Sains Malaysiana 44(10)(2015): 1417–1422

Outlier Detection using Generalized Linear Model in Malaysian Breast Cancer Data

(Pengesanan Nilai Tersisih menggunakan Model Linear Teritlak dalam Data Kanser Payudara Malaysia)

M. Nawama, A.I.N. Ibrahim*, I.B. Mohamed, M.S. Yahya & N.A.M. Taib

ABSTRACT

We consider the problem of outlier detection in bivariate exponential data fitted using the generalized linear model via Bayesian approach. We follow closely the work outlined by Unnikrishnan (2010) and present every step of the detection procedure in details. Due to the complexity of the resulting joint posterior distribution, we obtain the information on the posterior distribution from samples generated by Markov Chain Monte Carlo sampling, in particular, using either the Gibbs sampler or the Metropolis-Hastings algorithm. We use local breast cancer patients' data to illustrate the implementation of the method.

Keywords: Bayesian; Gibbs sampler; Metropolis-Hastings algorithm; Outlier

ABSTRAK

Kami mempertimbangkan masalah pengesanan nilai tersisih dalam data bivariat eksponen dengan menggunakan model linear teritlak melalui pendekatan Bayesian. Kami mengikuti secara rapat kajian yang digariskan oleh Unnikrishnan (2010) dan membentangkan setiap langkah prosedur pengesanan secara terperinci. Disebabkan kerumitan taburan posterior tercantum yang terhasil, kami mendapatkan maklumat mengenai taburan posterior tersebut daripada sampel yang dijana oleh pensampelan Markov Chain Monte Carlo, khususnya, menggunakan sama ada kaedah pensampelan Gibbs atau algoritma Metropolis-Hastings yang umum. Kami menggunakan data tempatan iaitu data pesakit kanser payudara untuk menggambarkan pelaksanaan kaedah tersebut.

Kata kunci: Algoritma Metropolis-Hastings; Bayesian; kaedah pensampelan Gibbs; nilai tersisih

Introduction

The existence of outliers in sample data is a common phenomenon in data analysis. Barnett and Lewis (1983) reviewed the literature on outliers in various types of statistical data. Outlier refers to an observation with abnormal properties compared to the others such as being surprisingly far from the main data set or having large residual (Anscombe & Guttman (1960) and Ferguson (1961)).

In recent years, there has been much interest in the development of outlier detection methods using Bayesian approach. The approach basically differs from the traditional non-Bayesian methods in their basic concepts and the use of relevant information such as prior probabilities. In the Bayesian set up, Freeman (1980) defined an outlier as 'any observation that has not been generated by the mechanism that generated the majority of the observation in the data set'. In other words, the detection of outliers in this framework is reduced to the problem of estimating the parameters of the distribution of the contaminated observations (Bayarri & Morales 2003). Other works along these lines can also be found (Marshall & Spiegelhalter 2007; Page & Dunson 2011; Pettit 1994). This idea can also be extended to identify outliers in bivariate data via generalized linear modeling (GLM). William (1987) employed the one-case deletion approach by looking at the changes in the deviance of a GLM when a single case is deleted from the data. Kuhnt and Pawlitschko (2003) assumed a particular GLM as the model under the null hypothesis for a regular data set and derived the rules for outlier identification. It is a fact that the resulting joint posterior distributions from these GLMs are often very complex and intractable. Therefore, a way to obtain information from these models is to generate samples by using simulation methods such as Markov chain Monte Carlo (MCMC) sampling. For a GLM, Bayesian inference using MCMC sampling allows simultaneous handling of the outlier detection and parameters estimation. Zeger and Karim (1991) presented a GLM with random effects model in Bayesian framework and used MCMC and Gibbs sampler to overcome the computational limitation, while Ishwaran (1999) applied the hybrid Monte Carlo for fitting Bayesian GLM with canonical link. Unnikrishnan (2010) applied the reversible jump MCMC and Metropolis-Hastings (MH) algorithm without giving details on the resulting posterior distributions of the multi-parameter set-up.

In this paper, we follow closely the work outlined by Unnikrishnan (2010) by presenting the steps in detail. We apply the proposed method to a local breast cancer study which motivates the methodological development of this research. The paper is organized as follows: The modified GLM with outlier is described in the next section; the local breast cancer study and the application and implementation of the outlier detection using GLM via Bayesian approach to this data set are discussed in detail in the following sections where we consider the case of single or multiple outliers; and the conclusions are given in the final section.

MODIFIED GLM WITH OUTLIER

We follow closely the general theory proposed by Unnikrishnan (2010). Let $\aleph = \{1, \ldots, N\}$ be a finite population with known N. For each unit $i \in \aleph$, we have the real valued response variable y_i and known $p \times 1$ vector of explanatory variables $\mathbf{x}_i$ where $\mathbf{w}'_i = (x_{i1} \ldots x_{ip})$.

Assume that a random sample of size n is obtained with a number of suspected outliers. Let v^k be the set of all outlying observations, where k denotes the number of outliers. We consider the models with/without outliers based on GLM such that

$$p\left(y_i \middle| \theta_i, \phi_i, \delta\right) = \begin{cases} \exp\left\{\phi_i\left(y_i\theta_i\right) + c_i\left(\theta_i, \phi_i\right) + d_i\left(y_i\right)\right\} & i \in \aleph - v^k \\ \exp\left\{\dfrac{\phi_i}{\delta}\left(y_i\theta_i\right) + c_i\left(\delta, \theta_i, \phi_i\right) + d_i\left(y_i\right)\right\} & i \in v^k \end{cases}, \tag{1}$$

where θ_i is a location parameter; ϕ_i and δ are scale parameters; and $c(.)$, $d(.)$ are known functions. The parameters θ_i are modelled through a specific link function $h(.)$ given by

$$h(\theta_i) = \mathbf{x}_i'\boldsymbol{\beta} + \varepsilon_i, \qquad i = 1, \ldots, N \tag{2}$$

where $\boldsymbol{\beta}' = (\beta_1 \ \ldots \ \beta_p)$ and the error components ε_i's are independently and normally distributed. Consequently, we can write $h(\theta_i)|\sigma^2 \sim N(\mathrm{x}_i' \ \boldsymbol{\beta},\ \sigma^2)$. Commonly, we usually assume that the model have the same mean for all observations but we expect to see higher variance for outlying observations, that is, when $\delta > 1$.

LOCAL BREAST CANCER DATA

Breast cancer is the most common cancer in Malaysia with the incidence rate for females of 47.4 per 100,000 women. Only recently the breast cancer specific survival information in Malaysia is available through the National Cancer Registry program under the purview of the Ministry of Health Malaysia. One of the established Breast Cancer Centre in the country is situated at the University of Malaya Medical Centre (UMMC) Kuala Lumpur. The UMMC is a 900 bed tertiary public hospital located in urban Kuala Lumpur. Prospective cohort studies of women with breast cancer treated in the UMMC are considered. The cohort comprises of patients who are diagnosed from year 1998 to 2002 and are followed up until March 2006. Patients underwent surgery and adjuvant chemotherapy under the care of general surgery and then followed by radiotherapy in UMMC. The information collected from the patients consists of race, age, date of diagnosis and pathological characteristics of tumour. In addition, the survival times and status of patients are recorded at the end of the study. The mortality information is confirmed by referring to the record in the National Registry of Births and Deaths. The data set has been used in several other papers including Taib *et al*. (2011, 2008). For our case, we consider the size of tumour as independent variable x and survival time to death in months from first diagnosis of the disease as dependent variable y. Only patients who registered in 2000 with age 60 years old and above are considered. The scatter plot of the data with the exponential fitted curve is given in Figure 1. It can be seen that the data appear to follow the exponential distribution, with one extreme observation a candidate to be an outlier. It is of interest to identify any outliers using the modified model (1) in the Bayesian framework.

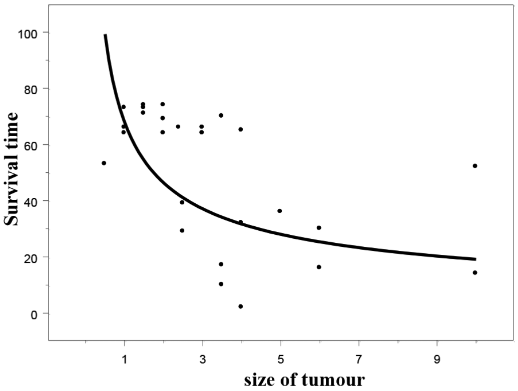

FIGURE 1. Plot of patients' survival time versus size of tumour

THE MODEL

By choosing the appropriate functions for $c(.)$, $d(.)$ and taking ϕ_i to be similar for all i, the appropriate model to study the exponential relationship between y and x in the present data set corresponding to the general model (1) is given by:

$$f\left(y_i \middle| \theta_i, \phi, \delta\right) = \begin{cases} \phi\theta_i \exp\left(-\phi\theta_i y_i\right) & \text{for } i \notin v^k \\ \dfrac{\phi\theta_i}{\delta} \exp\left(-\dfrac{\phi\theta_i}{\delta} y_i\right) & \text{for } i \in v^k \end{cases}, \tag{3}$$

with the link function $\log \theta_i = \beta(x_i - \bar{x}) + \varepsilon_i$, where $\bar{x}$ is the mean of the sample. We intend to detect outlying observation using the hierarchical Bayesian approach. Hence, we consider the following hierarchical prior distributions of the parameters:

$$\left.\begin{aligned} &\sigma^{-2}\Big| a_\sigma, b_\sigma \sim \text{gamma}\left(\frac{a_\sigma}{2}, \frac{b_\sigma}{2}\right), \quad \beta\Big| a_\beta, b_\beta \sim \text{gamma}\left(\frac{a_\beta}{2}, \frac{b_\beta}{2}\right), \\ &\delta \sim \text{Uniform}\left(1, \delta_{\max}\right), \qquad\qquad \phi\Big| a_\phi, b_\phi \sim \text{gamma}\left(\frac{a_\phi}{2}, \frac{b_\phi}{2}\right), \\ &\log\ \theta_i\Big|\sigma^2,\ \beta \sim \mathrm{N}(x_i^*\beta,\ \sigma^2) \text{ where } x_i^* = (x_i - \bar{x}), \\ &\text{and } p(v^k|k) = \binom{N}{k}^{-1}. \end{aligned}\right\} \tag{4}$$

Now, the joint likelihood function is given by:

$$L\left(\mathbf{y}\middle|\boldsymbol{\theta},\phi,\delta,v^k\right)=\prod_{i\notin v^k}\phi\theta_i\exp\left(-\phi\theta_i y_i\right)\times \prod_{i\in v^k}\frac{\phi\theta_i}{\delta}\exp\left(-\frac{\phi\theta_i}{\delta}y_i\right). \tag{5}$$

Correspondingly, from the result obtained in (4) and (5), the full joint posterior distribution for the parameters $\boldsymbol{\theta}$, $\sigma^2, \beta, \phi, \delta, v^k$ is given by:

$$\begin{aligned}p\left(\boldsymbol{\theta},\phi,\delta,\beta,\sigma^2,v^k\,\mathrm{I}\,\mathrm{y}\right)\propto&\prod_{i\notin v^k}\phi\theta_i\exp\left(-\phi\theta_i y_i\right)\times\\ &\prod_{i\in v^k}\frac{\phi\theta_i}{\delta}\exp\left(-\frac{\phi\theta_i}{\delta}y_i\right)\times\\ &\prod_{i=1}^{n}\frac{1}{\theta_i\left(2\pi\sigma^2\right)^{1/2}}\exp\left\{-\frac{1}{2}\left(\frac{\log\theta_i-x_i^*\beta}{\sigma}\right)^2\right\}\times\\ &\left(\sigma^{-2}\right)^{\left(\frac{a_\sigma}{2}-1\right)}\exp\left(-\frac{b_\sigma}{2\sigma^2}\right)\times\beta^{\frac{a_\beta}{2}-1}\exp\left(-\beta\frac{b_\beta}{2}\right)\times\\ &\phi^{\frac{a_\phi}{2}-1}\exp\left(-\phi\frac{b_\phi}{2}\right)\times\frac{1}{\delta_{\max}-1}\times\left(\frac{N!}{(N-k)!k!}\right)\end{aligned} \tag{6}$$

It is clear that the full joint posterior distribution is intractable. Hence, we employ the MCMC sampling method, in particular, using Gibbs sampler with MH algorithm in the outlier detection procedure.

SAMPLING METHODS OF THE PARAMETERS

Note that model (3) involves multiple parameters that are structured hierarchically such that the dependency of the parameters is reflected in the joint probability distribution. The exact conditional posterior distributions of some of the parameters can be specified directly from the resulting joint posterior distribution (6). Therefore, we can sample directly from these conditional posterior distributions; in other words, using Gibbs sampling. For the rest, we employ the MH algorithm for sampling purposes. The sampling methods for each of the parameters $\boldsymbol{\theta}$, $\sigma^2, \beta, \phi, \delta, v^k$ are given in detail below. The sequence of the process is chosen to satisfy the hierarchical dependency of model (3).

PARAMETER σ^2

Looking at (6) and letting $\gamma = \sigma^{-2}$, the conditional posterior distribution for parameter γ given $\boldsymbol{\theta}$, $\sigma^2, \beta, \phi, \delta, v^k$ is given by:

$$f\left(\gamma\middle|\boldsymbol{\theta},\beta,\phi,\delta,v^k\right)\propto\gamma^{\frac{n}{2}+\frac{a_\sigma}{2}-1}\exp\left\{-\gamma\left(\sum_{i=1}^{n}\frac{\left(\log\theta_i-x_i^*\beta\right)^2}{2}+\frac{b_\sigma}{2}\right)\right\}.$$

The posterior conditional distribution for γ is therefore gamma$\left(\frac{n}{2}+\frac{a_\sigma}{2},\frac{b_\sigma}{2}+\frac{1}{2}\sum_{i=1}^{n}\left(\log\theta_i-x_i^*\beta\right)^2\right)$. Hence, the conditional posterior distribution for σ^2 is the inverse gamma with the same parameters as those for γ. Therefore, we can sample σ^2 directly from the conditional posterior distribution.

PARAMETER β

From (6), the conditional posterior distribution for parameter β given σ^2, $\boldsymbol{\theta}$, ϕ, δ, v^k is given by:

$$f\left(\beta\middle|\sigma^2,\boldsymbol{\theta},\phi,\delta,v^k\right)\propto\beta^{\frac{a_\beta}{2}-1}\exp\left(-\frac{1}{2}\left[\beta b_\beta\sum_{i=1}^{n}\left(\frac{\left(\log\theta_i-x_i^*\beta\right)^2}{\sigma}\right)\right]\right).$$

Here, we use the MH algorithm to sample from the conditional posterior distribution of β as it is intractable. We sample the proposal value for β, says β_{prop}, from gamma $\left(\frac{a_\beta}{2},\frac{b_\beta}{2}\right)$, where $g_\beta(.)$ denote the density of this proposal distribution. Then, using the MH algorithm, the candidate β_{prop} is then accepted with probability

$$\alpha\left(\beta,\beta_{\text{prop}}\right)=\min\left(1,\frac{f\left(\beta_{prop}\middle|\sigma^2,\boldsymbol{\theta},\phi,\delta,v^k\right)g_\beta(\beta)}{f\left(\beta\middle|\sigma^2,\boldsymbol{\theta},\phi,\delta,v^k\right)g_\beta\left(\beta_{\text{prop}}\right)}\right). \tag{7}$$

The full formula of the acceptance probability above is easily obtained by substituting the relevant functions into this equation.

PARAMETER δ

In this case, the conditional posterior distribution for parameter δ given $\boldsymbol{\theta}$, $\phi, \sigma^2, \beta, v^k$ is given by:

$$f\left(\delta\middle|\boldsymbol{\theta},\phi,\sigma^2,\beta,v^k\right)\propto\prod_{i\in v^k}\frac{\phi\theta_i}{\delta}\exp\left(-\frac{\phi\theta_i}{\delta}y\right)\times\frac{1}{\delta_{\max}-1}.$$

We choose the proposal density for δ_{prop} as $g_\delta(\delta) = 1/(\delta_{\max} - 1)$, so that δ_{prop} has a uniform distribution. When using the MH algorithm, the acceptance probability for candidate δ_{prop} can be obtained by replacing the corresponding functions in (7).

PARAMETER ϕ

We now look at the conditional posterior distribution for parameter ϕ given σ^2, β, δ, $\boldsymbol{\theta}$, v^k. From (6), the conditional posterior distribution is given by:

$$\begin{aligned}f\left(\phi\middle|\sigma^2,\beta,\delta,\theta,v^k\right)\propto&\prod_{i\notin v^k}\phi\theta_i\exp\left(-\phi\theta_i y_i\right)\times\\ &\prod_{i\in v^k}-\frac{\phi\theta_i}{\delta}\exp\left(-\frac{\phi\theta_i}{\delta}y_i\right)\times\\ &\phi^{\frac{a_\phi}{2}-1}\exp\left(-\phi\frac{b_\phi}{2}\right).\end{aligned}$$

Here, we introduce a function ω_i where:

$$\omega_i = \begin{cases} 1 & \text{for } i \in v^k \\ 0 & \text{for } i \notin v^k \end{cases}, \qquad (8)$$

so that

$$f\left(\phi \middle| \sigma^2, \beta, \delta, \boldsymbol{\theta}, v^k\right) \propto \phi^{n+\frac{a_\phi}{2}-1} \exp\left\{-\phi\left[\frac{b_\phi}{2} + \sum_{i=1}^{n} \theta_i y_i \left(1 - \omega_i + \frac{\omega_i}{\delta}\right)\right]\right\}.$$

Thus, $\phi | \sigma^2, \beta, \delta, \boldsymbol{\theta}, v^k \sim \text{gamma}\left(n + \frac{a_\phi}{2}, \frac{b_\phi}{2} + \sum_{i=1}^{n} \theta_i y_i \left(1 - \omega_i + \frac{\omega_i}{\delta}\right)\right)$. Therefore, we can sample ϕ directly from this conditional distribution.

PARAMETER $\boldsymbol{\theta}$

Next, we obtained the conditional posterior distribution for parameter $\boldsymbol{\theta}$ given $\sigma^2, \beta, \delta, \phi, v^k$ such that:

$$f\left(\theta \middle| \sigma^2, \beta, \delta, \phi, v^k\right) \propto \prod_{i \notin v^k} \phi\theta_i \exp\left(-\phi\theta_i y_i\right) \times \prod_{i \in v^k} \frac{\phi\theta}{\delta} \exp\left(-\frac{\phi\theta}{\delta} y\right)$$
$$\times \prod_{i=1}^{n} \frac{1}{\theta_i \left(2\pi\sigma^2\right)^{1/2}} \exp\left\{-\frac{1}{2}\left(\frac{\log\theta_i - x_i^* \beta}{\sigma}\right)^2\right\}.$$

Furthermore, for each θ_i, we use the function ω_i as defined in (8) so that:

$$f_i\left(\theta_i \middle| \sigma^2, \beta, \delta, \phi, v^k\right) \propto \theta_i \exp\left\{-\theta_i\left[\phi y_i \left(1 - \omega_i + \frac{\omega_i}{\delta}\right)\right]\right\} \times$$
$$\frac{1}{\theta_i} \exp\left\{-\frac{1}{2}\left(\frac{\log\theta_i - x_i^* \beta}{\sigma}\right)^2\right\}.$$

Thus, we propose to use a proposal distribution for each θ_i which is lognormal $\left(x_i^* \beta, \sigma^2\right)$, where $g_i(.)$ denotes the proposal density. Therefore, using MH algorithm, for parameter $\boldsymbol{\theta}$, we update θ_1 to θ_n one at a time, where for each θ_i, the acceptance probability for candidate θ_{iprop} can be obtained by replacing the corresponding functions in (7).

PARAMETER v^k

Finally, using (6), the conditional posterior distribution for parameter v^k given $\sigma^2, \beta, \delta, \phi, \boldsymbol{\theta}$ is:

$$f\left(v^k \middle| \sigma^2, \beta, \delta, \phi, \boldsymbol{\theta}\right) \propto \prod_{i \notin v^k} \phi\theta_i \exp\left(-\phi\theta_i y_i\right) \times$$
$$\prod_{i \in v^k} \frac{\phi\theta_i}{\delta} \exp\left(-\frac{\phi\theta_i}{\delta} y_i\right).$$

For a given value of k, let $v^k = v = \{v_1, \ldots, v_k\}$. Under the priors given in (4), any of the k distinct units are equally likely to become outliers. Then, in order to find a new value of v, first select a unit at random from v, say v_i^*, and select a unit at random from the complement set v^c, say v_{prop}. If the proposal is accepted, then v_i^* goes out of v and v_{prop} replaces the value v_i^*; this will become a new value denoted by $\tilde{v}$. Next, using MH algorithm, this state is accepted with probability:

$$\alpha\left(v, \tilde{v}\right) = \min\left(1, \frac{\prod_{i \notin \tilde{v}} f\left(y_i \middle| \theta_i, \phi\right) \times \prod_{j \in \tilde{v}} f\left(y_j \middle| \theta_j, \phi, \delta\right)}{\prod_{i \notin v} f\left(y_i \middle| \theta_i, \phi\right) \times \prod_{j \in v} f\left(y_j \middle| \theta_j, \phi, \delta\right)}\right).$$

The sampling methods involving the sampling of values of parameters discussed earlier are repeated for a large number of iterations. In order to improve the consistency and remove the effects of the initial values, a burn-in of reasonable number of iterations is employed. From the last part of the sampling method, we have the number of times each k combination of observations is in the set v^k of possible outlier. Hence, we can calculate the proportion of iterations such that each k combination of observations is in the set v^k. We can then regard the proportion as the probability of each k combination of observations being a set of outliers for a given k. The set of observations with a large value of the proportion is identified to be an outlier.

RESULTS AND DISCUSSION

We now apply the proposed method on the local breast cancer data. We use 100000 iterations, with a burn-in of 50000 iterations. Here we consider the case when there are one (k = 1) or two (k = 2) outliers. The simulated values of the parameters σ^2, β, and ф when k = 1 are shown in Figures 2-4, respectively. It can be seen that the shape of the histograms of the three parameters resemble that of the gamma distributions. Note that the parameters σ^2, β, and ф show similar behaviour for the case when k = 2. Figure 5 shows the estimated probability of being an outlier for observations 1 to 26, given that k = 1. Given that there is one outlier, we identify observation 26, which has the highest probability (close to 0.20), as an outlier as this probability is distinctly higher if compared to the other probabilities corresponding to the other patients in the data. Given that there are two outliers, there are 325 possible sets of observations. We found that the probabilities of a set of two observations being outliers are small, ranging from 0.008 to 0.001; this indicates that the existence of two outliers in this data set is highly unlikely. Note that sets involving observation 26 are at the higher spectrum of the probabilities, with observations {21, 26} and {14, 26} having the highest probability. We can see that given that there are outliers in the data, with high probability observation 26 is an outlier in both cases.

In survival data, many authors have tried to give specific meaning to the outlier due to the special features of the data. Collet (2003) referred an outlier in survival as an individual who has extremely long survival time, but the values of the explanatory variables suggested the individual should have died earlier and vice versa. Nardi and Schemper (1999) and Therneau et al. (1990) associated outlier to individuals who 'died too soon' or 'lived too long', while Maller and Zhou (1994) identified

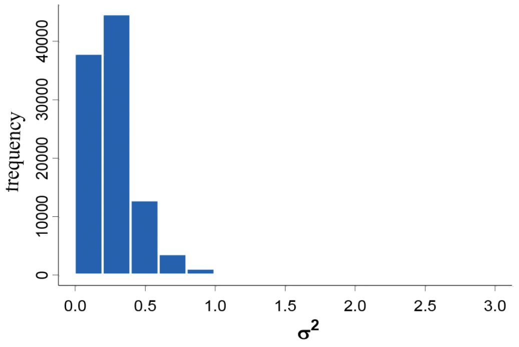

FIGURE 2. Histogram of the marginal posteriors for σ^2

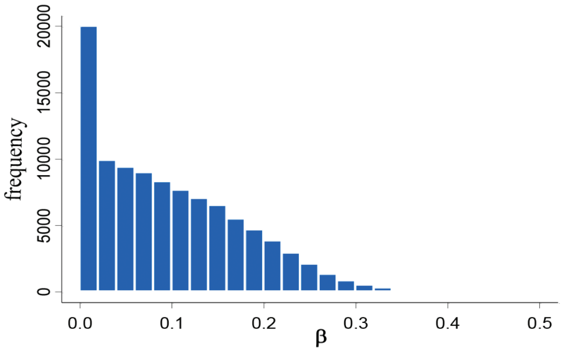

FIGURE 3. Histogram of the marginal posteriors for β

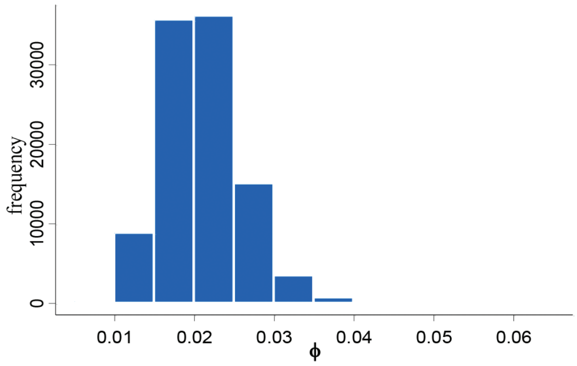

FIGURE 4. Histogram of the marginal posteriors for ϕ

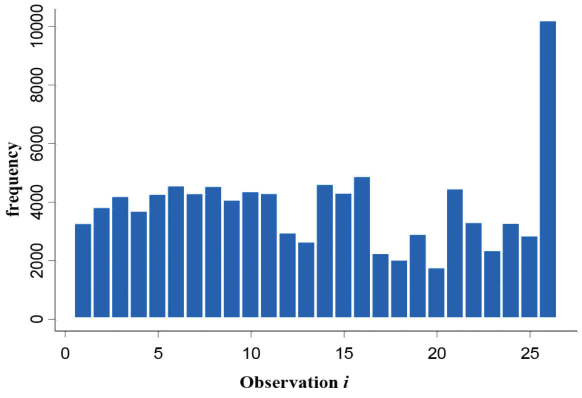

FIGURE 5. Probability for an observation being outlier in Breast Cancer data

outlier as individual who is already 'immune' or 'cured'. Using these definitions, patient 26 fits into the definition of outliers such that the patient's survival time is rather long even though the size of tumour for this patient is amongst the largest in the data set. Such identification enables the breast cancer specialists to monitor the background of such patients in finding the insight on factors that contribute to the improved survival life times for patients with similar prognosis.

CONCLUSION

In this paper, we have considered the problem of detecting outlier using Bayesian approach in generalized linear model. We have shown that with the choice of prior distribution for the parameters, we can obtain the information from samples generated using MCMC sampling, in particular using either the Gibbs sampler or the general MH algorithm. When applied to the local breast cancer data, observation 26 who has a large size of tumour but with long survival time which is 52 months from diagnosed time, is identified as an outlier.

ACKNOWLEDGEMENT

The authors were most grateful to the editors and referees for their suggestions, which led to a substantial improvement of the article. This research was financialy supported by the Research Grant (No. RG188-11AFR, RP009C-13AFR) from the University of Malaya.

REFERENCES

Anscombe, F.J. & Guttman, I. 1960. Rejection of outliers. *Technometrics* 2: 123-147.

Barnett, V. & Lewis T. 1983. *Outliers in Statistical Data*, Chichester: John Wiley & Sons .

Bayarri, M.J. & Morales, J. 2003 Bayesian measures of surprise for outlier detection. *Journal of Statistical Planning and Inference* 111: 3-22.

Collet, D. 2003. *Modelling Survival Data in Medical Research*. Boca Raton, FL: Chapman & Hall / CRC.

Ferguson, T.S. 1961. Rules for rejection of outliers. *Review of the International Statistical Institute* 29: 29-43.

Freeman, P.R. 1980. On the number of outliers in data from a linear model. In *Bayesian Statistics*, edited by Bernardo, J.M., DeGroot, M.H., Lindley, D.V. & Smith, A.F.M. pp. 349-65. Valencia: University Press.

Ishwaran, H. 1999. Applications of hybrid Monte Carlo to Bayesian generalized linear models: quasicomplete separation and neural networks. *Journal of Computational and Graphical Statistics* 8: 779-799.

Kuhnt, S. & Pawlitschko, J. 2003. Outlier Identification Rules for Generalized Linear Models. *Technical Report* no 12, Department of Statistics, University of Dortmund.

Maller, R.A. & Zhou, S. 1994. Testing for sufficient follow-up and outliers in survival data. *Journal of the American Statistical Association* 89: 1499-509.

Marshall, E.C. & Spiegelhalter, D.J. 2007. Identifying outliers in Bayesian hierarchical models: A simulation-based approach. *Bayesian Analysis* 2: 409-444.

Nardi, A. & Schemper, M. 1999. New residuals for Cox regression and their application to outlier screening. *Biometrics* 55(2): 523-529.

Page, G.L. & Dunson, D.B. 2011. Bayesian local contamination models for multivariate outliers. *Technometrics* 53: 152-162.

Pettit, L.I. 1994. Bayesian approaches to the detection of outliers in Poisson samples. *Communication in Statistics-Theory and Methods* 23: 1785-1795.

Taib, N.A., Akmal, M.N., Mohamed, I.B. & Yip, C.H. 2011 Improvement in survival of breast cancer patients trends in survival over two time periods in a single institution in an Asia Pacific Country Malaysia. *Asian Pacific J. of Can. Prev.* 12: 345-349.

Taib, N.A., Yip, C.H. & Mohamed, I. 2008. Survival analysis of Malaysian women with breast cancer: Results from the University of Malaya Medical Centre. *Asian Pacific J. of Can. Prev.* 9: 197-202.

Therneau, T.M., Grambcsh, P.M. & Fleming, T.R. 1990. Martingale-based residuals for survival models. *Biometrika* 77(1): 147-60.

Unnikrishnan, N.K. 2010. Bayesian analysis for outliers in survey sampling. *Computational Statist. and Data Analysis* 54: 1962-1974.

Williams, A.D. 1987. Generalized linear model diagnostic using the deviance and single case deletions. *Appl. Statistics* 36: 181-191.

Zeger, L.S. & Karim, M.R. 1991. Generalized linear models with random effects: A Gibbs sampling approach. *Journal of the American Statistical Association* 86: 79-86.

Mardziah Nawama, Adriana Irawati Nur Ibrahim*,
Ibrahim Mohamed & Mohd Sahar Yahya
Institute of Mathematical Sciences
University of Malaya
59100 Kuala Lumpur
Malaysia.

Nur Aishah Mohd Taib
Department of Surgery
University of Malaya Medical Centre
59100 Kuala Lumpur
Malaysia

*Corresponding author; email: adrianaibrahim@um.edu.my

Received: 22 March 2013
Accepted: 15 June 2015