Sains Malaysiana 48(4)(2019): 893–899
http://dx.doi.org/10.17576/jsm-2019-4804-22

Outlier Detection in 2 × 2 Crossover Design using Bayesian Framework

(Pengesanan Titik Terpencil dalam 2 × 2 Reka Bentuk Pindah Silang Menggunakan Rangka Kerja Bayesian)

F.P. LIM, I.B. MOHAMED*, A.I.N. IBRAHIM, S.L. GOH & N.A. MOHAMED @ A. RAHMAN

ABSTRACT

We consider the problem of outlier detection method in 2×2 crossover design via Bayesian framework. We study the problem of outlier detection in bivariate data fitted using generalized linear model in Bayesian framework used by Nawama. We adapt their work into a 2×2 crossover design. In Bayesian framework, we assume that the random subject effect and the errors to be generated from normal distributions. However, the outlying subjects come from normal distribution with different variance. Due to the complexity of the resulting joint posterior distribution, we obtain the information on the posterior distribution from samples by using Markov Chain Monte Carlo sampling. We use two real data sets to illustrate the implementation of the method.

Keywords: Bayesian; crossover design; Markov Chain Monte Carlo; outlier

ABSTRAK

Kami mengambil kira masalah kaedah pengesanan nilai terpencil dalam kajian pindah silang 2×2 melalui rangka kerja Bayesian. Kami mengkaji masalah pengesanan titik tersisih bagi data bivariat yang disuaikan dengan model linear teritlak dalam rangka kerja Bayesian yang digunakan oleh Nawama. Kami menyesuaikan kerja-kerja tersebut ke dalam 2×2 kajian pindah silang. Dalam rangka kerja Bayesian, kami menganggap bahawa kesan subjek rawak dan ralat akan dijana daripada taburan normal. Walau bagaimanapun, nilai terpencil pula tertabur normal dengan varians yang berbeza. Disebabkan taburan posterior tercantum yang kompleks, kami mendapatkan maklumat mengenai taburan posterior daripada sampel yang dijana melalui pensampelan Markov Chain Monte Carlo (MCMC). Kami menggunakan dua set data sebenar untuk menggambarkan pelaksanaan kaedah tersebut.

Kata kunci: Bayesian; Markov Chain Monte Carlo; reka bentuk pindah silang; titik terpencil

INTRODUCTION

In a standard 2×2 crossover design, we assume that there are two different groups of subjects. Each group receives two treatments in a different order and the trial is to last for two treatment periods, with the order of treatments reversed in the second period. A common problem in crossover trials is the occurrence of extremely large or small observations. These extraordinary observations are called outliers and they may influence the conclusion drawn from the data set. An outlier is a data point which is significantly different from the remaining data (Aggarwal 2013). Hawkins (1980) formally defined the concept of outlier as an observation which deviates so much from the other observations as to arouse suspicions that it was generated by a different mechanism.

Chow and Tse (1990) proposed two procedures based on Cook's likelihood distance and the estimated distance for the detection of outliers in crossover studies. Liu and Weng (1991) carried out procedures based on Hotelling T^2 statistics and residuals for the same purpose. Wang and Chow (2003) presented a general test procedure based on a mean-shift model. Furthermore, Ramsay and Elkum (2005) compared different outlier detection methods proposed by Chow and Tse (1990) and Liu and Weng (1991) via simulation study. They concluded that the estimated distance test performs better than other tests. Most recently, Karasoy and Daghan (2012) applied these existing methods to a real data set in order to investigate outliers. In crossover studies, Enachescu and Enachescu (2009) initially used principal components for the identification of outliers. Meanwhile, Singh et al. (2014) provided details regarding a studentized residual test and the Lund test for identification of outlier subjects. It is therefore important that methods of identifying outliers in 2×2 crossover design are developed for proper handling of the data in studies.

Lim et al. (2016) carried out two outlier detection procedures based on residuals in non-Bayesian framework. Under a simplified model of 2×2 crossover design, we present a classical calculation of studentized residual and propose a new studentized residual using median absolute deviation to identify possible outliers. The performances of both procedures are compared via simulation. With the availability of data set provided by the University of Malaya Medical Centre (UMMC), the results showed that the procedure using performs better in detecting outliers. In this study, we extend our previous works and consider the problem of outlier detection method in 2×2 crossover design via Bayesian framework. We study the problem of outlier detection in bivariate data fitted using generalized linear model (GLM) in Bayesian framework presented by Nawama et al. (2015) and Unnikrishnan (2010). We follow closely their works but adapt them into a 2×2 crossover design. In Bayesian framework, we assume that the random subject effect and the errors to be generated from normal distributions. However, the outlying subjects come from normal distribution with different variance. Due to the complexity of the resulting joint posterior distribution, we obtain the information on the posterior distribution from samples by using Markov Chain Monte Carlo (MCMC) sampling. The paper is organized as follows: The concept of standard 2×2 crossover design is described in the next section; two real data sets and the application and implementation of the outlier detection using Bayesian approach to these data set are discussed in detail in the following sections where we consider the case of single outlier; the conclusions are given in the final section.

THE 2 × 2 CROSSOVER DESIGN

Let Y_{ijk} be the response of the th subject in sequence i during period j under the $d[i, j]$th treatment, where $i, j = 1,2$; m_i is the size of group with treatment $d[i, j]$ and $k = 1,2,\ldots, m_i$. From Jones and Kenward (1989), the full model is

$$Y_{ijk} = \mu + p_j + \tau_{d[i,j]} + \lambda_{d[i,j-1]} + S_{ik} + e_{ijk} \tag{1}$$

where μ is the overall mean; p_j the fixed effect of the jth period; $\tau_{d[i,j]}$ the fixed effect of the treatment administered in period j of sequence i; $\lambda_{d[i,j-1]}$ is the fixed effect of the carryover of the treatment administered in period j – 1 of sequence i where $\lambda_{[i,0]} = 0$, S_{ik} is the random effect of the th subject; and e_{ijk} the random error. The variance components $\{S_{ik}\}$ and $\{e_{ijk}\}$ are assumed to be independent and normally distributed with mean 0 and variances σ_s^2 and σ_e^2, respectively.

DATA DESCRIPTION

There are two data sets considered for this study: Clayton and Leslie's data (1981) and kinesiology data. For the first data set, Clayton and Leslie (1981) considered the blood concentration-time curve (AUC) data from two erythromycin formulations in a bioavailability study. In their study, a standard 2×2 crossover experiment was conducted with 18 subjects to compare a new erythromycin formulation (erythromycin stearate) with a reference formulation (erythromycin base). As no sequence identification of each subject was provided in Clayton and Leslie (1981), we adapt the order of periods given in Weiner (1989) and assign subject 1 through 9 to sequence 1 and the remaining subjects to sequence 2.

For the second data set, kinesiology comes from the Greek word *kinesis*, which means motion. In the medical sciences, it is the name given to the study of muscles and the movement of the body, the mechanics of body movements. Kinesiology data for this study are obtained from UMMC Sport Medicine Clinic. The UMMC is a government-funded medical institution located in Petaling Jaya, southwest corner of Kuala Lumpur, which was established in 1962. A two period crossover and randomized placebo-controlled trial of AB (treatment followed by sham taping)/BA (placebo followed by treatment taping) design is conducted. There are 77 subjects, from eighty-one subjects volunteered, completed the study (AB = 37, BA = 40) which observed a minimum washout period of one week. Pre and post measurements of peak oxygen consumption or peak (in mL/kg/min) recorded from a six-minute A strand submaximal cycling exercise test conducted at least one week apart. peak is mainly used to gauge cardiorespiratory fitness of an individual.

Since there is none of the musculoskeletal outcome measures demonstrated convincing association with kinesiotape (KT) use, we therefore propose to investigate the effect of KT on the VO_2 peak. Thus far, no study has explored the effect of KT on measurements of VO_2 peak.

THE MODEL

Under the full model (1), for case without outliers, the expected value of y_{ijk} is

$$E(y_{ijk}) = E(\mu + p_j + \tau_{d[i,j]} + \lambda_{d[i,j-1]} + S_{ik} + e_{ijk})$$
$$= \mu + p_j + \tau_{d[i,j]} + \lambda_{d[i,j-1]}$$

while the variance of y_{ijk} is

$$\mathrm{Var}(y_{ijk}) = \mathrm{Var}(\mu + p_j + \lambda_{d[i,j-1]} + S_{ik} + e_{ijk})$$
$$= \mathrm{Var}\,(S_{ik}) + \mathrm{Var}(e_{ijk}) = \sigma_s^2 + \sigma_e^2.$$

Hence,

$$y_{ijk} \sim N\,(\mu + p_j + \tau_{d[i,j]} + \lambda_{d[i,j-1]}, \sigma_s^2 + \sigma_e^2).$$

On the other hand, for the case with outliers, the variance component $\{S_{ik}\}$ is assumed to be independent and normally distributed with mean 0 and variances $\delta^2\sigma_s^2$. Therefore,

$$y_{ijk} \sim N\,(\mu + p_j + \tau_{d[i,j]} + \lambda_{d[i,j-1]}, \delta^2\sigma_s^2 + \sigma_e^2).$$

Assume that a random sample size $n = \sum m_i$ of is obtained with a number of suspected outliers. Define $y_{ijk} = (y_{111}, y_{112}, \ldots, y_{11n})$. Let v^h be the set of all outlying observations, where h denotes the number of outliers. We consider the model with/without outliers such that

$$f\,(y_{ijk}|\mu, p_j, \tau_{d[i,j]}, \lambda_{d[i,j-1]}, \delta)$$
$$= \begin{cases} \left[2\pi(\sigma_s^2 + \sigma_e^2\right]^{-\frac{1}{2}} \exp\left[-\dfrac{1}{2(\sigma_s^2 + \sigma_e^2)}(y_{ijk} - \mu - p_j - \tau_{d[i,j]} - \lambda_{d[i,j-1]})^2\right] & \text{for } k \notin v^h \\ \left[2\pi(\delta^2\sigma_s^2 + \sigma_e^2\right]^{-\frac{1}{2}} \exp\left[-\dfrac{1}{2(\delta^2\sigma_s^2 + \sigma_e^2)}(y_{ijk} - \mu - p_j - \tau_{d[i,j]} - \lambda_{d[i,j-1]})^2\right] & \text{for } k \in v^h \end{cases} \tag{2}$$

Using the Bayesian approach, we consider normal prior distributions for the overall mean, μ, the period effect, p_j, the treatment effect, $\tau_{d[i,\,j]}$, and the carryover effect, $\lambda_{d[i,\,j-1]}$, as suggested by Chen and Huang (2015). For the parameter δ, Unnikrishnan (2010) assume that this extra variance component of the outlying observations is bounded above by a known constant δ_{max}, so that $1 < \delta < \delta_{max} < \infty$ and therefore Uniform$(1, \delta_{max})$ prior is assigned to it. According to the suggestions in Unnikrishnan (2010), we shall assume that any distinct -tuples are equally likely to be outliers and prior for v^h assigns equal probability of $\binom{N}{h}^{-1}$. In other words, we assume that

$$\mu \sim N(\mu_0, \sigma_\mu^2)$$

$$p_j \sim N(\mu_p, \sigma_p^2), \quad j = 1, 2$$

$$\tau_{d[i,\,j]} \sim N(\mu_\tau, \sigma_\tau^2), \quad i = 1, 2; \; j = 1, 2$$

$$\lambda_{d[i,\,j-1]} \sim N(\mu_\lambda, \sigma_\lambda^2)$$

$$\delta \sim Uniform(1, \delta_{max})$$

$$p(v^h|h) = \binom{N}{h}^{-1} \tag{3}$$

where the hyperparameters μ_0, σ_μ^2, μ_p, σ_p^2, u_τ, σ_τ^2, u_λ, σ_λ^2, δ_{max}, N, h are all pre-specified. Then, the joint likelihood function is given by

$$L(\boldsymbol{y}|\mu, \boldsymbol{p}, \boldsymbol{\tau}, \boldsymbol{\lambda}, \delta, v^h)$$

$$= \prod_{i=1}^{2}\prod_{j=1}^{2}\prod_{k\in v^h}\left[2\pi(\sigma_s^2+\sigma_e^2)\right]^{-\frac{1}{2}}$$

$$\exp\left[-\frac{1}{2(\sigma_s^2+\sigma_e^2)}(y_{ijk}-\mu-p_j-\tau_{d[i,j]}-\lambda_{d[i,j-1]})^2\right]$$

$$\times\Pi_{i=1}^{2}\Pi_{j=1}^{2}\Pi_{k\in v^h}\left[2\pi(\delta^2\sigma_s^2+\sigma_e^2)\right]^{-\frac{1}{2}}$$

$$\exp\left[-\frac{1}{2(\delta^2\sigma_s^2+\sigma_e^2)}(y_{ijk}-\mu-p_j-\tau_{d[i,j]}-\lambda_{d[i,j-1]})^2\right]. \tag{4}$$

Consequently, from the result obtain in (3) and (4), the full joint posterior distribution for the parameters μ, $\boldsymbol{p}$, $\boldsymbol{\tau}$, $\boldsymbol{\lambda}$, δ, v^h is given by

$$f(\mu, \boldsymbol{p}, \boldsymbol{\tau}, \boldsymbol{\lambda}, \delta, v^h|\boldsymbol{y})$$

$$\propto \prod_{i=1}^{2}\prod_{j=1}^{2}\prod_{k\in v^h}\left[2\pi(\sigma_s^2+\sigma_e^2)\right]^{-\frac{1}{2}}$$

$$\exp\left[-\frac{1}{2(\sigma_s^2+\sigma_e^2)}(y_{ijk}-\mu-p_j-\tau_{d[i,j]}-\lambda_{d[i,j-1]})^2\right]$$

$$\times\prod_{i=1}^{2}\prod_{j=1}^{2}\prod_{k\in v^h}\left[2\pi(\delta^2\sigma_s^2+\sigma_e^2)\right]^{-\frac{1}{2}}$$

$$\exp\left[-\frac{1}{2(\delta^2\sigma_s^2+\sigma_e^2)}(y_{ijk}-\mu-p_j-\tau_{d[i,j]}-\lambda_{d[i,j-1]})^2\right]$$

$$\times\left(2\pi\sigma_\mu^2\right)^{-\frac{1}{2}}\exp\left[-\frac{(\mu-\mu_0)^2}{2\sigma_\mu^2}\right]$$

$$\times\left(2\pi\sigma_p^2\right)^{-\frac{1}{2}}\exp\left[-\frac{(p_j-\mu_p)^2}{2\sigma_p^2}\right]$$

$$\times\left(2\pi\sigma_\tau^2\right)^{-\frac{1}{2}}\exp\left[-\frac{(\tau_{(d[i,j]}-\mu_\tau)^2}{2\sigma_\tau^2}\right]$$

$$\times\left(2\pi\sigma_\lambda^2\right)^{-\frac{1}{2}}\exp\left[-\frac{(\tau_{(d[i,j-1]}-\mu_\lambda)^2}{2\sigma_\lambda^2}\right]$$

$$\times\frac{1}{\delta_{\max}-1}\times\left(\frac{N!}{(N-h)!h!}\right). \tag{5}$$

Since this posterior distribution is intractable, sampling is carried out using the MCMC sampling method, in particular using Metropolis-Hastings (MH) algorithm (Tierney 1994).

SAMPLING METHODS OF THE PARAMETERS

Note that model (2) involves multiple parameters that are structured hierarchically such that the dependency of the parameters is reflected in the joint probability distribution. The conditional posterior distributions of the parameters are intractable and therefore we use the MH algorithm for sampling purposes. The sampling methods for each of the parameters μ, $\boldsymbol{p}$, $\boldsymbol{\tau}$, $\boldsymbol{\lambda}$, δ, v^h are given in detail as below.

PARAMETER δ

Based on the full joint posterior distribution (5), the conditional posterior distribution for parameter given is given by

$$f(\delta|\mu, \boldsymbol{p}, \boldsymbol{\tau}, \boldsymbol{\lambda}, v^h)$$

$$\propto \prod_{i=1}^{2}\prod_{j=1}^{2}\prod_{k\in v^h}\frac{[2\pi(\delta^2\sigma_s^2+\sigma_e^2)]^{-\frac{1}{2}}}{\delta_{\max}-1}\exp\left[-\frac{1}{2(\delta^2\sigma_s^2+\sigma_e^2)}\right.$$

$$\left.(y_{ijk}-\mu-p_j-\tau_{d[i,j]}-\lambda_{d[i,j-1]})^2\right].$$

Here we propose to use a proposal density for δ_{prop} as

$$g(\delta) = \frac{1}{\delta_{max}-1}$$

so that δ_{prop} has a uniform $(1, \delta_{max})$ distribution. Using the MH algorithm, candidate point δ_{prop} is accepted with probability

$$\alpha(\delta,\, \delta_{prop}) = \min\left(1, \frac{f(\delta_{prop}\big|\mu, \boldsymbol{p}, \boldsymbol{\tau}, \boldsymbol{\lambda}, v^h)g(\delta)}{f(\delta\big|\mu, \boldsymbol{p}, \boldsymbol{\tau}, \boldsymbol{\lambda}, v^h)g(\delta_{prop})}\right). \tag{6}$$

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

PARAMETER μ

Based on the full joint posterior distribution (5), the conditional posterior distribution for parameter μ given $\boldsymbol{p}, \boldsymbol{\tau}, \boldsymbol{\lambda}, \delta, v^h$ is given by

$$f(\mu \mid \boldsymbol{p}, \boldsymbol{\tau}, \boldsymbol{\lambda}, \delta, v^h)$$

$$\propto \prod_{i=1}^{2}\prod_{j=1}^{2}\prod_{k\notin v^h}\left[2\pi\left(\sigma_s^2+\sigma_e^2\right)\right]^{-\frac{1}{2}}$$

$$\exp\left[-\frac{1}{2\left(\sigma_s^2+\sigma_e^2\right)}\left(y_{ijk}-\mu-p_j-\boldsymbol{\tau}_{d[i,j]}-\boldsymbol{\lambda}_{d[i,j-1]}\right)^2\right]$$

$$\times\prod_{i=1}^{2}\prod_{j=1}^{2}\prod_{k\in v^h}\left[2\pi\left(\delta^2\sigma_s^2+\sigma_e^2\right)\right]^{-\frac{1}{2}}$$

$$\exp\left[-\frac{1}{2\left(\delta^2\sigma_s^2+\sigma_e^2\right)}\left(y_{ijk}-\mu-p_j-\boldsymbol{\tau}_{d[i,j]}-\boldsymbol{\lambda}_{d[i,j-1]}\right)^2\right]$$

$$\times\left(2\pi\sigma_\mu^2\right)^{\frac{1}{2}}\exp\left[-\frac{\left(\mu-\mu_0\right)^2}{2\sigma_\mu^2}\right].$$

Here, we introduce a function ω_k where

$$\omega_k = \begin{cases} 1 & \text{for } k \in v^h \\ 0 & \text{for } k \notin v^h \end{cases} \tag{7}$$

so that

$$f(\mu|\boldsymbol{p}, \boldsymbol{\tau}, \boldsymbol{\lambda}, \delta, v^h)$$

$$\propto \prod_{i=1}^{2}\prod_{j=1}^{2}\prod_{k=1}^{n}\left\{\left[2\pi(\sigma_s^2+\sigma_e^2)\right]^{-\frac{(1-\omega_k)}{2}}\right.$$

$$\left.\exp\left[-\frac{(1-\omega_k)}{2(\sigma_s^2+\sigma_e^2)}(y_{ijk}-\mu-p_j-\boldsymbol{\tau}_{d[i,j]}-\boldsymbol{\lambda}_{d[i,j-1]})^2\right]\right\}$$

$$\times\prod_{i=1}^{2}\prod_{j=1}^{2}\prod_{k=1}^{n}\left\{\left[2\pi(\delta^2\sigma_s^2+\sigma_e^2)\right]^{-\frac{\omega_k}{2}}\right.$$

$$\left.\exp\left[-\frac{\omega_k}{2(\delta^2\sigma_s^2+\sigma_e^2)}(y_{ijk}-\mu-p_j-\boldsymbol{\tau}_{d[i,j]}-\boldsymbol{\lambda}_{d[i,j-1]})^2\right]\right\}$$

$$\times\left(2\pi\sigma_\mu^2\right)^{-\frac{1}{2}}\exp\left[-\frac{(\mu-\mu_0)^2}{2\sigma_\mu^2}\right].$$

Here we choose the proposal density for μ_{prop} as $g(\mu) = \left(2\pi\sigma_\mu^2\right)^{-\frac{1}{2}}\exp\left[-\frac{(\mu-\mu_0)^2}{2\sigma_\mu^2}\right]$, so that μ_{prop} has a $N(\mu_0,\sigma_\mu^2)$ distribution. Using the MH algorithm, the acceptance probability for candidate point μ_{prop} then can be obtained by replacing the corresponding functions in (6).

PARAMETER p

Based on the full joint posterior distribution (5), the conditional posterior distribution for parameter p_j, j = 1, 2, given μ, $\boldsymbol{\tau}$, $\boldsymbol{\lambda}$, δ, v^h is given by

$$f_j(p_j \mid \mu, \boldsymbol{\tau}, \boldsymbol{\lambda}, \delta, v^h, \boldsymbol{p})$$

$$\propto \prod_{i=1}^{2}\prod_{j=1}^{2}\prod_{k\notin v^h}^{n}\left[2\pi(\sigma_s^2+\sigma_e^2)\right]^{-\frac{1}{2}}$$

$$\exp\left[-\frac{1}{2(\sigma_s^2+\sigma_e^2)}(y_{ijk}-\mu-p_j-\boldsymbol{\tau}_{d[i,j]}-\boldsymbol{\lambda}_{d[i,j-1]})^2\right]$$

$$\times\prod_{i=1}^{2}\prod_{j=1}^{2}\prod_{k\in 1}^{n}\left[2\pi(\delta^2\sigma_s^2+\sigma_e^2)\right]^{-\frac{1}{2}}$$

$$\exp\left[-\frac{1}{2(\delta^2\sigma_s^2+\sigma_e^2)}(y_{ijk}-\mu-p_j-\boldsymbol{\tau}_{d[i,j]}-\boldsymbol{\lambda}_{d[i,j-1]})^2\right]$$

$$\times\left(2\pi\sigma_p^2\right)^{-\frac{1}{2}}\exp\left[-\frac{(p_j-\mu_p)^2}{2\sigma_p^2}\right].$$

We use the function ω_k as defined in equation (7) so that

$$f_j(p_j \mid \mu, \boldsymbol{\tau}, \boldsymbol{\lambda}, \delta, v^h, \boldsymbol{p})$$

$$\propto \prod_{i=1}^{2}\prod_{j=1}^{2}\prod_{k\notin v^h}^{n}\left\{\left[2\pi(\sigma_s^2+\sigma_e^2)\right]^{-\frac{(1-\omega_k)}{2}}\right.$$

$$\left.\exp\left[-\frac{(1-\omega_k)}{2(\sigma_s^2+\sigma_e^2)}(y_{ijk}-\mu-p_j-\boldsymbol{\tau}_{d[i,j]}-\boldsymbol{\lambda}_{d[i,j-1]})^2\right]\right\}$$

$$\times\prod_{i=1}^{2}\prod_{j=1}^{2}\prod_{k=1}^{n}\left\{\left[2\pi(\delta^2\sigma_s^2+\sigma_e^2)\right]^{-\frac{\omega_k}{2}}\right.$$

$$\left.\exp\left[-\frac{\omega_k}{2(\delta^2\sigma_s^2+\sigma_e^2)}(y_{ijk}-\mu-p_j-\boldsymbol{\tau}_{d[i,j]}-\boldsymbol{\lambda}_{d[i,j-1]})^2\right]\right\}$$

$$\times\left(2\pi\sigma_p^2\right)^{-\frac{1}{2}}\exp\left[-\frac{(p_j-\mu_p)^2}{2\sigma_p^2}\right].$$

Here we propose to use a proposal density for p_j as $g_j(p_j) = \left(2\pi\sigma_p^2\right)^{-\frac{1}{2}}\exp\left[-\frac{(p_j-\mu_p)^2}{2\sigma_p^2}\right]$. Therefore, by using MH algorithm, for parameter $\boldsymbol{p}$, we update p_1 first and followed by p_2, where for each p_j, the acceptance probability for candidate point $p_{j_{prop}}$ can be obtained by replacing the corresponding functions in (6).

PARAMETER τ

Based on the full joint posterior distribution (5), the conditional posterior distribution for parameter $\tau_{d[i,j]}$, i, j, = 1, 2, given μ, $\boldsymbol{p}$, $\boldsymbol{\lambda}$, δ, v^h is given by

$$f_{d[i,j]}(\tau_{d[i,j]} \mid \mu, \boldsymbol{p}, \boldsymbol{\lambda}, \delta, v^h, \boldsymbol{\tau})$$

$$\propto \prod_{i=1}^{2}\prod_{j=1}^{2}\prod_{k\notin v^h}^{n}[2\pi(\sigma_s^2+\sigma_e^2)]^{-\frac{1}{2}}$$
$$\exp\left[-\tfrac{1}{2(\sigma_s^2+\sigma_e^2)}(y_{ijk}-\mu-p_j-\tau_{d[i,j]}-\lambda_{d[i,j-1]})^2\right]$$
$$\times\prod_{i=1}^{2}\prod_{j=1}^{2}\prod_{k\in v^h}\left[2\pi(\delta^2\sigma_s^2+\sigma_e^2)\right]^{-\frac{1}{2}}$$
$$\exp\left[-\tfrac{1}{2(\delta^2\sigma_s^2+\sigma_e^2)}(y_{ijk}-\mu-p_j-\tau_{d[i,j]}-\lambda_{d[i,j-1]})^2\right]$$
$$\times\left(2\pi\sigma_\tau^2\right)^{-\frac{1}{2}}\exp\left[-\tfrac{(\tau_{d[i,j]}-\mu_\tau)^2}{2\sigma_\tau^2}\right].$$

We use the function ω_t as defined in equation (7) so that

$$f_{d[i,j]}(\tau_{d[i,j]}\,|\,\mu,\boldsymbol{p},\boldsymbol{\lambda},\delta,v^h,\boldsymbol{\tau})$$
$$\propto \prod_{i=1}^{2}\prod_{j=1}^{2}\prod_{k=1}^{n}\left\{\left[2\pi(\sigma_s^2+\sigma_e^2)\right]^{-\frac{(1-\omega_k)}{2}}\right.$$
$$\left.\exp\left[-\tfrac{(1-\omega_k)}{2(\sigma_s^2+\sigma_e^2)}(y_{ijk}-\mu-p_j-\tau_{d[i,j]}-\lambda_{d[i,j-1]})^2\right]\right\}$$
$$\times\prod_{i=1}^{2}\prod_{j=1}^{2}\prod_{k=1}^{n}\left\{\left[2\pi(\delta^2\sigma_s^2+\sigma_e^2)\right]^{-\frac{\omega_k}{2}}\right.$$
$$\left.\exp\left[-\tfrac{\omega_k}{2(\delta^2\sigma_s^2+\sigma_e^2)}(y_{ijk}-\mu-p_j-\tau_{d[i,j]}-\lambda_{d[i,j-1]})^2\right]\right\}$$
$$\times\left(2\pi\sigma_p^2\right)^{-\frac{1}{2}}\exp\left[-\tfrac{(\tau_{d[i,j]}-\mu_\tau)^2}{2\sigma_\tau^2}\right].$$

Here we propose to use a proposal density for as Therefore, by using MH algorithm, for parameter , we update and one by one, where for each , the acceptance probability for candidate point can be obtained by replacing the corresponding functions in (6).

PARAMETER λ

Based on the full joint posterior distribution (5), the conditional posterior distribution for parameter $\lambda_{d[i,\,j-1]}$, $i, j = 1, 2$, given $\mu, \boldsymbol{p}, \boldsymbol{\tau}, \delta, v^h$ is given by

$$f_{d[i,\,j-1]}(\lambda_{d[i,\,j-1]}\,|\,\mu,\boldsymbol{p},\boldsymbol{\tau},\delta,v^h,\boldsymbol{\lambda})$$
$$\propto \prod_{i=1}^{2}\prod_{j=1}^{2}\prod_{k\notin v^h}^{n}[2\pi(\sigma_s^2+\sigma_e^2)]^{-\frac{1}{2}}$$
$$\exp\left[-\tfrac{1}{2(\sigma_s^2+\sigma_e^2)}(y_{ijk}-\mu-p_j-\tau_{d[i,j]}-\lambda_{d[i,j-1]})^2\right]$$
$$\times\prod_{i=1}^{2}\prod_{j=1}^{2}\prod_{k\in v^h}\left[2\pi(\delta^2\sigma_s^2+\sigma_e^2)\right]^{-\frac{1}{2}}$$
$$\exp\left[-\tfrac{1}{2(\delta^2\sigma_s^2+\sigma_e^2)}(y_{ijk}-\mu-p_j-\tau_{d[i,j]}-\lambda_{d[i,j-1]})^2\right]$$
$$\times\left(2\pi\sigma_p^2\right)^{-\frac{1}{2}}\exp\left[-\tfrac{(\lambda_{d[i,j-1]}-\mu_\lambda)^2}{2\sigma_\lambda^2}\right].$$

We use the function ω_t as defined in equation (7) so that

$$f_{d[i,\,j-1]}(\lambda_{d[i,\,j-1]}\,|\,\mu,\boldsymbol{p},\boldsymbol{\tau},\delta,v^h,\boldsymbol{\lambda})$$
$$\propto \prod_{i=1}^{2}\prod_{j=1}^{2}\prod_{k\notin v^h}^{n}\left\{[2\pi(\sigma_s^2+\sigma_e^2)]^{\frac{(1-\omega_k)}{2}}\right.$$
$$\left.\exp\left[-\tfrac{(1-\omega_k)}{2(\sigma_s^2+\sigma_e^2}(y_{ijk}-\mu-p_j-\tau_{d[i,j]}-\lambda_{d[i,j-1]})^2\right]\right\}$$
$$\times\prod_{i=1}^{2}\prod_{j=1}^{2}\prod_{k=1}^{n}\left\{\left[2\pi(\delta^2\sigma_s^2+\sigma_e^2\right]^{\frac{\omega_k}{2}}\right.$$
$$\left.\exp\left[-\tfrac{\omega_k}{2(\delta^2\sigma_s^2+\sigma_e^2}(y_{ijk}-\mu-p_j-\tau_{d[i,j]}-\lambda_{d[i,j-1]})^2\right]\right\}$$
$$\times\left(2\pi\sigma_p^2\right)^{-\frac{1}{2}}\exp\left[-\tfrac{(\lambda_{d[i,j-1]}-\mu_\lambda)^2}{2\sigma_\lambda^2}\right].$$

Here we propose to use a proposal density for $\lambda_{d[i,j-1]}$ as $g_{d[i,\,j-1]}(\lambda_{d[i,\,j-1]}) = \left(2\pi\sigma_p^2\right)^{-\frac{1}{2}}\exp\left[-\frac{\left(\lambda_{d[i,j=1]}-\mu_\lambda\right)^2}{2\sigma_\lambda^2}\right]$. Therefore, by using MH algorithm, for parameter $\boldsymbol{\lambda}$, we have $\lambda_{10} = \lambda_{20} = 0$ and update λ_{11} and λ_{21} one by one, where for each $\lambda_{d[i,\,j-1]}$, the acceptance probability for candidate point $\lambda_{d[i,j-1]_{prop}}$ can be obtained by replacing the corresponding functions in (6).

PARAMETER v^h

Based on the full joint posterior distribution (5), the conditional posterior distribution for parameter v^h given $\mu, \boldsymbol{p}, \boldsymbol{\tau}, \boldsymbol{\lambda}, \delta$ is given by

$$f(v^h\,|\,\mu,\boldsymbol{p},\boldsymbol{\tau},\boldsymbol{\lambda},\delta)$$
$$\propto \prod_{i=1}^{2}\prod_{j=1}^{2}\prod_{k\notin v^h}\left[2\pi(\sigma_s^2+\sigma_e^2\right]^{\frac{1}{2}}$$
$$\exp\left[-\tfrac{1}{2(\sigma_s^2+\sigma_e^2}(y_{ijk}-\mu-p_j-\tau_{d[i,j]}-\lambda_{d[i,j-1]})^2\right]$$
$$\times\prod_{i=1}^{2}\prod_{j=1}^{2}\prod_{k\in v^h}\left[2\pi(\delta^2\sigma_s^2+\sigma_e^2\right]^{\frac{1}{2}}$$
$$\exp\left[-\tfrac{1}{2(\delta^2\sigma_s^2+\sigma_e^2}(y_{ijk}-\mu-p_j-\tau_{d[i,j]}-\lambda_{d[i,j-1]})^2\right].$$

For the case of $h = 1$, we let $v^1 = v = \{v_1\}$. To find a new value of v_1, we select a unit at random from v^c, say v_{prop}. If the proposal is accepted, then v_1 goes out and v_{prop} replace the value v_1 as the current outlier. Then, using MH algorithm, this state is accepted with probability

$$\alpha(v_1,v_{prop})=\min\left(1,\frac{\begin{matrix}\Pi_{i=1}^{2}\Pi_{j=1}^{2}\Pi_{v_{prop}\notin v}f(y_{ijk}|\mu,p_j,\tau_{d[i,j]},\lambda_{d[i,j-1]})\\ \times\Pi_{i=1}^{2}\Pi_{j=1}^{2}\Pi_{v_{prop}\in v}f(y_{ijk}|\mu,p_j,\tau_{d[i,j]},\lambda_{d[i,j-1]},\delta)\end{matrix}}{\begin{matrix}\Pi_{i=1}^{2}\Pi_{j=1}^{2}\Pi_{k\notin v}f(y_{ijk}|\mu,p_j,\tau_{d[i,j]},\lambda_{d[i,j-1]})\\ \times\Pi_{i=1}^{2}\Pi_{j=1}^{2}\Pi_{k\in v}f(y_{ijk}|\mu,p_j,\tau_{d[i,j]},\lambda_{d[i,j-1]},\delta)\end{matrix}}\right). \tag{8}$$

RESULTS AND DISCUSSION

The method as described in the previous section is now applied to Clayton and Leslie's data (1981) and kinesiology data. The values of μ_0, μ_π, μ_τ and μ_λ equal to 0 while the values of $\sigma_\mu^2, \sigma_p^2, \sigma_\tau^2, \sigma_\lambda^2, \sigma_s^2$ and σ_e^2 equal to 1000; these values are suggested by Chen and Huang (2015), and δ_{max} equals to 10.

Using Clayton and Leslie's data (1981), we run the method for 1000 iterations, with a burn-in of 500. We are especially interested in estimating the probability of a subject being an outlier. The probability of subject i being an outlier in this model can be estimated using the proportion of iterations that $v = \{i\}$. Figure 1 shows the estimated probability of being an outlier for subjects 1 to 18. Given that there is one outlier, subject 11 (blood concentration are 7.14 µg.h/mL and 9.83 µg.h/mL, respectively, in period 1 and period 2) has the highest probability of being an outlier with the probability of approximately 0.30. As can be seen, subject 11 (that is, subject 2 from group 2) have high value of blood concentration in period 2 compared to their means (3.51 µg.h/mL in group 1 and 4.72 µg.h/mL in group 2) indicating the subject 11 is candidate to be outliers. Note that subject 11 is one of the outlier, identified by the procedure using *SR*2 as described in Lim et al. (2016) with non-Bayesian framework.

The kinesiology data of peak oxygen consumption or VO_2 peak also is used as illustration in this section. However, only 74 subjects who completed the study (AB = 37, BA = 37) are included for analyses. We run the method for 10000 iterations, with a burn-in of 5000. With the same interest as in the previous section, Figure 2 shows the estimated probability of being an outlier for subjects 1 to 74 using the proportion of iterations that $v = \{i\}$. Given that there is one outlier, subject 50 has the highest probability of being an outlier with the probability of approximately 0.32. This is likely because for subject 50 (that is, subject 13 from group 2), the VO_2 is unusually large in the data set for period 1. Therefore, we may conclude that subject 13 from group 2 is likely an outlier. Note that subject 50 is one of the outliers identified by the procedures using *SR*1 and *SR*2 as described in Lim et al. (2016) with non-Bayesian framework.

Conclusion

In this chapter, we have considered the problem of detecting outlier using Bayesian approach in 2 × 2 crossover design. We have shown that with the chosen prior distributions for the parameter, we can obtain the information from samples generated by MCMC sampling, in particular using the MH algorithm. When applied to both Clayton and Leslie's data (1981) and kinesiology data, this method is able to detect an unusual large observation as being an outlier with the highest probability as compared to the other observations.

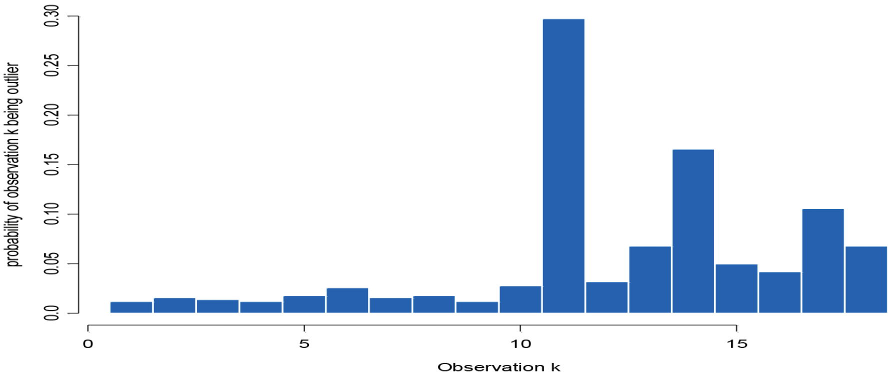

FIGURE 1. Clayton and Leslie's data: Probability for an observation being outlier

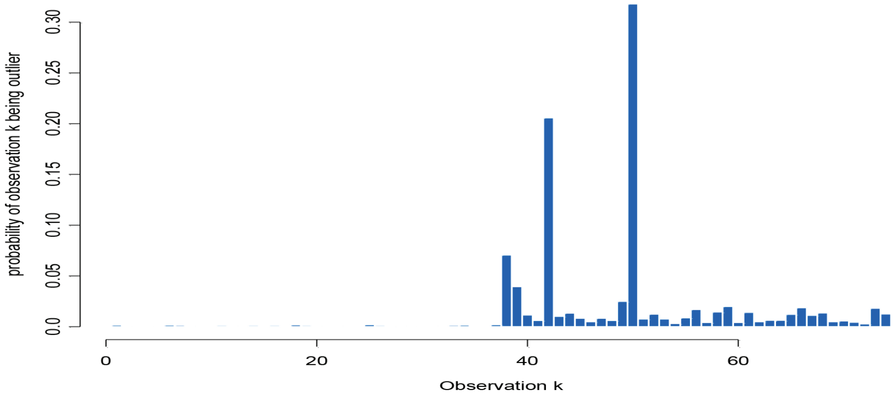

FIGURE 2. Kinesiology data: Probability for an observation being outlier

ACKNOWLEDGEMENTS

We thank University of Malaya Medical Center (UMMC) Sport Medicine Clinic for providing the data. This research was partially financed by UMRG (No. RP014C-15SUS), FRGS (FP037-2014B), UM-BKP (BK 017 - 2012) and GP-IPM (GP-IPM/2017/9587400).

REFERENCES

Aggarwal, C.C. 2013. *Outlier Analysis*. New York: Springer.

Chen, Y.I. & Huang, C.S. 2015. Bayesian bioequivalence test based on skew normal model for bioavailability measure. *Journal of the Chinese Statistical Association* 53: 38-54.

Chow, S.C. & Tse, S.K. 1990. Outliers detection in bioavailability/bioequivalence studies. *Statistics in Medicine* 9: 549-558.

Clayton, D. & Leslie, A. 1981. The bioavailability of erythromycin stearate versus enteric-coated erythromycin based when taken immediately before and after food. *Journal of International Medical Research* 9: 470-477.

Enachescu, D. & Enachescu, C. 2009. A new approach for outlying records in bioequivalence trials. Paper presented at the *XIII International Conference Applied Stochastic Models and Data Analysis* 13: 250-257.

Hawkins, D.M. 1980. *Identification of Outliers*. London: Chapman and Hall.

Jones, B. & Kenward, M.G. 1989. *Design and Analysis of Cross-Over Trials*. 1st edition. London: Chapman and Hall.

Karasoy, D. & Daghan, G. 2012. Examination of outliers in bioequivalence studies. *Bulletin of Clinical Psychopharmacology* 22(4): 307-312.

Lim, F.P., Mohamed, I., Daud, N. & Goh, S.L. 2016. Comparison of outlier detection methods in standard 2×2 crossover design. *Sains Malaysiana* 45(3): 499-506.

Liu, J.P. & Weng, C.S. 1991. Detection of outlying data in bioavailability/bioequivalence studies. *Statistics in Medicine* 10: 1375-1389.

Nawama, M., Ibrahim, A.I.N., Mohamed, I., Yahya, M.S. & Taib, N.A. 2015. Outlier detection using generalized linear model in Malaysian breast cancer data. *Sains Malaysiana* 44(10): 1417-1422.

Ramsay, T. & Elkum, N. 2005. A comparison of four different methods for outlier detection in bioequivalence studies. *Journal of Biopharmaceutical Statistics* 15(1): 43-52.

Singh, R., Namdev, K.K. & Chilkoti, D. 2014. An assessment on handling of outlier subjects in bioequivalence study - A review. *Journal of Pharmaceutical and Bioanalytical Science* 3(1): 14-20.

Tierney, L. 1994. Markov chains for exploring posterior distributions. *The Annals of Statistics* 22(4): 1701-1728.

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

Wang, W. & Chow, S.C. 2003. Examining outlying subjects and outlying records in bioequivalence trials. *Journal of Biopharmaceutical Statistics* 13(1): 43-56.

Weiner, D. 1989. *Bioavailability. Notes on the Training Course for New Clinical Statisticians* sponsored by the Biostatistics Subsections of Pharmaceutical Manufacture Association, March, 1989, Washington, DC.

F.P. Lim, I.B. Mohamed*, A.I.N. Ibrahim & N.A. Mohamed @ A. Rahman
Institute of Mathematical Sciences
University of Malaya
50603 Kuala Lumpur, Federal Territory
Malaysia

F.P. Lim
Faculty of Sciences
Universiti Putra Malaysia
43400 UPM Serdang, Selangor Darul Ehsan
Malaysia

S.L. Goh
Sport Medicine Clinic
University of Malaya Medical Centre
50603 Kuala Lumpur, Federal Territory
Malaysia

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

Received: 23 April 2016
Accepted: 20 January 2019