doi:10.3969/j.issn.1001-4616.2016.04.004

缺失数据下 Logistic 回归多变点模型的贝叶斯估计

何朝兵

(安阳师范学院数学与统计学院,河南 安阳 455000)

[摘要] 利用随机的方法填充了缺失数据,获得了 Logistic 回归多变点模型的完全数据似然函数. 研究了变点位置等未知参数的满条件分布. 利用筛选法和 Metropolis-Hastings 算法对参数进行抽样,把 Gibbs 样本的均值作为参数的贝叶斯估计. 随机模拟的结果表明估计的精度较高.

[关键词] 完全数据似然函数,满条件分布,筛选法,Gibbs 抽样,Metropolis-Hastings 算法
[中图分类号]O212.4;O212.8 [文献标志码]A [文章编号]1001-4616(2016)04-0014-05

Bayesian Estimation of Logistic Regression Model with Multiple Change Points for Missing Data

He Chaobing

(School of Mathematics and Statistics, Anyang Normal University, Anyang 455000, China)

Abstract: The missing data is filled in by a random way. The complete-data likelihood function of logistic regression model with multiple change points is obtained. The full conditional distributions of change-point positions and other unknown parameters are studied. All the parameters are sampled by screening method and Metropolis-Hastings algorithm, and the means of Gibbs samples are taken as Bayesian estimations of the parameters. Random simulation results show that the estimations are fairly accurate.

Key words: complete-data likelihood function, full conditional distribution, screening method, Gibbs sampling, Metropolis-Hastings algorithm

变点模型在数理统计中非常重要,在工业质量控制、水文统计等领域应用非常广泛[1-4]. 随着统计计算技术的快速发展,贝叶斯方法的应用越来越广泛,特别是其中的 Markov Chain Monte Carlo(MCMC)方法应用尤其广泛[5-7]. MCMC 方法中的 Gibbs 抽样和 Metropolis-Hastings 算法使变点模型的参数估计变得非常方便. Logistic 回归模型是非线性回归模型,是最流行的二分数据的广义线性模型[8-10]. 在实际问题中,当研究的因变量是二分类时,通常首选 Logistic 回归进行建模. 文献[11-14]对完全数据下 Logistic 回归单变点模型进行了研究,但对缺失数据下多变点情形的研究还不多见.

本文主要利用 MCMC 方法研究了缺失数据下 Logistic 回归多变点模型的参数估计问题. 利用随机的方法填充了缺失数据,利用筛选法和 Metropolis-Hastings 算法对参数进行 Gibbs 抽样,把 Gibbs 样本的均值作为参数的贝叶斯估计. 随机模拟的结果表明估计的精度较高.

1 缺失数据下的 Logistic 回归模型

Logistic 回归模型描述如下:

$$y_i \sim b(1,p_i),\ i=1,2,\cdots,n.\ \text{且诸}\ y_i\ \text{独立},$$

式中,$p_i=F(\alpha+\beta x_i)=[1+e^{-(\alpha+\beta x_i)}]^{-1}$,$F(\cdot)$为分布 Logistic(0,1)的分布函数. 其中 $\boldsymbol{\theta}=(\alpha,\beta)$为模型参数向量.

收稿日期:2015-09-20.
基金项目:河南省高等学校重点科研项目(16A110001).
通讯联系人:何朝兵,硕士,讲师,研究方向:概率统计. E-mail:chaobing5@163.com

完全数据下 Logistic 回归模型的似然函数为

$$L(\boldsymbol{\theta})=\prod_{i=1}^{n}p_i^{y_i}(1-p_i)^{1-y_i}=\prod_{i=1}^{n}\{\mathrm{e}^{-(1-y_i)(\alpha+\beta x_i)}[1+\mathrm{e}^{-(\alpha+\beta x_i)}]^{-1}\}.$$

假设观察数据(y_i,x_i)中有部分y_i缺失,但对应的x_i能观察到,但为了充分利用数据x_i,不妨添加缺失的$y_i=y_{1i}$,由于$y_{1i}\sim b(1,F(\alpha+\beta x_i))$,所以可以通过随机抽样产生$y_{1i}$.

引入示性变量$\delta_i=I(y_i$ 没有缺失$)$,令$\boldsymbol{x},\boldsymbol{y},\boldsymbol{u}_1,\boldsymbol{\delta}$分别表示由$x_i,y_i,y_{1i},\delta_i$组成的向量,则添加数据后的似然函数为

$$L(\boldsymbol{x},\boldsymbol{y},\boldsymbol{u}_1,\boldsymbol{\delta}\mid\boldsymbol{\theta})=\prod_{i=1}^{n}\{[p_i^{y_i}(1-p_i)^{1-y_i}]^{\delta_i}[p_i^{y_{1i}}(1-p_i)^{1-y_{1i}}]^{(1-\delta_i)}\}=\prod_{i=1}^{n}\{\mathrm{e}^{-d_i(\alpha+\beta x_i)}[1+\mathrm{e}^{-(\alpha+\beta x_i)}]^{-1}\},\quad(1)$$

式中,$d_i=\delta_i(1-y_i)+(1-\delta_i)(1-y_{1i})$.

2 多变点模型

Logistic 回归多变点模型如下:

$$p_i=\varphi(x_i;\boldsymbol{\gamma})=\begin{cases}F(\alpha_1+\beta_1x_i),i=1,2,\cdots,k_1;\\F(\alpha_2+\beta_2x_i),i=k_1+1,\cdots,k_2;\\F(\alpha_3+\beta_3x_i),i=k_2+1,\cdots,n.\end{cases}\quad(2)$$

式中,$\boldsymbol{\gamma}=(k_1,k_2,\alpha_1,\alpha_2,\alpha_3,\beta_1,\beta_2,\beta_3)$,$(\alpha_1,\beta_1)\neq(\alpha_2,\beta_2)$,$(\alpha_2,\beta_2)\neq(\alpha_3,\beta_3)$,$1\leqslant k_1<k_2\leqslant n$. k_1、k_2 称为变点位置参数.

3 贝叶斯估计

下面讨论缺失数据下 Logistic 回归变点模型中参数的贝叶斯估计.

令$D_1=\{1,2,\cdots,k_1\}$,$D_2=\{k_1+1,\cdots,k_2\}$,$D_3=\{k_2+1,\cdots,n\}$,由式(1)和式(2)可得此变点模型的似然函数为

$$L(\boldsymbol{x},\boldsymbol{y},\boldsymbol{u}_1,\boldsymbol{\delta}\mid\boldsymbol{\gamma})=\prod_{m=1}^{3}\prod_{i\in D_m}\{\mathrm{e}^{-d_i(\alpha_m+\beta_mx_i)}[1+\mathrm{e}^{-(\alpha_m+\beta_mx_i)}]^{-1}\}.$$

下面确定参数的先验分布.

(1)对于(k_1,k_2)取无信息先验分布:$\pi(k_1,k_2)=\dfrac{1}{C_n^2}=\dfrac{2}{n(n-1)}$,$1\leqslant k_1<k_2\leqslant n$.

(2)取α_m,β_m的先验分布分别为正态分布$N(\mu_{1m},\sigma_{1m}^2)$,$N(\mu_{2m},\sigma_{2m}^2)$,$m=1,2,3$.

假设(k_1,k_2),$\alpha_1,\alpha_2,\alpha_3,\beta_1,\beta_2,\beta_3$相互独立,则

$$\pi(\boldsymbol{\gamma}\mid\boldsymbol{x},\boldsymbol{y},\boldsymbol{u}_1,\boldsymbol{\delta})\propto L(\boldsymbol{x},\boldsymbol{y},\boldsymbol{u}_1,\boldsymbol{\delta}\mid\boldsymbol{\gamma})\pi(k_1,k_2)\prod_{m=1}^{3}[\pi(\alpha_m)\pi(\beta_m)]\propto$$

$$\prod_{m=1}^{3}\left\{\mathrm{e}^{-\frac{(\alpha_m-\mu_{1m})^2}{2\sigma_{1m}^2}}\mathrm{e}^{-\frac{(\beta_m-\mu_{2m})^2}{2\sigma_{2m}^2}}\prod_{i\in D_m}\mathrm{e}^{-d_i(\alpha_m+\beta_mx_i)}[1+\mathrm{e}^{-(\alpha_m+\beta_mx_i)}]^{-1}\right\}.$$

当$\delta_i=0$时,

$$\pi(y_{1i}\mid\boldsymbol{\gamma},\boldsymbol{x},\boldsymbol{y},\boldsymbol{u}_{-1i},\boldsymbol{\delta})\propto\psi(y_{1i};\boldsymbol{\gamma},x_i)=\begin{cases}b(1,F(\alpha_1+\beta_1x_i)),i=1,2,\cdots,k_1;\\b(1,F(\alpha_2+\beta_2x_i)),i=k_1+1,\cdots,k_2;\\b(1,F(\alpha_3+\beta_3x_i)),i=k_2+1,\cdots,n.\end{cases}$$

式中,$\boldsymbol{u}_{-1i}=\{y_{1j}:j\neq i\}$.

下面求各参数的满条件分布. 简记α_m的满条件分布为$\pi(\alpha_m\mid\cdot)$.

$$\pi(\alpha_m\mid\cdot)\propto\mathrm{e}^{-\frac{(\alpha_m-\mu_{1m})^2}{2\sigma_{1m}^2}}\prod_{i\in D_m}\{\mathrm{e}^{-d_i\alpha_m}[1+\mathrm{e}^{-(\alpha_m+\beta_mx_i)}]^{-1}\}\propto\mathrm{e}^{-\frac{[\alpha_m-(\mu_{1m}-\sigma_{1m}^2s_m)]^2}{2\sigma_{1m}^2}}\prod_{i\in D_m}\{[1+\mathrm{e}^{-(\alpha_m+\beta_mx_i)}]^{-1}\}\leqslant$$

$$\mathrm{e}^{-\frac{[\alpha_m-(\mu_{1m}-\sigma_{1m}^2s_m)]^2}{2\sigma_{1m}^2}}\propto N(\mu_{1m}-\sigma_{1m}^2s_m,\sigma_{1m}^2),$$

式中,$s_m=\sum_{i\in D_m}d_i$. 可以利用筛选法随机产生α_m,具体步骤如下:

(1)由均匀分布 $U(0,1)$ 抽取 u,由 $N(\mu_{1m}-\sigma_{1m}^2 s_m,\sigma_{1m}^2)$ 抽取 z_m;

(2)如果 $u\leqslant\prod\limits_{i\in D_m}\{[1+\mathrm{e}^{-(z_m+\beta_m x_i)}]^{-1}\}$,则 $\alpha_m=z_m$,停止;否则转到步骤(1).

$$\pi(\beta_m\mid\cdot)\propto\mathrm{e}^{-\frac{(\beta_m-\mu_{2m})^2}{2\sigma_{2m}^2}}\prod_{i\in D_m}\{\mathrm{e}^{-d_i\beta_m x_i}[1+\mathrm{e}^{-(\alpha_m+\beta_m x_i)}]^{-1}\}\propto\mathrm{e}^{-\frac{[\beta_m-(\mu_{2m}-\sigma_{2m}^2 t_m)]^2}{2\sigma_{2m}^2}}\prod_{i\in D_m}\{[1+\mathrm{e}^{-(\alpha_m+\beta_m x_i)}]^{-1}\}\leqslant$$

$$\mathrm{e}^{-\frac{[\beta_m-(\mu_{2m}-\sigma_{2m}^2 t_m)]^2}{2\sigma_{2m}^2}}\propto N(\mu_{2m}-\sigma_{2m}^2 t_m,\sigma_{2m}^2),$$

式中,$t_m=\sum\limits_{i\in D_m}(d_i x_i)$. 可以利用筛选法随机产生 β_m,具体步骤如下:

(1)由均匀分布 $U(0,1)$ 抽取 u,由 $N(\mu_{2m}-\sigma_{2m}^2 t_m,\sigma_{2m}^2)$ 抽取 w_m;

(2)如果 $u\leqslant\prod\limits_{i\in D_m}\{[1+\mathrm{e}^{-(\alpha_m+w_m x_i)}]^{-1}\}$,则 $\beta_m=w_m$,停止;否则转到步骤(1).

$$\pi(k_1\mid\cdot)\propto\prod_{m=1}^{2}\Big(\prod_{i\in D_m}\{\mathrm{e}^{-d_i(\alpha_m+\beta_m x_i)}[1+\mathrm{e}^{-(\alpha_m+\beta_m x_i)}]^{-1}\}\Big),k_1=1,2,\cdots,k_2-1,$$

$$\pi(k_2\mid\cdot)\propto\prod_{m=2}^{3}\Big(\prod_{i\in D_m}\{\mathrm{e}^{-d_i(\alpha_m+\beta_m x_i)}[1+\mathrm{e}^{-(\alpha_m+\beta_m x_i)}]^{-1}\}\Big),k_2=k_1+1,\cdots,n.$$

$y_{1i},\alpha_1,\alpha_2,\alpha_3,\beta_1,\beta_2,\beta_3$ 都可以直接 Gibbs 抽样,但 k_1、k_2 的满条件分布比较复杂,我们利用 MCMC 方法中的 Metropolis-Hastings 算法对其抽样.

下面介绍 MCMC 方法的具体步骤.

在给出起始点 $\boldsymbol{\gamma}^{(0)}=(k_1^{(0)},k_2^{(0)},\alpha_1^{(0)},\alpha_2^{(0)},\alpha_3^{(0)},\beta_1^{(0)},\beta_2^{(0)},\beta_3^{(0)})$ 后,假定第 t 次迭代开始时的估计值为 $\boldsymbol{\gamma}^{(t-1)}$,则第 t 次迭代分为如下几步:

(1)$\delta_i=0$ 时,由分布 $\psi(y_{1i};\boldsymbol{\gamma}^{(t-1)},x_i)$ 随机产生 $y_{1i}^{(t)}$,令 $\boldsymbol{u}_1^{(t)}$ 表示 $y_{1i}^{(t)}$ 组成的向量;

(2)由 $\pi(\alpha_1\mid k_1^{(t-1)},k_2^{(t-1)},\alpha_2^{(t-1)},\alpha_3^{(t-1)},\beta_1^{(t-1)},\beta_2^{(t-1)},\beta_3^{(t-1)},\boldsymbol{x},\boldsymbol{y},\boldsymbol{u}_1^{(t)},\boldsymbol{\delta})$ 抽取 $\alpha_1^{(t)}$;

(3)由 $\pi(\alpha_2\mid k_1^{(t-1)},k_2^{(t-1)},\alpha_1^{(t)},\alpha_3^{(t-1)},\beta_1^{(t-1)},\beta_2^{(t-1)},\beta_3^{(t-1)},\boldsymbol{x},\boldsymbol{y},\boldsymbol{u}_1^{(t)},\boldsymbol{\delta})$ 抽取 $\alpha_2^{(t)}$;

(4)由 $\pi(\alpha_3\mid k_1^{(t-1)},k_2^{(t-1)},\alpha_1^{(t)},\alpha_2^{(t)},\beta_1^{(t-1)},\beta_2^{(t-1)},\beta_3^{(t-1)},\boldsymbol{x},\boldsymbol{y},\boldsymbol{u}_1^{(t)},\boldsymbol{\delta})$ 抽取 $\alpha_3^{(t)}$;

(5)由 $\pi(\beta_1\mid k_1^{(t-1)},k_2^{(t-1)},\alpha_1^{(t)},\alpha_2^{(t)},\alpha_3^{(t)},\beta_2^{(t-1)},\beta_3^{(t-1)},\boldsymbol{x},\boldsymbol{y},\boldsymbol{u}_1^{(t)},\boldsymbol{\delta})$ 抽取 $\beta_1^{(t)}$;

(6)由 $\pi(\beta_2\mid k_1^{(t-1)},k_2^{(t-1)},\alpha_1^{(t)},\alpha_2^{(t)},\alpha_3^{(t)},\beta_1^{(t)},\beta_3^{(t-1)},\boldsymbol{x},\boldsymbol{y},\boldsymbol{u}_1^{(t)},\boldsymbol{\delta})$ 抽取 $\beta_2^{(t)}$;

(7)由 $\pi(\beta_3\mid k_1^{(t-1)},k_2^{(t-1)},\alpha_1^{(t)},\alpha_2^{(t)},\alpha_3^{(t)},\beta_1^{(t)},\beta_2^{(t)},\boldsymbol{x},\boldsymbol{y},\boldsymbol{u}_1^{(t)},\boldsymbol{\delta})$ 抽取 $\beta_3^{(t)}$;

(8)$k_1^{(t)}\sim\pi(k_1\mid k_2^{(t-1)},\alpha_1^{(t)},\alpha_2^{(t)},\alpha_3^{(t)},\beta_1^{(t)},\beta_2^{(t)},\beta_3^{(t)},\boldsymbol{x},\boldsymbol{y},\boldsymbol{u}_1^{(t)},\boldsymbol{\delta})\triangleq\pi(k_1\mid\cdot)$,选建议分布 $q(k_1^{(t-1)},k_1')$ 为取值 $1,2,\cdots,k_2^{(t-1)}-1$ 的离散型均匀分布,即 $q(k_1^{(t-1)},k_1')=(k_2^{(t-1)}-1)^{-1}$,

令

$$\alpha(k_1^{(t-1)},k_1')=\min\left\{\frac{\pi(k_1'\mid\cdot)}{\pi(k_1^{(t-1)}\mid\cdot)},1\right\},$$

从 $1,2,\cdots,k_2^{(t-1)}-1$ 中任意抽取一个 k_1',然后产生一个随机数 u,若 $u\leqslant\alpha(k_1^{(t-1)},k_1')$,则 $k_1^{(t)}=k_1'$,否则 $k_1^{(t)}=k_1^{(t-1)}$;

(9)$k_2^{(t)}\sim\pi(k_2\mid k_1^{(t)},\alpha_1^{(t)},\alpha_2^{(t)},\alpha_3^{(t)},\beta_1^{(t)},\beta_2^{(t)},\beta_3^{(t)},\boldsymbol{x},\boldsymbol{y},\boldsymbol{u}_1^{(t)},\boldsymbol{\delta})\triangleq\pi(k_2\mid\cdot)$,选建议分布 $q(k_2^{(t-1)},k_2')$ 为取值 $k_1^{(t)}+1,\cdots,n$ 的离散型均匀分布,即 $q(k_2^{(t-1)},k_2')=(n-k_1^{(t)})^{-1}$,

令

$$\alpha(k_2^{(t-1)},k_2')=\min\left\{\frac{\pi(k_2'\mid\cdot)}{\pi(k_2^{(t-1)}\mid\cdot)},1\right\},$$

从 $k_1^{(t)}+1,\cdots,n-1$ 中任意抽取一个 k_2',然后产生一个随机数 u,若 $u\leqslant\alpha(k_2^{(t-1)},k_2')$,则 $k_2^{(t)}=k_2'$,否则 $k_2^{(t)}=k_2^{(t-1)}$. $\boldsymbol{\theta}^{(t)}=(k_1^{(t)},k_2^{(t)},\alpha_1^{(t)},\alpha_2^{(t)},\alpha_3^{(t)},\beta_1^{(t)},\beta_2^{(t)},\beta_3^{(t)})$ 称为 θ 的一个 Gibbs 样本.

假设 Gibbs 样本的容量为 M,并且第 B 次迭代以后抽样收敛,则把后 $M-B$ 个迭代值的算术平均作为参数的贝叶斯估计.

4 随机模拟

下面进行随机模拟试验.

取 $n=300$,$\boldsymbol{\gamma}=(k_1,k_2,\alpha_1,\alpha_2,\alpha_3,\beta_1,\beta_2,\beta_3)=(60,200,-4,6,3,2,-0.5,1.5)$. 首先确定 $x_1,x_2,\cdots,x_n$,然后选择大部分 x_i,根据 $y_i\sim b(1,p_i)$,$p_i=\varphi(x_i;\boldsymbol{\gamma})$ 随机产生 y_i,而剩余的一小部分 x_i 对应的 y_i 作为缺失数据. 则 x_i,y_i 即为随机模拟出来的观察数据,根据这些数据进行参数估计.

取 $\alpha_1,\alpha_2,\alpha_3$ 的先验分布分别为正态分布 $N(-4.2,0.5)$,$N(5.6,1.4)$,$N(3.5,0.1)$,取 β_1,β_2,β_3 的先验分布分别为正态分布 $N(1.6,0.4)$,$N(-0.4,1.2)$,$N(1.8,0.5)$.取 $B=10\ 000$,$M=20\ 000$. 参数估计结果见表 1.

表 1　各参数的贝叶斯估计

Table 1　Bayesian estimations of the parameters

参数	真值	均值	相对误差	MC 误差	2.5%分位数	中位数	97.5%分位数
k_1	60	59.996 00	0.000 07	0.007 49	59	60	62
k_2	200	201.082 30	0.005 41	0.018 01	197	201	205
α_1	−4	−4.076 37	0.019 09	0.002 35	−4.458 42	−4.080 54	−3.684 74
α_2	6	6.045 06	0.007 51	0.003 51	5.476 59	6.041 10	6.625 75
α_3	3	3.035 19	0.011 73	0.001 75	2.747 14	3.035 61	3.322 96
β_1	2	2.048 07	0.024 04	0.001 19	1.854 01	2.047 83	2.243 89
β_2	−0.5	−0.522 13	0.044 25	0.000 30	−0.571 73	−0.522 13	−0.472 79
β_3	1.5	1.548 14	0.032 10	0.000 90	1.400 30	1.547 05	1.694 93

k_1,k_2 的 Gibbs 抽样迭代过程见图 1 和图 2.

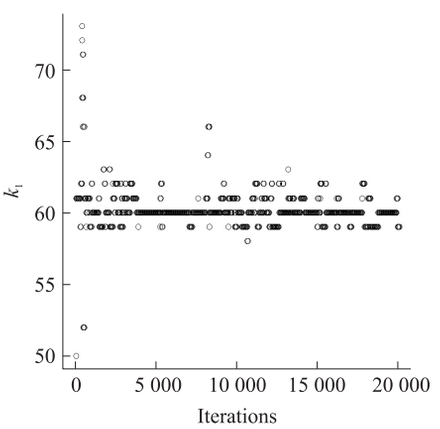

图 1　k_1 的 Gibbs 抽样迭代

Fig. 1　Gibbs sampling iterations of k_1

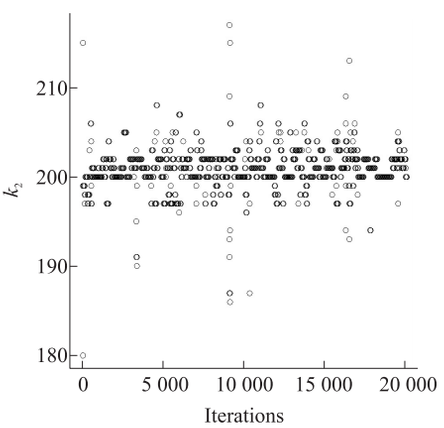

图 2　k_2 的 Gibbs 抽样迭代

Fig. 2　Gibbs sampling iterations of k_2

Gibbs 抽样收敛性诊断最常用的方法是同时产生多条马氏链,如果迭代过程中这几条马氏链逐渐稳定且趋于重合,则抽样收敛. 我们产生 2 条马氏链,k_1,k_2 的 2 条迭代链见图 3 和图 4.

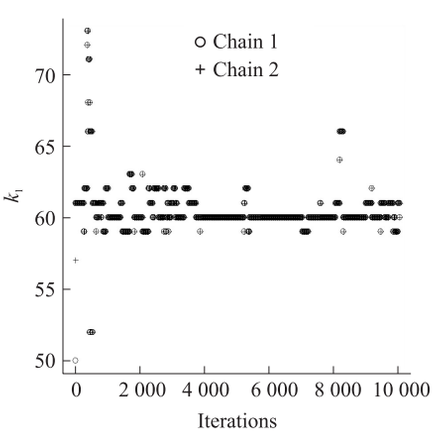

图 3　k_1 的两条迭代链

Fig. 3　Two iterative chains of k_1

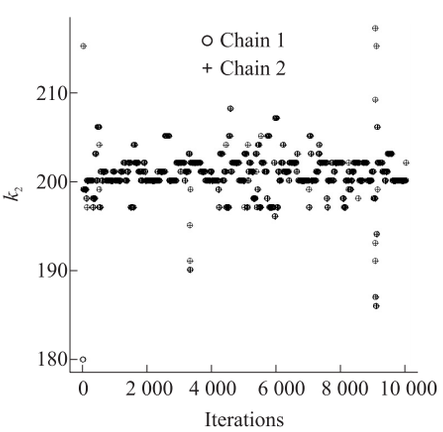

图 4　k_2 的两条迭代链

Fig. 4　Two iterative chains of k_2

由表 1 可以看出 k_1、k_2 的相对误差小于 1%,其他参数不超过 5%,估计精度较高,MC 误差都很小. 由图 1 和图 2 可以看出迭代的波动很小. 由图 3 和图 4 可以看出 k_1、k_2 的两条迭代链都趋于重合,所以抽样收敛. 综上分析,随机模拟试验的效果较好.

[参考文献]

[1] PAGE E S. Continuous inspection schemes[J]. Biometrika,1954,41(1):100-115.

[2] CHERNOFF H,ZACKS S. Estimating the current mean of a normal distribution which is subjected to changes in time[J]. The annals of mathematical statistics,1964,35(3):999-1 018.

[3] FEARNHEAD P. Exact and efficient Bayesian inference for multiple changepoint problems[J]. Statistics and computing,2006,16(2):203-213.

[4] YUAN T,KUO Y. Bayesian analysis of hazard rate,change point,and cost-optimal burn-in time for electronic devices[J]. IEEE transactions on reliability,2010,59(1):132-138.

[5] LIANG F M,WONG W H. Real-parameter evolutionary Monte Carlo with applications to Bayesian mixture models[J]. Journal of the American statistical association,2001,96(454):653-666.

[6] LAVIELLE M,LEBARBIER E. An application of MCMC methods for the multiple change-points problem[J]. Signal processing,2001,81(1):39-53.

[7] KIM J,CHEON S. Bayesian multiple change-point estimation with annealing stochastic approximation Monte Carlo[J]. Computational statistics,2010,25(2):215-239.

[8] POLSON N G,SCOTT J G,WINDLE J. Bayesian inference for logistic models using Pólya-Gamma latent variables[J]. Journal of the American statistical association,2013,108(504):1 339-1 349.

[9] HOSMER JR D W,LEMESHOW S,STURDIVANT R X. Applied logistic regression[M]. New York:John Wiley & Sons,2013.

[10] PEDUZZI P,CONCATO J,KEMPER E,et al. A simulation study of the number of events per variable in logistic regression analysis[J]. Journal of clinical epidemiology,1996,49(12):1 373-1 379.

[11] CHEN J,GUPTA A K. Parametric statistical change point analysis:with applications to genetics,medicine,and finance[M]. Berlin:Springer Science & Business Media,2011.

[12] GUREVICH G,VEXLER A. Change point problems in the model of logistic regression[J]. Journal of statistical planning and inference,2005,131(2):313-331.

[13] PASTOR B R,GUALLAR E,CORESH J. Transition models for change-point estimation in logistic regression[J]. Statistics in medicine,2003,22(7):1 141-1 162.

[14] MUGGEO V M R. Estimating regression models with unknown break-points[J]. Statistics in medicine,2003,22(19):3 055-3 071.

[责任编辑:陆炳新]