

文章编号:1007-6735(2013)03-0256-05

稳健的变量选择方法及其应用

樊亚莉¹, 徐群芳²

(1. 上海理工大学 理学院, 上海 200093; 2. 浙江农林大学 理学院, 临安 311300)

摘要: 在已有的变量选择方法和稳健估计方法的基础上, 提出了一种针对纵向数据的稳健的变量选择方法, 通过模拟衡量其稳健性, 并将其应用到一组实际的纵向数据分析中. 模拟和实例分析结果表明, 提出的稳健的变量选择方法在选择变量、估计变量系数的同时, 对数据中可能存在的异常值有明显的抵抗作用.

关键词: 惩罚函数; 估计方程; 稳健性; 变量选择
中图分类号: O 212.1 **文献标志码:** A

Robust Variable Selection Method and Its Application

FAN Ya-li¹, XU Qun-fang²

(1. College of Science, University of Shanghai for Science and Technology, Shanghai 200093, China;
2. College of Science, Zhejiang Agriculture and Forestry University, Linan 311300, China)

Abstract: Based on the variable selection methods and robust estimation methods presnted in literatures, a robust variable selection method was proposed for longitudinal data, and the robustness of the proposed method was evaluated by simulation study. The proposed method was then applied in the analysis of a real data set. It is found from the simulation and analysis of the real data that the proposed robust variable selection method can estimate correctly the regression parameters and select properly the important covariant variables, and it is also robust against outliers in the data set.

Key words: *penalty function; estimation equation; robustness; variable selection*

统计学的中心任务是对观测到的因变量和自变量进行建模, 从而对因变量在未来的取值进行预测和一些统计推断. 在现实问题的数据中往往会含有多个自变量, 如果建立的模型中含有多余的自变量, 不仅会引起过度拟合, 还会影响估计的精度和估计的效率^[1]. 因此, 如何选择出真正对因变量有影响的自变量是建模过程中值得关注的问题. 另一方面, 如果数据中存在异常值, 这些个别的异常值也会对模

型中参数的估计造成很大的影响^[2].

针对上述两个问题, 统计学家们作了大量的研究. 一方面, 对于变量选择问题, 统计学方法已经由古典的子集选择方法发展到近年来的系数压缩方法^[3]. 文献[4]提出的带有光滑截断绝对差的惩罚似然方法(SCAD)能够将变量选择和对回归系数的估计同时进行, 这大大缩减了传统的 AIC (Akaikes infotmation critetion)、BIC (Bayesian information

criterion)等方法的计算量和计算时间. 另外, 文献[5]提出的最小绝对压缩选择算子方法(LASSO)和文献[6]提出的硬截断惩罚方法(Hard), 以及之后文献[7]提出的弹性网方法(EN)和文献[8]的自适应的 LASSO 方法(ALASSO)都属于同一类型的变量选择方法. 这些方法的共同点是将在一个损失函数(如最小二乘损失、似然函数、估计方程等)加上一个惩罚项, 作为新的目标函数加以优化. 它们的区别在于所加的惩罚函数各有不同^[9]. 另一方面, 在稳健性问题上, 许多学者研究了异常点对估计的影响以及对异常点的诊断^[2,10]. 文献[11-12]等对稳健的估计方法进行了研究.

本文将上述两方面的工作结合起来, 提出了一个稳健的变量选择方法, 具体来说就是首先对估计方程稳健化, 用一个依赖于自变量的权重函数使得杠杆点对估计方程的影响降低, 并且对残差项使用一个有界的得分函数来限制因变量中异常点的影响; 然后在稳健化的估计方程中加上一个连续的惩罚函数项, 将稳健估计和变量选择同步进行. 同样考虑将变量选择方法稳健化的研究还有文献[13-14]. 文献[13]是基于稳健化的拟似然方程建立一系列检验统计量来进行变量选择, 文献[14]是基于一个加权二次预测风险函数将传统的 Mallows 的 C_p 准则(期望在 p 附近的准则)加以推广来进行变量选择. 但是, 这些工作都需要利用蒙特卡洛模拟来估计所拟合的模型的自由度, 计算量比较大. 本文的方法综合了稳健估计方程和系数压缩方法的优点, 计算方便. 模拟效果和实际应用效果显示, 本文方法能在变量选择和估计的同时有效抵抗数据中异常点的影响.

1 稳健的惩罚估计方程

考虑一组纵向数据, 包含 n 个个体, 每个个体有 m_i 次观测, $i = 1, 2, \dots, n$. 每一组观测包含一个因变量 y_{ij} 和一个自变量 x_{ij} , 这里 x_{ij} 为一个 d 维向量, 设 $\sum_{i=1}^n m_i = N$. 对这组数据 $\{(x_{ij}, y_{ij}), i = 1, 2, \dots, n; j = 1, 2, \dots, m_i\}$ 拟合一个边际线性模型

$$Y_i = X_i \beta + \varepsilon_i, \quad i = 1, 2, \dots, n \quad (1)$$

式中, $Y_i = (y_{i1}, \dots, y_{im_i})^T$; $X_i = (x_{i1}, \dots, x_{im_i})^T$; β 是 d 维回归系数向量; $\varepsilon_i = (\varepsilon_{i1}, \dots, \varepsilon_{im_i})^T$ 表示随机误差向量. 假设 ε_{ij} 的期望为 0, 方差为 σ^2 . 假设不同个体之间的观测是相互独立的.

受文献[11]中稳健估计方法的启发, 考虑稳健

的估计方程

$$U(\beta) = \sum_{i=1}^n X_i^T R_i^{-1}(\alpha) W_i \psi\left(\frac{Y_i - X_i \beta}{\sigma}\right) = 0 \quad (2)$$

式中, $R_i(\alpha)$ 为工作相关阵, 它依赖一个未知参数 α . 常见的工作相关阵包括等相关、一阶自回归相关等^[13,15]. 函数 $\psi(\cdot)$ 是一个有界得分函数, 用来限制因变量中异常点的影响, 本文选取 Huber 型函数 $\psi_c(x) = \min\{c, \max\{-c, x\}\}$, 这里 c 是一个平衡稳健性和效率之间的调节参数. c 越大, 估计的效率越高, 估计的稳健性就越差; 相反, c 越小, 估计的稳健性越好, 但效率越低. 权重矩阵 $W_i = \text{diag}(w_{i1}, \dots, w_{im_i})$ 用来降低自变量中杠杆点的影响. 这里, 类似于文献[12], 使用如下定义的权重

$$w_{ij} = w(x_{ij}) =$$

$$\min\left\{1, \left[\frac{b_0}{(x_{ij} - m_x)^T S_x^{-1} (x_{ij} - m_x)}\right]^{\frac{r}{2}}\right\} \quad (3)$$

式中, r 为大于 1 的常数; b_0 是自由度为 d 的卡方分布的 0.95 分位数; m_x, S_x 为 x_{ij} 的中心和离差的某种稳健估计, 本文取 m_x 为 x_{ij} 的中位数, 即 m_x 的第 k 个分量取为 $X = (X_1^T, \dots, X_n^T)^T$ 第 k 列的中位数; $d \times d$ 维的对角阵 S_x 的第 k 个对角元取为 $1.483(\text{median}|\mathbf{X}(k) - m_x(k) \otimes \mathbf{I}_N|)$, $\mathbf{X}(k)$ 表示 $\mathbf{X}$ 的第 k 列, $\otimes$ 表示 kronecker 乘积, $\mathbf{I}_N$ 表示 N 维元素全为 1 的列向量.

如果取 $\psi(x) = x$, $w_{ij} = 1$, 则式(2)退化为普通的估计方程. 文献[11-12]主要考虑了模型参数的稳健估计方法, 并没有考虑变量选择问题.

为了使变量选择和估计能同时进行, 现考虑惩罚估计方程

$$U^P(\beta) = \sum_{i=1}^n X_i^T R_i(\alpha)^{-1} W_i \psi\left(\frac{Y_i - X_i \beta}{\sigma}\right) - nq_\lambda(|\beta|) \text{sgn}(\beta) = 0 \quad (4)$$

式中, $q_\lambda(|\beta|) = (q_{\lambda 1}(|\beta_1|), \dots, q_{\lambda d}(|\beta_d|))^T$ 为某些常用惩罚函数的导数; $\text{sgn}(\beta)$ 表示 β 的符号; λ 为大于 0 的调节参数, 它的不同取值表示惩罚的程度不同.

现考虑几种近年来常用的一些惩罚函数. 第 1 种是 LASSO 惩罚函数^[5], 这是近年来广泛使用的一种惩罚函数. 第 2 种是硬截断 Hard 惩罚函数^[6], 这种惩罚函数来自于小波截断准则, 在某种程度上对应于最优子集选择. 第 3 种是 SCAD 惩罚函数^[4], 某种程度上 SCAD 可以看成是 LASSO 和 Hard 的综合. 第 4 种是弹性网惩罚函数, 记为 EN^[7]. 还有一种是自适应的 LASSO 惩罚, 记为 ALASSO^[8].

2 计算过程

在模型(1)中, β 的最终估计是方程(4)的解. 本文采用文献[14]提出的 MM(majorize-minorize)算法, 具体计算过程如下:

a. 对一个给定的 λ 值, 计算 β 的初始值 $\beta^{(0)}$. 本文取 $R_i(\alpha)$ 为单位阵, $\psi(x) = x$ 时估方程(2)的解作为初始值.

b. 用 $\beta^{(k)}$ 表示第 k 次迭代所得的解, $k \geq 0$, 则 $\beta^{(k+1)} = \beta^{(k)} + [D\beta^{(k)} + \Delta_\lambda(\beta^{(k)})]^{-1} U^P(\beta)$ 式中, D 为 d 阶导数阵 $\partial(U(\beta))/\partial\beta$.

$$\Delta_\lambda(\beta) = \text{diag}(q_\lambda(|\beta_1|)/(\epsilon + |\beta_1|), \dots, q_\lambda(|\beta_d|)/(\epsilon + |\beta_d|))$$

式中, ϵ 为一个很小的正数, 本文计算时取 $\epsilon = 10^{-6}$.

c. 对当前给定的 λ 值, 当 $\|\beta^{(k+1)} - \beta^{(k)}\| \leq 10^{-4}$ 时, 终止迭代, 记最后一次迭代的结果为 $\hat{\beta}_\lambda$.

d. 将 $\hat{\beta}_\lambda$ 带入 GCV 统计量来衡量当前模型和数据的拟合效果.

$$\text{GCV}(\lambda) = \frac{\text{RSS}_R(\lambda)/n}{\{1 - d(\lambda)/n\}^2}$$

式中, $\text{RSS}_R(\lambda)$ 是稳健化的残差和 $W(\psi(Y - X\hat{\beta}_\lambda/\sigma))^2$, $Y = (Y_1, \dots, Y_n)^T$, $X = (X_1^T, \dots, X_n^T)^T$, $W = \text{diag}(W_1, \dots, W_n)$, $d(\lambda) = \text{tr}[(D + \Delta_\lambda(\hat{\beta}_\lambda)^{-1})D]$.

e. 取 $\lambda = \arg \min_\lambda \text{GCV}(\lambda)$ 为最终用的 λ , 然后再进行第 a—e 步, 算出最终的 $\hat{\beta}$.

λ 的选择一般在一个事先规定的范围内. 本文在计算中让 λ 跑遍从 $\frac{1}{\sqrt{N}} \sim 4$ 的均匀间隔的 50 个数. λ 的范围一般没有统一规定, 要根据具体数据和模型来调节^[4-5].

本文主要求 β 的估计, 但计算过程中涉及到多余参数 α 和 σ 的估计问题. 本文对 σ 的估计采用和文献[11]类似的方法, 取

$$\hat{\sigma} = 1.483 \text{median} \left\{ \left| (y_{ij} - x_{ij}^T \beta^*) - \text{median}(y_{ij} - x_{ij}^T \beta^*) \right| \right\}$$

α 的估计依赖于工作相关阵 $R(\alpha)$ 的形式, 对各种常见的工作相关阵, 采取和文献[13]相同的估计方法.

3 数值模拟

从下列真实模型中产生数据:

$$y_{ij} = x_{ij} \beta_0 + \epsilon_{ij}, \quad i = 1, 2, \dots, n, \quad j = 1, 2, \dots, m$$

这里 $\beta_0 = (3, 1.5, 0, 0, 2, 0, 0, 0)^T$, $m = 5$, x_{ij} 从均值为 0 的多元正态分布中抽取, 其中, 第 k 个与第 l 个分量的相关系数取为 $0.5^{|l-k|}$, $\epsilon_i = (\epsilon_{i1}, \dots, \epsilon_{im})^T$ 服从 $N(0, \sigma^2 R_i(\alpha))$. 在模拟中, 将权重函数 w_{ij} 中的 r 取为 1, Huber 函数中的常数 c 取为 2.

为了考察本文所提方法的稳健性, 通过 3 种污染数据方式产生数据中的异常值. 污染 1 是随机挑选 1 个 x_{ij} 把它替换成 $x_{ij} + 5$. 污染 2 是随机挑选 1 个 x_{ij} 把它替换成 $x_{ij} + 5$, 并且同时将随机挑选的另外 1 个 y_{ij} 替换成 $y_{ij} + 10$. 污染 3 是将污染 2 中污染的个数变为 3 个.

衡量变量选择方法优劣的标准有 3 个. 第一个是模型误差 ME , 采用文献[4]中的定义方式.

$$ME = (\hat{\beta} - \beta_0)^T E(XX^T)(\hat{\beta} - \beta_0)$$

式中, $X = (X_1^T, \dots, X_n^T)^T$, $X_i = (x_{i1}, \dots, x_{im})^T$, 由于 X 均值为 0, 用 X 的样本协方差阵估计 $E(XX^T)$. 用 RME 表示所选模型的模型误差和最优模型(即在事先知道 β 中哪些分量为 0 的情况下, 通过式(2)对模型参数直接估计)的模型误差之比. 最后比较各种方法所得出 RME 的中位数, 记作 MRME, 该指标越小越好. 还用多次模拟中的平均选对比例(选对指把该是 0 的识别为 0, 不是 0 的识别为非 0), 记作 C-F, 以及平均正确压缩为 0 的系数个数(记为 C-S)来衡量各种方法变量选择的效果. C-F 最好的情况是 1(即百分百选对), C-S 最好的情况是 5(因为真实模型中有 5 个系数为 0). 取 $n = 50$, $\sigma = 1$, 工作相关阵取为真实的协方差阵(等相关, $\alpha = 0.5$). 表 1 给出了模拟结果. 其中, 后缀“-R”表示稳健方法, “-NR”表示非稳健方法, 非稳健方法是指在式(2)中取 $\psi(x) = x$, $w_{ij} = 1$.

在表 1 中, SCAD-R 表示在惩罚函数 SCAD 下稳健方法, SCAD-NR 表示在惩罚函数 SCAD 下非稳健方法, 其余符号的含义类似. 从表 1 可以看出, 在使用各种惩罚函数的情况下, 当数据集中含有异常值的时候, 通常的非稳健方法会受到强烈的影响, 尤其是使用 EN 惩罚函数的时候. 由于非稳健方法对方差估计的不稳定, 造成其在变量选择方面的波动. 相比之下, 稳健方法对异常值就有比较强的抵抗力, 引入权重函数和有界的得分函数能够大大降低异常值引起的模型误差, 尤其是在污染 3 的情况下. 同时, 稳健方法在数据存在异常值的情况下, 能够在一定程度上挽回完全选对的比例.

表 1 不同惩罚函数下稳健方法与非稳健方法表现的比较
Tab.1 Comparison between robust and non-robust methods under different penalty functions

方 法	污染 1			污染 2			污染 3		
	MRME	C-F	C-S	MRME	C-F	C-S	MRME	C-F	C-S
SCAD-R	2.12	0.52	4.18	2.52	0.60	4.45	2.58	0.70	4.53
SCAD-NR	5.00	0.42	3.34	9.05	0.54	4.28	21.20	0.65	4.53
HARD-R	2.12	0.95	4.95	2.08	0.96	4.96	2.10	0.96	4.96
HARD-NR	7.86	0.94	4.83	7.22	0.95	4.98	11.50	0.94	4.98
ALASSO-R	1.92	0.96	4.96	2.20	0.97	4.97	1.95	0.96	4.97
ALASSO-NR	5.26	0.25	3.80	9.96	0.22	3.86	22.20	0.40	4.00
LASSO-R	4.32	0.35	3.82	4.42	0.31	3.99	3.52	0.41	3.99
LASSO-NR	5.26	0.25	3.80	9.96	0.22	3.86	22.20	0.40	4.00
EN-R	4.59	0.28	3.29	4.88	0.23	3.39	3.94	0.21	3.64
EN-NR	21.20	0.20	3.15	23.60	0.17	2.86	23.40	0.16	3.55

4 实例分析

将本文提出的稳健变量选择方法应用到文献[16]的妇女黄体酮纵向数据集中,这批数据共收集了 34 位妇女在 1 个月经周期内的尿样,并检测黄体酮水平.将每位妇女的每次尿检看作 1 次观测,则这批数据共有 34 个个体,每个个体的观测次数从 11~28 不等,共收集了 $N = 492$ 次观测数据.将每次观测到黄体酮水平的对数作为因变量,自变量包括测量时间、被测者年龄和被测者身体质量指数.

用 Y_{ij} 表示第 i 个妇女的第 j 次观测因变量值, T_{ij} 表示观测时间, A_i 表示第 i 位妇女年龄, B_i 表示第 i 位妇女的身体质量指数.该组数据已有多位作者研究过.文献[17]对该组数据拟合了一个半参数混合效应模型,研究的结论是年龄 A 和身体质量指数 B 对因变量的影响不显著,而时间 T 对因变量的影响是显著的,而且时间 T 对因变量的影响是非线性的.文献[2]用半参数混合模型来检测该批数据的异常值.本文对该批数据拟合线性回归模型为

$$Y_{ij} = \beta_0 + \beta_1 A_i + \beta_2 B_i + \beta_3 T_{ij} + \beta_4 A_i B_i + \beta_5 A_i T_{ij} + \beta_6 B_i T_{ij} + \beta_7 T_{ij}^2 + \epsilon_{ij} \quad (5)$$

式中, ϵ_{ij} 为独立同分布的随机误差,同时服从 $N(0, \sigma^2)$ 分布.假定不同的个体内部观测值的相关结构相同,都取为一阶自回归相关结构,即 $R(\alpha)$ 的 (k, l) 元为 $\alpha^{|l-k|}$.

为了检测这批数据中的异常值,计算了每个观测值的权重,用 s_{ij} 表示第 i 个个体的第 j 次观测的权重, s_{ij} 按公式计算.

$$s_{ij} = w_{ij} \frac{\psi((Y_{ij} - X_{ij}\boldsymbol{\beta}^{(0)})/\hat{\sigma})}{(Y_{ij} - X_{ij}\boldsymbol{\beta}^{(0)})/\hat{\sigma}}$$

式中, $\boldsymbol{\beta}^{(0)}$ 为式(2)的解; $X_{ij} = (1, A_i, B_i, T_{ij}, A_i B_i, A_i T_{ij}, B_i T_{ij}, T_{ij}^2)$.

各个观测值的权重如图 1 所示.

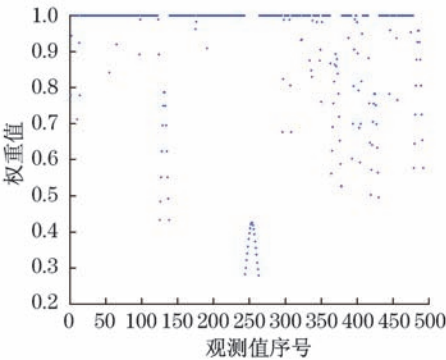


图 1 492 个观测值的权重
Fig.1 Weights of 492 observations

从图 1 可以看出,第 18 个个体观测值的权重较低,低于 0.4,尤其是第 244 个和第 263 个观测值的权重相对较低,低于 0.3.这表明第 18 个个体,尤其是第 244 个和第 263 个观测值是异常值.这个结果在某种程度上和文献[2]的结果一致.根据文献[2]的研究结果,第 18 个个体中的观测值是强影响点.

现将本文的稳健变量选择方法应用到该批数据,所得结果如表 2 所示(见下页).

表 2 中的数据表示模型(5)中回归系数的估计值,0 表示该系数被估计为 0,或者说模型不选择该系数所对应的自变量,认为它对因变量影响不显著.从表 2 可以看出,5 种惩罚函数对应的惩罚估计都认为模型中应该包含截距项、时间 T 和时间 T 的平方项,而认为年龄 A 和身体质量指数 B 以及交互效应都对因变量没有显著影响,不应包含在模型中.这个分析结果和文献[2,17]的分析结果类似.

为了考量本文方法的稳健性,进一步用交互验证的方法比较稳健和非稳健两种方法的表现.每次剔除 1 个个体的数据,用剩余的 33 个个体的数据估计式(5)中的回归系数,并且用平均均方误差 MSE

来衡量各个方法在交互验证过程的稳定性,定义

MSE = \frac{1}{34} \sum_{i=1}^{34} \|Y_i - X_i \hat{\beta}_{(-i)}\|^2

式中, \hat{\beta}_{(-i)} 表示去掉第 i 个个体之后回归系数的估计.

表 2 不同惩罚函数下黄体酮数据的分析结果
Tab.2 Results for progesterone data under different penalty functions

系数	SCAD	Hard	ALASSO	LASSO	EN
\beta_0(截距)	0.876 8	0.986 4	0.943 1	0.871 8	0.870 2
\beta_1(A)	0	0	0	0	0
\beta_2(B)	0	0	0	0	0
\beta_3(T)	0.480 4	0.560 5	0.568 5	0.497 7	0.493 6
\beta_4(A \cdot B)	0	0	0	0	0
\beta_5(A \cdot T)	0	0	0	0	0
\beta_6(B \cdot T)	0	0	0	0	0
\beta_7(T^2)	0.021 7	0.026 7	0.046 6	0.030 5	0.028 6

MSE 越小,表示该方法越稳健. 计算结果如表 3 所示.

表 3 中的数值是 MSE 的值,这里非稳健方法指的是计算的过程中取 \psi(x) = x, w_{ij} = 1. 从表 3 可以看出,非稳健方法的 MSE 比稳健方法对应的 MSE 要大,仅有 LASSO 方法除外. 因此,从大体上来说,稳健方法得到的结果比较稳定,分析结果不易受到异常值的影响.

表 3 稳健和非稳健方法的比较
Tab.3 Comparison between robust and non-robust methods

EN	SCAD	Hard	ALASSO	LASSO
3.083 2	3.009 4	3.190 1	3.017 6	3.110 4
3.110 7	3.051 5	3.377 1	3.025 9	3.105 8

5 结束语

将文献中已有的变量选择方法和稳健估计方法结合起来,提出了一类针对纵向数据线性回归模型的稳健变量选择方法,该方法先将通常的估计方程稳健化,然后结合一些常用的惩罚函数,构成稳健的惩罚估计方程,通过解该方程获得模型回归系数的估计. 加惩罚函数的作用是能将很小的回归系数估计为零,从而达到在估计的同时进行变量选择,选出真正对因变量有影响的自变量. 将估计方程稳健化的作用是当数据中含有异常点时,估计结果能保持相对稳定性. 将稳健的变量选择方法应用在模拟实

验和实际数据上的应用结果显示,当数据中含有异常点时,稳健方法比非稳健方法更加稳定.

参考文献:

[1] Miller A. Subset selection in regression[M]. London: Chapman and Hall,2002.

[2] Fung W K, Zhu Z Y, Wei B C, et al. Influence diagnostics and outlier tests for semiparametric mixed models[J]. Journal of Royal Statistical Society, Ser B, 2002,64(2):565-579.

[3] 王大荣,张忠占. 线性回归模型中变量选择方法综述[J]. 数理统计与管理,2010,29(4):615-627.

[4] Fan J,Li R. Variable selection via nonconcave penalized likelihood and its oracle properties[J]. Journal of the American Statistical Association,2001,96(456):1348-1360.

[5] Tibshirani R. Regression shrinkage and selection via the lasso [J]. Journal of Royal Statistical Society, 1996,58(1):267-288.

[6] Donoho D L. Ideal spatial adaptation by wavelet shrinkage[J]. Biometrika,1994,81(3):425-455.

[7] Zou H. Regularization and variable selection via the elastic net[J]. Journal of Royal Statistical Society, Ser B,2005,67(2):301-320.

[8] Zou H. The adaptive lasso and its oracle properties [J]. Journal of American Statistical Association,2006,101(475):1418-1429.

[9] Fan J. A selective overview of variable selection in high dimensional feature space[J]. Statistica Sinica, 2010,20(1):101-148.

[10] Preisser J S. Robust regression for clustered data with application to binary responses[J]. Biometrics, 1999,55(2):574-579.

[11] He X. Robust estimation in generalized partial linear models for clustered data [J]. Journal of American Statistical Association,2005,100(472):1176-1184.

[12] Sinha S K. Robust inference in generalized linear models for longitudinal data[J]. Canadian Journal of Statistics,2006,34(2):261-278.

[13] Liang K Y. Longitudinal data analysis using generalized linear models[J]. Biometrika,1986,73(1):13-22.

[14] Hunter D. Variable selection using MM algorithms[J]. Annals of Statistics,2005,33(4):1617-1642.

[15] 樊亚莉,王林平. AR(1)模型下 F 检验的性质[J]. 上海理工大学学报,2005,27(4):295-299.

[16] Sowers M F. Urinary ovarian and gonadotrophin hormone levels in premenopausal women with low bone mass[J]. J Bone Min Res,1998,13(7):1191-1202.

[17] Zhang D W. Semiparametric stochastic mixed models for longitudinal data [J]. Journal of American Statistical Association,1998,93(442):710-719.

(编辑:石 瑛)