

纵向数据下部分非线性模型的广义经验似然推断

肖燕婷, 孙晓青, 孙 瑾

(西安理工大学学院应用数学系, 陕西 西安 710054)

摘 要: 本文研究了纵向数据下部分非线性模型中未知参数的置信域的构造. 利用经验似然方法, 构造了非线性函数中未知参数的广义对数经验似然比统计量, 证明了其渐近于卡方分布. 同时, 得到了未知参数的最大经验似然估计, 并证明了其渐近正态性.

关键词: 纵向数据; 部分非线性模型; 经验似然; 置信域

MR(2010) 主题分类号: 62G05

中图分类号: O212.7

文献标识码: A

文章编号: 0255-7797(2016)06-1238-07

1 引言

纵向数据是生物医学、流行病学和计量经济学等学科中经常出现的一类复杂数据. 由于其具有个体间观测值独立, 个体内观测值相关的特性, 不同于一般的独立数据, 对其相关研究已成为统计学界的研究热点之一.

本文考虑如下的纵向数据下的部分非线性模型

$$Y_{ij} = g(X_{ij}, \beta) + m(T_{ij}) + e_{ij}, i = 1, \dots, n, j = 1, \dots, m_i, \quad (1.1)$$

其中观测数据来自 n 个个体, 第 i 个个体具有 m_i 次观测, 总的观测次数为 $N = \sum_{i=1}^n m_i$. Y_{ij} 和 $(X_{ij} \in R^p, T_{ij})$ 分别为第 i 个个体的第 j ($j = 1, \dots, m_i$) 次观测的响应变量和协变量; β 为 $d \times 1$ 维未知参数向量; $g(\cdot, \cdot)$ 为已知的可测函数; $m(\cdot)$ 为定义在 $[0, 1]$ 上的未知光滑函数; e_{ij} 为随机误差. 记第 i 个个体的随机误差向量为 $e_i = (e_{i1}, e_{i2}, \dots, e_{im_i})^T$, $\{e_i, i = 1, \dots, n\}$ 相互独立, 且 $E(e_i) = 0$, $\text{Var}(e_i) = \Sigma_i$ 并要求 Σ_i 为正定阵.

如果不考虑纵向数据, 部分非线性模型 (1.1) 在独立数据下, 已有不少学者研究了模型中未知参数和未知函数的估计, 如 Li 和 Nie^[1] 提出可以采用轮廓非线性最小二乘方法和局部线性逼近技术估计模型参数; Xiao 等^[2] 将经验似然方法应用于该模型, 构造了未知参数和未知函数的经验对数似然比统计量, 从而可以得到未知参数的置信域和未知函数的同时置信带.

对于复杂数据下的部分非线性模型, 冯三营等^[3] 考虑了当非参数协变量具有可加测量误差时, 采用逆卷积方法, 构造了模型中未知参数的经验对数似然比统计量, 并证明了其具有渐近卡方分布. 接着, 文献 [4] 给出了模型中回归系数, 光滑函数以及误差方差的最大经验似然估计. 肖燕婷等^[5] 借助核实数据, 给出了协变量带测量误差的部分非线性模型中未知参数的两种估计方法. 武大勇等^[6] 在响应变量随机缺失的情形下, 给出了未知参数的最大经验似

*收稿日期: 2015-10-23

接收日期: 2016-02-25

基金项目: 国家自然科学基金 (61303223; 11501443); 陕西省自然科学基金 (2015JM1039); 陕西省教育厅基金 (2016JK1545); 西安理工大学校基金 (2015CX009)

作者简介: 肖燕婷 (1981-), 女, 陕西西安, 讲师, 主要研究方向: 非 (半) 参数统计.

然估计, 并证明了估计的渐近正态性. 刘强^[7]在解释变量带有测量误差且响应变量随机缺失的复杂情形下, 利用核实数据, 给出了未知参数和非参数函数的两种估计. 但在纵向数据下, 对该模型的研究还比较少.

本文针对纵向数据下的部分非线性模型 (1.1), 利用 Owen^[8]等提出的经验似然方法, 构造了模型中未知参数 β 的广义对数经验似然比统计量, 证明了所提统计量的渐近 χ^2 分布, 所得结果可以构造未知参数的置信域. 同时, 利用对数经验似然比函数, 得到了未知参数的最大经验似然估计, 证明了其渐近正态性.

2 方法与主要结果

假设观测数据 $\{(X_{ij}, T_{ij}, Y_{ij}), i = 1, \dots, n, j = 1, \dots, m_i\}$ 由模型 (1.1) 产生. 对 (1.1) 式两边在给定 $T_{ij} = t$ 时求条件期望, 可以得到

$$m(t) = m_2(t) - m_1(t, \beta), \quad (2.1)$$

其中 $m_1(t, \beta) = E(g(X_{ij}, \beta)|T_{ij} = t)$, $m_2(t) = E(Y_{ij}|T_{ij} = t)$. 利用核估计方法, 可以得到 $m_1(t, \beta)$ 和 $m_2(t)$ 的估计, 分别由下式给出

$$\hat{m}_1(t, \beta) = \sum_{i=1}^n \sum_{j=1}^{m_i} W_{ij}(t) g(X_{ij}, \beta), \quad \hat{m}_2(t) = \sum_{i=1}^n \sum_{j=1}^{m_i} W_{ij}(t) Y_{ij}, \quad (2.2)$$

其中 $W_{ij}(t) = K_h(T_{ij} - t) / \sum_{k=1}^n \sum_{l=1}^{m_k} K_h(T_{kl} - t)$ 为核权函数, $K_h(\cdot) = K(\cdot/h)$ 且 $K(\cdot)$ 为核函数, h 为带宽. 因此 $m(t)$ 的估计为

$$\hat{m}(t) = \hat{m}_2(t) - \hat{m}_1(t, \beta). \quad (2.3)$$

为研究方便起见, 先给出一些符号说明. 令 $\tilde{Y}_{ij} = Y_{ij} - \sum_{k=1}^n \sum_{l=1}^{m_k} W_{kl}(T_{ij}) Y_{kl}$, $\tilde{g}_{ij}(\beta) = g(X_{ij}, \beta) - \sum_{k=1}^n \sum_{l=1}^{m_k} W_{kl}(T_{ij}) g(X_{kl}, \beta)$. 如果再令 $g^{(1)}(X_{ij}, \beta) = \frac{\partial g(X_{ij}, \beta)}{\partial \beta}$ 为非线性函数 $g(\cdot, \cdot)$

关于参数 β 的偏导数, 则有 $\tilde{g}_{ij}^{(1)}(\beta) = g^{(1)}(X_{ij}, \beta) - \sum_{k=1}^n \sum_{l=1}^{m_k} W_{kl}(T_{ij}) g^{(1)}(X_{kl}, \beta)$. 引入向量的记号, $\tilde{Y}_i = (\tilde{Y}_{i1}, \dots, \tilde{Y}_{im_i})^T$, $\tilde{g}_i(\beta) = (\tilde{g}_{i1}(\beta), \dots, \tilde{g}_{im_i}(\beta))^T$, $\tilde{D}_i(\beta) = (\tilde{g}_{i1}^{(1)}(\beta), \dots, \tilde{g}_{im_i}^{(1)}(\beta))^T$.

为了构造未知参数 β 的对数经验似然比函数, 需要引入如下的辅助随机向量

$$Z_i(\beta) = \tilde{D}_i^T(\beta) V_i^{-1} (\tilde{Y}_i - \tilde{g}_i(\beta)), \quad (2.4)$$

其中 V_i 可以为任意指定的工作协方差阵. 但在实际应用中, 为了避免由于错误指定协方差阵而引起的估计效率的降低, 可以用其估计值 $\hat{V}_i$ 代替真实值 V_i , 且 $\hat{V}_i = n^{-1} \sum_{i=1}^n \tilde{e}_i \tilde{e}_i^T$, 其中 $\tilde{e}_i = \tilde{Y}_i - \tilde{g}_i(\hat{\beta})$, $\hat{\beta}$ 是根据 $\frac{1}{n} \sum_{i=1}^n Z_i(\beta) = 0$, 且在 $Z_i(\beta)$ 中令 $V_i = I$ 得到的未知参数 β 的初估计.

类似于文献 [9], 定义如下广义的经验对数似然比函数

$$L(\beta) = -2 \max \left\{ \sum_{i=1}^n \log(np_i) | p_i \geq 0, \sum_{i=1}^n p_i = 1, \sum_{i=1}^n p_i Z_i(\beta) = 0 \right\}. \quad (2.5)$$

根据 Lagrange 乘子方法, $L(\beta)$ 可以写成

$$L(\beta) = 2 \sum_{i=1}^n \log[1 + \lambda^T Z_i(\beta)], \quad (2.6)$$

其中 $\lambda = \lambda(\beta)$ 由下式决定

$$\frac{1}{n} \sum_{i=1}^n \frac{Z_i(\beta)}{1 + \lambda^T Z_i(\beta)} = 0. \quad (2.7)$$

通过极小化 $L(\beta)$, 可以得到参数 β 的最大经验似然估计, 记做 $\tilde{\beta}$. 将 $\tilde{\beta}$ 代入 (2.3) 式, 可以得到未知函数 $m(t)$ 的最终估计, 为 $\hat{m}(t) = \hat{m}_2(t) - \hat{m}_1(t, \tilde{\beta})$.

3 渐近性质

令 $g^{(1)}(X_{ij}, \beta) = h(T_{ij}, \beta) + u_{ij}(\beta)$, $i = 1, \dots, n, j = 1, \dots, m_i$, 其中 $h(T_{ij}, \beta) = E(g^{(1)}(X_{ij}, \beta)|T_{ij})$, 且 $E(u_{ij}(\beta)) = 0$. 令 $h_a(T_{ij}, \beta)$ 为 $h(T_{ij}, \beta)$ 的第 a 个分量, $a = 1, \dots, d$. 类似于 (2.2) 式, $h_a(t, \beta)$ 的核估计可以定义为 $\hat{h}_a(t, \beta) = \sum_{i=1}^n \sum_{j=1}^{m_i} W_{ij}(t) g_a^{(1)}(X_{ij}, \beta)$, 其中 $g_a^{(1)}(X_{ij}, \beta)$ 为 $g^{(1)}(X_{ij}, \beta)$ 的第 a 个分量.

为了得到估计量的渐近性质, 需要给出以下假设条件.

C1: 带宽满足 $h = h_0 N^{-1/5}$, 对某个 $h_0 > 0$.

C2: 核函数 $K(\cdot)$ 为对称的概率密度函数, 且在它的支撑 $[-1, 1]$ 上有界.

C3: $\sup_{0 \leq t \leq 1} E(e_{ij}^4 | T_{ij} = t) < \infty, E(u_{ija}^4) < \infty$, 其中 u_{ija} 为 u_{ij} 的第 a 个分量, $a = 1, \dots, d$.

C4: T 的密度函数 $f(t)$ 在 $(0, 1)$ 是连续可微的, 且有界.

C5: 对任意的 β , 非线性函数 $g(x, \beta)$ 具有二阶连续导数.

C6: 函数 $m(t), h_a(t, \beta), a = 1, \dots, d$ 在 $(0, 1)$ 上是二次连续可微的.

C7: 定理 2.2 中的矩阵 Γ 为正定矩阵.

注 条件 C1–C7 是文献中经常用到的条件. 条件 C1 说明了在估计 $m(\cdot)$ 时不必欠光滑; C2 为核函数的一般性条件; 条件 C3–C6 为部分非线性模型中常见的条件; 条件 C7 保证了最大经验似然估计 $\tilde{\beta}$ 的渐近方差的存在.

定理 2.1 假设条件 C1–C6 成立, 如果 β 为真实参数值, 则有

$$L(\beta) \xrightarrow{d} \chi_d^2,$$

其中 $\xrightarrow{d}$ 表示依分布收敛, χ_d^2 表示自由度为 d 的卡方分布.

记 $\chi_d^2(\alpha)$ 为 χ_d^2 的 α ($0 < \alpha < 1$) 分位数. 利用定理 2.1, 可以得到参数 β 的近似 $1 - \alpha$ 置信域为

$$R_\alpha(\check{\beta}) = \{\check{\beta} | L(\check{\beta}) \leq \chi_d^2(\alpha)\}.$$

定理 2.2 假设条件 C1–C7 成立, 则有

$$\sqrt{n}(\tilde{\beta} - \beta) \xrightarrow{d} N(0, \Gamma^{-1} \Lambda \Gamma^{-1}),$$

其中 $\Gamma = \lim_{n \rightarrow \infty} n^{-1} \sum_{i=1}^n u_i^T V_i^{-1} u_i$, $\Lambda = \lim_{n \rightarrow \infty} n^{-1} \sum_{i=1}^n u_i^T V_i^{-1} \Sigma_i V_i^{-1} u_i$.

4 定理的证明

为了完成定理的证明, 需要首先给出以下引理.

引理 1 假设条件 C1–C6 成立, 则有

$$\begin{aligned} \sup_{a \leq t \leq b} \{|m_2(T_{ij}) - \hat{m}_2(T_{ij})|^2 | T_{ij} = t\} &= O((nh)^{-1} + h^4), \\ \sup_{a \leq t \leq b} \{|m_1(T_{ij}; \beta) - \hat{m}_1(T_{ij}; \beta)|^2 | T_{ij} = t\} &= O((nh)^{-1} + h^4), \\ \sup_{a \leq t \leq b} \{|h_a(T_{ij}; \beta) - \hat{h}_a(T_{ij}; \beta)|^2 | T_{ij} = t\} &= O((nh)^{-1} + h^4), \quad a = 1, \dots, d. \end{aligned}$$

证 仅给出以上第 1 个等式的证明, 其余两个等式可用同样的方法证明. 根据不等式 $(A + B)^2 \leq 2A^2 + 2B^2$ 可以得到

$$\begin{aligned} E\{|m_2(T_{ij}) - \hat{m}_2(T_{ij})|^2 | T_{ij} = t\} &\leq 2E\{|m_2(T_{ij}) - \sum_{k=1}^n \sum_{l=1}^{m_k} W_{kl}(T_{ij}) m_2(T_{kl})|^2 | T_{ij} = t\} \\ &\quad + 2E\{|\sum_{k=1}^n \sum_{l=1}^{m_k} W_{kl}(T_{ij}) (m_2(T_{kl}) - Y_{kl})|^2 | T_{ij} = t\} \\ &= I_1(t) + I_2(t). \end{aligned} \quad (4.1)$$

类似于文献 [10] 中引理 1 中的证明, 可以得到

$$I_1(t) \leq cn^{-1}h + ch^4, \quad (4.2)$$

基于这样的事实 $E[Y_{kl} - m_2(T_{kl})] = 0$, 由条件 C2–C4 可得

$$I_2(t) \leq c(nh)^{-1}. \quad (4.3)$$

结合 (4.1)–(4.3) 式, 该结论得证.

引理 2 假设条件 C1–C7 成立, 如果 β 为参数的真值, 则有

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n Z_i(\beta) \xrightarrow{d} N(0, \Lambda),$$

其中 Λ 的定义见定理 2.2.

证 记 σ_i^{kl} 为矩阵 V_i^{-1} 的第 (k, l) 元素, 通过简单计算有

$$\begin{aligned} \frac{1}{\sqrt{n}} \sum_{i=1}^n Z_i(\beta) &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \sum_{k=1}^{m_i} \sum_{l=1}^{m_i} u_{ik} \sigma_i^{kl} e_{il} + \frac{1}{\sqrt{n}} \sum_{i=1}^n \sum_{k=1}^{m_i} \sum_{l=1}^{m_i} \sigma_i^{kl} e_{il} \check{h}(T_{ik}, \beta) \\ &\quad + \frac{1}{\sqrt{n}} \sum_{i=1}^n \sum_{k=1}^{m_i} \sum_{l=1}^{m_i} \sigma_i^{kl} \check{m}(T_{il}) u_{ik} + \frac{1}{\sqrt{n}} \sum_{i=1}^n \sum_{k=1}^{m_i} \sum_{l=1}^{m_i} \sigma_i^{kl} \check{m}(T_{il}) \check{h}(T_{ik}, \beta) \\ &= \frac{1}{\sqrt{n}} \sum_{i=1}^n R_{1i} + \frac{1}{\sqrt{n}} \sum_{i=1}^n R_{2i} + \frac{1}{\sqrt{n}} \sum_{i=1}^n R_{3i} + \frac{1}{\sqrt{n}} \sum_{i=1}^n R_{4i} \\ &= A_1 + A_2 + A_3 + A_4, \end{aligned} \quad (4.4)$$

其中 $\check{h}(T_{ik}, \beta) = h(T_{ik}, \beta) - \hat{h}(T_{ik}, \beta)$, $\check{m}(T_{il}) = m(T_{il}) - \hat{m}(T_{il})$.

对 A_1 , 很容易得到 $E(A_1) = 0$ 和 $\text{cov}(A_1) = \Lambda$. 根据 Lindeberg-Feller 中心极限定理可以得到 $A_1 \xrightarrow{d} N(0, \Lambda)$.

根据引理 1 可以得到 $E\|A_2\|^2 \leq c\{(nh)^{-1} + h^4\} \rightarrow 0$, 因此得到 $A_2 \xrightarrow{p} 0$. 类似的还有 $E\|A_3\|^2 \leq c\{(nh)^{-1} + h^4\} \rightarrow 0$, 即 $A_3 \xrightarrow{p} 0$. 进一步, 根据引理 1 和 Cauchy-Schwarz 不等式, 可以得到 $E\|A_4\| \leq c\sqrt{n}\{(nh)^{-1} + h^4\} \rightarrow 0$, 由此推得 $A_4 \xrightarrow{p} 0$.

综合以上讨论, 这就完成了该引理的证明.

引理 3 假设条件 C1-C7 成立, 如果 β 为参数的真值, 则有

$$\frac{1}{n} \sum_{i=1}^n Z_i(\beta) Z_i^T(\beta) \xrightarrow{p} \Lambda.$$

证 仍然使用引理 2 中的记号, 并记 $J_i = R_{2i} + R_{3i} + R_{4i}$, 则有

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n Z_i(\beta) Z_i^T(\beta) &= \frac{1}{n} \sum_{i=1}^n R_{1i} R_{1i}^T + \frac{1}{n} \sum_{i=1}^n R_{1i} J_i^T + \frac{1}{n} \sum_{i=1}^n J_i R_{1i}^T + \frac{1}{n} \sum_{i=1}^n J_i J_i^T \\ &= B_1 + B_2 + B_3 + B_4. \end{aligned}$$

根据大数定律, 有 $B_1 \xrightarrow{p} \Lambda$. 令 $B_{2,st}$ 为 B_2 的第 (s, t) 元素, $R_{1i,s}$ 为 R_{1i} 的第 s 个元素, $J_{i,t}$ 为 J_i 的第 t 个元素, 则有 $|B_{2,st}| \leq (\frac{1}{n} \sum_{i=1}^n R_{1i,s}^2)^{1/2} (\frac{1}{n} \sum_{i=1}^n J_{i,t}^2)^{1/2}$, 这就证明了 $B_2 \xrightarrow{p} 0$. 同理还可以证明 $B_l \xrightarrow{p} 0, l = 3, 4$, 因此该引理得证.

定理 1 的证明 根据引理 2-3 和 Owen^[8] 的思想, 可以得到

$$|\lambda| = O_p(n^{-1/2}). \quad (4.5)$$

对 (2.6) 式作用 Taylor 展式, 并采用引理 2-3, 可以得到

$$L(\beta) = 2 \sum_{i=1}^n \left\{ \lambda^T Z_i(\beta) - \frac{1}{2} [\lambda^T Z_i(\beta)]^2 \right\} + o_p(1). \quad (4.6)$$

根据引理 2 和 (4.5) 式, (4.6) 式, 可以得到

$$\lambda = \left\{ \sum_{i=1}^n Z_i(\beta) Z_i^T(\beta) \right\}^{-1} \sum_{i=1}^n Z_i(\beta) + o_p(n^{-\frac{1}{2}}). \quad (4.7)$$

将 (4.7) 式代入 (4.6) 式, 可以得到

$$L(\beta) = \left[\frac{1}{\sqrt{n}} \sum_{i=1}^n Z_i(\beta) \right]^T \left[\frac{1}{n} \sum_{i=1}^n Z_i(\beta) Z_i^T(\beta) \right]^{-1} \left[\frac{1}{\sqrt{n}} \sum_{i=1}^n Z_i(\beta) \right] + o_p(1). \quad (4.8)$$

再结合 (4.8) 式, 引理 2-3, 该定理得证.

定理 2 的证明 注意到 $\tilde{\beta}$ 是 β 的最大经验似然估计, 也就意味着 $\frac{\partial L(\beta)}{\partial \beta} \big|_{\beta=\tilde{\beta}} = 0$. 根据 (2.6) 式和 (2.7) 式可以得到

$$\frac{\partial L(\beta)}{\partial \beta} \big|_{\beta=\tilde{\beta}} = \sum_{i=1}^n \frac{\frac{\partial \lambda^T}{\partial \beta} Z_i(\beta) + \left(\frac{\partial Z_i(\beta)}{\partial \beta} \right)^T \lambda}{1 + \lambda^T Z_i(\beta)} \big|_{\beta=\tilde{\beta}} = \sum_{i=1}^n \frac{\left(\frac{\partial Z_i(\beta)}{\partial \beta} \right)^T \lambda}{1 + \lambda^T Z_i(\beta)} \big|_{\beta=\tilde{\beta}} = 0.$$

因此 $\tilde{\beta}$ 和 $\tilde{\lambda} = \lambda(\tilde{\beta})$ 满足 $Q_{1n}(\tilde{\beta}, \tilde{\lambda}) = 0$ 和 $Q_{2n}(\tilde{\beta}, \tilde{\lambda}) = 0$, 其中

$$Q_{1n}(\beta, \lambda) = \frac{1}{n} \sum_{i=1}^n \frac{Z_i(\beta)}{1 + \lambda^T Z_i(\beta)}$$

和

$$Q_{2n}(\beta, \lambda) = \frac{1}{n} \sum_{i=1}^n \frac{(\frac{\partial Z_i(\beta)}{\partial \beta})^T \lambda}{1 + \lambda^T Z_i(\beta)}.$$

对 $Q_{1n}(\tilde{\beta}, \tilde{\lambda})$ 和 $Q_{2n}(\tilde{\beta}, \tilde{\lambda})$ 在点 $(\beta, 0)$ 作用 Taylor 展式, 可以得到

$$\begin{aligned} 0 &= Q_{1n}(\tilde{\beta}, \tilde{\lambda}) = Q_{1n}(\beta, 0) + \frac{\partial Q_{1n}(\beta, 0)}{\partial \beta}(\tilde{\beta} - \beta) + \frac{\partial Q_{1n}(\beta, 0)}{\partial \lambda^T}(\tilde{\lambda} - 0) + o_p(\delta_n), \\ 0 &= Q_{2n}(\tilde{\beta}, \tilde{\lambda}) = Q_{2n}(\beta, 0) + \frac{\partial Q_{2n}(\beta, 0)}{\partial \beta}(\tilde{\beta} - \beta) + \frac{\partial Q_{2n}(\beta, 0)}{\partial \lambda^T}(\tilde{\lambda} - 0) + o_p(\delta_n), \end{aligned}$$

其中 $\delta_n = \|\tilde{\beta} - \beta\| + \|\tilde{\lambda}\|$, 因此可以得到

$$\begin{pmatrix} \tilde{\lambda} \\ \tilde{\beta} - \beta \end{pmatrix} = S_n^{-1} \begin{pmatrix} -Q_{1n}(\beta, 0) + o_p(\delta_n) \\ o_p(\delta_n) \end{pmatrix},$$

$$\text{其中 } S_n = \begin{pmatrix} \frac{\partial Q_{1n}(\beta, \lambda)}{\partial \lambda^T} & \frac{\partial Q_{1n}(\beta, \lambda)}{\partial \beta} \\ \frac{\partial Q_{2n}(\beta, \lambda)}{\partial \lambda^T} & \frac{\partial Q_{2n}(\beta, \lambda)}{\partial \beta} \end{pmatrix}_{(\beta, 0)} = \begin{pmatrix} -\frac{1}{n} \sum_{i=1}^n Z_i(\beta) Z_i^T(\beta) & \frac{1}{n} \sum_{i=1}^n \frac{\partial Z_i(\beta)}{\partial \beta} \\ \frac{1}{n} \sum_{i=1}^n \frac{\partial Z_i(\beta)}{\partial \beta} & 0 \end{pmatrix}.$$

根据 (4.5) 式并注意到 $Q_{1n}(\beta, 0) = \frac{1}{n} \sum_{i=1}^n Z_i(\beta) = O_p(n^{-1/2})$ 可以得到 $\delta_n = o_p(n^{-1/2})$. 经过简单计算可以得到

$$\tilde{\beta} - \beta = \left(\frac{1}{n} \sum_{i=1}^n \frac{\partial Z_i(\beta)}{\partial \beta} \right)^{-1} \frac{1}{n} \sum_{i=1}^n Z_i(\beta) + o_p(n^{-1/2}). \quad (4.9)$$

根据引理 1 可以证明

$$\frac{1}{n} \sum_{i=1}^n \frac{\partial Z_i(\beta)}{\partial \beta} = \frac{1}{n} \sum_{i=1}^n \tilde{D}_i^T(\beta) V_i^{-1} \tilde{D}_i(\beta) \xrightarrow{d} \Gamma, \quad (4.10)$$

此式连同 (4.9) 式及引理 2 和 Slutsky 定理, 可以证得该定理.

参 考 文 献

- [1] Li R, Nie L. Efficient statistical inference procedures for partially nonlinear models and their applications[J]. Biometrics, 2008, 64(3): 904–911.
- [2] Xiao Y T, Tian Z, Li F X. Empirical likelihood-based inference for parameter and nonparametric function in partially nonlinear models[J]. J. Korean Stat. Soc., 2014, 43(4): 367–379.
- [3] 冯三营, 李高荣, 薛留根, 陈放. 非线性半参数 EV 模型的经验似然置信域 [J]. 高校应用数学学报, 2010, 25(1): 53–63.

- [4] 冯三营, 薛留根. 非线性半参数 EV 模型的最大经验似然估计 [J]. 数学物理学报, 2012, 32(4): 729–743.
- [5] 肖燕婷, 田铮, 孙瑾. 核实数据下非线性半参数 EV 模型的估计 [J]. 数学杂志, 2015, 35(5): 1075–1085.
- [6] 武大勇, 李锋. 随机缺失下半参数回归模型的最大经验似然估计 [J]. 山东大学学报, 2015, 50(4): 20–23.
- [7] 刘强. 缺失数据下非线性半参数 EV 模型的估计 [J]. 系统科学与数学, 2010, 30(9): 1236–1250.
- [8] Owen A. Empirical likelihood ratio confidence intervals for a single function[J]. Biometrika, 1988, 75(2): 237–249.
- [9] Li G R, Tian P, Xue L G. Generalized empirical likelihood inference in semiparametric regression model for longitudinal data[J]. Acta Math. Sinica, Engl. Ser., 2008, 24(12): 2029–2040.
- [10] 薛留根, 朱力行. 纵向数据下部分线性模型的经验似然推断 [J]. 中国科学, 2007, 37(1): 31–44.

GENERALIZED EMPIRICAL LIKELIHOOD INFERENCE FOR PARTIALLY NONLINEAR MODELS WITH LONGITUDINAL DATA

XIAO Yan-ting, SUN Xiao-qing, SUN Jin

(*Department of Applied Mathematics, Xi'an University of Technology, Xi'an 710054, China*)

Abstract: In this paper, we study the construction of confidence region for unknown parameter in partially nonlinear models with longitudinal data. By empirical likelihood method, the generalized empirical log-likelihood ratio for parameter in nonlinear function is proposed and shown to be asymptotically chi-square distribution. At the same time, the maximum empirical likelihood estimator of the parameter in nonlinear function is obtained and asymptotic normality is proved.

Keywords: longitudinal data; partially nonlinear models; empirical likelihood; confidence region

2010 MR Subject Classification: 62G05