gamma prior distribution, Jacobian transformation, and least-squares estimates. The IE-Reg, IE-BReg, and IE-EBReg models were compared to find which model predicted better. To deal with outlier problems, IE-BReg based on Huber's and biweight functions, and the adopted algorithm based on IRLS to find the estimates, were proposed. For distributional assessment, Q-Q, ECDF, box plots, and the KS test were applied. Some criteria, namely, AIC, BIC, D, D/df, and MSE, were also computed for all regression models.

According to the results of the application, it was concluded that IE-BReg and IE-EBReg with a log link function performed the best in terms of the AIC, BIC, D, D/df, and MSE statistics, so they are recommended for these data. In contrast, IE-Reg showed poor results compared with those of the other models. Results indicated that the IE-EBReg model is highly capable of improving the performance of regression models to a greater extent in the prediction of daily confirmed COVID-19 cases in Saudi Arabia. Finally, it is found the following regressors are significant for the model: Explanatory variables are: X1, daily recovered COVID-19 cases; X3, daily active COVID-19 cases; X4, tests per million (PCR tests); X5, curfew hours per day; X6, maximal temperature in Celsius per day; X7, maximal relative humidity (%); and X9, maximal pressure in hectopascal (hPa).

Acknowledgments

The authors would like to thank the editor and referees for their helpful comments, which improved the presentation of the paper. Also, the authors would like to extend their sincere appreciation to the Deanship of Scientific Research at King Saud University for its funding this Research Group (RG -1435-056).

Conflict of interest

The authors declares that they have no conflicts of interest.

Data availability

The data is freely available at: https://covid19.moh.gov.sa/ https://twitter.com/SaudiMOH?s=20/ https://www.worldometers.info/coronavirus/ https://www.wunderground.com/.

References

1. E. Tello-Leal, B. A. Macias-Hernandez, Association of environmental and meteorological factors on the spread of COVID-19 in Victoria, Mexico, and air quality during the lockdown, *Environ. Res.*, (2020), 110442.
2. S. Kodera, E. A. Rashed, A. Hirata, Correlation between COVID-19 morbidity and mortality rates in Japan and local population density, temperature, and absolute humidity, *Int. J. Env. Res. Pub. He.*, **17** (2020), 5477.
3. S. A. Meo, A. A. Abukhalaf, A. A. Alomar, N. M. Alsalame, T. Al-Khlaiwi, A. M. Usmani, Effect of temperature and humidity on the dynamics of daily new cases and deaths due to COVID-19 outbreak in Gulf countries in Middle East Region, *Eur. Rev. Med. Pharmacol. Sci.*, **24** (2020), 7524–7533.

4. L. A. Casado-Aranda, J. Sanchez-Fernandez, M. I. Viedma-del-Jesus, Analysis of the scientific production of the effect of COVID-19 on the environment: A bibliometric study, *Environ. Res.*, (2020), 110416.
5. B. Dogan, M. B. Jebli, K. Shahzad, T. H. Farooq, U. Shahzad, Investigating the effects of meteorological parameters on COVID-19: Case study of New Jersey, United States, *Environ. Res.*, **191** (2020), 110148.
6. S. A. Meo, A. A. Abukhalaf, A. A. Alomar, O. M. Alessa, W. Sami, D. C. Klonoff, Effect of environmental pollutants PM-2.5, carbon monoxide, and ozone on the incidence and mortality of SARS-COV-2 infection in ten wildfire affected counties in California, *Sci. Total Environ.*, **757** (2021), 143948.
7. J. Yuan, Y. Wu, W. Jing, J. Liu, M. Du, Y. Wang, et al., Non-linear correlation between daily new cases of COVID-19 and meteorological factors in 127 countries, *Environ. Res.*, **193** (2021), 110521.
8. P. McCullagh, J. A. Nelder, *Generalized Linear Models*, 1^{st} edition, Chapman and Hall, London, 1983.
9. J. A. Nelder, D. Pregibon, An extended quasi-likelihood function, *Biometrika*, **74** (1987), 221–232.
10. J. A. Nelder, R. W. M. Wedderburn, Generalized linear models, *J. R. Stat. Soc. Ser. A*, **135** (1972), 370–384.
11. K. H. Yuan, P. M. Bentler, Improving the convergence rate and speed of Fisher-scoring algorithm: ridge and anti-ridge methods in structural equation modeling, *Ann. Inst. Stat. Math.*, **69** (2017), 571–597.
12. P. De Jong, G. Z. Heller, *Generalized linear models for insurance data*, 1^{st} edition, Cambridge Books, 2008.
13. T. F. Liao, *Interpreting Probability Models: Logit, Probit, and Other Generalized Linear Models*, No 07–101, SAGE Publications, Thousand Oaks, 1994.
14. R. Richardson, B. Hartman, Bayesian nonparametric regression models for modeling and predicting healthcare claims, *Insur. Math. Econ.*, **83** (2018), 1–8.
15. C. Song, Y. Wang, X. Yang, Y. Yang, Z. Tang, X. Wang, et al., Spatial and Temporal Impacts of Socioeconomic and Environmental Factors on Healthcare Resources: A County-Level Bayesian Local Spatiotemporal Regression Modeling Study of Hospital Beds in Southwest China, *Int. J. Env. Res. Pub. He.*, **17** (2020), 5890.
16. Y. Mohamadou, A. Halidou, P. T. Kapen, A review of mathematical modeling, artificial intelligence and datasets used in the study, prediction and management of COVID-19, *Appl. Intell.*, **50** (2020), 3913–3925.
17. T. A. Trunfio, A. Scala, A. D. Vecchia, A. Marra, A. Borrelli, Multiple Regression Model to Predict Length of Hospital Stay for Patients Undergoing Femur Fracture Surgery at "San Giovanni di Dio e Ruggi d'Aragona" University Hospital, *In European Medical and Biological Engineering Conference*, Springer, Cham, (2020), 840–847.
18. A. Z. Keller, A. R. R. Kamath, U. D. Perera, Reliability analysis of CNC machine tools, *Reliab. Eng.*, **3** (1982), 449–473.

19. Y. Abdel-Aty, A. Shafay, M. M. M. El-Din, M. Nagy, Bayesian inference for the inverse exponential distribution based on pooled type-II censored samples, *J. Stat. Appl. Pro.*, **4** (2015), 239.
20. S. Dey, Inverted exponential distribution as a life distribution model from a Bayesian viewpoint, *Data Sci. J.*, **6** (2007), 107–113.
21. S. K. Singh, U. Singh, A. S. Yadav, P. K. Vishwkarma, On the estimation of stress strength reliability parameter of inverted exponential distribution, *IJSW*, **3** (2015), 98–112.
22. L. Fahrmeir, G. Tutz, *Multivariate Statistical Modelling Based on Generalized Linear Models*, 2nd edition, Springer Science and Business Media, Berlin/Heidelberg, 2013.
23. E. Cepeda, D. Gamerman, Bayesian methodology for modeling parameters in the two parameter exponential family, *Rev. Estad.*, **57** (2015), 93–105.
24. D. K. Dey, S. K. Ghosh, B. K. Mallick, *Generalized Linear Models: A Bayesian Perspective*, 1st edition, CRC Press, New York, 2000.
25. U. Olsson, *Generalized Linear Models, An Applied Approach*, 1st edition, Student Litteratur Lund., Sweden, 2002.
26. N. Sano, H. Suzuki, M. Koda, A robust ensemble learning using zero-one loss function, *J. Oper. Res. Soc. Japan*, **51** (2008), 95–110.
27. H. Robbins, An empirical Bayes approach to statistics, *In Breakthroughs in statistics*, Springer, (1955), 388-394.
28. L. Wei, Empirical Bayes test of regression coefficient in a multiple linear regression model, *Acta Math. Appl. Sin–E*, **6** (1990), 251–262.
29. R. S. Singh, Empirical Bayes estimation in a multiple linear regression model, *Ann. Inst. Stat. Math.*, **37** (1985), 71–86.
30. W. M. Houston, D. J. Woodruff, Empirical Bayes Estimates of Parameters from the Logistic Regression Model, *ACT Res. Report Ser.*, (1997), 97–96.
31. S. L. Wind, An empirical Bayes approach to multiple linear regression, *Ann. Stat.*, **1** (1973), 93–103.
32. S. Y. Huang, Empirical Bayes testing procedures in some nonexponential families using asymmetric Linex loss function, *J. Stat. Plan. Infer.*, **46** (1995), 293–309.
33. R. J. Karunamuni, Optimal rates of convergence of empirical Bayes tests for the continuous one-parameter exponential family, *Ann. Stat.*, (1996), 212–231.
34. M. Yuan, Y. Lin, Efficient empirical Bayes variable selection and estimation in linear models, *J. Am. Stat. Assoc.*, **100** (2005), 1215–1225.
35. L. S. Chen, Empirical Bayes testing for a nonexponential family distribution, *Commun. Stat., Theor. M.*, **36** (2007), 2061–2074.
36. B. Efron, *Large-scale inference: Empirical Bayes methods for estimation, testing, and prediction*, 1st, Cambridge University Press, 2012.
37. M. Shao, An empirical Bayes test of parameters for a nonexponential distribution family with Negative Quadrant Dependent random samples, *In 2013 10th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD)*, IEEE, (2013), 648–652.

38. J. E. Kim, D. A. Nembhard, Parametric empirical Bayes estimation of individual time-pressure reactivity, *Int. J. Prod. Res.*, **56** (2018), 2452–2463.

39. K. Jampachaisri, K. Tinochai, S. Sukparungsee, Y. Areepong, Empirical Bayes Based on Squared Error Loss and Precautionary Loss Functions in Sequential Sampling Plan, *IEEE Access*, **8** (2020), 51460–51469.

40. Y. Li, L. Hou, Y. Yang, J. Tong, Huber's M-Estimation-Based Cubature Kalman Filter for an INS/DVL Integrated System, *Math. Probl. Eng.*, (2020), 2020.

41. B. Sinova, S. Van Aelst, Advantages of M-estimators of location for fuzzy numbers based on Tukey's biweight loss function, *Int. J. Approx. Reason.*, **93** (2018), 219–237.

42. P. McCullagh, J. A. Nelder, *Generalized Linear Models*, 2^{nd} edition, Chapman and Hall/CRC, 1989.

43. S. Das, D. K. Dey, On Bayesian analysis of generalized linear models using the Jacobian technique, *Am. Stat.*, **60** (2006), 264–268.

44. S. Ferrari, F. Cribari-Neto, Beta regression for modelling rates and proportions, *J. Appl. Stat.*, **31** (2004), 799–815.

45. S. Das, D. K. Dey, On Bayesian analysis of generalized linear models: A new perspective, Technical Report, *Statistical and Applied Mathematical Sciences Institute*, Research Triangle Park, (2007), 33.

46. P. J. Huber, Robust estimation of a location parameter, *Ann. Math. Stat.*, **35** (1964), 73–101.

47. P. J. Rousseeuw, A. M. Leroy, *Robust Regression and Outlier Detection*, 1^{st} edition, John Wiley and Sons, NY, 1987.

48. L. Chang, B. Hu, G. Chang, A. Li, Robust derivative-free Kalman filter based on Huber's M-estimation methodology, *J. Process Control*, **23** (2013), 1555–1561.

49. P. J. Huber, *Robust Statistics*, 1^{st} edition, John Wiley and Sons, NY, 1981.

50. R. A. Maronna, R. D. Martin, V. J. Yohai, *Robust Statistics: Theory and Methods*, 1^{st} edition, John Wiley and Sons, West Sussex, 2006.

51. F. Wen, W. Liu, Iteratively reweighted optimum linear regression in the presence of generalized Gaussian noise, *In 2016 IEEE International Conference on Digital Signal Processing (DSP)*, IEEE, (2016), 657–661.

52. H. Kikuchi, H. Yasunaga, H. Matsui, C. I. Fan, Efficient privacy-preserving logistic regression with iteratively Re-weighted least squares, *In 2016 11th Asia Joint Conference on Information Security (AsiaJCIS)*, IEEE, (2016), 48–54.

53. J. Tellinghuisen, Least squares with non-normal data: Estimating experimental variance functions, *Analyst*, **133(2)** (2008), 161–166.

54. R. M. Leuthold, On the use of Theil's inequality coefficients, *Am. J. Agr. Econ.*, **57** (1975), 344–346.

55. T. Niu, L. Zhang, B. Zhang, B. Yang, S. Wei, An Improved Prediction Model Combining Inverse Exponential Smoothing and Markov Chain, *Math. Probl. Eng.*, **2020** (2020), 11.

56. J. J. Faraway, *Extending the linear model with R: Generalized linear, mixed effects and nonparametric regression models*, 2^{nd} edition, CRC press, 2016.

57. J. Fox, S. Weisberg, *An R companion to applied regression*, 3^{rd} edition, Sage publications, Inc., 2018.

58. E. Dikici, F. Orderud, B. H. Lindqvist, Empirical Bayes estimator for endocardial edge detection in 3D+ T echocardiography, *In 2012 9th IEEE International Symposium on Biomedical Imaging (ISBI)*, IEEE, (2012), 1331–1334.

59. A. Coluccia, F. Ricciato, Improved estimation of instantaneous arrival rates via empirical Bayes, *In 2014 13th Annual Mediterranean Ad Hoc Networking Workshop*, IEEE, (2014), 211–216.

Appendices

5.1. Proof of Lemma 2.1:

Suppose that, in $y_i \sim f(y_i;\gamma)$, as in (2.3), the log-likelihood function based on y_i, $i = 1, 2, ..., n$ is given as in (2.5). Link function $\log(\tilde{\mu}_i)$ connecting the $\tilde{\mu}_i$ with linear model $x_i'\beta$ in this case is given as in (2.6). Score function U_r for log likelihood is written from one observation as

$$U_r(\beta) = \frac{\partial l(\beta)}{\partial \beta_r} = \sum_{i=1}^{n} \frac{\partial l_i}{\partial \tilde{\mu}_i}\frac{\partial \tilde{\mu}_i}{\partial \eta_i}\frac{\partial \eta_i}{\partial \beta_r}, \quad r = 1, 2, ..., p. \tag{5.1}$$

From (2.5) and (2.6), we have

$$U_r(\beta) = \sum_{i}^{n} \left(\frac{1}{\tilde{\mu}_i} - \frac{\log(2)}{y_i}\right)\tilde{\mu}_i x_{ir}, \tag{5.2}$$

which can be written in matrix notation as

$$U\left(\hat{\beta}\right) = X^t Q\left(y, \tilde{\mu}\left(\hat{\beta}\right)\right). \tag{5.3}$$

Taking the second derivatives of $l(\beta)$, we have

$$\frac{\partial U_r(\beta)}{\partial \beta_j} = \frac{\partial^2 l(\beta)}{\partial \beta_j \partial \beta_r} = \sum_{i}^{n}\left(-\frac{\log(2)}{y_i}\frac{\partial \tilde{\mu}_i}{\partial \beta_j}\right)x_{ir}, \quad j = 1, 2, ..., p;$$

hence,

$$-E\left(\frac{\partial U_r(\beta)}{\partial \beta_j}\right) = \sum_{i=1}^{n} E\left(\frac{1}{y_i}\right)\frac{\partial \tilde{\mu}_i}{\partial \beta_j}\log(2)x_{ij}.$$

Since $E\left(\frac{1}{y_i}\right) = \frac{1}{\tilde{\mu}_i \log(2)}$ and $\frac{\partial \tilde{\mu}_i}{\partial \beta_j} = x_{ij}\tilde{\mu}_i$, then $I_{rj} = -E\left(\frac{\partial U_r(\beta)}{\partial \beta_j}\right) = \sum_i^n x_{ij}x_{ir}$, and Fisher's information matrix is given as

$$I\left(\hat{\beta}\right) = X' W\left(\hat{\beta}\right)X, \tag{5.4}$$

where $W\left(\hat{\beta}\right)$ is the unit matrix. Fisher's scoring process can be applied to obtain

$$I\left(\hat{\beta}^{(s-1)}\right)\hat{\beta}^{(s)} = I\left(\hat{\beta}^{(s-1)}\right)\hat{\beta}^{(s-1)} + U\left(\hat{\beta}^{(s-1)}\right), \quad s = 1, 2, 3, ...$$

From (5.3) and (5.4), we have

$$\left(X'W\left(\hat{\beta}^{(s-1)}\right)X\right)\hat{\beta}^{(s)} = \left(X'W\left(\hat{\beta}^{(s-1)}\right)X\right)\hat{\beta}^{(s-1)} + X'Q\left(y, \tilde{\mu}\left(\hat{\beta}^{(s-1)}\right)\right), \tag{5.5}$$

and

$$(X'X)\hat{\beta}^{(s)} = X'\left[X\hat{\beta}^{(s-1)} + Q\left(y, \tilde{\mu}\left(\hat{\beta}^{(s-1)}\right)\right)\right].$$

Thus, the estimated coefficients $\hat{\beta}$ are given by

$$\hat{\beta}^{(s)} = (X'X)^{-1}X'\left[X\hat{\beta}^{(s-1)} + Q\left(y, \tilde{\mu}\left(\hat{\beta}^{(s-1)}\right)\right)\right] = (X'X)^{-1}X'Z,$$

as given in (2.7). To derive the MLS of β, IRLS algorithm is used. Under regularity conditions on the likelihood function, the MLE $\hat{\beta}^{(s)}$ is asymptotically normal, unbiased, and efficient, with covariance matrix equal to the inverse of Fisher's information matrix (Houston and Woodruff [30]). Therefore, asymptotically,

$$\hat{\beta} \equiv N\left[\beta, (X'WX)^{-1}\right],$$

where $(X'WX)^{-1}$ is the inverse of Fisher's information matrix.

5.2. *Proof of Lemma 3.1:*

Suppose that $y_i \sim f(y_i; \gamma)$ is as in (2.3), $\gamma = \tilde{\mu}\log(2)$; then, the density function of y_i is given by

$$f(y_i; \tilde{\mu}_i) = \frac{\log(2)\tilde{\mu}_i}{y_i^2} e^{-\frac{\log(2)\tilde{\mu}_i}{y_i}}. \tag{5.6}$$

Consider a gamma prior for $\tilde{\mu}_i$, which can be written as in (3.6). Posterior distribution of $\tilde{\mu}_i$ is given by

$$\pi(\tilde{\mu}_i|y_i) = \frac{\lambda^\alpha \log(2)\tilde{\mu}_i^\alpha}{\Gamma(\alpha)y_i^2} e^{-\tilde{\mu}_i\left(\lambda + \frac{\log(2)}{y_i}\right)}. \tag{5.7}$$

Using Jacobian transformation from $\tilde{\mu}_i$ to η_i, and using the log link function that is given as in (2.6), we have

$$\pi(\eta_i|y_i) \propto [g^*(\eta_i)]^\alpha e^{-g^*(\eta_i)\left(\lambda + \frac{\log(2)}{y_i}\right)} \frac{\partial g^*(\eta_i)}{\partial \eta_i}, \tag{5.8}$$

where $g^*(\eta_i) = e^{\eta_i} = \tilde{\mu}_i$ and $\frac{\partial g^*(\eta_i)}{\partial \eta_i} = e^{\eta_i}$. Then,

$$\pi(\eta_i|y_i) \propto e^{(1+\alpha)\eta_i} e^{-e^{\eta_i}\left(\lambda + \frac{\log(2)}{y_i}\right)}. \tag{5.9}$$

Taking the derivative of the log posterior, we have

$$\frac{\partial \log(\pi(\eta_i|y_i))}{\partial \eta_i} \propto (1+\alpha) - e^{\hat{\eta}_i}\left(\lambda + \frac{\log(2)}{y_i}\right) = 0, \tag{5.10}$$

hence, the posterior mode of η_i is given as in (3.7). Thus, the estimated coefficients $\hat{\beta}^*$, IE-BReg, and IE-EBReg models in this case, are given as in (3.5) and (3.8), respectively.

In the case that prior distribution parameter λ is unknown, for this estimation task, $\hat{\lambda}$ estimates are obtained via numerical maximization of the following marginal likelihood (Shao [37], Dikici et al. [58], and Coluccia et al. [59])

$$f(y|\tilde{\mu}) = \prod_{i=1}^{n} \int_0^{\infty} f(y_i;\tilde{\mu}_i)\pi(\tilde{\mu}_i|y_i)d\mu_i. \tag{5.11}$$

Thus, the IE-EBReg estimate is found by placing these estimated prior distribution parameter into Equation (3.7) by $\hat{\lambda}$.

5.3. Proof of Lemma 3.2:

Suppose that $y_i \sim f(y_i;\gamma)$ is as in (2.3), $\gamma = \tilde{\mu}\log(2)$; then, the density function of y_i is given as in (5.6). Consider a gamma prior $G(\alpha,\lambda)$ for $\tilde{\mu}_i$, which can be written as in (3.6). Posterior distribution of $\tilde{\mu}_i$ is given as in (5.7). Using Jacobian transformation from $\tilde{\mu}_i$ to η_i, zero–one loss function and using the log link function as in (2.6), we have

$$(1+\alpha) = e^{\hat{\eta}_i}\left(\lambda + \frac{\log(2)}{y_i}\right). \tag{5.12}$$

Because of $0 < \alpha < \infty$, thus,

$$1 < e^{\hat{\eta}_i}\left(\lambda + \frac{\log(2)}{y_i}\right) < \infty, \tag{5.13}$$

hence, we get

$$\lambda > \frac{1}{e^{\hat{\eta}_i}} - \left(\frac{\log(2)}{y_i}\right). \tag{5.14}$$

Since $0 < y_i < \infty$ for every $i = 1, ..., n$, thus

$$-\infty < \frac{1}{e^{\hat{\eta}_i}} - \left(\frac{\log(2)}{y_i}\right) < \frac{1}{e^{\hat{\eta}_i}}. \tag{5.15}$$

From (5.14) and (5.15), we find $\lambda \le \frac{1}{e^{\hat{\eta}_i}}$, but $e^{\hat{\eta}_i} \le 1$, because of equation (5.13) and the fact $e^{\hat{\eta}_i} < e^{\hat{\eta}_i}\left(\lambda + \frac{\log(2)}{y_i}\right)$. Now, by using $\frac{1}{e^{\hat{\eta}_i}} \ge 1$ and $\lambda \le \frac{1}{e^{\hat{\eta}_i}}$, then we obtain $\lambda \le 1$. Hence, $\frac{1}{\lambda} \ge 1$ and $\frac{\alpha}{\lambda^2} \ge \frac{\alpha}{\lambda}$, thus the variance of $\tilde{\mu}$ is greater than or equal the mean.

5.4. *Proof of Lemma 3.3:*

Suppose that $y_i \sim f(y_i; \gamma)$ is as it is in (2.3), $\gamma = \tilde{\mu}\log(2)$, and the density function of y_i is given as in (2.3). Consider $\tilde{\mu}_i$ has a gamma prior with density function, which can be written as in (3.6). Using the posterior distribution of η_i that is given in (5.9), we have

$$E(e^{-\alpha\eta_i}) = \int_{-\infty}^{\infty} e^{-\alpha\eta_i}\pi(\eta_i|y_i)d\eta_i, \tag{5.16}$$

$$= \frac{\lambda^\alpha \log(2)}{\Gamma(\alpha)y_i^2}\int_{-\infty}^{\infty} e^{\eta_i}e^{-e^{\eta_i}\left(\lambda+\frac{\log(2)}{y_i}\right)}d\eta_i, \tag{5.17}$$

and

$$E(e^{-\alpha\eta_i}) = \frac{\lambda^\alpha \log(2)}{\Gamma(\alpha)y_i^2\left(\lambda + \frac{\log(2)}{y_i}\right)}. \tag{5.18}$$

Using the LINEX loss function, we have

$$\frac{\lambda^\alpha \log(2)}{\Gamma(\alpha)y_i^2\left(\lambda + \frac{\log(2)}{y_i}\right)} = e^{-\alpha\hat{\eta}_i}, \tag{5.19}$$

Thus, the posterior Bayes estimates of η_i by using the LINEX loss function are given as in (3.9).

In the case that prior distribution parameter λ is unknown for this estimation task, $\hat{\lambda}$ estimates are obtained via numerical maximization of the marginal likelihood of Equation (5.11). As a result, the IE-EBReg estimates are found by placing these estimated prior distribution parameters into Equation (3.9) by $\hat{\lambda}$.

5.5. *Proof of Lemma 3.4:*

Suppose that $y_i \sim f(y_i; \gamma)$ is as in (2.3), $\gamma = \tilde{\mu}\log(2)$; then, the density function of y_i is given as in (5.6). Consider a gamma prior $G(\alpha, \lambda)$ for $\tilde{\mu}_i$, which can be written as in (3.6). Posterior distribution of $\tilde{\mu}_i$ is given as in (5.7). Using Jacobian transformation from $\tilde{\mu}_i$ to η_i, LINEX loss function and using the log link function as in (2.6), we have

$$\frac{\lambda^\alpha \log(2)}{\Gamma(\alpha)y_i^2\left(\lambda + \frac{\log(2)}{y_i}\right)} = e^{-\alpha\hat{\eta}_i}, \tag{5.20}$$

Because of $0 < \alpha < \infty$, thus $0 < e^{-\alpha\hat{\eta}_i} < 1$, and

$$0 < e^{-\alpha\hat{\eta}_i}\frac{\Gamma(\alpha)y_i^2}{\log(2)}\left(\lambda + \frac{\log(2)}{y_i}\right) < \frac{\Gamma(\alpha)y_i^2}{\log(2)}\left(\lambda + \frac{\log(2)}{y_i}\right). \tag{5.21}$$

Using (5.20) and (5.21), we get

$$\lambda^\alpha < \frac{\Gamma(\alpha)y_i^2}{\log(2)}\left(\lambda + \frac{\log(2)}{y_i}\right), \tag{5.22}$$

hence

$$\lambda^{\alpha} < \left(\lambda + \frac{\log(2)}{y_i}\right), \tag{5.23}$$

and,

$$\lambda\left(\lambda^{\alpha-1} - 1\right) < \frac{\log(2)}{y_i}. \tag{5.24}$$

Thus, $\lambda < \frac{\log(2)}{y_i}$ and $\left(\lambda^{\alpha-1} - 1\right) < \frac{\log(2)}{y_i}$. Hence, $\frac{1}{\lambda} < 1 + \frac{\log(2)}{y_i}$. Since $0 < y_i < \infty$ for every $i = 1, ..., n$, thus $1 < 1 + \frac{\log(2)}{y_i} < \infty$. Now, by using $\frac{1}{\lambda} < 1 + \frac{\log(2)}{y_i}$ and $1 < 1 + \frac{\log(2)}{y_i}$, then we obtain $\frac{1}{\lambda} \leq 1$, $\lambda \geq 1$ and $\frac{\alpha}{\lambda^2} \leq \frac{\alpha}{\lambda}$. Thus, the variance of $\tilde{\mu}$ is smaller than or equal the mean.